Álgebra Linear Computacional

Alexandre Salles da Cunha
Ana Paula Couto da Silva

25 de setembro de 2025

Álgebra Linear Computacional

Alexandre Salles da Cunha
Ana Paula Couto da Silva

Sumário

Capítulo 1

Aritmética de Ponto Flutuante e Erros Numéricos

Este documento apresenta os principais conceitos de Álgebra Linear necessários para a disciplina DCC639 - Álgebra Linear Computacional (ALC), ofertada pelos Profs. Alexandre Salles da Cunha e Ana Paula Couto, para os cursos de graduação em Ciência da Computação, Sistemas de Informação e Matemática Computacional, da Universidade Federal de Minas Gerais.

O material que aqui apresentamos revisa alguns conceitos vistos na disciplina MAT038 - Geometria Analı́tica e Álgebra Linear, assim como introduz conceitos novos, indispensáveis para o curso de ALC que ministramos. Não temos a pretensão de subsitituir os excelentes livros textos [2, 4, 5, 1, 3] que nos ajudaram a produzir estas notas. Assumimos que o público que lê este documento tenha sido exposto ao conteúdo de MAT038.

1.1 Erros numéricos

Vamos iniciar nossos estudos de erros numéricos apresentando um exemplo do impacto do efeito de empregarmos aritmética de precisão finita em resultados numéricos avaliados pelo computador. Para tanto, vamos discutir o comportamento das funções $f_1(x)$ e $f_2(x)$ abaixo indicadas, quando $x \to 0$.

$$f_1(x) = \frac{e^x - e^{-2x}}{x}$$

$$f_2(x) = \frac{x - \sin(x)}{x^3}$$

Observe que nos dois casos, para avaliarmos o limite, precisamos levantar a indeterminação, usando a regra de L'Hôpital. Feito isso, verificamos que

$$\lim_{x \to 0} f_1(x) = 3$$

e

$$\lim_{x \to 0} f_2(x) = \frac{1}{6}.$$

Devemos esperar que, os valores retornados pelo computador ao avaliar estas funções, para valores de x cada vez mais próximos de zero, sejam compatíveis com os valores calculados para o limite. Então, vamos agora verificar numericamente o comportamento destas funções quando o argumento x tende a 0_+ (pela direita). Para isso, considere o seguinte trecho de código apresentado na Figura 1.1. Verifique os resultados para as duas funções quando x assume valores no intervalo de 10^{-1} a 10^{-18}; os valores são reduzidos por um fator de 10.

```
n = 18
x = np.zeros(n,dtype = 'float64')
f1 = np.zeros(n,dtype = 'float64')
f2 = np.zeros(n,dtype = 'float64')
for i in range(n):
    x[i] = np.power(0.1,i+1)
    f1[i] = (math.exp(x[i])-math.exp(-2.*x[i]))/x[i]
    f2[i] = (x[i]-math.sin(x[i]))/(np.power(x[i],3.0))
```

Figura 1.1: Trecho de código que implementa as funções $f_1(x)$ e $f_2(x)$ exatamente como são definidas.

Observe que, de início, quando os valores de x são reduzidos, os valores numéricos de $f_1(x)$ e $f_2(x)$ tendem ao valor do limite das duas funções. Porém, a partir de um determinado valor pequeno de x estes valores começam a oscilar e, à partir de $x = 10^{-17}$ e $x = 10^{-8}$, os valores retornados pelo procedimento são zero em ambos os casos, e em nada concordam com o valor esperado dado pelos limites, 3 e $\frac{1}{6}$. Por quê isso ocorre ? Há uma combinação de razões:

1. Nem todos os números reais são representáveis em um computador digital.

2. As operações aritméticas envolvendo representações ou aproximações de números reais estão sujeitas à erros numéricos.

3. e, finalmente, a combinação dos fatores acima faz com que diferenças muito pequenas entre grandezas não sejam corretamente representáveis (isso não significa que não consigamos representar números muito pequenos).

```
x  =
   0.1
   0.0100000
   0.0010000
   0.0001000
   0.0000100
   0.0000010
   0.0000001
   1.000D-08
   1.000D-09
   1.000D-10
   1.000D-11
   1.000D-12
   1.000D-13
   1.000D-14
   1.000D-15
   1.000D-16
   1.000D-17
   1.000D-18

f1  =

    2.8644016
    2.9851494
    2.9985015
    2.99985
    2.999985
    2.9999985
    2.9999998
    3.
    3.
    3.0000002
    3.0000002
    3.0000447
    2.9987124
    2.9976022
    3.1086245
    2.220446
    0.
    0.
```

Retornaremos ao estudo das funções $f_1(x), f_2(x)$ e de formas de implementá-las que reduzem os erros numéricos em breve. Antes disso, vamos formalizar algumas definições e um modelo de computação usando aritmética de ponto flutuante, visando compreender a razão dos desvios acima identificados.

```
f2  =
    0.1665834
    0.1666658
    0.1666667
    0.1666667
    0.1666673
    0.1666537
    0.1720536
    0.
    0.
    0.
    0.
    0.
    0.
    0.
    0.
    0.
    0.
    0.
```

1.2 Representação de números de ponto flutuante

Neste texto, empregamos o conceito de palavra de comprimento finito para fazer referência ao fato de que os números reais são armazenados em um conjunto finito de unidades básicas de informações, bits, que armazenam dois valores, 0 ou 1. Este número finito de unidades básicas limita os números que são representáveis com exatidão no computador. O modelo que discutimos a seguir não é limitado a unidades que armazenam apenas 0's ou 1's, mas qualquer conjunto discreto de valores no alfabeto $\{0, 1, \ldots, \beta - 1\}$, onde β é a base empregada.

Um computador digital que emprega palavra de comprimento finito não permite a representação de todos os números reais. Na medida em que operações aritméticas são realizadas com estas grandezas armazenadas, há acúmulo de erros numéricos, de arredontamento.

O conjunto dos números representáveis na máquina, aqui chamado de F, consiste no conjunto dos valores f que podem ser escritos na forma abaixo

$$
\begin{aligned}
& f = \pm .d_1 d_2 \ldots d_p \times \beta^e && (1.1)\\
& 0 \leq d_i < \beta && i = 1, \ldots, p\\
& 0 \neq d_i &&\\
& L \leq e \leq U &&
\end{aligned}
$$

em conjunto com o número zero.

O formato (1.1) é chamado de **representação de número de ponto flutuante normalizado.** Os valores $d_1, d_2, \ldots, d_p$ são chamados dígitos. Os dígitos devem pertencer ao alfabeto $\{0, 1, \ldots, \beta - 1\}$, onde β é um número inteiro que define a base empregada no computador. p é o número de dígitos do computador. O número $.d_1 d_2 \ldots d_p$ é fracionário e é chamado de **mantissa ou significado** de f. O expoente é e e deve pertencer a um interalo de valores inteiros L, U.

Diferentemente dos números reais, há limites superiores e inferiores para as magnitudes dos números representáveis com o modelo (1.1). Veja que para qualquer $f \in F$, $f \neq 0$, temos como estabelecer limites para a magnitude dos números representáveis,

$$m \leq |f| \leq M,$$

onde

$$m = \beta^{L-1}$$

$$M = \beta^{U}(1 - \beta^{-p}).$$

O fato de apenas um pequeno subconjunto dos nḿeros reais poderem ser representados com exatidão no modelo (1.1) tem efeitos importantes nos resultados de algoritmos numéricos, que envolvem grandezas operações de ponto flutuante.

Para um determinado valor real $a \in \mathbb{R}$, se $|a| < m$ ou $|a| > M$, temos *underflow* e *overflow* respectivamente. Nesse caso, a não pode ser aproximado por algum número em F, pois extrapola os limites das magnitudes dos números representáveis internamente. Normalmente esse não é um problema, pois uma mudança de escala na representação dos dados normalmente resolve a questão.

Por outro lado, se esse não é o caso (overflow ou underflow) e se não existe um $\overline{f}$ satisfazendo exatamente as condições definidas em (1.1) tal que $\overline{f} = a$, a representação de a é obtida por meio de seu arredondamento para algum número f que satisfaça estas condições, evitando sempre o arrendondamento para o valor zero. Em outras palavras $f = fl(a)$, f é a representação em ponto flutuante de a. A quantidade real $a - fl(a)$ é um erro de representação interna, erro de arredondamento.

Para mensurar os efeitos destes erros, vamos considerar o subconjunto G dos números reais definidos como

$$G = \{x \in \mathbb{R} : m \leq |x| \leq M\} \cup \{0\}$$

e o operador $fl : G \to F$ (lê-se *float*) que já introduzimos, mas que agora é formalmente definido como:

$$fl(x) = \left\{ \begin{array}{c} \text{número } c \in F \text{ mais próximo de } x \text{ com empates} \\ \text{sendo arbitrariamente tratados de forma a afastar } c \text{ de zero.} \end{array} \right\} \quad (1.2)$$

O operador $fl(\cdot)$ satisfaz a relação

$$fl(x) = x(1+\epsilon), \epsilon \leq \mathbf{u} = \frac{1}{2}\epsilon_m, \quad (1.3)$$

onde $\mathbf{u}$ é a chamado de *erro unitário de arredondamento* e ϵ_m é chamado de precisão da máquina.

Na verdade, se *op* representa uma das quatro operações básicas, e x, y são dois números reais, a relação (1.3) que é conhecida como axioma básico da aritmética de ponto flutuante, pode ser generalizado para

$$\frac{|fl(x \; op \; y) - (x \; op \; y)|}{|x \; op \; y|} \leq \mathbf{u}. \quad (1.4)$$

Note que esta relação apresenta um limite para a distância entre x e sua representação interna segundo o modelo (1.1). A precisão da máquina é um conceito fundamental cujo entendimento extrapola a utilidade das expressões (1.3) e (1.4) acima. O valor de $\mathbf{u}$ corresponde à metade distância entre 1 e o menor número positivo maior que 1, representável pela máquina (esta diferença é a precisão da máquina).

Para apresentar a precisão da máquina, propomos o algoritmo da Figura 1.2, que avalia ϵ_m. O algoritmo é executado em uma máquina que emprega palavra de 64 bits, seguindo o padrão IEEE754. Veja o resultado da execução do algoritmo na Figura 1.3.

Observe que, se todos os números reais fossem representáveis pela máquina e se as operações aritméticas do algoritmo da Figura 1.2 fossem realizadas com precisão infinita ou aritmética exata, o algoritmo jamais terminaria. Esse não é o caso e o algoritmo termina após um número finito de passos, pois em alguma iteração o teste

```
b != a
```

será falso e o laço não será executado. Isso ocorre pois iremos comparar $fl(1+u^{-k})$ e $fl(1)$. Quando u^{-p} for *suficientemente pequeno*, teremos $fl(fl(1+u^{-k})-fl(1)) = 0$ e não $fl(fl(1+u^{-k}) - fl(1)) = u^{-k}$ como poderíamos esperar. O valor $k-1$ imediamente anterior ao momento em que $fl(fl(1+u^{-k}) - fl(1)) = 0$ se observa define a precisão da máquina como $\epsilon_m = 2^{-(k-1)}$.

Assim sendo, informalmente, a precisão da máquina é a menor diferença representável entre dois números de ponto flutuante armazenados na máquina. Formalmente, sua definição é a seguinte: *a precisão da máquina é o menor número ϵ no formato 2^{1-p} para algum p inteiro positivo, tal que $fl(1+\epsilon) > fl(1)$.*

Para $\epsilon \leq \epsilon_m$, $fl(1+\epsilon) - fl(1) = 0$. Este fato tem efeitos catastróficos e explica o comportamento numérico que observamos para f_1 e f_2, qual seja, dos valores numéricos diferirem dos limites das funções quando $x \to 0$. Veja o exemplo da Figura 1.4 e verifique que o resutlado esperado para $z = 10^5 \times \epsilon_m \approx 10^{-11}$, porém o resultado obtido foi 0.

```
def PrecisaodaMaquina():
    a = 1.0
    u = 1.0
    b = a + u
    k = 0
    while (b != a):
        u = u / 2.0
        b = a + u
        k = k + 1
    r = a + u
    p = 2*u
    print('r = ',r,'k = ',k,'Precisao = ',"{:.16E}".format(p))
    return r,k,p
```

Figura 1.2: Algorimo que avalia a precisão da máquina

.

```
r,k,p = PrecisaodaMaquina()
r =  1.0 k =  53 Precisao =  2.2204460492503131E-16
```

Figura 1.3: Resultado da execução do algoritmo da Figura 1.2, para uma máquina que emprega 64 bits para representar expoente e mantissa, segundo o padrão IEEE754.

```
-->x = 10;
-->y = 10 + %eps;
->%eps
%eps  =
   2.220D-16
-->c = 1E5;
-->z = (x - y)*c
 z  =
   0.
```

Figura 1.4: Mais um exemplo dos efeitos de erros de arredonamento.

1.2.1 Como implementar funções do tipo $f_1(x)$, $f_2(x)$?

É nosso papel identificar os argumentos para os quais o resultado numérico de expressões matemáticas não deverá concordar em nada com o resultado da mesma função, avaliada com aritmética exata. Para casos como o que apresentamos aqui, precisamos encontrar meios de reescrever as funções, de forma que sua avaliação possa ser implementada de forma diferente (mas matematicamente equivalente) da definição da função.

Para ilustrar a ideia, vamos considerar a função $f_2(x)$ e vamos usar a expansão de $\sin(x)$ em série de potências. Exceto pelas operações aritméticas básicas, as demais funções matemáticas não são nativas no computador. Quando desejamos, por exemplo, avaliar $\sin(\frac{\pi}{4})$, a máquina recorre a algum procedimento implementado, disponível em alguma biblioteca que possa ser acessada, e que avalia $\sin(x)$ para $x = \frac{\pi}{4}$. O mesmo ocorre com outras funções como e^x, $\log(x)$, etc.

A expansão de $\sin(x)$ é dada por:

$$\begin{aligned}\sin(x) &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} \end{aligned} \tag{1.5}$$

Substituindo (1.5) na definição de $f_2(x) = \frac{x-\sin(x)}{x^3}$, reescrevemos

$$\begin{aligned}f_2(x) &= \frac{1}{x^3}\left(x - \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots\right)\right) \\ &= \frac{1}{6} - \sum_{n=2}^{\infty} (-1)^n \frac{x^{2(n-1)}}{(2n+1)!} \end{aligned} \tag{1.6}$$

Portanto, podemos usar uma aproximação para o valor da série (1.6) ao invés de $f_2(x)$ conforme foi definida na origem. Basta somarmos tantos quantos termos forem necessários em (1.6) para que alguma precisão para a quantidade a ser calculada seja obtida. Além disso, vamos precisar tomar cuidado com o cálculo da função fatorial. Veja o algoritmo apresentado na Figura (1.5) e os resultados obtidos.

Vamos dar mais um exemplo de formas alternativas de se avaliar funções sujeitas à condições extremas de erros numéricos. As duas grandes fontes de erros numéricos em operações de ponto flutuante são:

Fonte 1 subtração de quantidades de magnitude muito próxima

Fonte 2 soma de quantidades de magnitudes muito díspares.

```
def AvaliaF2(x,kmax,tol):
    v = 1.0/6.0
    k = 2
    convergiu = 0
    while (k <= kmax) & (convergiu == 0):
        f = math.factorial(2*k+1)
        termo = np.power(-1,k)*np.power(x,2*(k-1))/f
        v = v - termo
        if abs(termo) <= tol:
            convergiu = 1
        print('k = ',k,'v = ',"{:.10f}".format(v),\
               '|termo| = ',"{:.10E}".format(abs(termo)))
        k = k + 1
```

Figura 1.5: Algoritmo ilustrando a implementação da reformulação da função $f_2(x)$. kmax e tol são parâmetros que controlam o número máximo de termos no somatório e a precisão desejada.

Suponha que desejemos avaliar a menor raiz da quadrática $f(x) = x^2 - 2px - q$ para $p, q > 0$. Aplicando a fórmula de báscara temos que $p \pm \sqrt{p^2 + q}$ são as duas raízes de $f(x)$, a menor delas sendo dada por

$$x_{min} = p - \sqrt{p^2 + q}. \tag{1.7}$$

Verifique que se $q <<< p^2$ teremos um caso típico de subtração de quantidades muito próximas ao avaliar x_{min}, também chamado de *cancelamento catastrófico.*

Chamando $x_{max} = p + \sqrt{p^2 + q}$ como a maior das duas raízes, observe que $-q = x_{min}x_{max}$, e logo $x_{min} = \frac{-q}{x_{max}}$ que nos permite reescrever x_{min} de forma mais conveniente para avaliação numérica:

$$x_{min} = \frac{-q}{p + \sqrt{p^2 + q}}. \tag{1.8}$$

Considere os valores $p = 12345678, q = 1$ e o resultado apresentado na Figura 1.6 da avaliação de x_{min} pelas duas expressões matematicamente equivalentes (1.7) e (1.8) acima. Observe que o resultado obtido pela expressão (1.7), $-4.0978193283081055e-08$, possui poucos dígitos significativos que concordam com a resposta correta, $-4.0500003321000205e - 08$ (sujeita a menos erros numéricos), produzida pela expressão (1.8). Observe que com o uso de (1.7) não conseguimos garantir precisão da ordem de 0.01 na resposta de um cálculo bastante simples.

A mensagem desta seção é clara: a forma como lemos uma função não é neces-

```
def funcao1(p,q):
    valor = p - math.sqrt(p*p + q)
    print(valor)
    return valor
def funcao2(p,q):
    valor = - q / (p + math.sqrt(p*p + q))
    print(valor)
    return valor

p = 12345678
q = 1

valor1 = funcao1(p,q)
valor2 = funcao2(p,q)

-4.0978193283081055e-08
-4.0500003321000205e-08
```

Figura 1.6: Avaliação da menor raiz de $f(x) = x^2 - 2px - q$ por duas alternativas matematicamente equivalentes, mas não numericamente equivalentes, para $p = 12345678, q = 1$.

sariamente a forma como devemos implementaá-la para que seja avaliada numericamente.

Capítulo 2

Fundamentos de Álgebra Linear

2.1 Vetores, operações entre vetores e combinação linear

Um vetor é uma lista ordenada de, por exemplo, números reais ou números complexos. Um vetor v de dimensão n possui n entradas ou coordenadas e sempre será representado como uma coluna, $v \in \mathbb{R}^{n\times 1}$ (uma matriz especial com apenas uma coluna e n linhas). A $i-$ésima coordenada do vetor v é representada por v_i. O conjunto $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) é a coleção de todos os vetores n dimensionais que podem ser obtidos ao se atribuir valores em $\mathbb{R}$ (resp. $\mathbb{C}^n$) para suas n entradas. Mais do que um simples conjunto, $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) é um *espaço vetorial*, conjunto cuja definição será apresentada em breve.

As principais operações que envolvem vetores no $\mathbb{R}^n$ são:

1. Soma.
   A operação de soma associa a todo par de vetores u, v, um novo vetor $z \in \mathbb{R}^n$, tal que:
   $z = u + v = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$. Veja que para a subtração temos $z = u - v = u + (-v) = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} -v_1 \\ -v_2 \\ \vdots \\ -v_n \end{bmatrix} = \begin{bmatrix} u_1 - v_1 \\ u_2 - v_2 \\ \vdots \\ u_n - v_n \end{bmatrix}$.
   O elemento neutro ou origem ou vetor nulo é representado por:

$0_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, que satisfaz $u + 0_n = u$ para qualquer $u \in \mathbb{R}^n$.

2. Multiplicação por escalar.

   Vetores podem ser multiplicados por quaisquer números ou escalares. Se $\alpha \in \mathbb{R}$ é o escalar que multiplica o vetor, distribuímos o produto por cada uma das coordenadas do vetor: $\alpha u = \alpha \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} \alpha u_1 \\ \alpha u_2 \\ \vdots \\ \alpha u_n \end{bmatrix}$.

3. Dados dois vetores $u, v \in \mathbb{R}^n$, a quantidade $\sum_{i=1}^n v_i u_i$ é chamada de produto interno ou produto escalar de v por u e é representada por $v^T u = \langle v, u \rangle$. Em particular, o produto interno $\langle \cdot, \cdot \rangle$ induz a norma Euclideana: $\langle v, v \rangle = v^T v = \sum_{i=1}^n v_i^2 = \|v\|_2^2$ é o quadrado da norma Euclideana de v.

   Dois vetores u, v tais que $u^T v = 0$ são **ortogonais**, isto é, o ângulo formando entre eles é $\frac{\pi}{2}$ radianos.

As operações descritas são centrais em Álgebra Linear. Em particular, com as duas primeiras, definimos combinações lineares de vetores. Dados escalares $\alpha, \beta \in \mathbb{R}$ e dois vetores $v, u \in \mathbb{R}^n$, a soma $\alpha v + \beta u$ é chamada de **combinação linear** de v, u com pesos α, β, respectivamente. Podemos generalizar a ideia para mais vetores. Por exemplo, dados m escalares $\{\alpha_i \in \mathbb{R} : i = 1, \ldots, m\}$ e uma coleção de vetores $C = \{x^1, x^2, \cdots, x^m\}$, todos em $\mathbb{R}^n$, o vetor x dado por

$$x = \sum_{i=1}^m \alpha_i x^i$$

é uma combinação linear dos vetores de C. Veja que o resultado x depende dos elementos do conjunto C, pois x pode ser escrito em função dos elementos de C, empregando-se os pesos adequados, isto é, os valores $\alpha_i : i = 1, \ldots, m$ na combinação.

2.2 Espaços e Subespaços Vetoriais

Um conjunto de vetores $\mathcal{V} \subseteq \mathbb{R}^n$ define um **espaço vetorial** se for fechado para a soma e multiplicação por escalar. Mais precisamente, $\mathcal{V}$ é um espaço vetorial se e somente se as duas propriedades de fechamento seguintes forem satisfeitas:

1. Dados quaisquer $\alpha \in \mathbb{R}$ e $v \in \mathcal{V}$, $\alpha v \in \mathcal{V}$.

   Nesse caso, dizemos que $\mathcal{V}$ é fechado na multiplicação por escalar.

2. Dados quaisquer $v, u \in \mathcal{V}$, $u + v \in \mathcal{V}$.
   Nesse caso, dizemos que $\mathcal{V}$ é fechado na soma de seus elementos.

O conjunto $\mathbb{R}^n$ é um espaço vetorial pois atende às duas propriedades de fechamento acima relacionadas. Um subconjunto C de um espaço vetorial é um **subespaço vetorial** se ele próprio satisfizer as duas propriedades de fechamento acima.

Exemplo 1 *Verifique que o conjunto*

$$X = \{(x_1, x_2)^T \in \mathbb{R}^2 : 0 \leq x_i \leq 1, i = 1, 2\}$$

não define um espaço vetorial.
Para mostrar que não define basta mostrarmos que pelo menos uma das duas propriedades de fechamento não é satisfeita. Tomando $x = (1, 1)^T$ *e* $\alpha = 10$*, verificamos que* $\alpha x \notin X$*.*

Exemplo 2 *Verifique que o conjunto*

$$\mathcal{V} = \{x \in \mathbb{R}^2 : x = \beta(1, 1)^T \text{ para todo } \beta \in \mathbb{R}\}$$

define um espaço vetorial.
Observe que a definição de $\mathcal{V}$ *é: toda a coleção de vetores do* $\mathbb{R}^2$ *que podem ser obtidos como múltiplos do vetor* $(1, 1)^T$*. Para mostrar que define devemos mostrar que as duas propriedades acima são satisfeitas tomando-se escalares e pontos quaisquer.*

1. *Tome um* $v \in \mathcal{V}$ *e veja que* $\alpha v = \alpha\beta(1, 1)^T$ *para algum* β*. Como* $\alpha\beta \in \mathbb{R}$ $\alpha v \in \mathcal{V}$*. Logo é fechado na multiplicação por escalar.*

2. *Tome* $v, u \in \mathcal{V}$*. De forma análoga, existem* α, β *tais que* $v = \alpha(1, 1)^T, u = \beta(1, 1)^T$*. Então* $u + v = (\alpha + \beta)(1, 1)^T$*. Como* $\alpha + \beta \in \mathbb{R}$*,* $u + v \in \mathcal{V}$ *e o conjunto é fechado na soma.*

Dois conceitos relacionados à combinação linear são os de **independência linear** e **dependência linear**. Uma coleção de vetores $\{x^1, \ldots, x^m\}$ de um espaço vetorial $\mathcal{X}$ é linearmente independente (LI) se e somente se o **sistema linear homogêneo**

$$\sum_{i=1}^{m} \alpha_i x^i = 0$$

somente admite solução trivial, isto é, se apenas $\alpha_i = 0$ para todo $i = 1, \ldots, m$ resolve o sistema. Se, por outro lado, existem $\alpha_i : i = 1, \ldots, m$, nem todos nulos, tais que $\sum_{i=1}^{m} \alpha_i x^i = 0$, os vetores são denominados *linearmente dependentes* (LD).

Com as definições previamente introduzidas, podemos analisar o que ocorre quando combinamos vetores. Vamos considerar os vetores $w, v, z \in \mathbb{R}^3$ e os escalares $\alpha, \beta, \gamma \in \mathbb{R}$. Qual a representação geométrica resultante de:

1. Todos os possíveis vetores αw, obtidos atribuindo-se todos os valores possíveis de $\alpha \in \mathbb{R}$?

2. Todas as combinações $\alpha w + \beta v$, para todos os possíveis $\alpha, \beta \in \mathbb{R}$?

3. Idem para $\alpha w + \beta v + \gamma z$?

Para responder à estas questões, temos que verificar se os vetores utilizados na combinação são linearmente independentes. Primeiramente, vamos considerar que os três vetores são LI. Assim, temos que as combinações em (1) geram uma linha, as combinações em (2) um plano do $\mathbb{R}^3$ e em (3) todo o espaço $\mathbb{R}^3$. Se v e w são LD, por exemplo, as combinações em (2) também geram uma linha e em (3) passam a gerar um plano. Ou seja, a dimensão dos espaços gerados a partir das combinações lineares é definida após a verificação de independência linear do conjunto de vetores a ser analisado.

Mais formalmente, a partir da definição de independência linear, segue a definição da **dimensão do espaço vetorial** associado ao conjunto $C = \{x^1, \ldots, x^m\}$. A dimensão do espaço gerado pelo conjunto é a cardinalidade do maior subconjunto de C composto por elementos linearmente independentes.

Vamos assumir que o conjunto de vetores $C = \{x^1, \ldots, x^m\}$ seja composto por vetores LI. O subespaço linear associado a este conjunto, representado por

$$span(\{x^1, x^2, \ldots, x^m\}),$$

é definido da seguinte forma:

$$span(\{x^1, x^2, \ldots, x^m\}) = \left\{ x \in \mathbb{R}^n : x = \sum_{i=1}^{m} \alpha_i x^i, \forall \alpha_i \in \mathbb{R}, i = 1, \ldots, m \right\}.$$

Veja que na definição acima, todos os (infinitos) valores de $\alpha_i \in \mathbb{R} : i = 1, \ldots, m$ devem ser considerados para a definição do subespaço. Se C possui apenas um vetor, x^1, $span(\{x^1\})$ é uma linha. Se existem $\alpha_i : i = 1, \ldots, m$ tais que $y = \sum_{i=1}^{m} \alpha_i x^i$, dizemos que o vetor $y \in span(\{x^1, x^2, \ldots, x^m\})$. Veja: $y \in span(\{x^1, x^2, \ldots, x^m\})$ pois pode ser escrito como uma combinação linear dos vetores usados para se caracterizar o subsepaço linear.

A dimensão do subespaço *span*$(\{x^1, \ldots, x^m\})$ corresponde ao número de vetores linearmente independentes em $\{x^1, \ldots, x^m\}$. Como assumimos que C possui m vetores linearmente independentes, a dimensão de *span*$(\{x^1, \ldots, x^m\})$ é m. Isso significa que, com m vetores convenientemente escolhidos (ou seja, LI), somos capazes de caracterizar algebricamente qualquer outro vetor y no subsepaço.

Caso uma combinação linear x^{m+1} de $\{x^1, \ldots, x^m\}$ fosse adicionada ao conjunto C não mudaríamos o subespaço vetorial *span*$(\{x^1, \ldots, x^m\})$. Um conjunto de vetores $\{x^1, \ldots, x^m\}$ LI define uma base para o espaço vetorial que gera. O conjunto $\{x^1, \ldots, x^m, x^{m+1}\}$ onde x^{m+1} é combinação linear dos demais não é uma base para *span*$(\{x^1, \ldots, x^m\})$, visto que uma base necessita ser composta por vetores LI.

Exemplo 3 *Qual é a dimensão do subespaço vetorial gerado pelo conjunto de vetores* $C = \{(1,0,0,0)^T, (0,0,1,0)^T, (1,1,1,0)^T, (1,1,2,0)^T\}$ *? O vetor* $z = (2,3,1,0)^T$ *pertence a este subespaço vetorial ?*
O subespaço vetorial é subespaço do $\mathbb{R}^4$. *Os vetores* $(1,0,0,0)^T, (0,0,1,0)^T, (1,1,1,0)^T$ *são claramente LI. Entretanto, o vetor* $(1,1,2,0)^T$ *é a soma dos outros três. Portanto, dim(span(C))* $= 3$. *O conjunto* C *não forma uma base para span(C); já o conjunto* $C' = \{(1,0,0,0)^T, (0,0,1,0)^T, (1,1,1,0)^T\}$ *forma uma base. Em resumo,* $C \neq C'$*, porém span(C) = span(C').*

Seja $\mathcal{X}$ um espaço (ou subespaço vetorial $\mathcal{V}$). As seguintes propriedades são válidas para quaisquer vetores $u, v, w \in \mathcal{X}$ (ou $u, v, w \in \mathcal{V}$):

- Associatividade da adição: $u + (v + w) = (u + v) + w$
- Comutabilidade da adição: $u + v = v + u$
- Existência de um elemento nulo: $0 + v = v + 0 = v$
- Existência do inverso aditivo: para todo $v \in \mathcal{X}$ existe $-v \in \mathcal{X}$ tal que $v + (-v) = (-v) + v = 0$
- Propriedades da multiplicação por escalar: $\alpha(u + v) = \alpha u + \alpha v$, $(\alpha + \beta)u = \alpha u + \beta u$, $(\alpha\beta)u = \alpha(\beta u), 1u = u$.

2.3 Inversas de matrizes

Uma matriz A quadrada n dimensional cujas colunas (ou linhas) são LD é chamada de *singular*. Isso significa que não existe uma matriz A^{-1} chamada de inversa de A tal que $AA^{-1} = A^{-1}A = I_n$ (onde I_n é a matriz identidade de ordem n). O determinante, *det*(A), de uma matriz singular é zero.

Para uma matriz inversível, que admite inversa, vale $A^{-1}A = AA^{-1} = I_n$. Veja que se $Ax = b$, o papel da inversa é representar o operação $x = A^{-1}b$, em sentido oposto.

2.4 Normas vetoriais

Assim como empregamos o módulo de um número real (ou complexo) como uma função que expressa o quão grande o número é, empregamos normas vetoriais para obter informação similar para vetores em espaços vetoriais. Quando atribuímos uma norma aos elementos de um espaço vetorial, dizemos que temos um *espaço vertorial normado*.

Mais precisamente, uma norma em um espaço vetorial $\mathcal{X}$ é uma função real que associa a todo elemento $x \in \mathcal{X}$ um valor $\|x\|$ satisfazendo as seguintes três propriedades:

1. $\|x\| \geq 0$ para todo $x \in \mathcal{X}$ e $\|x\| = 0 \iff x = 0$.
2. $\|x + y\| \leq \|x\| + \|y\|$ para todo $x, y \in \mathcal{X}$.
3. $\|\alpha x\| = |\alpha| \|x\|$ para todo $\alpha \in \mathbb{R}, x \in \mathcal{X}$.

De uma forma geral, a função $\|x\|_p$ definida abaixo, para valores de p satisfazendo $1 \leq p \leq \infty$, define uma função que satisfaz aos critérios necessários para ser chamada de norma vetorial. É a chamada norma p do vetor x:

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}.$$

Alguns casos particulares da norma vetorial p merecem ser destacados. Quando $p = 2$, temos a norma Euclideana, tão empregada em Geometria Analítica. Quando $p = 1$, a norma representa a soma de valores absolutos das entradas do vetor e, quando $p = \infty$, a norma *infinito* é dada pelo módulo da maior coordenada do vetor em módulo, isto é:

$$\|x\|_\infty = \lim_{p \to \infty} \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} = \max\{|x_1|, \ldots, |x_n|\}.$$

As normas vetoriais (e matriciais que serão estudadas posteriormente) são instrumentos importantes para se caracterizar a convergência de algoritmos em Álgebra Linear Computacional, Otimização e Aprendizado de Máquina.

Uma pergunta recorrente quando se inicia o estudo de normas vertoriais e seu uso em algoritmos iterativos é qual norma deve ser empregada, ou seja, qual valor de p deve ser considerado. A aplicação considerada deve ditar qual é a norma mais adequada a ser considerada. O conhecimento do especialista no problema considerado deve ser levado em conta para se determinar o valor apropriado de p. Mas de qualquer forma, cabe mencionar que se um algoritmo iterativo produz uma sequência $\{x^k\}_{k=0}^{\infty}$ tal que $\lim_{k\to\infty}\|x^k\|_p = 0$, então $\lim_{k\to\infty}\|x^k\|_{p_1} = 0$ para qualquer outro valor $p_1 : 1 \leq p_1 \leq \infty$. Ou seja, se um algoritmo produz uma sequência de resultados que converge mediante uma norma, o mesmo pode ser dito para qualquer outra norma.

Posteriormente, ainda neste capítulo, estudaremos uma classe particular de matrizes simétricas, denominadas simétrias positivas definidas (SPD). Naquele momento, introduziremos uma classe adicional de normas, as chamadaqs *normas vetoriais induzidas por matrizes SPD.*

As normas vetoriais nos permitem caracterizar conjuntos de pontos *suficientemente próximos de outros.* Para tanto, vamos definir a *bola*

$$\mathcal{B}_p(y, r) = \{x \in \mathbb{R}^n : \|x - y\|_p \leq r\}$$

como o conjunto dos pontos do $\mathbb{R}^n$ que distam de $y \in \mathbb{R}^n$ não mais do que $r \in \mathbb{R}_+$. Os argumentos $y \in \mathbb{R}^n$ e $r \in \mathbb{R}_+$ são chamados de centro e raio da bola, respectivamente.

Na Figura 2.1, representamos $\mathcal{B}_p(0, 1)$, para três valores de p. Essas bolas unitárias do $\mathbb{R}^2$ são centradas na origem $y = (0, 0)^T$. Então, veja que para definir uma bola, precisamos de três argumentos: a norma p que define a distância considerada, o ponto de referência y em torno do qual as distâncias são consideradas e o valor do raio empregado r. Empregamos o termo raio mesmo quando $p \neq 2$, em analogia ao caso em que se considera a norma Euclideana.

À bola $\mathcal{B}_p(y, r)$, associamos dois subconjuntos disjuntos, a saber: sua fronteira

$$bd(\mathcal{B}_p(y, r)) = \mathcal{B}_p(y, r) \cap \{x \in \mathbb{R}^n : \|x - y\|_p = r\}$$

e seu interior

$$int(\mathcal{B}_p(y, r)) = \mathcal{B}_p(y, r) \cap \{x \in \mathbb{R}^n : \|x - y\|_p < r\}.$$

Veja que $\mathcal{B}_p(y, r) = bd(\mathcal{B}_p(y, r)) \cup int(\mathcal{B}_p(y, r))$.

2.5 Normas matriciais

Da mesma forma como empregamos normas vetorais para conferir uma noção de magnitude aos vetores, faremos algo semelhante para matrizes, definindo para elas

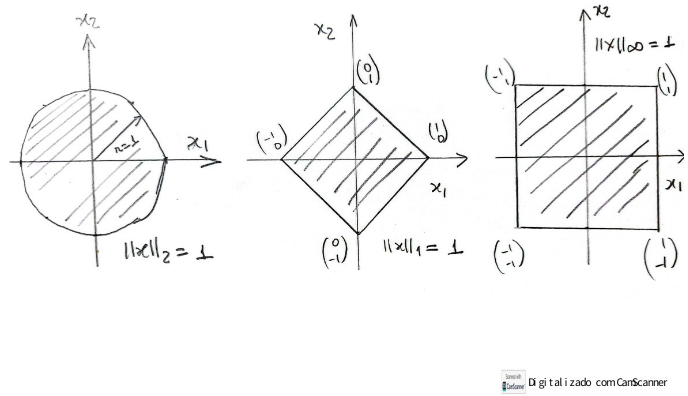

Figura 2.1: Bolas unitárias centradas na origem para $p = 2, 1, \infty$, nesta ordem, da esquerda para a direita.

normas apropriadas. Associaremos a elas, uma função chamada norma da matriz, que por um lado quantifica: (a) o quão grande a matriz é, e mais importante ainda, (b) em quanto a transformação linear que a matriz induz pode transformar a magnitude (ou melhor dizendo, a norma) do vetor de entrada. Ou seja, a norma matricial tem o poder de informar o quanto a matriz pode alterar a magnitude dos vetores sobre os quais a matriz é aplicada. Do ponto de vista algorítmico e de aplicações, esse conceito é fundamental. As normas matriciais também são de fundamental importância para a caracterização de sistemas lineares malcondicionados, propensos ao acúmulo de erros numéricos de grande monta.

Um primeiro aspecto que deve ser mencionado é que qualquer matriz $A \in \mathbb{R}^{m \times n}$ pode ser entendida como um vetor em $\mathbb{R}^{nm}$ cujas entradas são organizadas de forma diferente. Não é surpresa, portanto, que qualquer norma vetorial pode ser utilizada para esta representação vetorial de A, sendo capaz de conferir a ela uma noção de magnitude para a matriz. Apesar disso, é conveniente o uso de outras normas, ditas *normais matriciais*, em substituição à norma do vetor nm dimensional correspondente.

Mais formalmente, uma norma matricial é uma função que atribui para toda matriz $A \in \mathbb{R}^{m \times n}$ a grandeza $\|A\|$ (lê-se norma da matriz A) que satisfaz as seguintes propriedades:

1. $\|A\| \geq 0$ e $\|A\| = 0$ apenas se A é identicamente nula.

2. $\|\alpha A\| = |\alpha| \|A\|$ para qualquer $\alpha \in \mathbb{R}, A \in \mathbb{R}^{m \times n}$

3. $\|A + B\| \leq \|A\| + \|B\|$.

4. Para o caso em que $m = n$, deve valer também $\|AB\| \leq \|A\|\|B\|$, chamada propriedade de submultiplicatividade.

Algumas normas matriciais muito importantes são as normas 1 e ∞ abaixo definidas. Para a apresentação das mesmas, recordamos que A_j e a_j representam a $j-$ésima coluna e linha de A, respectivamente.

- $\|A\|_1 := \max_{j=1,\ldots,n}\|A_j\|_1$. Esta norma é chamada de **norma de máxima coluna**, pois todas as colunas de A são comparadas e a norma da coluna de maior norma 1 é a norma 1 da matriz.

- $\|A\|_\infty := \max_{i=1,\ldots,m}\|a_i^T\|_1$. De modo análogo, esta norma é chamada de **norma de máxima linha**, pois todas as linhas de A são comparadas e a norma ∞ de A é norma 1 de sua linha de maior norma 1.

Além das normas 1 e ∞ acima apresentadas, há uma terceira que é análoga à norma vetorial Euclideana, a norma de Frobenius, $\|\cdot\|_F$, definida como

$$\|A\|_F := \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{\frac{1}{2}}.$$

Além das três anteriormente citadas, uma quarta norma fundamental é a norma espectral, pois se relaciona ao espectro (conjunto de autovalores) de A ou de A^TA, dependendo da matriz ser simétrica ou não simétrica. A norma espectral de A, representada por $\|A\|_2$ não deve ser confundida com a norma de Frobenius de A, cuja expressão analítica é análoga à da norma Euclideana do vetor nm dimensional que contém as entradas de A empilhadas por linhas ou colunas, por exemplo. Veja a definição da norma espetral:

$$\|A\|_2 := \begin{cases} \lambda_{max}(A) & \text{se } A = A^T \\ \sigma_{max}(A) = \sqrt{\lambda_{max}(A^TA)} & \text{caso contrário} \end{cases}$$

Na definição acima, $\lambda_{max}(A)$ é o maior autovalor em módulo da matriz A. Quando a matriz é simétrica, seus autovalores são reais. Porém, podem ser negativos, caso a matriz não seja positiva semidefinida. Assim, os módulos dos autovalores devem ser considerados para determinação da norma da matriz simétrica.

Já a grandeza σ é chamada de valor singular de A. Assim sendo, σ_{max} é o maior valor singular da matriz em questão. O maior valor singular de A não simétrica é a raiz quadrada do maior autovalor de A^TA. Como a matriz A^TA é simétrica positiva semidefinida, A^TA sempre terá pelo menos um autovalor positivo (a não ser que seja identicamente nula) e, assim, $\sigma_{max}(A) > 0$. Sendo mais preciso, A^TA terá exatamente $posto(A)$ autovalores positivos e seus os demais $n - posto(A)$ autovalores serão nulos.

Exemplo 4 *Avaliar as normas* $1, \infty$ *de Frobenius e espectral de* A. *O comando* `spec` *do Scilab fornece os autovalores e autovetores de uma matriz. Veja que como a matriz* A *não é simétrica, aplicamos a definição de* $\|A\|_2$ *e calculamos os autovalores de* $A^T A$. *Com isso avaliamos que* $\|A\|_2 \approx 2.9209096$. *Os valores das normas* 1 *e* ∞, *4 e 3, respectivamente, podem ser avaliados por inspeção de suas colunas e linhas. Finalmente,* $\|A\|_F = \sqrt{1^2 + 2^2 + 2^2} = +3$.

```
A  =
   1.   2.
   0.   2.
--> [v,l] = spec(A'*A)
 v  =
  -0.9664996   0.2566679
   0.2566679   0.9664996
 l  =
   0.4688711   0.
   0.          8.5311289
--> sqrt(l)
 ans  =
   0.6847416   0.
   0.          2.9208096
```

- $\|A\|_1 = \max\{1, 4\} = 4$
- $\|A\|_\infty = \max\{3, 2\} = 3$
- $\|A\|_2 = 2.9208096$, *raiz quadrada do maior autovalor de* $A^T A$, *que é* 8.5311289.

2.5.1 Normas matriciais subordinadas e induzidas por normas vetoriais

Toda norma vetorial pode ser empregada para se definir uma norma matricial chamada **norma matricial induzida** por norma vetorial. Para uma norma vetorial $\|\cdot\|_v$, a norma matricial induzida por $\|\cdot\|_v$ é uma função de $\mathbb{R}^{m \times n}$ em $\mathbb{R}_+$, definida da seguinte forma:

$$\|A\|_M = \max_{x \neq 0} \frac{\|Ax\|_v}{\|x\|_v}. \tag{2.1}$$

Não é difícil mostrar que uma função definida como (2.1) satisfaz às quatro propriedades (a quarta se $n = m$) que definem funções de matrizes que podem ser chamadas de normas matriciais.

Tão importante quanto conhecer a definição é captar a potência da informação que a norma matricial induzida por norma vetorial revela. Por isso, vamos interpretar o significado da expressão dada por (2.1). Veja que, para um dado $x \neq 0$, a razão $\frac{\|Ax\|_v}{\|x\|_v}$ informa em quanto o *tamanho de* x, medido pela norma vetorial v, pode ser alterado (aumentado ou reduzido) quando a transformação linear Ax ocorre e sua magnitude é medida, também segundo a norma vetorial v.

Veja que a norma matricial induzida considera todos os possíveis vetores x distintos de zero por meio do operador max. Assim, ela de fato nos informa o quanto a matriz pode alterar a magnitude de um vetor de entrada x, medido pela norma v. Veja que ao considerar o denominador, a expressão relativiza a magnitude do vetor x ao qual é aplicada a transformação linear.

Na verdade, a expressão (2.1) pode ser simplificada, pois não precisamos considerar todos os vetores x distintos de zero. Para efeito do operador max, basta considerarmos todos os vetores x em $bd(\mathcal{B}_v(0,1))$, isto é, que tem norma v unitária. Esta simplificação pode ser feita pois qualquer vetor $x \in \mathbb{R}^n$ pode ser escrito como $x = \alpha y$ para algum $\alpha \neq 0$ e $y \in \mathbb{R}^n : \|y\|_v = 1$, de forma que $\|\alpha x\|_v = |\alpha| \|y\|_v$. Assim sendo, podemos reescrever:

$$\begin{aligned} \|A\|_M &= \max_{x \neq 0} \frac{\|Ax\|_v}{\|x\|_v} \\ &= \max_{x=\alpha y, \alpha \neq 0, \|y\|_v = 1} \frac{\|A(\alpha y)\|_v}{\|\alpha y\|_v} \\ &= \max_{\|y\|_v=1} \frac{|\alpha|}{|\alpha|} \frac{\|Ay\|_v}{\|y\|_v} \\ &= \max_{\|y\|_v=1} \frac{\|Ay\|_v}{\|y\|_v} \end{aligned} \tag{2.2}$$

Reintepretando agora o conceito da norma matricial de A induzida pela norma vetorial $\|\cdot\|_v$, temos que $\|A\|_v$, consiste na menor quantidade L para a qual a desigualdade seguinte vale para qualquer vetor $x \in \mathbb{R}^n, \|x\|_v = 1$:

$$\|Ax\|_v \leq L \|x\|_v.$$

Reinterpretando mais um pouco. Quando a norma matricial e a norma vetorial satisfazem a desigualdade acima e, além disso, quando existe um $x : \|x\|_v = 1$ que faz a desigualdade ser satisfeita de forma justa (na igualdade), a norma matricial é chamada de induzida pela norma vetorial. Quando a desigualdade é sempre satisfeita entre um par de norma matricial e de norma vetorial, a norma é chamada de **norma matricial subordinada** à norma vetorial. Toda norma matricial induzida por norma vetorial é também subordinada àquela norma vetorial. O inverso não é verdadeiro.

Um exemplo de subordinação que não é acompanhado por indução é dado pelo par: norma matricial Frobenius e norma Euclideana, com $v = 2$. Já a norma matricial espectral é induzida pela norma vetorial $v = 2$ (Euclideana). Veja: para caracterizar a indução é usada a igualdade em (2.1). Para caracterizar apenas a subordinação, podemos substituir a igualdade por $\leq$ e usar o valor $\|A\|_M$ da norma.

Exemplo 5 *Vamos ilustrar o conceito das normas induzidas, retornando à matriz do Exemplo 4. Considere o vetor* $v_2 = V(:,2)$*, dado pela segunda coluna de* V*, que é o autovetor de* $A^T A$ *associado ao seu maior autovalor* $\lambda_2 = 8.5311289$*. Veja o resultado de* Av_2 *calculado abaixo. Veja que a magnitude do vetor* $z = Av_2$*, medida pela norma Euclideana é* $\sqrt{8.5311289} = 2.9208096 = \sigma_{max}(A)$*. Assim sendo, a imagem* Ax *para qualquer vetor* x *de norma Euclideana unitária é limitada superiormente por* 2.9208096 *e existe um vetor,* v_2 *dado abaixo, que faz* $\|Av_2\|_2 = 2.9208096$ *(isto é a desigualdade é satisfeita na igualdade, de forma justa).*

```
 -->[V,l] = spec(A'*A)
 V  =
  -0.9664996   0.2566679
   0.2566679   0.9664996
 l  =
   0.4688711   0.
   0.          8.5311289
-->z = A*V(:,2)
 z  =
   2.1896672
   1.9329993
-->norm(z,2)
 ans  =
   2.9208096
-->u = z / norm(z,2)
 u  =
   0.7496782
   0.6618026
```

Os vetores v_1, v_2 *são chamados de vetores singulares à direita de* A*. Quando* v *é* *um vetor singular à direita de A, o resultado Av, após normalização, é chamado de vetor singular à esquerda de A. Veja que o vetor u calculado acima é o vetor singular à esquerda de A, associado a v_2. Esse assunto será explorado em detalhes quando estudarmos a fatoração SVD de matrizes.*

Exercício 2.5.1 *É sabido que a norma matricial* 1 *é induzida pela norma vetorial* 1 *e que a norma matricial* ∞ *é induzida pela norma vetorial* ∞*. Verifique que existem vetores x unitários nestas normas que fazem a igualdade* (2.1) *ser observada.*

2.6 Determinante

A toda matriz A quadrada de ordem $n \geq 1$, associa-se uma função $\det(A) : A \to \mathbb{R}$ denominada determinante da matriz A. Se a matriz $A = (a)$ possui ordem 1, isto é, se é um escalar, $\det(A) = a$. Para as demais matrizes, a expressão analítica da função determinante é recursiva, e é dada pela expressão da expansão do determinante de Laplace:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \tag{2.3}$$

onde o índice i é um índice qualquer das linhas de A, A_{ij} é a submatriz quadrada de ordem $n-1$ de A, obtida quando a $i-$ésima linha e a $j-$ésima coluna de A são removidas de A.

Observe que esta expressão faz uma expansão da função do determinante ao longo da linha i da matriz A, calculando recursivamente expressões para determinantes de submatrizes de A e, de suas submatrizes, até que o problema de calcular o determinante seja trivialmente resolvido, pois trata-se do determinante de uma matriz escalar (de ordem 1).

As seguintes propriedades são válidas para determinantes:

- $\det(A) = \det(A^T)$
- $\det(AB) = \det(A)\det(B)$ para $A, B \in \mathbb{R}^{n \times n}$
- $\det(\alpha A) = \alpha^n \det(A)$ para $\alpha \in \mathbb{R}, A \in \mathbb{R}^{n \times n}$.
- $\det(I) = 1$
- $\det(A)\det(A^{-1}) = 1$

A expressão de Laplace (2.3) para o cálculo do determinante tem muito mais valor teórico do que prático, pois o determinante de uma matriz é raramente calculado usando essa expressão. Via de regra, quando necessário, o determinante é calculado através de alguma fatoração da matriz.

Um resultado fundamental em Álgebra Linear é

$$\det(A) = 0 \iff A \text{ é singular} .$$

Logo, se $\det(A) = 0$, as linhas (e colunas) de A são LD. Nesse caso, A não admite inversa A^{-1}.

2.7 Transformações Lineares

Vamos recordar as operações de produto de uma matriz $A \in \mathbb{R}^{m \times n}$ por um vetor $x \in \mathbb{R}^n$ e interpretar o resultado. Na sequência, estaremos em condições de associar a toda matriz A quatro espaços fundamentais, de grande importância no estudo tanto de Álgebra Linear quanto de Álgebra Linear Computacional.

Considere os três vetores $u, v, w \in \mathbb{R}^3$ $u = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $v = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$, $w = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, e sua combinação linear $b = x_1 u + x_2 v + x_3 w$ para pesos $x_1, x_2, x_3 \in \mathbb{R}$. Veja que podemos sintetizar esta operação de combinação linear de u, v, w como um produto de uma matriz A, cujas colunas são os vetores u, v, w, por um vetor $x = (x_1, x_2, x_3)^T$ que encapsula os pesos na combinação linear desejada: $Ax = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = b$. Em sentido oposto, podemos interpretar o vetor b como o resultado da combinação linear das colunas de A com pesos dados pelas entradas de x (para que a operação seja conformável, o número de colunas de A e de linhas de x devem ser idênticas). Veja que o resultado b pode ser entendido como um vetor em $span(\{A_1, \ldots, A_n\})$ onde A_i representa a $i-$ésima coluna de A.

O espaço $span(\{A_1, \ldots, A_n\})$ é um dos quatro espaços fundamentais de A, chamado espaço coluna de A, aqui representado por $C(A)$. Sabemos que $b \in span(\{A_1, \ldots, A_n\}) = C(A)$ pois existe um vetor x que permite escrever $b = Ax$. Em outras palavras, x é o certificado de que $b \in C(A)$: o sistema linear $Ax = b$ é consistente, admitindo solução. Guarde isso: quando $Ax = b$, x é o certificado de pertinência de b em $C(A)$.

Usando os conceitos de normas vetoriais, vamos verificar qual é o efeito de aplicarmos a matriz A em todos os vetores do $\mathbb{R}^n$ que possuem norma $p = 2$ unitárias. Ou seja, vamos investigar o efeito da **transformação linear** $\sum_{i=1}^{n} A_i x_i$ para vetores

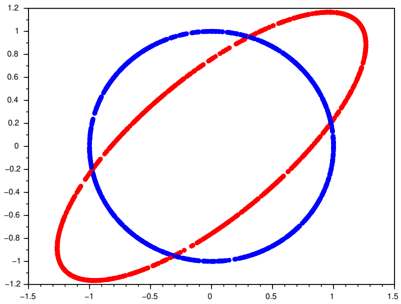

Figura 2.2: Imagem da transformação linear Ax para x no disco unitário na norma $p = 2$.

$x = (x_1, \ldots, x_n)^T$ tais que $\|x\|_2 = 1$.

Visando simplificar a representação geométrica, mas sem comprometer a generalidade da apresentação, vamos considerar o caso $n = 2$ e a matriz $A = \begin{bmatrix} 1.2 & 0.4 \\ 0.6 & 1 \end{bmatrix}$. A imagem do círculo unitário mediante a transformação Ax (isto é a imagem de $bd(\mathcal{B}_2(0, 1))$) é apresentada na Figura 2.2. O domínio considerado é indicado em azul na figura, enquanto a imagem da função Ax é indicada em vermelho. Veja que ao aplicarmos A em x de norma Euclideana unitária, deformamos o círculo, transformando-o em uma elipse.

O conjunto de pontos indicados em vermelho na Figura 2.2 define uma elipse ou uma *hiper-elipse* quando $n > 2$:

A imagem da bola unitária mediante a transformação linear Ax é uma elipse.

Essa elipse possui $n = 2$ eixos principais, pois a matriz A considerada no nosso exemplo possui duas colunas linearmente independentes. Caso possuísse uma coluna linearmente dependente das demais (no caso $n \geq 2$), pelo menos um dos eixos da elipse seriam degenerados, isto é, deixariam de existir e a elipse perderia pelo menos uma dimensão: seria um sólido com dimensão inferior a n. A hiperelipse n dimensional colapsaria em um sólido de dimensão inferior a n.

Ao aplicarmos A em alguns vetores x particulares (chamados de vetores singulares à direita de A) de norma unitária obtemos outros vetores, distintos do vetor identicamente nulo, cujas normas são os chamados valores singulares da matriz A.

Retornando ao caso da matriz A considerada em nosso exemplo, os eixos da elipse são associados a dois vetores u^1, u^2, obtidos com a fatoração SVD da matriz A. Esse assunto será examinado cuidadosamente ao longo do curso. Nessa seção, apresentamos os vetores e valores singulares sem nos preocuparmos como são calculados e porque existem.

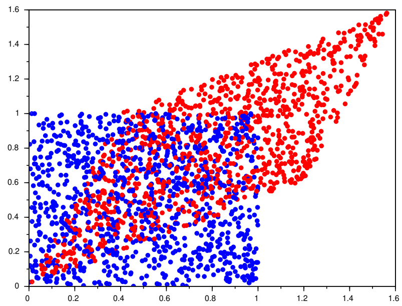

Figura 2.3: Imagem da transformação linear Ax para x pontos aleatoriamente escolhidos no conjunto $\{x \in \mathbb{R}^2 : x \in [0,1]^2\}$.

Verifique usando o `scilab` que se aplicarmos A no vetor:

- $v^1 = (-0.8012766 \quad -0.5982941)^T$ de norma Euclideanda unitária, obtemos o vetor $\tilde{u}^1 = (-1.2008495 \quad -1.0790601)^T$ que pode ser escrito como $\tilde{u}^1 = 6.7933741u^1$ para $u^1 = (-0.7438189 \quad -0.6683812)^T$, onde $\|u^1\|_2 = 1$. O valor $\sigma_1 = 6.7933741$ é chamado de primeiro valor singular da matriz A.

- $v^2 = (0.5982941 \quad -0.8012766)^T$ também de norma Euclideana unitária, obtemos o vetor $\tilde{u}^2 = (0.3974423 \quad -0.4423001)^T$ que pode ser escrito como $\tilde{u}^2 = 0.5946342u^1$ para $u^2 = (0.6683812 \quad -0.7438189)^T$, onde $\|u^2\|_2 = 1$. O valor $\sigma_2 = 0.5946342$ é chamado de segundo valor singular da matriz A. Note que $\sigma_1 \geq \sigma_2$.

Os vetores u^1, u^2 (assim como seus simétricos $-u^1, -u^2$) definem os eixos principais da elipse e os chamados *vetores singulares a esquerda de* A. Veja que a transformação linear Ax da bola unitária resulta em vetores com normas Euclideanas que pertencem ao invervalo $[\sigma_2, \sigma_1]$. Se a matriz tivesse colunas linearemente dependentes, as transformações lineares resultariam em vetores com normas Euclideanas no conjunto $\{0\} \cup [\sigma_{min}, \sigma_1]$ onde $\sigma_{min} > 0$ é o menor valor singular da matriz A.

Complementando nosso estudo de transformações lineares associadas à matriz A definida acima, vamos considerar a transformação do quadrado $\{x \in [0,1]^2\}$, indicada na Figura 2.3. Para tanto, aplicamos a matriz A em alguns pontos aleatoriamente escolhidos no quadrado. Observe que a imagem desse quadrado é um losango. Os vértices do losango consistem no resultado Ax para pontos x que são os vértices do quadrado considerado. A área do losango é o módulo do determinante de A, $|a_{11}a_{22} - a_{21}a_{12}| = |\det(A)|$.

A matriz A considerada nas Figuras 2.2 e 2.3 é não singular, tendo determinante distinto de zero. Vamos agora considerar uma matriz singular, a matriz $B = \begin{bmatrix} 1.5 & 0.8 \\ -0.75 & -0.4 \end{bmatrix}$. Verificamos que as colunas de B são linearmente dependentes, logo seu determinante é nulo. Na Figura 2.4, apresentamos a transformação linear Bx para pontos x aleatoriamente escolhidos no quadrado $\{x \in \mathbb{R}^2 : x \in [0,1]^2\}$. Veja que a imagem da transformação linear é a linha $span(B_1)$. A dimensão deste subespaço é 1 e não dois pois qualquer combinação linear das colunas de B resulta em algum vetor que pode ser escrito como um escalar α por B_1. Não há como representar outros pontos do $\mathbb{R}^2$ usando apenas as colunas de B. Diferentemente do exemplo associado à matriz A da Figura 2.2, $C(B) \neq \mathbb{R}^2$. Assim, nesse caso, deve haver algum outro subespaço do $\mathbb{R}^2$ que, em conjunto com $C(B)$, permita escrever qualquer ponto $z \in \mathbb{R}^2$. Veja que a área obtida pela imagem dos vértices do quadrado é zero, pois este losango foi degenerado em um segmento de reta.

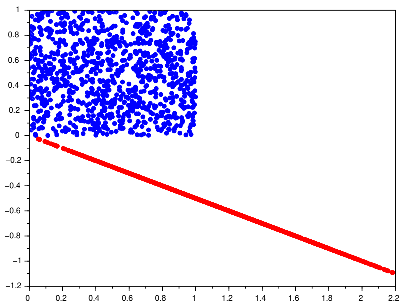

Figura 2.4: Imagem da transformação linear Bx para x pontos aleatoriamente escolhidos no conjunto $\{x \in \mathbb{R}^2 : x \in [0,1]^2\}$.

2.8 Visões complementares sobre representação de matrizes e seus produtos

Dada uma matriz $A \in \mathbb{R}^{m \times n}$, vamos adotar a convenção de representar suas colunas por $A_1, A_2, \ldots, A_n$ (empregando maiúsculas) e as colunas associadas às linhas de A por $a_1, a_2, \ldots, a_m$. Diante dessa notação, a_i^T é um vetor linha (uma matriz de dimensão $1 \times m$). A menos que seja indicado explicitamente, usaremos maiúsculas para representar colunas de matrizes e minúsculas para representar suas linhas.

Diante dessa notação, podemos representar uma matriz A com m linhas e n colunas por três formas:

- Visão elemento a elemento:
$$A = (a_{ij})_{i=1,\dots,m,j=1,\dots,n} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

- Visão por linhas:
$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix}$ de forma que a coluna associada à $i-$ésima linha de A é $a_i = \begin{bmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{bmatrix}$ e a linha propriamente é $a_i^T = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$.

- Visão por colunas: $A = \begin{bmatrix} A_1 & A_2 & \dots & A_n \end{bmatrix}$

À partir dessas visões distintas da representação das matrizes, podemos interpretar o produto de matrizes de formas distintas, que podem ser mais ou menos convenientes, dependendo da natureza do algoritmo que essas transformações (produtos de matrizes) venham a sintetizar. Sim, algoritmos em Álgebra Linear Computacional são representados por transformações lineares associadas a matrizes.

De início, vamos considerar como podemos algebricamente representar o produto de uma matriz A por um vetor x, obtendo um vetor b. Isso é, vamos considerar o produto $Ax = b$ e as visões (algebricamente equivalentes) seguintes sobre como o produto pode ser avaliado ou calculado:

- Visão elemento a elemento, em que cada elemento b_i de b é o produto escalar ou produto interno dos vetores coluna a_i por x:
$$b_i = a_i^T x = \langle a_i, x \rangle = \sum_{j=1}^{n} a_{ij} x_j.$$
Vetorialmente temos:
$$b = Ax = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_m^T x \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n a_{1j} x_j \\ \sum_{j=1}^n a_{2j} x_j \\ \vdots \\ \sum_{j=1}^n a_{mj} x_j \end{bmatrix}$$

- Visão de que b é o resultado da combinação linear das colunas de A com pesos dados pelas entradas de x:

$$b = Ax = \begin{bmatrix} A_1 & A_2 & \dots & A_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A_1x_1 + A_2x_2 + \cdots + A_nx_n =$$

$\sum_{i=1}^{n} A_i x_i$.

Exemplo 6 *Para ilustrar a visão da combinação linear das colunas de A para o produto $Ax = b$, considere o produto:*

$$b = Ax = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 3 \\ 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + 4 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}.$$

Por outro lado, temos que $b_1 = [1 \;\; 1 \;\; 2] \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix} = 1 + 4 - 4 = 1$, $b_2 = [0 \;\; 2 \;\; 3] \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix} = 8 - 6 = 2$, $b_3 = [3 \;\; 1 \;\; -1]^T \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix} = 3 + 4 + 2 = 9$.

Nosso próximo passo consiste em verificar como podemos agora interpretar e escrever de formas diferentes o produto de duas matrizes: $C = AB, A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}$, $C \in \mathbb{R}^{m \times p}$.

- Considerando a visão de produto interno, $c_{ij} = a_i^T B_j$, podemos escrever:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \text{ para todo } i = 1, \dots, m \text{ e } j = 1, \dots, p$$

  que matricialmente resulta na seguinte representação:

$$C = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix} \begin{bmatrix} B_1 & B_2 & \dots & B_p \end{bmatrix} = \begin{bmatrix} a_1^T B_1 & \cdots & a_1^T B_p \\ a_2^T B_1 & \cdots & a_2^T B_p \\ \vdots & \cdots & \vdots \\ a_m^T B_1 & \cdots & a_m^T B_p \end{bmatrix}$$

- Visão de que a coluna C_j de C resulta da combinação linear das colunas de A

pelas entradas da coluna B_j. Veja:

$$\begin{aligned}
C &= AB \\
&= A\left[\begin{array}{cccc} B_1 & B_2 & \dots & B_p \end{array}\right] \\
&= \left[\begin{array}{cccc} AB_1 & AB_2 & \dots & AB_p \end{array}\right] \\
&= \left[\begin{array}{cccc} \sum_{k=1}^n A_k b_{k1} & \sum_{k=1}^n A_k b_{k2} & \dots & \sum_{k=1}^n A_k b_{kp} \end{array}\right] \\
&= \left[\begin{array}{cccc} C_1 & C_2 & \dots & C_p \end{array}\right]
\end{aligned}$$

- E, finalmente, a visão de que a $i-$ésima linha de C, c_i^T, é a combinação linear das linhas de B, com pesos dados elementos na $i-$ésima linha de A:

$$\begin{aligned}
C &= AB \\
&= \left[\begin{array}{c} a_1^T \\ a_2^T \\ \dots \\ a_m^T \end{array}\right] B \\
&= \left[\begin{array}{c} a_1^T B \\ a_2^T B \\ \dots \\ a_m^T B \end{array}\right] \\
&= \left[\begin{array}{c} \sum_{k=1}^n b_k^T a_{1k} \\ \sum_{k=1}^n b_k^T a_{2k} \\ \vdots \\ \sum_{k=1}^n b_k^T a_{mk} \end{array}\right] \\
&= \left[\begin{array}{c} c_1^T \\ c_2^T \\ \vdots \\ c_m^T \end{array}\right]
\end{aligned}$$

A seguir, ilustramos as duas últimas visões por meio dos dois próximos exemplos.

Exemplo 7 $C = AB =$

$$\left[\begin{array}{cc} 2 & 3 \\ 4 & 5 \end{array}\right]\left[\begin{array}{cc} 1 & 2 \\ 3 & 4 \end{array}\right] = \left[\begin{array}{cc} \left[\begin{array}{cc} 2 & 3 \\ 4 & 5 \end{array}\right]\left[\begin{array}{c} 1 \\ 3 \end{array}\right] & \left[\begin{array}{cc} 2 & 3 \\ 4 & 5 \end{array}\right]\left[\begin{array}{c} 2 \\ 4 \end{array}\right] \end{array}\right] =$$

$$= \left[\begin{array}{cc} 1\left[\begin{array}{c} 2 \\ 4 \end{array}\right] + 3\left[\begin{array}{c} 3 \\ 5 \end{array}\right] & 2\left[\begin{array}{c} 2 \\ 4 \end{array}\right] + 4\left[\begin{array}{c} 3 \\ 5 \end{array}\right] \end{array}\right] = \left[\begin{array}{cc} 11 & 16 \\ 19 & 28 \end{array}\right]$$

Exemplo 8 $C = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \\ \begin{bmatrix} 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \end{bmatrix} =$

$$\begin{bmatrix} 2\begin{bmatrix} 1 & 2 \end{bmatrix} + 3\begin{bmatrix} 3 & 4 \end{bmatrix} \\ 4\begin{bmatrix} 1 & 2 \end{bmatrix} + 5\begin{bmatrix} 3 & 4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 11 & 16 \\ 19 & 28 \end{bmatrix}$$

2.8.1 Produto externo

Anteriormente recordamos o conceito de produto interno de dois vetores, uma função $u^T v = \langle u, v \rangle$ que retorna um número associado ao par de vetores. No produto interno, os vetores u, v precisam ter dimensões conformáveis, isto é, os dois vetores devem pertencer a um espaço vetorial de mesma dimensão, $\mathbb{R}^n$, por exemplo. Nesse momento, vamos definir uma outra operação com dois vetores, o chamado **produto externo de** u **por** v, que define uma matriz e não um escalar. O produto externo de $u \in \mathbb{R}^m$ por $v \in \mathbb{R}^n$ é a matriz $A = uv^T \in \mathbb{R}^{m \times n}$. Note que no produto externo, os vetores u, v não precisam ser conformáveis, isto é, podemos ter $m \neq n$.

Veja o detalhamento do produto externo à partir da definição do produto de duas matrizes $m \times 1$ por outra $1 \times n$, a seguir:

$$A = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} =$$

$$\begin{bmatrix} u_1v_1 & u_1v_2 & \cdots & u_1v_n \\ u_2v_1 & u_2v_2 & \cdots & u_2v_n \\ \vdots & \vdots & \vdots & \vdots \\ u_mv_1 & u_mv_2 & \cdots & u_mv_n \end{bmatrix} = \begin{bmatrix} u_1 \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \\ u_2 \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \\ \vdots \\ u_m \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \end{bmatrix} =$$

$$\begin{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} v_1 & \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} v_2 & \cdots & \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} v_n \end{bmatrix}$$

Note que pela expressão que obtivemos, todas as colunas de A são múltiplas de u. Ou seja, $C(A) = span(\{u\})$. Por outro lado, temos também que todas as linhas de A são múltiplas de v^T. Veja um exemplo numérico do produto externo na sequência e verifique que a matriz A possui apenas uma linha e uma coluna linearmente independentes.

Exemplo 9 $A = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 5 \\ 1 & 2 & 5 \\ 1 & 2 & 5 \end{bmatrix}$

2.8.2 Particionamento em blocos nas transformações lineares

Para tirar proveito das visões que apresentamos das transformações lineares, muitas vezes é conveniente fazer um particionamento das matrizes em blocos. Veja como a matriz A abaixo foi particionada em blocos, cada um deles correspondendo a uma matriz identidade de ordem 2.

$$A = \left[\begin{array}{cc|cc|cc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{array}\right] = \begin{bmatrix} I_2 & I_2 & I_2 \\ I_2 & I_2 & I_2 \end{bmatrix}$$

Observe que a matriz A dada possui 4 linhas e 6 colunas e o particionamento empregado para A foi de uma matriz em 2×3 blocos I_2. Se implementamos um particionamento em blocos em A e desejamos realizar a soma $A + B$, para uma matriz B de ordem 4×6, a soma pode ser feita bloco por bloco, separadamente.

O particionamento em blocos também facilita explicitar partes relevantes em um produto de matrizes. Veja o exemplo do produto AB abaixo, onde A foi particionada em 2×2 blocos. Observe que o número de linhas no particionamento em blocos da matriz B deve ser conformável com o número de colunas no particionamento em blocos de A.

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix}$$

O produto acima envolve a matriz A particionada em 2×2 blocos por uma matriz B particionada em 2×1 blocos. Desta forma, a matriz resultante é uma matriz cujo particionamento em blocos é 2×1.

O primeiro bloco de linhas da matriz resultante corresponde ao produto da primeira linha de blocos de A pelo correspondente bloco de colunas de B: $A_{11}B_{11} + A_{12}B_{21}$. De forma análoga, para a segunda linha de blocos de B temos: $A_{21}B_{11} + A_{22}B_{21}$, em analogia direta com o que obteríamos no caso de um produto escalar de uma linha de A por uma coluna de B.

2.9 Os quatro espaços fundamentais associados a $A \in \mathbb{R}^{m\times n}$

Nessa seção, vamos apresentar quatro subespaços vetoriais associados à uma matriz A qualquer, de ordem $m \times n$. São eles:

1. Espaço coluna de A: $C(A) = \{y = Ax | x \in \mathbb{R}^n\}$. Corresponde ao subespaço vetorial do $\mathbb{R}^m$ gerado pela combinação linear das colunas de A, $A_1, \ldots, A_n$. O espaço coluna de A também é conhecido como **espaço imagem** de A.

2. Espaço linha de A: $C(A^T) = \{x = A^T y | y \in \mathbb{R}^m\}$. Corresponde ao subespaço do $\mathbb{R}^n$ gerado pela combinação linear das linhas $a_1^T, a_2^T, \ldots, a_m^T$ de A.

3. Espaço nulo de A: $N(A) = \{x \in \mathbb{R}^n : Ax = 0\}$. Corresponde ao subespaço do $\mathbb{R}^n$ formado pelas soluções do sistema linear homogêneo $Ax = 0$. O espaço nulo de A também é chamado de **núcleo** ou **kernel** de A.

4. Espaço nulo de A^T (ou espaço nulo à esquerda de A): $N(A^T) = \{y \in \mathbb{R}^m : A^T y = 0\}$. Corresponde ao subespaço do $\mathbb{R}^m$ formado pelas soluções do sistema linear homogêneo $A^T y = 0$.

Recorde-se que o primeiro deles, o espaço coluna de A, $C(A)$, já foi introduzido anteriormente. Por completude, voltamos a enunciá-lo e a discutí-lo. O **espaço coluna** de A, $C(A)$, é definido como o subespaço vetorial do $\mathbb{R}^n$ que pode ser obtido por meio de todas as possíveis combinações lineares das colunas de A, isto é,

$$C(A) = \{y = Ax | x \in \mathbb{R}^n\}.$$

Veja que só podemos dizer que $y \in C(A)$ se existe um $x \in \mathbb{R}^n$ que, quando empregado para combinar as colunas de A, permite sintetizar ou escrever y. O vetor $x : y = Ax$ é o certificado de que $y \in C(A)$.

Uma base para $C(A)$ é um conjunto minimal de r vetores $\{y^1, \ldots, y^r\}$ tal que $C(A) = span(\{y^1, \ldots, y^r\})$. A dimensão de $C(A)$ é r que é o número de colunas LI de A.

Veja o exemplo do espaço coluna da matriz $A = \begin{bmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{bmatrix}$ indicado na Figura 2.5. O subespaço $C(A)$ possui dimensão 2, pois A possui 2 colunas LI, de forma que $C(A)$ é um subespaço *imerso* no $\mathbb{R}^3$, porém não coincide com $\mathbb{R}^3$. Dessa forma, exitem vetores $z \in \mathbb{R}^3$ que não podem ser escritos como uma combinação linear das colunas de A. O caso ilustrado na figura é do vetor $b \in \mathbb{R}^3, b \in C(A)$ dado por

$b = \begin{bmatrix} 0.4 \\ 2.5 \\ 1.7 \end{bmatrix}$. Sabemos que $b \in C(A)$ pois existe $x = (0.4, 0.3)^T$ que certifica isso: $b = Ax$. Veja então que quando discutimos a existência de solução para um sistema linear $Ax = b$, estamos efetivamente discutindo se existe um certificado x de que $b \in C(A)$.

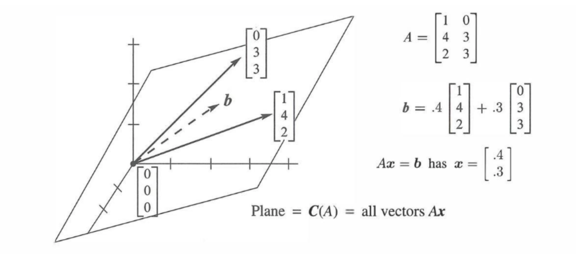

Figura 2.5: Exemplo de espaço coluna. Figura extraída de [2]

O **espaço linha** de A nada mais é do que o espaço coluna de A^T. Produzindo combinações lineares dos vetores colunas associados às linhas de A

$$x = \sum_{i=1}^{m} y_i a_i,$$

obtemos elementos x em $C(A^T)$. Estas operações podem ser equivalentemente entendidas por meio da seguinte visão $y^T A = x^T$:

$$\begin{aligned} x^T &= \begin{bmatrix} y_1 & y_2 & \cdots & y_m \end{bmatrix} \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix} \\ &= \sum_{i=1}^{m} y_i a_i^T \end{aligned}$$

No caso da matriz A indicada na Figura 2.5, os vetores a_1, a_2 associados às duas primeiras linhas de A são LI e fornecem uma base para $C(A^T)$. Resumidamente, podemos escrever para esse exemplo que $C(A^T) = span(\{a_1, a_2\})$. Nesse caso, $C(A^T) = \mathbb{R}^2$ e a dimensão de $C(A^T)$ é 2. Não por acaso, temos nesse exemplo que $dim(C(A)) = dim(C(A^T)) = 2$. Mostraremos em breve que sempre temos

$$dim(C(A)) = dim(C(A^T))$$

e que a dimensão destes espaços é chamada de posto ou rank de A: ou seja, o número de linhas e de colunas LI de qualquer matriz $A \in \mathbb{R}^{m\times n}$ é sempre igual.

O **espaço nulo** de A corresponde ao conjunto de todas as soluções para o sistema linear homogêneo $Ax = 0$. O sistema linear $Ax = b$ é chamado homogêneo se e somente se o vetor de termos independentes no sistema, b, é um m dimensional de zeros, isto é, $b = 0_m$. Observe que sempre haverá pelo menos uma solução $x = 0_n$ para o sistema linear $Ax = 0$.

Veja também que o conjunto $\mathcal{V} = \{0_n\}$ atende a todos os requisitos necessários para que possa ser chamado de subespaço vetorial: é fechado na soma e na multiplicação por escalar. Suponha agora que exista $x \neq 0$ tal que $Ax = 0$. Claramente αx para $\alpha \in \mathbb{R}$ também é uma solução pois $A(\alpha x) = \alpha Ax = 0$. Por outro lado, dados $x, y : Ax = Ay = 0$, temos que $A(x + y) = A(x - y) = 0$, de forma que $x + y, x - y \in N(A)$. Portanto, as soluções x de $Ax = 0$ de fato formam um subespaço vetorial. No caso da matriz A do exemplo, $N(A) = \{0_n\}$. Não há solução não trivial (distinta do vetor identicamente nulo) para $N(A)$ no caso desse exemplo.

O **espaço nulo à esquerda** de A, $N(A^T)$ é definido de forma análoga ao $N(A)$, correspondendo aos vetores $y \in \mathbb{R}^m$ que resolvem o sistema linear $A^T y = 0$. Para o caso do exemplo dado, observe que o vetor $y = (-2, 1, -1)^T$ resolve o sistema linear $A^T y = 0_3$. Assim sendo, temos que $span(\{(-2, 1, -1)^T\}) \subseteq N(A^T)$. Na verdade, temos que $span(\{(-2, 1, -1)^T\}) = N(A^T)$ e, em breve, mostraremos como calcular os subespaços $N(A), N(A^T)$ e como relacionar suas dimensões com as de $C(A^T), C(A)$ respectivamente. Por agora, verifique que não poderíamos ter $dim(N(A^T)) > 1$, pois nesse caso teríamos $dim(N(A^T)) + dim(C(A)) > 3 = dim(\mathbb{R}^3)$.

2.10 Posto de $A \in \mathbb{R}^{m\times n}$

O **posto** ou **rank** de uma matriz, $posto(A)$ ou $r(A)$, é o número de linhas ou de colunas linearmente independentes da matriz A. Estes valores são iguais. Logo $r(A) = dim(C(A)) = dim(C(A^T))$. Apresentaremos um algoritmo que fornecerá uma demonstração construtiva para esse fato.

Antes disso, vamos enunciar (sem demonstrar) algumas propriedades importantes do posto $r(A)$:

1. $r(A) \leq min(m, n)$.
2. $r(A) = r(A^T)$.
3. $r(A^T A) = r(AA^T) = r(A) = r(A^T)$.
4. $r(AB) \leq min(r(A), r(B))$.

5. $r(A+B) \leq r(A) + r(B)$.

6. Se $A \in \mathbb{R}^{m \times r}$, $B \in \mathbb{R}^{r \times n}$, $r(A) = r(B) = n$ então $r(AB) = n$.

Dizemos que uma matriz $A \in \mathbb{R}^{m \times n}$ possui **posto ou rank completo** quando $r(A) = \min(m, n)$. Quando isso não ocorre, dizemos que possui **deficiência de posto**. A deficiência é o valor dado por $\min(m, n) - r(A)$.

Para mostrar que o número de linhas e de colunas LI de qualquer matriz é o mesmo (e esse número recebe o nome de posto ou rank da matriz), vamos apresentar um procedimento que produz uma fatoração $A = CR$ para a matriz A. A fatoração produz dois fatores: a matriz C, cujas colunas são LI e definem $C(A)$ e a matriz R, cujas linhas são LI e definem $C(A^T)$. Atenção aqui: Não confundir na exposição que segue C, $C(A)$ e $C(C)$. Este último, em abuso de linguagem, aqui representa o espaço coluna da matriz C.

A demonstração é construtiva, isto é, utiliza um procedimento ou algoritmo para a comprovação do resultado. Ao final do mesmo, teremos que o número de colunas de C e de linhas R são iguais (como é necessário para que o produto CR seja conformável) e igual ao posto $r = r(A)$.

Assumimos que a matriz A não possui colunas ou linhas de zero, caso contrário as mesmas podem ser removidas da matriz. De início, assumimos que o número de colunas de C e de linhas de R será n. O número final de colunas de C e de linhas de R poderá ser distinto de n, se $r(A) < n$. Assim sendo, de início assumimos que $C \in \mathbb{R}^{m \times n}, R \in \mathbb{R}^{n \times n}$.

O procedimento consiste nos seguintes passos.

- Inicializamos a matriz C com a primeira coluna de A. Para expressar o fato de que $A_1 = C_1$, a primeira coluna de R é a primeira coluna da identidade. Fazemos $r = 1$. Ao longo do algoritmo, a variável r armazenará um limite inferior válido para o posto de A que desejamos descobrir. Ao final, r será o posto de A.

- Para cada coluna de índice $k = 2, \ldots, n$ de A, verificamos se $A_k \notin span(\{C_1, \cdots, C_r\})$.

  - Em caso positivo, $\{C_1, \cdots, C_r, A_k\}$ é um conjunto de vetores LI. Inserimos a coluna A_k na coluna C_{r+1} de C. A coluna k de R consiste em um vetor de zeros, exceto pela entrada correspondente à coluna A_k, cuja entrada é 1. Incrementamos $r = r + 1$, indicando que encontramos mais uma coluna de A que é LI.

  - Em caso negativo, isto é, $A_k \in span(\{C_1, \cdots, C_{k-1}\})$, existe um certificado z desta dependência linear. Veja: $[C_1, \cdots, C_r]z = A_k$. Veja que este certificado z envolve apenas r entradas. Estes r valores devem ser usados para preencher as r primeiras entradas da coluna k de R; as demais $n-r$ linhas da coluna k de R são nulos. Como a coluna A_k pode ser escrita como combinação linear das demais, a linha k de R seria formada por zeros. Desta forma, podemos simplesmente não inserir a coluna A_k em C e eliminamos a correspondente linha de zeros de R. Nesse caso, não incrementamos r, pois caracterizamos dependência linear nessa iteração.

O resumo da lógica deste procedimento é o seguinte: quando caracterizamos dependência linear de uma coluna A_k de A com as colunas anteriores de A já inseridas em C, não inserimos a coluna A_k em C e podemos remover uma linha de zeros de R. Sempre que a independência linear é caracterizada inserimos uma coluna de A em C e preservamos uma linha de R. Nesse caso, incrementamos a variável r que guarda o número de colunas LI de R. Ao final do processo, as r linhas de R também são LI, pois há r colunas de uma identidade em R. Vamos ilustrar o procedimento com um exemplo.

Exemplo 10 *Obter a fatoração $A = CR$ que revela o posto da matriz* $A = \begin{bmatrix} 1 & 1 & 2 & 4 \\ 1 & 2 & 2 & 5 \\ 1 & 3 & 2 & 6 \end{bmatrix}$.

- *Com a inicialização do procedimento temos:* $C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$, $R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

  $r = 1$.

- *Para $k = 2$, vericamos que $A_2 \notin span(\{C_1\})$. Portanto, inserimos A_2 na coluna $2 = r + 1$ de C, incrementamos $r \leftarrow r + 1 = 2$. A correspondente coluna 2 de R é a coluna da identidade. Veja:* $C = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 3 & 0 & 0 \end{bmatrix}$, $R =$

  $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

- *Para $k = 3$, verificamos que $A_3 = 2C_1 + 0C_2$. Portanto, $A_3 \in span(\{C_1, C_2\})$ e A_3 não é necessária para caracterizar $C(A)$. Não precisamos incluir A_3 em C. Removemos uma coluna de C e a linha $3 = r+1$ de R. Veja:* $C = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 3 & 0 \end{bmatrix}$,

$$R = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

- *Para $k = 4$, verificamos que $A_4 = 3C_1 + 1C_2$. Portanto, $A_4 \in span(\{C_1, C_2\})$ e A_4 não é necessária para caracterizar $C(A)$. Não precisamos incluir A_4 em C. Removemos mais uma coluna de C e a linha $r + 1 = 3$ de R. Veja:* $C = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, R = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$

Ao final do algoritmo temos, por construção, que as r linhas de R são LI pois conseguimos extrair uma submatriz identidade de ordem r de $R : r \times n$, escolhendo os índices das colunas de R em que incrementamos a variável r ao longo do algoritmo. Além disso, o número de linhas LI de R e de colunas LI de C é exatamente $r =$ posto(A). Como resultado, escrevemos:

$$A = \begin{bmatrix} 1 & 1 & 2 & 4 \\ 1 & 2 & 2 & 5 \\ 1 & 3 & 2 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

Exercício 2.10.1 *Utilizando o algoritmo que apresentamos acima, verifique as fatorações para as matrizes abaixo indicadas.*

- $A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix}, C = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}, R = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix}$

- $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}, C = A, R = I$

- $A = \begin{bmatrix} 1 & 2 & 5 \\ 1 & 2 & 5 \\ 1 & 2 & 5 \end{bmatrix}, C = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, R = \begin{bmatrix} 1 & 2 & 5 \end{bmatrix}$

Exercício 2.10.2 *Observe pelo exemplo abaixo, que a fatoração $A = CR$ não é única e que a fatoração que apresentamos é distinta daquela que seria obtida com a aplicação do algoritmo. Qual seria a fatoração obtida com a aplicação do algoritmo ? Fatoração alternativa àquela obtida com o algortimo que apresentamos:*
$A = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 0 & 4 \end{bmatrix}, C = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, R = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \end{bmatrix}$

Para concluirmos esta seção vamos fazer algumas observações muito importantes que decorrem da fatoração $A = CR$ que revela o posto da matriz A:

1. O algoritmo fornecido para fornecer a fatoração $A = CR$ que revela o posto de A demonstra que a dimensão dos espaços coluna e linha de uma matriz qualquer são sempre iguais ao seu posto.

2. O espaço coluna de A é o espaço coluna da matriz C na fatoração.
   Veja que se $b \in C(A)$, existe $x : Ax = b$. Logo $CRx = b$ e, consequentemente $b \in C(C)$ pois $z = (Rx)$ certifica isso. Resultado:

$$b \in C(A) \iff b \in C(C).$$

3. O espaço linha de A é o espaço linha de R na fatoração.
   De forma análoga, se $z \in C(A^T)$ existe $u : A^T u = z$. Portanto $z = (R^T C^T)u = R^T(C^T u)$ e $x = (C^T u)$ certifica que $z \in C(R^T)$. Resultado

$$z \in C(A^T) \iff z \in C(R^T).$$

2.10.1 Matrizes de rank-1

Considere a matriz $A = \begin{bmatrix} 1 & 2 & 5 \\ 1 & 2 & 5 \\ 1 & 2 & 5 \end{bmatrix}$ que foi tratada no Exercício 2.10.1, e sua fatoração CR: $A = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 5 \end{bmatrix}$. Veja que a fatoração de A revela que $posto(A) = 1 = dim(C(A)) = dim(C(A^T))$. Além disso, a fatoração também mostra que $A = uv^T$, isto é, A é o produto externo entre $u = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ e $v = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$.
O produto externo entre dois vetores gera uma matriz de posto-1.

Matrizes de posto-1 são elementos fundamentais para se escrever matrizes mais gerais. Uma matriz $A \in \mathbb{R}^{m\times n}$ qualquer de posto $r \geq 1$ pode ser escrita como uma soma de r.

Vamos assumir que o algoritmo que produz a fatoração CR tenha sido aplicado e que, para fins didáticos, a matriz C seja renomeada por U e a matriz R por V, de forma que as r colunas de U e as r linhas de V são LI. Isto é, $posto(A) = r$ e $A = UV$ para $U \in \mathbb{R}^{m\times r}, V \in \mathbb{R}^{r\times n}$. Vamos escrever a fatoração usando a representação conveniente:

$$\begin{aligned}
A &= UV \\
&= \begin{bmatrix} U_1 & U_2 & \cdots & U_r \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_r^T \end{bmatrix} \\
&= \sum_{k=1}^{r} U_k v_k^T
\end{aligned}$$

Verifique que a expressão do termo genérico $a_{ij} = \langle u_i^T, V_j \rangle$ é exatamente $(\sum_{k=1}^{r} U_k v_k^T)_{ij}$.

Exemplo 11 *Considere as fatorações que revelam o posto das matrizes apresentadas no Exercício 2.10.1 e verifique como as matrizes podem ser escritas como somas de matrizes de rank-1.*

- $A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix}, U = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}, V = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix}.$

  $A = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}$

- $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}, U = A, V = I.$

  $A = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 5 \\ 6 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}.$

2.10.2 O Teorema Fundamental da Álgebra Linear

Já demonstramos que, dada uma matriz $A \in \mathbb{R}^{m \times n}$ de posto $r \leq \min(m, n)$, $dim(C(A)) = dim(C(A^T)) = r$. Nosso objetivo agora é enunciar e mostrar alguns resultados adicionais sobre os quatro espaços fundamentais de A, que são conhecidos como o Teorema Fundamental da Álgebra Linear (TFAL), enunciado a seguir.

Teorema 2.10.1 *Teorema Fundamental da Álgebra Linear. Para uma matriz $A \in \mathbb{R}^{m \times n}$ de posto $r \leq \min(m, n)$, valem os seguintes resultados para as dimensões dos 4 espaços fundamentais de A:*

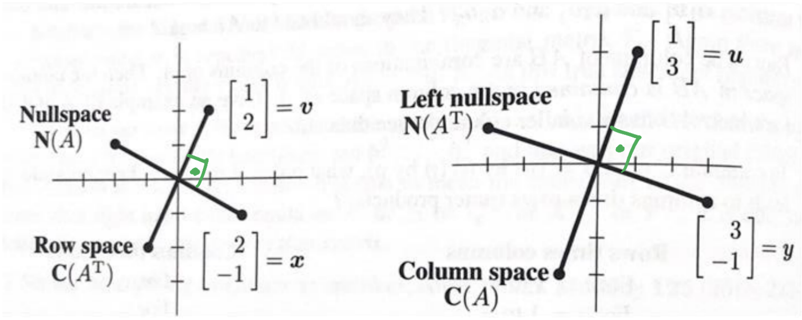

Figura 2.6: Exemplos dos quatro espaços fundamentais associados a uma matriz de posto 1. Figura extraída de [2]

- $dim(C(A)) = r$
- $dim(C(A^T)) = r$
- $dim(N(A)) = n - r$
- $dim(N(A^T)) = m - r$

Veja o seguinte exemplo que demonstra os quatro espaços associados uma matriz de posto 1 e suas dimensões.

Exemplo 12 $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = uv^T = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix}$, $m = 2$, $n = 2$.

1. $C(A) = span(\left\{\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right\})$.
2. $C(A^T) = span(\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\})$.
3. *Nulo* $N(A) = span(\left\{\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right\})$.
4. $N(A^T) = span(\left\{\begin{bmatrix} 3 \\ -1 \end{bmatrix}\right\})$.

Veja a Figura 2.6 que ilustra estes quatro subespaços, cada um deles com bases de dimensão 1.

Para demonstrar os resultados discutidos nessa seção, vamos empregar a definição de **espaços ortogonais**. Sabemos que dois vetores $u, v \in \mathbb{R}^n$ são ortogonais se e somente se $u^T v = 0$. Podemos generalizar a ideia de ortogonalidade para subespaços vetoriais.

Dados dois subespaços $\mathcal{V}, \mathcal{Y}$ de um mesmo espaço vetorial $\mathbb{R}^n$, dizemos que $\mathcal{V} \perp \mathcal{Y}$, isto é, $\mathcal{V}$ e $\mathcal{Y}$ são ortogonais, se e somente se qualquer $v \in \mathcal{V}$ e $y \in \mathcal{Y}$ satisfizerem $v^T y = 0$, isto é, se qualquer par de vetores em cada um destes espaços for ortogonal.

À partir do conceito de subespaços ortogonais, segue outro bastante importante. Dado um subsepaço vetorial $\mathcal{V} \subseteq \mathbb{R}^n$, definimos $\mathcal{V}^{\perp} \subseteq \mathbb{R}^n$ como o **complemento ortogonal** de $\mathcal{V}$ como a coleção de todos os vetores do $\mathbb{R}^n$ que são ortogonais a todos os vetores em $\mathcal{V}$. Ou seja, matematicamente temos:

$$\mathcal{V}^{\perp} = \{x \in \mathbb{R}^n : x^T y = 0, \text{ para qualquer } y \in \mathcal{V}\}.$$

Nesse momento cabe enfatizar que espaços ortogonais não são necessariamente complementos ortogonais. Considere os espaços $\mathcal{V} = span(\{[1,0,0]^T\})$ e $\mathcal{Y} = span(\{[0,1,0]^T\})$. $\mathcal{V} \perp \mathcal{Y}$ mas $\mathcal{V} \neq \mathcal{Y}^{\perp}$. Veja também que a soma da dimensão destes dois subespaços é 2 e não 3, a dimensão do $\mathbb{R}^3$, algo que deveria ser observado para qualquer subespaço e seu complemento ortogonal do $\mathbb{R}^3$.

De posse dessas definições, vamos demonstrar o TFAL bem como discutir os resultados seguintes, usualmente conhecidos como a segunda parte do TFAL:

- Qualquer par de vetores x, z: $x \in N(A)$ e $z \in C(A^T)$ satisfazem $x^T z = 0$ e $N(A) = (C(A^T))^{\perp}$.
- Qualquer par de vetores u, v: $u \in N(A^T)$ e $v \in C(A)$ satisfazem $u^T v = 0$ e $N(A^T) = (C(A))^{\perp}$.

Veja que um vetor $x \in N(A)$ deve satisfazer $a_i^T x = 0, i = 1, \ldots, m$. Logo $x \in N(A) \to x \perp a_i$ para todo $i = 1, \ldots, m$, e portanto $N(A) \subseteq C(A^T)^{\perp}$. Por outro lado, tomando qualquer $x \in C(A^T)^{\perp}$ temos que $Ax = 0$ e $x \in N(A)$, mostrando que $N(A) \supseteq C(A^T)^{\perp}$. Combinando os dois resultados temos $N(A) = C(A^T)^{\perp}$.

Uma outra maneira de se chegar a esse resultado faz uso da fatoração da matriz A que vimos. Suponha que a matriz A tenha sido fatorada na forma $A = UV$, onde V possui $r = posto(A)$ linhas LI e U possui r colunas LI. Substituindo $A = UV$ em $Ax = 0$ temos que $UVx = 0$. Para o sistem linear homogêneo admitir solução, $Vx = 0$ deve valer, pois sendo LI as colunas de U não é possível esperar $Uw = 0$ para $w \neq 0_r$. Assim sendo, $v_i^T x = 0$ para todo $i = 1, \ldots, r$.

Como todo $z \in C(A^T)$ pode ser escrito como $z = \sum_{i=1}^{r} \beta_i v_i$, (os vetores v_i : $i = 1, \ldots, r$ formam uma base para o espaço) temos que para $x \in N(A)$, $z^T x = \sum_{i=1}^{r} \beta_i (v_i^T x) = 0$. Como x e z são quaisquer em $N(A)$ e $C(A^T)$ temos

$$N(A) = C(A^T)^{\perp}.$$

Das observações acima temos o resultado que envolve a dimensão dos subespaços:

$$dim(C(A^T)) + dim(N(A)) = n.$$

Raciocínio análogo pode ser empregado para mostrar que

$$N(A^T) = C(A)^{\perp}$$

e

$$dim(C(A)) + dim(N(A^T)) = m.$$

Veja que os quatro subespaços fundamentais associados a uma matriz $A \in \mathbb{R}^{m\times n}$ separam os espaços $\mathbb{R}^m$ e $\mathbb{R}^n$ em dois pares de subespaços que formam complementos ortogonais. Veja a Figura 2.7. O único vetor comum a cada um dos subespaços em cada um dos pares é o vetor nulo. Assim, qualquer ponto $y \in \mathbb{R}^m$ pode ser escrito da forma segundo $y = y^1 + y^2$, onde $y^1 \in N(A^T), y^2 \in C(A)$ são únicos nos respectivos subespaços complementares. Os vetores y^1, y^2 são as projeções de y nos subespaços $N(A^T), C(A)$ respectivamente, isto é, os pontos de $N(A^T), C(A)$ mais próximos de y na norma Euclideana. Analogamente, qualquer $x \in \mathbb{R}^n$ pode ser escrito como $x = x^1 + x^2$, onde $x^1 \in N(A), x^2 \in C(A^T)$ também são únicos. Ao longo do curso, aprenderemos como calcular as projeções de um ponto nos subespaços complementares ortogonais.

Por hora, resumimos estes resultados:

- $\mathbb{R}^n = N(A) \oplus C(A^T)$, isto é, a soma direta de $N(A)$ e $C(A^T)$ resulta em $\mathbb{R}^n$.
- $\mathbb{R}^m = N(A^T) \oplus C(A)$, isto é, a soma direta de $N(A^T)$ e $C(A)$ resulta em $\mathbb{R}^m$.

Sempre que qualquer ponto $x \in \mathbb{R}^n$ pode ser decomposto $x = x^1 + x^2$, $x^1 \in \mathcal{V}, x^2 \in \mathcal{Y}$ dizemos que $\mathbb{R}^n = \mathcal{V} \oplus \mathcal{Y}$ (soma direta).

2.10.3 Existência e unicidade de soluções para $Ax = b$ à luz dos quatro espaços fundamentais

Munidos dos resultados que apresentamos sobre os quatro espaços fundamentais, vamos agora discutir a existência e a unicidade de soluções para um sistema linear $Ax = b$, em m restrições e n variáveis. Acompanhe a Figura 2.8 ao longo dos pontos que abordamos nesta seção.

Quando o sistema linear $Ax = b$ admite solução, temos que $b \in C(A)$. O vetor x certifica o fato: combinamos as colunas $A_1, \ldots, A_n$ de A com os pesos $x_1, \ldots, x_n$ e obtemos b. Assumindo que $b \in C(A)$, a solução (ou o certificado) x é única ? A

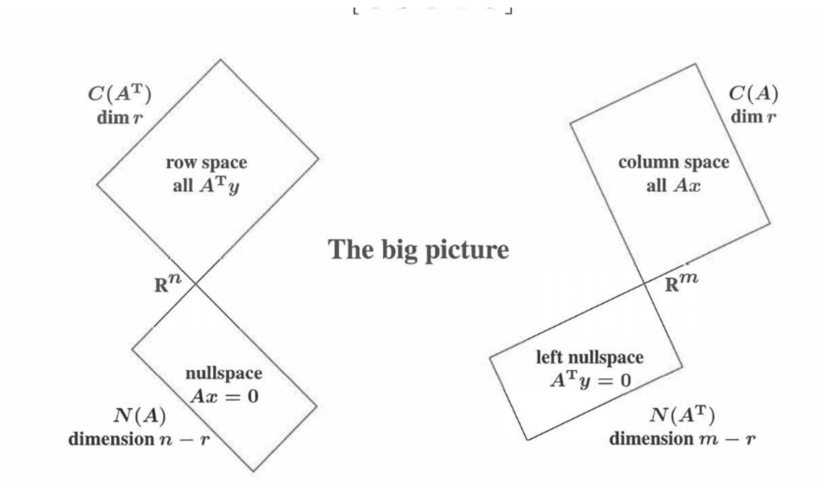

Figura 2.7: Representação didática dos quatro espaços fundamentais associados a uma matriz $A \in \mathbb{R}^{m \times n}$ qualquer. Figura extraída de [2]

resposta depende do subespaço $N(A)$. Vamos supor que exista $x^1 \neq 0$, $x^1 \in N(A)$. Veja que

$$\begin{aligned} Ax &= b \\ Ax^1 &= 0 \\ A(x + x^1) &= b \end{aligned}$$

Como $x^1 \neq 0_n$, $x + x^1$ é uma solução alternativa para $Ax = b$. Qualquer $x + y$ para $y \in N(A)$ também será. Portanto, sempre que $N(A) \neq \{0_n\}$, isto é, quando este subespaço não se resumir ao vetor 0_n, quando o sistema $Ax = b$ admitir solução, admitirá infinitas soluções. Por outro lado, o sistema admitirá solução única quando $posto(A) = r = n$, ou seja, o posto coluna é completo. Nesse caso, $N(A) = \{0_n\}$ e o sistema não admitirá solução (quando $b \notin C(A)$) ou admitirá apenas uma solução.

É importante destacar que a matriz A pode possuir posto $r < n$ e mesmo assim o sistema linear $Ax = b$ não admitir solução. Só é possível empregar um vetor $y \in N(A)$ para produzir soluções alternativas para o sistema quando $b \in C(A)$, ou seja, quando houver pelo menos um certificado x de que isso ocorre.

2.11 Autovalores e autovetores

Assumimos nesta seção que A seja uma matriz real quadrada de ordem n. Um dos objetivos centrais do curso de ALC é apresentar algoritmos para encontrar os autovalores e os autovetores de A. Por hora, apenas revisamos os conceitos mais importantes de autovalores e autovetores.

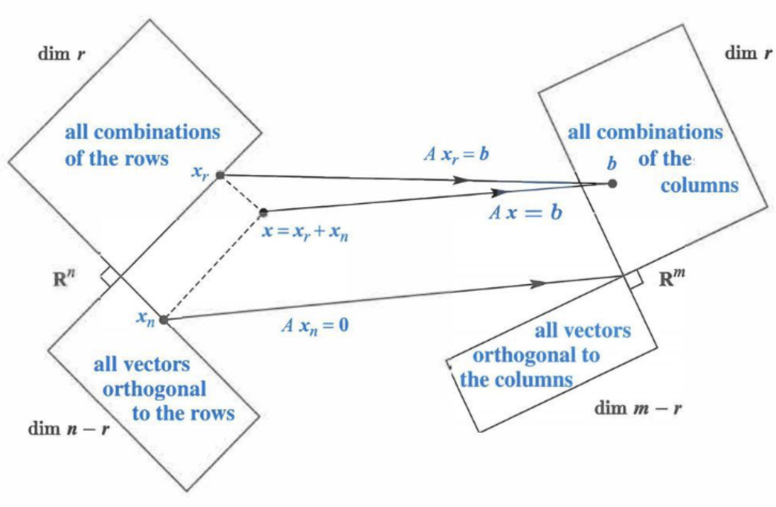

Figura 2.8: Representação das transformações lineares $Ax = b$ à luz dos quatro espaços fundamentais. Figura extraída de [2]

Um par $(\lambda, x) : \lambda \in \mathbb{C}, x \in \mathbb{C}^n, x \neq 0_n$ é um autopar (autovalor + autovetor) para A se e somente se

$$Ax = \lambda x.$$

Veja que para um autovetor x, a transformação linear Ax resulta sempre em algum vetor em $span(\{x\})$, possivelmente com norma e direção distintos das de x. Atenção aqui: uma matriz real pode ter autovalores e autovetores complexos.

Da definição temos:

$$Ax = \lambda x$$
$$Ax - \lambda x = 0$$
$$(A - \lambda I)x = 0.$$

Veja que a equação $Ax = \lambda x$ é trivialmente satisfeita para $x = 0$, que portanto não nos interessa. Assim sendo, como $x \neq 0$, $N(A - \lambda I) \neq \{0\}$ e portanto $\det(A - \lambda I) = 0$: a matriz $A - \lambda I$ é singular. A expressão $\det(A - \lambda I) = 0$ é a expressão de um polinômio, **o polinômio característico** de A. Ou seja, qualquer autovalor de A deve ser raiz para o polinômio que se obtem ao impormos $\det(A - \lambda I) = 0$.

Algumas propriedades importantes dos autovalores de A são as seguintes:

1. $\det(A) = \prod_{i=1}^{n} \lambda_i$

2. traço$(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$.

3. Uma matriz A de ordem n possui n autovalores (contando suas multiplicidades)

e caso, $a + bi$ seja um autovalor complexo, seu par conjugado $a - bi$ também é autovalor de A.

4. Matrizes reais simétricas possuem autovalores e autovetores reais.

Exemplo 13 *Encontre os autovalores e autovetores da matriz* $A = \begin{bmatrix} 8 & 3 \\ 2 & 7 \end{bmatrix}$. $A - \lambda I = \begin{bmatrix} 8-\lambda & 3 \\ 2 & 7-\lambda \end{bmatrix}$. $\det(A-\lambda I) = 0 \rightarrow \lambda^2 - 15\lambda + 50 = (\lambda - 10)(\lambda - 5) = 0$. *Portanto as raízes do polinômio característico de A e seus correspondentes autovetores são:* $\lambda_1 = 10, x^1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ *e* $\lambda_2 = 5, x^2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Salientamos que para fins práticos, não recorremos ao polinômio característico da matriz A para calcularmos seus autovalores. A obtenção dos autovalores se dá por meio de fatorações da matriz. Ao final do curso de ALC vamos apresentar um algoritmo que identifica os autovalores de A, e eventualmente seus autovetores. Esse algoritmo produz uma fatoração espectral de A, quando a mesma admite admite uma fatoração desse tipo, ou produz uma fatoração de Schur, quando A não admite uma fatoração espectral. Ambas fornecem os autovalores de A, mas apenas a fatoração espectral também fornece todos os n autovetores da matriz (quando ela os tem).

Uma fatoração espectal de A é do tipo $A = Q\Lambda Q^T$ onde Q é uma matriz ortogonal ($QQ^T = Q^TQ = I$) e Λ é uma matriz diagonal, armazenando em sua diagonal os autovalores de A. Uma fatoração de Schur de A é do tipo $A = QTQ^T$, onde T é uma matriz triangular superior que armazena os autovalores de A em sua diagonal.

Para interpretar a fatoração espectral, vamos assumir que A possua n autovetores linearmente independentes. Seja x^i o autovetor de A associado ao autovalor λ_i. Podemos escrever:

$$
\begin{aligned}
Ax^1 &= \lambda_1 x^1 \\
Ax^2 &= \lambda_2 x^2 \\
\vdots \;&\; \vdots \\
Ax^n &= \lambda_1 x^n
\end{aligned}
$$

$$
A[x^1 \;\; x^2 \;\; \cdots x^n] = [x^1 \;\; x^2 \;\; \cdots x^n] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}
$$

Representando a matriz $[x^1 \ x^2 \ \cdots x^n]$ por X e $\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$ por Λ e observando que X^{-1} existe (assumimos que A possui n autovetores LI), podemos escrever de forma sucinta a relação entre os n autopares e A

$$\begin{aligned} AX &= X\Lambda \\ A &= X\Lambda X^{-1} \end{aligned}$$

A fatoração espectral $A = Q\Lambda Q^T$ (onde Q é ortogonal) é um caso especial da fatoração $A = X\Lambda X^{-1}$ em que: (1) a matriz A não apenas possui n autovetores LI mas também (2) as colunas de X que fornecem os autoveores são ortonormais e, assim, $X^{-1} = X^T$.

Cabe enfatizar que nem toda matriz quadrada admite uma fatoração $A = X\Lambda X^{-1}$ ou ainda $A = Q\Lambda Q^T$. Porém toda matriz A admite uma fatoração de Schur $A = QTQ^T$ que revela seus autovalores (e apenas um autovetor).

2.11.1 Transformações lineares associadas às potências de matrizes

Vamos considerar as potências A^k para $k \geq 1$ e inteiro da matriz A. Veja que os autovetores da matriz A também são autovetores da matriz A^k. Veja a dedução algébrica e a Figura 2.9.

$$\begin{array}{rl} k = 1 & Ax = \lambda x \\ k = 2 & A(Ax) = A(\lambda x) = \lambda(Ax) = \lambda^2 x \\ k = 3 & A(A^2x) = A(\lambda^2 x) = \lambda^2(Ax) = \lambda^3 x \\ \vdots & \vdots \\ k \text{ qualquer} & A^k x = \lambda^k x \end{array}$$

Os autovalores também são importantes para predizer o comportamento assintótico ou mesmo facilitar o cálculo da transformação linear dada por A^k. Para discutir essa observação, vamos assumir que A possua n autovetores LI. Sendo $\{x^1, x^2, \ldots, x^n\}$ o conjunto destes autovetores, veja que $\mathbb{R}^n = span(\{x^1, x^2, \ldots, x^n\})$. Portanto, qualquer $v \in \mathbb{R}^n$ pode ser escrito a partir da base $\{x^1, x^2, \ldots, x^n\}$ para o

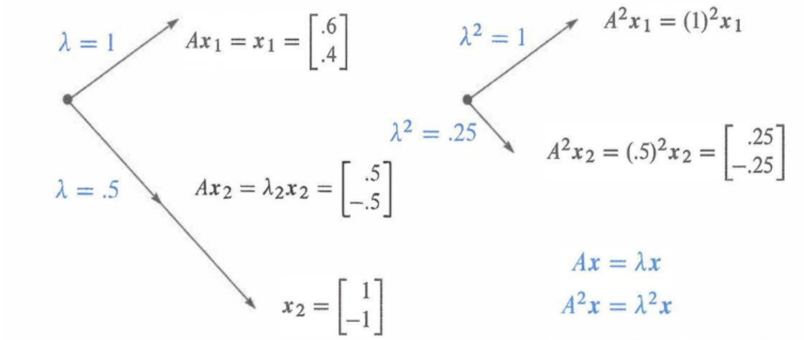

Figura 2.9: Os autopares de A e de A^2. Figura extraída de [2]

$\mathbb{R}^n$:

$$v = c_1 x^1 + c_2 x^2 + \cdots + c_n x^n.$$

Adicionalmente, vamos supor que os autovetores de A sejam associados a autovalores distintos. Isto é, $x^1, \cdots, x^n$ têm os autovalores $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$, respectivamente. Então para matrizes A satisfazendo a primeira destas premissas, podemos escrever

$$\begin{aligned}
Av &= A(c_1 x^1 + c_2 x^2 + \cdots + c_n x^n) \\
Av &= c_1 A x^1 + c_2 A x^2 + \cdots + c_n A x^n \\
Av &= c_1 \lambda_1 x^1 + c_2 \lambda_2 x^2 + \cdots + c_n \lambda_n x^n \\
A^k v &= c_1 \lambda_1^k x^1 + c_2 \lambda_2^k x^2 + \cdots + c_n \lambda_n^k x^n
\end{aligned}$$

Vamos agora considerar $\lim_{k\to\infty} A^k v$. Note que se $|\lambda_1| > 1$, o componente $c_1 \lambda_1^k x^1$ crescerá com o aumento de k. Nesse caso, no limite $A^k v$ é um vetor com entradas de magnitudes muito grandes. Por outro lado, se $|\lambda_1| < 1$ todos os termos de $A^k v$ tendem a zero quando $k \to \infty$.

Considere o seguinte procedimento, que é inicializado com $k = 0$ e $v^0 = v$. O procedimento explora a ideia que discutimos mas normaliza o resultado da transformação linear a cada iteração, indexada por k.

- Inicialização: $k = 0, v^0 = v$.

- Repita:

  - $v^{k+1} \leftarrow A v^k$
  - $v^{k+1} \leftarrow \frac{v^{k+1}}{\|v^{k+1}\|}$

– $k \leftarrow k+1$

Veja que $v^\infty = \lim_{k\to\infty} A^k v \in span(\{x^1\})$. Isso significa que ao aplicarmos A muitas vezes em v, o resultado tende a ser um vetor ao longo da direção de x^1, o autovetor associado ao maior autovalor de A em módulo.

Exemplo 14 *Considere os números de Fibonacci dados pela sequência infinita:* $0, 1, 1, 2, 3, 5, 8, 13, 21, \ldots$. *Os números* F_k *de Fibonacci são definidos recursivamente, da seguinte forma:*

- $F_0 = 0, F_1 = 1$ *são os dois primeiros.*
- *Para os demais vale a equação de diferenças:* $F_{k+2} = F_{k+1} + F_k$.

Veja que, dado um vetor contendo dois números de Fibonacci consecutivos, $v^k = (F_{k+1}, F_k)^T$, *podemos escrever o sistema de Equações de diferenças para calcular o vetor* $v^{k+1} = (F_{k+2}, F_{k+1})^T$ *como*

$$\begin{array}{ccccc} F_{k+2} & = & F_{k+1} & + & F_{k+1} \\ F_{k+1} & = & F_{k+1} & & \end{array}$$

Então o processo de geração dos números de Fibonacci à partir dos 2 números iniciais pode ser sintetizado pela equação de diferenças:

$$v^{k+1} = Av^k$$

$$v^{k+1} = \begin{bmatrix} F_{k+2} \\ F_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}.$$

Claramente, se desejamos o p-ésimo número de Fibonaccci, podemos gerar os números de Fibonacci necessários, iterando o processo. Começamos com $v^0 = (0, 1)^T$ *e calculamos os* $v^k : k = 1, \ldots, p$ *necessários. A segunda entrada de* v^p *fornece* F_p. *Podemos também calcular a potência p de A e usar o fato de que* $v^p = A^p v^0$.*Isso é o melhor que podemos fazer ? Não.*

Veja que a matriz A que rege o processo de diferenças é simétrica e portanto admite uma fatoração espectral $A = Q\Lambda Q^T$. *Então o sistema de equações de diferenças pode ser reescrito de uma forma bastante mais conveniente, à partir de combinações lineares dos autovetores de A, com pesos que dependem das potências dos autovalores.*

Veja $v^k = (Q\Lambda Q^T)^k v^0 = Q\Lambda^k Q^T v^0$. *Substituindo* $Q^T v^0$ *por c, temos:*

$$v^{k+1} = \begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} \lambda_1^k & \\ & \lambda_2^k \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \sum_{i=1}^2 c_i(\lambda_i^k q_i).$$

Os autovalores de A são $\lambda_1 = \frac{1+\sqrt{5}}{2}$ *e* $\lambda_2 = \frac{1-\sqrt{5}}{2}$ *e* $c = \frac{1}{\sqrt{5}}(1,1)^T$. *Com estes valores, temos que a segunda entrada* F_k *de* v^{k+1} *é dada por*

$$F_k = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^k - \left(\frac{1-\sqrt{5}}{2}\right)^k\right).$$

Os números de Fibonacci são inteiros, de forma que esta expressão que é exata (e envolve números irracionais) deve retornar números inteiros. Isso é verdade desde que aritmética de precisão infinita seja empregada.

Como esse não é o caso em situações práticas de interesse, vale observar que: (1) $\lambda_2 < 0$ *(2) a potência* λ_2^k *é positiva para k par e negativa para k ímpar. Por outro lado,* $\left|\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^k\right| < \frac{1}{2}$. *Assim sendo, podemos calcular o número de Fibonacci empregando apenas o primeiro termo e arredondando o resultado para o inteiro mais próximo. Para calcularmos, por exemplo,* F_{200} *basta calcularmos* $\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{200}$ *e arrendondarmos o resultado para o inteiro mais próximo.*

2.11.2 Observações adicionais sobre autovalores

Para concluir esta seção introdutória sobre autovalores e autovetores, considere as seguintes observações adicionais.

- Vamos assumir que A admita inversa. Com os autopares (λ, x) de A, dispomos também dos autopares $(\frac{1}{\lambda}, x)$ de sua inversa A^{-1}. Veja:

$$\begin{aligned} A^{-1}Ax &= A^{-1}(\lambda x) \\ x &= A^{-1}(\lambda x) \\ A^{-1}x &= \frac{1}{\lambda}x \end{aligned}$$

  Ou seja, se x é autovetor de A com autovalor λ, x é autovetor de A^{-1} com autovalor $\frac{1}{\lambda}$.

- Se somamos a A a quantidade sI para um escalar $s \in \mathbb{R}$, os autovalores de $A + sI$ são a soma dos autovalores de A com s:

$$(A + sI)x = Ax + sx = \lambda x + sx = (\lambda + s)x.$$

- Para toda matriz B inversível, os autovalores de $C = BAB^{-1}$ são iguais aos autovalores de A.

$$(BAB^{-1})(Bx) = BAx = B\lambda x = \lambda(Bx).$$

As matrizes A e $C = BAB^{-1}$ são denominadas *similares* e, como tal, tem o mesmo espectro (conjunto de autovalores). Os autovetores de matrizes similares não são iguais, a não ser que $B = I$ e, nesse caso, $A = C$.

- Se B (inversível) é a matriz que caracteriza a similaridade de A e de C, um autovetor de C é Bx, onde x é autovetor de A. A existência de B, B^{-1} satisfazendo $C = BAB^{-1}$ garante a similaridade entre A e C. A consequência é que A e C possuem os mesmos autovalores.

Na sequência, vamos discutir a importância de algumas matrizes especiais, dentre elas as matrizes ortogonais que já definimos e empregamos.

2.12 Algumas matrizes especiais

2.12.1 Matrizes ortogonais e unitárias

Conforme definimos na Seção 2.11 sobre autovalores e autovetores, matrizes ortogonais são matrizes quadradas em que toda coluna tem norma Euclideana unitária e todo par de colunas distintas são ortogonais. Usualmente empregamos Q para representar uma matriz ortogonal. Assim sendo temos: $QQ^T = Q^TQ = I$.

Observe que a inversa de uma matriz ortogonal é sua transposta: $Q^{-1} = Q^T$.

O conceito equivalente ao de uma matriz real ortogonal nos complexos é o de **matriz unitária**. Uma matriz $Q \in \mathbb{C}^{n\times n}$ é unitária quando $QQ^* = Q^*Q = I_n$, onde Q^* é a matriz obtida ao se transpor e conjugar a parte complexa das entradas de Q.

Quando a matriz possui colunas ortonormais, isto é, $Q_i^TQ_j = 0$ para $i \neq j$ e $Q_i^TQ_i = 1$, mas não é quadrada (o número de colunas é menor que o número de linhas), a matriz é chamada de ortonormal. Supondo então ser esse o caso, $Q \in \mathbb{R}^{m\times n}$, $n < m$, $Q^TQ = I_n$. Porém $QQ^T \neq I_m$.

Exemplo 15 *Veja os exemplos de matrizes com colunas ortonormais abaixo indicadas. Nenhuma das duas é ortogonal pois não são quadradas:* $Q_1^TQ_1 = 1 = I_1$ *e* $Q_2^TQ_2 = I_2$*.* $Q_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$, $Q_2 = \frac{1}{3}\begin{bmatrix} 2 & 2 \\ 2 & -1 \\ -1 & 2 \end{bmatrix}$

Matrizes ortogonais (unitárias) são de enorme importância em Computação Científica e em Álgebra Linear Computacional. Os motivos são relacionados às seguintes propriedades destas matrizes:

1. A norma Euclideana e os ângulos formados entre vetores não são alterados pela matriz Q. Veja:

- $\|Qx\|_2^2 = x^T Q^T Qx = x^T x = \|x\|_2^2$, portanto Q preserva a norma Euclideana.
- $\langle Qx, Qy \rangle = x^T Q^T Qy = x^T y = \langle x, y \rangle$, portanto os ângulos formados por x, y são os mesmos formados por Qx e Qy.

2. As colunas de uma matriz ortogonal são uma base ortonormal para $\mathbb{R}^n$.
3. As linhas de uma matriz ortogonal são uma base ortonormal (possivelmente diferente) para $\mathbb{R}^n$.
4. Usar transformações lineares induzidas por Q não acarreta erro numérico substancial.
5. Usar transformações lineares induzidas por Q não acarreta *overflow*.

2.12.2 Matrizes simétricas positivas definidas

Uma matriz é simétrica quanto $A = A^T$. O conceito equivalente nos complexos é o de matriz Hermitiana: $A \in \mathbb{C}^{n \times n}$ é Hermitiana se $A = A^*$. Nesta seção, vamos estudar algumas classes de matrizes reais simétricas.

Definimos como $\mathcal{S}_n$ o espaço de todas as matrizes reais simétricas (verifique que $\mathcal{S}_n$ define um espaço vetorial). Algumas matrizes simétricas são de particular interesse: as matrizes simétricas positivas definidas e as semidefinidas positivas. Para apresentar a definição destas, considere a seguinte função $f : \mathbb{R}^n \to \mathbb{R}$:

$$f(x) = x^T Ax = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{2.4}$$

A função $f(x)$ apresentada em (2.4) é chamada de **energia da matriz**. Considere agora as seguintes definições derivadas da função f:

- $A \in \mathcal{S}_n$ é **simétrica positiva definida** (SPD) se e somente se $f(x) > 0$ para qualquer $x \in \mathbb{R}^n, x \neq 0_n$.
- $A \in \mathcal{S}_n$ é **simétrica semipositiva definida** se e somente se $f(x) \geq 0$ para qualquer $x \in \mathbb{R}^n, x \neq 0_n$.
- $A \in \mathcal{S}_n$ é **simétrica negativa definida** se e somente se $f(x) < 0$ para qualquer $x \in \mathbb{R}^n, x \neq 0_n$. Veja que se A é negativa definida, $-A$ é positiva definida.
- $A \in \mathcal{S}_n$ é **simétrica seminegativa definida** se e somente se $f(x) \leq 0$ para qualquer $x \in \mathbb{R}^n, x \neq 0_n$. De forma análoga, se A é seminegativa definida, $-A$ é semipositiva definida.

- $A \in \mathcal{S}_n$ é **indefinida** se nenhuma das classificações acima se aplicar, ou seja, se existirem x, y tais que $f(x) > 0, f(y) < 0$.

O conceito de positividade da matriz, isto é, matrizes para as quais $f(x) > 0$ para todo $x \neq 0_n$, não é exclusivo para matrizes simétricas. Podemos tratar de matrizes positivas definidas (ou semipositivas, negativas, seminegativas) não necessariamente simétricas. Porém, normalmente os dois conceitos, simetria e positividade, vêm juntos. Esse é o caso que será tratado ao longo do curso de ALC.

Designamos os conjuntos das matrizes de ordem n simétricas positivas, semipositivas definidas, negativas, seminegativas definidas como $\mathcal{S}_n^{++}, \mathcal{S}_n^{+}, \mathcal{S}_n^{--}, \mathcal{S}_n^{-}$, respectivamente.

Vamos nos concentrar agora nas matrizes simétricas positivas definidas. Estas matrizes são de fundamental importância, pois surgem em diversas aplicações em Ciênica da Computação e nas Engenharias. Muitos modelos de problemas físicos levam a sistemas de equações caracterizados por matrizes simétricas positivas definidas.

Veja que uma matriz $A \in \mathcal{S}_n^{++}$ não pode ser singular (não pode ter determinante nulo). Suponha o contrário: suponha que exista $x \neq 0$ tal que $Ax = 0$. Então $x^T A x = 0$ contrariando a hipótese inicial de que A é positiva definida.

Para demonstrarmos que uma matriz não é positiva definida, basta encontrarmos um vetor x tal que $f(x) \leq 0$. Entretanto, a caracterização da positividade à partir da definição que apresentamos, isto é usando o conceito da energia da matriz, não é trivial. Vamos considerar um caso simples, em que empregamos a energia para chegar à conclusão de que a matriz de interesse é positiva.

Exemplo 16 *Considere a matriz S e sua energia:*
$x^T S x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 8x_1x_2 + 9x_2^2.$

Para mostrar que S é SPD, precisamos mostrar que a expressão algébrica $2x_1^2 + 8x_1x_2 + 9x_2^2$ assume valores positivos para quaisquer valores x_1 e x_2, não simultaneamente nulos. Para mostrar isso vamos reorganizar os quadrados da seguinte forma: $x^T S x = 2x_1^2 + 8x_1x_2 + 9x_2^2 = 2(x_1 + 2x_2)^2 + x_2^2 > 0$. Portanto S é SPD.

À seguir, apresentamos um conjunto de testes equivalentes para se caracterizar a positividade de uma matriz. Se a matriz atender à qualquer um destes testes, atenderá a todos. Portanto, cada um deles serve como instrumento para a caracterização de positividade de uma matriz. Assim sendo, temos liberdade de empregar aquele que nos parecer mais conveniente, dependendo da situação envolvida. Os testes são resumidos na forma do seguinte teorema:

Teorema 2.12.1 *$S \in \mathcal{S}_n^{++}$ (isto é simétrica positiva definida) se e somente se:*

1. *a energia $f(x)$ de S é sempre positiva para qualquer $x \neq 0_n$.*
2. *todos os autovalores de S são positivos.*
   (Recorde-se que toda matriz simétrica possui autovalores reais - nunca seus autovalores são complexos).
3. *S admitir uma fatoração $S = M^T M$, onde M é uma matriz de posto completo n. Em particular, S admite uma fatoração de Cholesky $S = LL^T$, onde L é triangular inferior, com todos os elementos ao longo de sua diagonal sendo positivos.*
4. *os determinantes das submatrizes principais de S são positivos.*
   (A submatriz principal $S_k : k = 1, \dots, n$ de A é a matriz formada pelas primeiras k linhas e colunas de S.)
5. *possui todos os pivôs positivos no processo de Eliminação de Gauss.*

Comentaremos mais sobre os dois últimos destes testes quando discutirmos fatorações básicas e revisarmos a Eliminação de Gauss. No momento, vamos discutir os primeiros três.

Mencionamos (sem ainda provar, isso será feito no último bloco de conteúdo do curso de ALC) que uma matriz real simétrica S admite a chamada fatoração espectral $S = Q\Lambda Q^T = \sum_{i=1}^n \lambda_i q_i q_i^T$ onde $(\lambda_i, q_i) : i = 1, \dots, n$ são os n autopares de S.

Suponha portanto que S seja SPD e possua um autovalor $\lambda_k \leq 0$. Então, tomando $x = q_k$ temos $x^T S x = x^T (\sum_{i=1}^n \lambda_i q_i q_i^T) x = q_k^T (\sum_{i=1}^n \lambda_i q_i q_i^T) q_k = \sum_{i=1}^n \lambda_i q_k^T q_i q_i^T q_k = \lambda_i$ pois $q_k^T q_k = 1, q_k^T q_i = 0, i \neq k$. De fato, os dois testes são equivalentes. Por outro lado, admita que $S = M^T M$, onde M possui posto completo. Claramente $x^T S x = x^T M^T M x = \|Mx\|_2^2 > 0$ para qualquer $x \neq 0$. Por transitividade, os três primeiros testes são equivalentes.

Exercícios Propostos

As questões de 6 a 11 foram adaptadas de [2].

Questão 01: Considere que B seja uma matriz 4×4 sobre a qual aplicamos as seguintes operações:

1. dobrar os valores da coluna 1
2. dividir os valores da linha 3
3. adicionar linha 3 à linha 1
4. trocar as linhas 1 e 4
5. subtrair a linha 2 de cada uma das outras linhas
6. substituir a coluna 4 pela coluna 3
7. eliminar a coluna 1, de forma que a dimensão da matriz resultante seja uma coluna a menos.

Escreva cada matriz utilizada para aplicar as operações descritas anteriormente.

Questão 02: Considere a matriz em blocos $\begin{bmatrix} A & I \\ I & C \end{bmatrix}$, onde I é uma matriz identidade e A possui dimensões $p \times q$. Quais as dimensões de C ?

Questão 03: Considere a matriz em blocos $K = \begin{bmatrix} I & A^T \\ A & 0 \end{bmatrix}$. Quais das seguintes afirmativas são necessariamente verdadeiras (necessariamente verdadeiras significa que são verdadeiras sem nenhuma consideração adicional).

a) K é simétrica.

b) A é quadrada ou larga (a matriz é larga quando não é alta, isto é, número de colunas maior que o número de linhas).

c) A submatriz identidade e a matriz de zeros em K possuem as mesmas dimensões.

d) A submatriz de zeros é quadrada.

Questão 04: Seja A uma matriz $m \times n$ e considere a matriz *empilhada* $S = \begin{bmatrix} A \\ \hline I \end{bmatrix}$, onde I é a matriz identidade.

a) Quando as colunas de S são linearmente independentes ?

b) Quando as linhas de S são linearmente independentes ? Obs: sua resposta pode depender de m, n ou do fato de A ter ou não linhas ou colunas linearmente independentes.

Questão 05: Considere que você necessite avaliar $z = (A + B)(x + y)$ onde A, B são matrizes conformáveis com os vetores x, y. Considere as seguintes alternativas e determine o número de operações de ponto flutuante de cada uma, indicando qual é a mais econômica ao final. Considere que A é $m \times n$ e que x, y são vetores n dimensionais.

a) Primeiro somar $A+B$, então somar $x+y$, e depois aplicar a soma $(A+B)$ na soma $(x+y)$.

b) Distribuir, avaliar cada termo e então somar: $z = Ax + Ay + Bx + By$.

Questão 06: Escolha uma única matriz B (3×3) tal que para toda matriz A:

a) BA $=$ 4A.

b) BA $=$ 4B.

c) BA possui as linhas 1 e 3 de A trocadas, preservando a linha 2 .

Questão 07: Descreva o espaço coluna (em termos de linhas ou planos) das seguintes matrizes:

a) $A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

b) $B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}$.

c) $C = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 0 \end{bmatrix}$.

Questão 08: Considere os vetores $v_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ e $v_2 = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$. Responda às questões abaixo:

a) Estes vetores são linearmente independentes?

b) Eles formam uma base para um espaço $\mathcal{V}$? Qual espaço eles geram?

c) Qual a dimensão do espaço gerado?

d) Quais matrizes A possuem $\mathcal{V}$ como espaço coluna?

d) Descreva todos os vetores v_3 que completam a base para $\mathbb{R}^3$.

Questão 09: As colunas de A são n vetores pertencentes à $\mathbb{R}^m$. Se estes vetores são linearmente independentes, qual é o rank de A? Se estes vetores geram $\mathbb{R}^m$ qual o rank de A? Se estes vetores que geram $\mathbb{R}^m$ são base para $\mathbb{R}^m$, qual a relação entre m, n e rank de A?

Questão 10: Encontre as bases e as dimensões para cada um dos quatro espaços fundamentais associados às matrizes A e B:

a) $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix}$.

b) $B = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 5 & 8 \end{bmatrix}$.

Questão 11: Se $Ax = b$ tem solução e $A^T y = 0$, $(y^T x = 0)$ ou $(y^T b = 0)$? Justifique.

Questão 12: Suponha que A seja uma matriz simétrica $(A^T = A)$. O espaç coluna de A é perpendicular ao espaço nulo de A? Justifique.

Questão 13: Considere a matriz $A = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$ e o vetor $b = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$. Uma solução para o sistema $Ax = b$ é o vetor $x = [1 \;\; 1 \;\; 2]^T$. Responda:

1. Esta solução é única ? Em caso positivo, justifique. Em caso negativo, justifique e apresente uma solução alternativa.

Questão 14: Uma matriz simétrica A possui os autovalores $3, -3$, com os respectivos autovetores $\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} \end{bmatrix}$ e $\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$. Qual é a matriz A ? Esta matriz é positiva definida, negativa definida ou indefinida ?

Questão 15: Considere a matriz V formada pelos autovetores da matriz A acima identificada. O que você pode dizer sobre os quatro espaços fundamentais da matriz A ? Isto é, caracterize todos os quatro espaços fundamentais com suas dimensões.

Questão 16: Suponha que a matriz A das duas questões anteriores tenha o seu autovalor -3 sustituído por 0, preservando os autovetores. O que você pode dizer sobre os quatro espaços fundamentais desta nova matriz A ? Isto é, caracterize todos os quatro espaços fundamentais com suas dimensões. Esta matriz é positiva definida, semi-positiva definida, negativa definida ou semi-negativa definida ?

Questão 17: Responda verdadeiro ou falso e justifique.

1. $\{(x, y) : y = |x|, x \in \mathbb{R}\}$ é um subespaço do $\mathbb{R}^2$.
2. $\{(x, y) : x^2 + y^2 = 0, x, y \in \mathbb{R}\}$ é um subespaço do $\mathbb{R}^2$.
3. $\{(x, y) : x^2 - y^2 = 0, x, y \in \mathbb{R}\}$ é um subespaço do $\mathbb{R}^2$.
4. $\{(x, y) : x - y = 1, x, y \in \mathbb{R}\}$ é um subespaço do $\mathbb{R}^2$.

Questão 18: Sejam W_1, W_2 dois subespaços de um espaço vetorial V e seja

$$W_1 + W_2 = \{w_1 + w_2 : w_1 \in W_1, w_2 \in W_2\}$$

a soma de W_1 e W_2.

1. Mostre que $W_1 \cap W_2$ e $W_1 + W_2$ são subespaços.
2. Mostre que $W_1 \cap W_2 \subseteq W_1 \cup W_2 \subseteq W_1 + W_2$.
3. $W_1 \cup W_2$ é um subespaço ? Justifique.
4. Quando $W_1 \cup W_2$ é um subespaço ?
5. Qual o menor subespaço de V contendo $W_1 \cup W_2$?

Questão 19: Sejam W_1 e W_2 subespaços vetoriais gerados respectivamente pelos v's e u's abaixo indicados.

- $v^1 = (1, 2, -1, -2)^T$, $v^2 = (3, 1, 1, 1)^T$ e $v^3 = (-1, 0, 1, -1)^T$

- $u^1 = (2, 5, -6, -5)^T$, $u^2 = (-1, 2, -7, 3)^T$.

Encontre as dimensões e bases para $W_1 \cap W_2$ e $W_1 + W_2$.

Observação Suponha que $U_1, \dots, U_m$ sejam subespaços de um espaço vetorial V. Cada elemento de $U_1 + U_2 + \cdots + U_m$ pode ser escrito como $u_1 + u_2 + \cdots + u_m$, onde $u_j \in U_j$. Estamos particularmente interessados em casos em que cada vetor em $U_1 + U_2 + \cdots + U_m$ pode ser representado na forma acima, de uma única forma (os u_j's são únicos). Neste caso, dizemos que o vetor é a soma direta destes m subespaços.

Definição: Suponha que $U_1, \dots, U_m$ sejam subespaços de V. A soma $U_1 + U_2 + \cdots + U_m$ é chamada de soma direta, se cada elemento u de $U_1 + U_2 + \cdots + U_m$ puder ser escrito de uma única forma $u_1 + u_2 + \cdots + u_m$, onde cada $u_j \in U_j$. Se $U_1 + U_2 + \cdots + U_m$ é uma soma direta, representamos como $U_1 \oplus U_2 \oplus \cdots \oplus U_m$.

Alguns resultados adicionais:

1. $U + U^\perp$ formam uma soma direta de V, se U é subespaço de V.
2. Se U, W são subespaços de V, então $U + W$ é uma soma direta se e somente se $U \cap W = \{0\}$.
3. Se $U_1, U_2, \dots, U_m$ são subespaços de V então $U_1 + U_2 + \cdots + U_m$ é uma soma direta se e somente se a única forma de escrevermos o vetor 0 (zero) como uma soma de $u_1 + u_2 + \cdots + u_m$ é tomando cada um dos u_j's como o próprio vetor 0.

Questão 20: Responda se a soma dos U's abaixo formam somas diretas.

1. $U_1 = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}$, $U_2 = \{(0, 0, z) \in \mathbb{R}^3 : z \in \mathbb{R}\}$.
2. $U_1 = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}$, $U_2 = \{(0, 0, z) \in \mathbb{R}^3 : z \in \mathbb{R}\}$, $U_3 = \{(0, y, y) \in \mathbb{R}^3 : y \in \mathbb{R}\}$.

Questão 21 Para $k \geq 2$ calcule A^k para:

1. $A = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}$
2. $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$

Questão 22: Responda verdadeiro ou falso e justifique.

1. $A^k = 0$ para todo inteiro positivo $k \geq 2$, então $A = 0$.
2. $A^k = 0$ para algum inteiro positivo k, então $\sum_i a_i i = 0$.
3. Se $\sum_i a_{ii} = 0$, então $|A| = 0$ (determinante de A é zero).
4. Se A, B são similares, $|A| = |B|$.
5. Se A, B são similares, então as duas matrizes possuem os mesmos autovalores.
6. Se A, B possuem os mesmos autovalores, então são similares.
7. Se A, B possuem o mesmo polinômio característico, então possuem os mesmos autovalores.
8. Se A, B possuem os mesmos autovalores, então possuem o mesmo polinômio característico.
9. $diag\{1, 2, \ldots, n\}$ é similar a $diag\{n, n-1, \ldots, 1\}$ (se verdadeira, encontre a matriz B e sua inversa que garantem a similaridade).
10. Se A possui autovalores repetidos, A é não diagonalizável.
11. Se A é unitariamente diagonalizável, então A é normal.
12. Se A possui r autovalores não nulos, então $rank(A) \geq r$.

Questão 23: Por que a matriz identidade I é a única matriz simétrica positiva definida com $\lambda_{min} = \lambda_{max} = 1$? Quais matrizes A são perfeitamente condicionadas, ou seja, $\kappa(A) = 1$? Importante: A matriz identidade é a matriz que possui o menor valor de $\kappa(A)$ possível.

Questão 24: Mostre que A e A^{-1} possuem o mesmo número de condição.

Questão 25: Matrizes ortogonais possuem norma $||Q||_2 = 1$. Se a matriz A pode ser fatorada como $A = QR$, mostre que $||A|| \leq ||R||$ e $||R|| \leq ||A||$. O que podemos concluir?

Capítulo 3

Fatorações Básicas

Nesta capítulo apresentamos algumas fatorações matriciais básicas. Independentemente da forma de $A \in \mathbb{R}^{m\times n}$ (quadrada, quando $m = n$, ou retangular esbelta se $n < m$ ou larga se $n > m$), fatorar uma matriz consiste no processo algorítmico que permite escrever A como produto de outras matrizes com alguma propriede ou mesmo topologia particular, mais convenientes para algum propósito específico.

Por topologia queremos dizer padrão de esparsidade, ou seja, a localização de uma região da matriz onde são autorizados a estarem localizados seus elementos não nulos. Para uma dada topologia, fora dessa região específica, todos os elementos da matriz devem ser nulos. Dois exemplos de topologias de matrizes são matrizes triangulares inferiores e superiores, definidas como as matrizes que tem zeros em todas suas entradas acima e abaixo da diagonal principal, respectivamente. Algumas propriedades de interesse dos fatores podem ser: ter posto completo, terem colunas ou linhas ortonormais, por exemplo.

Ao longo de todo o curso de ALC, vamos discutir diversas fatorações e algoritmos para computá-las. São elas:

1. Fatoração $PA = LU$, onde L é triangular inferior, com a diagonal unitária, U é uma triangular superior e P é uma matriz de permutação.

2. Fatoração de Cholesky $A = R^T R$, onde R é triangular inferior, com a diagonal positiva. A precisa ser SPD.

3. Fatoração Espectral A é simétrica e $A = Q\Lambda Q^T$, onde Λ é uma matriz diagonal com os autovalores de A e Q é ortogonal, como os autovetores de A em suas colunas.

4. Fatoração de Schur $A = QTQ^T$ onde Q é ortogonal, T é uma triangular superior. A matriz A não precisa ser simétrica nem diagonalizável. A diagonal de T armazena os autovalores de A.

5. Fatoração $A = QR$, onde Q é uma matriz com colunas ortonormais e R é uma triangular superior.

6. Decomposição em Valores Singulares (Singular Value Decomposition - SVD): $A = U\Sigma V^T$, $A \in \mathbb{R}^{m\times n}, U \in \mathbb{R}^{m\times n}$, $V \in \mathbb{R}^{n\times n}, \Sigma \in \mathbb{R}^{n\times n}$, $V^TV = I_n, U^TU = I_m$ e a matriz Σ é uma matriz de zeros, exceto pelas r primeiras entradas de sua diagonal, que guarda valores positivos $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_r$, sendo r o posto de A.

Nesta seção vamos nos concentrar nas duas primeiras: fatoração $PA = LU$ e de Cholesky. Assim, os algoritmos que vamos desenvolver fatoram ou decompõem uma matriz A, cujas linhas foram trocadas de ordem por P, na forma $PA = LU$ onde L é uma triangular superior e U uma trinangular superior. A fatoração de Cholesky é um caso particular de $PA = LU$, onde os fatores $L = U^T$ e $P = I$. A Fatoração de Cholesky só se aplica para matrizes simétricas positivas definidas, podendo ser adaptada para fatoração de matrizes de posto incompleto, desde que sejam simétricas positivas semidefinidas. Isto é, podem ser adaptadas para se fatorar $A = R^TR$ onde $R \in \mathbb{R}^{n\times r}$ é triangular inferior com $R_{ii} > 0, i = 1, \ldots, r$. Nessa seção vamos tratar primordialmente o caso em que A é quadrada, embora a fatoração $PA = LU$ pode ser adaptada para produzir fatores para uma matriz retangular.

Nas fatorações que desejamos computar (seja $PA = LU$ ou de Cholesky), as matrizes L, U devem possuir posto completo, exatamente o posto de A (esta sim, pode ter posto incompleto). Ou seja, as fatorações devem revelar o posto da matriz A e apresentar bases para $C(A), C(A^T)$.

Nossa opção é por denominar as duas fatorações estudadas nesta seção como básicas, pelas seguintes razões:

1. Os elementos algorítmicos que empregam são bastante simples,

2. As bases fornecidas para $C(A), C(A^T)$ não são ortonormais,

3. Com estas fatorações somos capazes de resolver boa parte dos sistemas lineares com os quais nos deparamos em aplicações, desde que sejam *bem condicionados*.

A fatoração $PA = LU$ e de Cholesky são adequadas para se resolver sistemas lineares que não sejam mal-condicionados e produzem resultados satisfatórios para tais sistemas. Informalmente, sistemas lineares *bem condicionados* são definidos por matrizes de coeficientes que, no processo de fatoração, não tendem a gerar fatores com grande acúmulo de erros numéricos. Erros numéricos grandes nos fatores se traduzem em erros numéricos grandes, por exemplo, nas soluções dos sistemas lineares onde as matrizes aparecem. Para fatorar matrizes mal-condicionadas, estudaremos fatorações específicas, como a fatoração QR ou SVD, na segunda metade do curso de ALC.

3.1 Razões para se fatorar matrizes

Antes de apresentarmos a primeira das fatorações discutidas aqui, é pertinente mencionar razões para que fatoremos matrizes. Essencialmente, a fatoração revela informação sobre a matriz A. No caso de sistemas dinâmicos lineares, a fatoração revela informações sobre o sistema físico que é representado pela matriz.

Algumas das razões mais importantes para se fatorar uma matriz $A \in \mathbb{R}^{m \times n}$ são:

1. Resolver um ou vários sistemas lineares, possivelmente definidos pela mesma matriz de coeficientes A.

2. Analisar a existência e unicidade das soluções de sistemas lineares.

3. Calcular o determinante de uma matriz quadrada.

4. Conhecer espaços vetoriais associados à matriz: $C(A), C(A^T), N(A), N(A^T)$. Eventualmente, podemos desejar que as bases para estes espaços sejam ortonormais. Para tanto, as fatorações empregadas devem levar estes aspectos em consideração.

5. Obter o espectro de A, ou seja seus autovetores e, eventualmente, seus autovetores.

6. Conhecer os valores singulares de A, assim como seus vetores singulares, de fundamental utilidade para o item abaixo.

7. Aproximar matrizes com muitas colunas ou muitas linhas por matrizes de posto baixo. Com isso podemos resolver problemas aplicados da Ciência da Computação em Otimização, em Inteligência Artificial, em Processamento de Imagens e de Sinais, apenas para citar algumas aplicações.

8. As fatorações de matrizes nos permitem **reformular problemas de Matemática Aplicada**, de uma forma mais conveniente, desde que o problema seja representado ou aproximado por um sistema linear.

Vamos brevemente discutir a primeira destas aplicações. Vamos supor que precisemos resolver vários sistemas lineares $Ax = b$ quadrados de ordem n, que diferem entre si apenas pelo vetor de termos independenes. Isto é, a matriz de coeficientes no sistema linear sempre é A, porém cada novo sistema linear possui um novo vetor b. Vamos supor que posto de A seja completo e que A tenha sido fatorada $PA = LU$. Vamos supor também que disponhamos de um algoritmo capaz de resolver um sistema linear triangular, seja ele triangular superior ou inferior. Para o desenvolvimento a seguir, recorde-se que $P^{-1} = P^T$, quando P é uma matriz de permutação. Veja que se dispomos da fatoração $PA = LU$ podemos substituir LU em PA e escrever:

$$\begin{aligned} Ax &= b \\ PAx &= Pb \\ LUx &= Pb \\ L(Ux) &= Pb \\ Ly &= Pb \quad (3.1) \\ Ux &= y \quad (3.2) \end{aligned}$$

No desenvolvimento acima, usamos o fato de que Ux é uma quantidade desconhecida. Chamamos esta quantidade de y e então resolvemos o sistema linear (3.1) com o algoritmo que supomos dispor. De posse desta quantidade y, resolvemos o sistema (3.2), no qual y é agora conhecido e define o vetor de termos independentes e x é a solução do sistema linear original.

Veja que, desde que disponhamos da fatoração $PA = LU$ e que sejamos capazes de resolver sistemas lineares triangulares, transformamos o problema de encontrar $Ax = b$, onde A é uma matriz quadrada sem nenhuma topologia particular, no problema de resolver dois sistemas lineares, todos os dois definidos por matrizes triangulares: primeiro (3.1) e depois (3.2).

Vamos mostrar ao longo desta seção que o custo computacional de se fatorar $PA = LU$ é $O(n^3)$ e o custo de se resolver um sistema linear onde a matriz de coeficientes é triangular (inferior ou superior, não importa) é $O(n^2)$. Então suponha agora, que precisemos resolver $k << n$ sistemas lineares distintos, definidos pela mesma A. Fatoramos a matriz uma vez, pagando o custo de $O(n^3)$. Para resolver os k sistemas lineares, resolvemos dois sistemas triangulares (3.1) e (3.2), para cada um. Então somamos $k(O(n^2) + O(n^2))$ ao custo computacional. Ao fim, o custo computacional total é $O(n^3 + 2kn^2)$ que é $O(n^3)$ para $2k << n$.

Com a discussão acima, mostramos que podemos usar as fatorações do tipo $PA = LU$ e $A = R^T T$ (Cholesky) para resolvermos sistemas lineares de forma eficiente.

Na próxima seção, vamos apresentar o algoritmo que resolve sistmas lineares triangulares inferiores e superiores, ingrediente importante para resolvermos sistemas lineares mais gerais.

3.1.1 Resolução de sistemas lineares triangulares

Vamos começar esta seção formalizando as definições de matrizes e sistemas lineares triangulares.

1. Uma matriz A, quadrada de ordem n, é triangular inferior se todos elementos acima da diagonal principal são nulos: $a_{ij} = 0$ para todo $i, j : 1 \leq i < j \leq n$.

2. Naturalmente, A é triangular superior se A^T é triangular inferior.

3. Um sistema linear $Ax = b$ é triangular, inferior ou superior, se a matriz de coeficientes A é triangular, inferior ou superior, respectivamente.

Veja dois exemplos de matrizes triangulares.

Exemplo 17 *U é triangular superior e L é triangular inferior.*

$$U = \begin{bmatrix} 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix} \quad L = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 3 & 1 & -1 & 0 \\ 4 & 1 & -3 & 3 \end{bmatrix}$$

Como resolveríamos o sistema linear $Lx = b$, onde b é um vetor qualquer, por exemplo $b = (2\ 3\ 2\ 9)^T$ e L é a matriz do exemplo acima ? Vamos explorar a esparsidade da matriz de coeficientes L e reescrever $Lx = b$ de forma mais conveniente, de forma que o algoritmo fique evidente. Usando o fato de que $l_{ij} = 0 : j > i$ para uma linha i do sistema linear $Lx = b$, temos:

$$\begin{aligned} \sum_{j=1}^{n} l_{ij} x_j &= b_i && i = 1, \ldots, n \\ \sum_{j=1}^{i} l_{ij} x_j + \sum_{j=i+1}^{n} l_{ij} x_j &= b_i && i = 1, \ldots, n \\ \sum_{j=1}^{i-1} l_{ij} x_j + l_{ii} x_i &= b_i && i = 1, \ldots, n \\ x_i &= \frac{b_i - \sum_{j=1}^{i-1} l_{ij} x_j}{l_{ii}} && k = 1, \ldots, n \end{aligned}$$

Portanto, se calcularmos $x_1, x_2, \ldots, x_{i-1}$ nesta ordem, usando a expressão (3.3) para $k = i$

$$x_k = \frac{b_k - \sum_{j=1}^{k-1} l_{kj} x_j}{l_{kk}}, \tag{3.3}$$

podemos calcular x_i com as entradas já calculadas anteriormente $x_1, x_2, \ldots, x_{i-1}$, necessárias na expressão (3.3). Este algoritmo, conhecido como **Algoritmo de Substituições Sucessivas**, é apresentado abaixo na Figura 3.1.

```
function [y] = SubsSucessivas(L,b,n)
  for i=1:n
    soma = 0.0;
    for k = 1:i-1
      soma = soma + L(i,k)*y(k)
    end
    if (L(i,i) <> 0.0)
       y(i) = (b(i)-soma)/L(i,i);
    else
       printf('Matriz L e singular \n')
       break
    end
  end
endfunction
```

Figura 3.1: Algoritmo de Substituições Sucessivas.

Observe que se $L_{ii} = 0$ para algum i, a matriz L e A são singulares. Posteriormente vamos discutir como tratar o caso singular. Aqui, caso isso ocorra, o algoritmo acusa a singularidade da matriz e interrompe sua execução. Veja o exemplo de aplicação do algoritmo.

Exemplo 18 `--> L,b`

```
 L  =
   2.   0.   0.   0.
  -1.   2.   0.   0.
   3.   1.  -1.   0.
   4.   1.  -3.   3.
 b  =
   2.
   3.
   2.
   9.
```

```
--> [y] = SubsSucessivas(L,b,size(L,1));
--> y'
 ans  =
   1.   2.   3.   4.
```

Vamos agora mostrar que a complexidade do algoritmo de Substituições Sucessivas é $O(n^2)$. O trecho de interesse do algoritmo, isto é, aquele que define sua complexidade é indicado na Figura (3.2).

```
for i=1:n
  soma = 0.0;
  for k = 1:i-1
    soma = soma + L(i,k)*y(k)
  end
  y(i) = (b(i)-soma)/L(i,i)
end
```

Figura 3.2: Trecho de interesse do algoritmo de Substituições Sucessivas.

Para avaliar a função de complexidade do algoritmo, vamos *contar* as operações aritméticas de ponto flutuante realizadas pelo algoritmo. Não consideramos as operações de incremento e comparação das variáveis inteiras, necessárias para as estruturas `for` ou `while`, por exemplo. Cada operação aritmética de ponto flutuante ($+$,-,$\times$,$\div$) tem o mesmo custo unitário, 1. Lembramos que $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$. Para um determinado valor de i fixo, a instrução

```
soma = soma + L(i,k)*y(k)
```

é executada $g(i) = (\sum_{k=1}^{i-1} 1)$ vezes, dentro da estrutura de controle

```
for k = 1:i-1
```

para este valor de i fixo. Veja que esta quantidade $g(i) = \sum_{k=1}^{i-1} 1$ é uma função de i e assim sendo, para cada valor de i distinto, teremos uma contribuição distinta. Cada vez que a instrução for executada, são realizadas uma soma e uma multiplicação, ou seja, incorremos em um custo de 2 operações.

Agora, como podemos escrever o número de vezes que a estrutura de controle

```
for k = 1:i-1
```

é chamada ? Veja que esta estrutura de controle ela *dentro* de uma estrutura de controle mais externa

```
for i=1:n
```

que controla os valores admissíveis de i. Se desejamos o custo computacional total relativo à instrução

```
soma = soma + L(i,k)*y(k)
```

precisamos avaliar $g(1)+g(2)+\cdots+g(n)$. Isso porque a estrutura mais interna será executada tantas vezes quantos forem os valores assumidos de i. Então podemos escrever a função de complexidade $f(n)$ do algoritmo como:

$$
\begin{aligned}
f(n) &= 2\sum_{i=1}^{n} g(i) \\
&= 2\sum_{i=1}^{n}\sum_{k=1}^{i-1} 1
\end{aligned}
$$

Para resolver um somatório como o acima, começamos a explicitar o resultado dos somatórios mais internos, pois estes assumem que valores fixos para as variáveis foram definidos nos somatórios anteriores. Então temos:

$$
\begin{aligned}
f(n) &= 2\sum_{i=1}^{n}\sum_{k=1}^{i-1} 1 \\
&= 2\sum_{i=1}^{n}(i-1) \\
&= 2\left(\frac{n(n+1)}{2} - n\right) \\
&= n(n+1) - 2n \\
&= n(n-1)
\end{aligned}
$$

Portanto, a instrução que estudadmos adiciona $n(n-1)$ operações aritméticas de ponto flutuante (flops) ao custo computacional do algoritmo.

O custo total do algoritmo deve levar em conta também o custo adicionado pela instrução

```
y(i) = (b(i)-soma)/L(i,i)
```

que é $\sum_{i=1}^{n} 2 = 2n$. Portanto, o custo total de substituições Sucessivas é $n(n-1) + 2n = n(n+1)$, ou seja, seu custo pertence à classe de complexidade $O(n^2)$.

Vamos agora discutir a resolução de **sistemas lineares triangulares superiores**. Uma vez que já discutimos os sistemas lineares inferiores e o algoritmo de substituições que o resolve detalhadamente, nossa exposição do caso triangular superior é mais breve. Tudo é análogo ao caso triangular inferior. Iniciamos com um exemplo de matriz U triangular superior.

$$U = \begin{bmatrix} 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

Observe que de forma análoga ao caso triangular inferior, U é triangular superior significa que $u_{ij} = 0 : j < i$. Feita esta observação, vamos deduzir a expressão do termo x_i que nos permite construir o algoritmo de resolução. Para tanto, considere uma linha i fixa do sistema:

$$\begin{aligned} \sum_{j=1}^{n} u_{ij}x_j &= y_i & i &= n, n-1, \ldots, 1 \\ \sum_{j=1}^{i-1} u_{ij}x_j + \sum_{j=i}^{n} u_{ij}x_j &= y_i & i &= n, n-1, \ldots, 1 \\ u_{ii}x_i + \sum_{j=i+1}^{n} u_{ij}x_j &= y_i & i &= n, n-1, \ldots, 1 \end{aligned}$$

Desta forma, a expressão que permite deduzir o algoritmo é dada por:

$$x_k = \frac{y_k - \sum_{j=k+1}^{n} u_{kj}x_j}{u_{kk}}$$

Cabe destacar que o algoritmo para resolvermos $Ux = y$ opera sobre o sistema linear na ordem inversa das linhas, isto é, primeiro na linha n, depois na linha $n-1$ e assim por diante até trabalhar a linha de índice 1. Isso porque para calcular a grandeza x_k é necessário dispor das incógnias x_n, x_{n-1}, x_{k+1} já calculadas. Por esta razão, o algoritmo é chamado de Algoritmo de Substituições Retroativas e é apresentado na Figura 3.3.

O exemplo abaixo ilustra o uso do algoritmo.

Exemplo 19 `--> U,y'`

```
 U  =

    3.    2.    1.    0.
    0.    1.    2.    3.
    0.    0.   -2.    1.
```

```
function [x] = SubsRetroativas(U,y,n)
  for i=n:-1:1
    soma = 0.0;
    for k = i+1:n
      soma = soma + U(i,k)*x(k)
    end
    if (U(i,i) <> 0)
     x(i) = (y(i) - soma)/U(i,i);
    else
     printf('Matriz U e singular \n');
    end
  end
endfunction
```

Figura 3.3: Algoritmo de Substituições Retroativas.

```
   0.    0.    0.    4.
 ans  =


  -10.   10.   1.   12.


--> x = SubsRetroativas(U,y,size(y,1));


--> x'
 ans  =


  -3.  -1.   1.   3.
```

3.1.2 Resolvendo sistemas lineares a partir de sistemas triangulares

Nesta seção, vamos ilustrar como podemos usar os algoritmos de Substitições Sucessivas (Figura 3.1) e Retroativas (Figura 3.3) para resolver, em duas etapas, um sistema linear cuja matriz de coeficientes tenha sido fatorada.

Como ainda não apresentamos como produzir a fatoração $PA = LU$, por hora, vamos empregar o algoritmo para fatoração $PA = LU$ disponível no `Scilab` para obtermos os fatores necessários. Para o exemplo que segue, empregamos a função `ResolveParTriangulares(A,b)`, descrita na Figura 3.4.

Exemplo 20 *Este exemplo ilustra a chamada da funçõ* `lu` *do scilab. Considere então a matriz A e seus fatores.*

```
function [x,L,U,P] = ResolveParTriangulares(A,b)
    [L,U,P] = lu(A)
    [m,n] = size(A)
    [y] = SubsSucessivas(L,P*b,n)
    [x] = SubsRetroativas(U,y,n)
endfunction
```

Figura 3.4: Resolução de um sistema linear $Ax = b$ por meio de dois sistemas lineares triangulares.

```
A  =
   6.     4.    2.     0.
  -3.     0.    3.     6.
   9.     7.    7.     2.
   12.    9.    12.    12.
b' =
   2.    3.    2.    9.
-->[x,L,U,P] = ResolveParTriangulares(A,b)
 x  =
   1.480D-16
   1.0000000
  -1.0000000
   1.0000000
 L  =
   1.      0.           0.    0.
  -0.25    1.           0.    0.
   0.5    -0.2222222    1.    0.
   0.75    0.1111111    1.    1.
 U  =
   12.    9.      12.          12.
   0.     2.25    6.           9.
   0.     0.     -2.6666667   -4.
   0.     0.      0.          -4.
 P  =
   0.    0.    0.    1.
   0.    1.    0.    0.
   1.    0.    0.    0.
   0.    0.    1.    0.
```

Agora vamos usar os fatores para resolver o sistema linear $Ax = b$, onde $b =$

$(2, 3, 2, 9)^T$.

3.2 Fatoração $A = LU$ e $PA = LU$

3.2.1 Eliminação de Gauss e Fatoração $A = LU$

Nesta seção, vamos recordar o método de Eliminação de Gauss, assumindo que não seja necessário efetuar trocas de linhas do sistema linear. Vamos mostrar que a Eliminação de Gauss produz os fatores desejados U e L que desejamos. Por razões didáticas, de início não faremos uso de trocas de linhas do sistema linear. Esta é uma hipótese não realista, é adotada aqui apenas para facilitar a exposição inicial. Com isso, faremos a fatoração $A = LU$ de A e, posteriormente, ao permitirmos a troca de linhas de A, faremos a fatoração $PA = LU$.

A ideia da Eliminação de Gauss é transformar o sistema linear $Ax = b$ em outro sistema linear, $Ux = y$, equivalente ao primeiro. Dois sistemas lineares equivalentes são indicados como

$$Ax = b \sim Ux = y,$$

que significa que toda solução de $Ax = b$ também é solução de $Ux = y$ e vice-versa.

Para transformar $Ax = b$ no equivalente $Ux = y$ podemos usar as seguintes operações linha elementares, isto é, operações realizadas sobre as linhas do sistema que não alteram seu conjunto de soluções:

(T1) Troca da ordem de duas linhas do sistema linear. Como mencionamos, nesta seção, vamos assumir que não será necessário aplicar T1.

(T2) Multiplicação de uma linha por uma constante não nula.

(T3) Substituição de uma linha do sistema pela soma da própria linha mais um múltiplo de outra linha do sistema linear.
Sendo i a linha que será substituída e j a linha que será multiplicada por m, temos que

$$\sum_{k=1}^{n} a_{ik} x_k = b_i$$

é substituída por

$$\sum_{k=1}^{n} (a_{ik} + m a_{jk}) x_j = b_i + m b_j.$$

Vamos recordar a Eliminação de Gauss por meio de um exemplo. O resultado será o sistema $Ux = y$, onde este vetor y é exatamente a solução do sistema linear $Ly = b$ que obteríamos, caso já dispuséssemos da fatoração $A = LU$ da matriz.

Todas as operações que fizermos sobre as linhas de A, replicaremos nas linhas de b. Isso é facultativo, pois podemos obter o resultado destas opearações sobre b, posteriormente, obtendo o y que resolve $Ly = b$.

Ao longo da aplicação da Eliminação de Gauss, adotaremos a notação de representar o sistema linear sendo transformado por $[A^j|b^j] \in \mathbb{R}^{n\times(n+1)}$, onde j indica o índice da operação de pivoteamento completa realizada até aquele momento do algoritmo (não confundir com a potência j da matriz - não é o caso aqui). No início do procedimento, antes de qualquer operação, temos $A^0 = A, b^0 = b$, de forma que $[A^0|b^0]$ nada mais é do que matriz A expandida em uma coluna por b. Ao longo da primeira iteração, $j = 1$, operamos sobre $[A^0|b^0]$ e ao final da operação completa obtemos então $[A^1|b^1]$.

Cada iteração do algoritmo será indexada por uma *op j*, *operação de pivoteamento j*, que compreenderá um conjunto de operações T2 e T3, para transformar uma coluna do sistema linear em uma coluna de um sistema triangular superior. Ao final da $j-$ésima op completa, temos o sistema equivalente $[A^j|b^j]$. Faremos uma op para cada um das colunas de A, exceto a última, visando transformar a matriz A^0 em uma triangular superior: ao final, A^{n-1} será triangular superior. Ou seja, indexaremos j de 1 até $n-1$, inclusive. A ideia de indexação se repete para as demais iterações.

Em cada operação de pivoteamento faremos uso de um multiplicador m_{ij} que significa: o mutiplicador associado à $i-$ésima linha de $[A^{j-1}|b^{j-1}]$, na $j-$ésima (op), ou equivalentemente na $j-$ésima coluna de A^j. Este multiplicador deve ser escolhido para criar zeros nas linhas abaixo da entrada $(A^{j-1})_{jj}$. Para um dado $j = 1, \ldots, n-1$, calculamos $m_{i,j}$ para $i = j+1, 2, \ldots, n$.

Para um dado $j = 1, \ldots, n-1$, m_{ij} é dado por:

$$m_{ij} = -\frac{a_{ij}^{j-1}}{a_{jj}^{j-1}}, i = j+1, \ldots, n. \tag{3.4}$$

Veja que este multiplicador é calculado de forma que a equação abaixo seja satisfeita:

$$m_{ij}a_{jj}^{j-1} + a_{ij}^{j-1} = 0, , i = j+1, \ldots, n. \tag{3.5}$$

A linha j de $[A^{j-1}|b^{j-1}]$ é chamada de pivot. O elemento pivot, a_{jj}^{j-1}, não pode ser zero, caso contrário o método falha. Aqui nesta seção, estamos assumindo a hipótese otimista que isso não ocorrerá. Múltiplos desta linha devem ser somados às linhas de índice $j+1, j+2, \ldots, n$ para que a j-ésima coluna de $[A^{j-1}|b^{j-1}]$ seja transformada em uma coluna de uma triangular superior. Essa é a ideia da operação T3 que faremos sobre as linhas do sistema.

Exemplo 21 *Usando a Eliminação de Gauss (sem pivoteamento parcial, ou troca de colunas), transformar o sistema linear* $Ax = b$ *em* $Ux = y$*, para* $A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix}$ $b = \begin{bmatrix} 1 \\ 3 \\ 7 \\ 3 \end{bmatrix}$.

op 0 , $j = 0$ *(representação do sistema original)*

$$[A|b] = [A^0|b^0] = \left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 1 \\ 4 & 3 & 3 & 1 & 3 \\ 8 & 7 & 9 & 5 & 7 \\ 6 & 7 & 9 & 8 & 3 \end{array}\right]$$

op 1 , $j = 1$*, primeira operação de pivoteamento.*

Veja que para $j = 1$ *precisamos multiplicar a linha 1 do sistema por* $m_{2,1} = -2, m_{3,1} = -4, m_{4,1} = -3$*, respectivamente, para criarmos zeros nas posições* $a^0_{2,1}, a^0_{3,1}, a^0_{4,1}$*, respectivamente. Empregando estes multiplicadores, ao final da primeira operação de pivoteamento (op) temos o sistema:* $[A^1|b^1] = \left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 3 & 5 & 5 & 3 \\ 0 & 4 & 6 & 8 & 0 \end{array}\right]$. *A primeira coluna de* A^1 *é uma coluna de uma triangular superior.*

op 2 , $j = 2$. *Para esta operação, temos* $m_{3,2} = -3, m_{4,2} = -4$. *Ao final da mesma, temos o sistema* $[A^2|b^2] = \left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 2 & 4 & -4 \end{array}\right]$. *As duas primeiras colunas de* A^2 *são colunas de uma triangular superior.*

op 3 , $j = 3$. *Para esta terceira operação, temos* $m_{4,3} = -1$. *Ao fina da op 3, temos:* $[A^3|b^3] = \left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 2 & -4 \end{array}\right]$.

Veja que ao final de $n-1 = 3$ *operações, obtivemos o sistema linear triangular* *superior acima. Para resolver o sistema linear, resolvemos por substituições sucessivas e encontramos* $x = \begin{bmatrix} -1 & 1 & 2 & -2 \end{bmatrix}^T$

3.2.2 Reinterpretando a Eliminação de Gauss como um conjunto de transformações lineares

Cada operação T3, que representa a multiplicação da linha pivot j por m_{ij}, a soma do resultado com a linha i e sua substituição pela soma, pode ser representado por um produto de uma matriz M_j por $[A^{j-1}, b^{j-1}]$. Atenção aqui: M_j é a matriz de multiplicadores e não uma coluna desta matriz.

A matriz M_j, chamada de matriz de multiplicadores na $j-$ésima op, é uma matriz que difere da matriz identidade apenas pela sua $j-$ésima coluna, nos elementos das linhas $i : i > j$, isto é (abaixo da diagonal principal), que recebem os multiplicadores m_{ij} calculados através de (3.4). A título de ilustração, veja a forma da matriz de multiplicadores na primeira op, para uma matriz A^{j-1}, com 4 linhas $M_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ m_{21} & 1 & 0 & 0 \\ m_{31} & 0 & 1 & 0 \\ m_{41} & 0 & 0 & 1 \end{bmatrix}$

No caso do exemplo que ilustramos, a instanciação da matriz M_1 corresponde à matriz

$$M_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{bmatrix}.$$

Observe que a matriz M_j é uma triangular inferior, com diagonal unitária. Seu determiante corresponde ao produto dos elementos em sua diagoanal, de forma que $\det(M_j) = 1$ e a matriz admite inversa.

Exercício 3.2.1 *Verifique que a matriz inversa de M_j, M_j^{-1} é uma matriz identidade, exceto pela $j-$ ésima coluna, que recebe o simétrico das entradas de M_j, nas linhas abaixo da diagonal principal:*

$$M_j^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -m_{j+1,j} & 1 & 0 & 0 \\ 0 & 0 & 0 & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & -m_{n,j} & 0 & 0 & 1 \end{bmatrix}$$

Definida a matriz M_j, veja que cada op pode ser represetada da seguinte forma:

$$M^j[A^{j-1}, b^{j-1}] = [A^{j-1}, b^{j-1}]. \tag{3.6}$$

No processo de Eliminação de Gauss que ilustramos anteriormente, não fizemos isso explicitamente. Nossa ideia é usar a representação (3.6), para obtermos os fatores.

Veja o processo completo associado à Eliminação de Gauss que ilustramos.

Exemplo 22 *O objetivo deste exemplo é ilustrar, passo a passo, as três transformações lineares que triangularizam a matriz A do Exemplo 21.*

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & m_{4,3} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & m_{2,3} & 1 & 0 \\ 0 & m_{2,4} & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ m_{2,1} & 1 & 0 & 0 \\ m_{3,1} & 0 & 1 & 0 \\ m_{4,1} & 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 1 \\ 4 & 3 & 3 & 1 & 3 \\ 8 & 7 & 9 & 5 & 7 \\ 6 & 7 & 9 & 8 & 3 \end{array}\right] =$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 1 \\ 4 & 3 & 3 & 1 & 3 \\ 8 & 7 & 9 & 5 & 7 \\ 6 & 7 & 9 & 8 & 3 \end{array}\right] =$$

$$\left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 2 & -4 \end{array}\right].$$

Vamos agora formalizar o processo de Eliminação de Gauss, de forma que obtenhamos os fatores L, U desejados na fatoração. Veja que realizamos $n-1$ operações, uma por coluna de A, e que por construção obtivemos:

$$M_{n-1}M_{n-2}\ldots M_2M_1[A|b] = [U|y]$$
$$\Rightarrow$$
$$M_{n-1}M_{n-2}\ldots M_2M_1A = U$$
$$M_{n-1}M_{n-2}\ldots M_2M_1b = y$$

Veja o processo de pré-multiplicar A pelas matrizes $M_1, M_2, \ldots, M_{n-1}$, nesta ordem, gerou uma matriz U. Em linguagem de ALC, *triangularizamos* A. O mesmo conjunto de transformações lineares aplicados em b gerou o vetor y que resolve o sistema (3.1) (até aqui a matriz P é a identidade, pois não trocamos linhas de A).

Vamos agora considerar a matriz $M = M_{n-1}M_{n-2}\dots M_2M_1$. Observe que M corresponde ao produto de $n-1$ matrizes de multiplicadores, cada qual com determinante igual a 1. Portanto $\det(M) = 1^{n-1} = 1$ e M admite inversa. Sua inversa é dada por

$L = M^{-1} = M_1^{-1}M_2^{-1}\dots M_{n-2}^{-1}M_{n-1}^{-1}$

A matriz L é uma triangular inferior, com diagonal unitária, que difere da matriz identidade, pois as entradas abaixo da diagonal principal recebem o simétrico dos multiplicadores, isto é, $L_{ij} = -m_{ij}$.

Exemplo 23 *Verificar que $M^{-1} = L$ é uma triangular inferior com diagonal unitária, e suas entradas abaixo da diagonal são os simétricos dos multiplicadores dados por* (3.4).

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -m_{21} & 1 & 0 & 0 & 0 \\ -m_{31} & -m_{32} & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & 1 & 0 \\ -m_{n,1} & -m_{n,2} & \cdots & -m_{n,n-1} & 1 \end{bmatrix} \qquad (3.7)$$

A verificação se dá analisando os produtos das matrizes que definem L, dos fatores mais à direita para os mais à esquerda. Veja que o resultado do primeiro produto $M_{n-2}^{-1}M_{n-1}^{-1}$:

$$\begin{aligned} M_{n-2}^{-1}M_{n-1}^{-1} &= \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -m_{n-1,n-2} & 1 & 0 \\ 0 & 0 & -m_{n,n-2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -m_{n,n-1} & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -m_{n-1,n-2} & 1 & 0 \\ 0 & 0 & -m_{n,n-2} & -m_{n,n-1} & 1 \end{bmatrix} \end{aligned}$$

Observe que a matriz resultante do produto $M_{n-2}^{-1}M_{n-1}^{-1}$ tem a seguinte forma:

- *Suas primeiras $n-3$ colunas são colunas de uma identidade de ordem n. A* *mesma observação se aplica para a última coluna, que também é uma coluna de uma identidade.*

- *As demais colunas, de índices $n-2$ a $n-1$, são as colunas de índices $n-2$ e $n-1$, respectivamente de M_{n-2}^{-1} e de M_{n-1}^{-1}.*

O padrão a ser observado é que a matriz correspondente ao produto $M_{n-k}^{-1}\cdots M_{n-1}^{-1}$ (para algum $k \geq 2$) possui suas primeiras $n-k-1$ primeiras colunas, assim como a última, de índice n, como colunas da identidade. Além disso, as colunas de índice $n-k$ até $n-1$ são as colunas de índices $(n-k),\cdots,(n-1)$ de $M_{n-k}^{-1},\ldots,M_{n-1}^{-1}$, respectivamente.

Assim, ao incorporamos o próximo fator no cálculo de L, isto é, ao calcularmos $M_{n-k-1}^{-1}(M_{n-k}^{-1}\cdots M_{n-1}^{-1})$ preservamos as primeiras $n-k-1$ colunas de M_{n-k-1}^{-1}, incluindo sua coluna de índice $n-k-1$, que é a única de suas colunas que difere da identidade e que armazena os simétricos dos multiplicadores da op $n-k-1$. Todas as demais colunas de M_{n-k-1}^{-1} são colunas da identidade. Portanto, o produto $M_{n-k-1}^{-1}(M_{n-k}^{-1}\cdots M_{n-1}^{-1})$ para as colunas de índice $n-k$ em diante será uma combinação linear das colunas da identidade por pesos que vêm das colunas de $(M_{n-k}^{-1}\cdots M_{n-1}^{-1})$. Já as colunas de $M_{n-k-1}^{-1}(M_{n-k}^{-1}\cdots M_{n-1}^{-1})$ de índice $n-k-1$ ou menor serão colunas da identidade. Esta é a invariante do processo, que se repete até incorporarmos o fator M_1^{-1} ao produto $M_2^{-1}\ldots M_{n-1}^{-1}$, para $k=n-2$.

Para ilustrar a incorporação de mais um fator aos já avaliados, considere o resultado de $M_{n-3}^{-1}M_{n-2}^{-1}M_{n-1}^{-1}$:

$$
\begin{aligned}
M_{n-3}^{-1}M_{n-2}^{-1}M_{n-1}^{-1} &= \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -m_{n-2,n-3} & 1 & 0 & 0 \\ 0 & 0 & -m_{n-1,n-3} & 0 & 1 & 0 \\ 0 & 0 & -m_{n,n-3} & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -m_{n-1,n-2} & 1 & 0 \\ 0 & 0 & 0 & -m_{n,n-2} & -m_{n,n-1} & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -m_{n-2,n-3} & 1 & 0 & 0 \\ 0 & 0 & -m_{n-1,n-3} & -m_{n-1,n-2} & 1 & 0 \\ 0 & 0 & -m_{n,n-3} & -m_{n,n-2} & -m_{n,n-1} & 1 \end{bmatrix}
\end{aligned}
$$

Em resumo, desde que armazenados os multiplicadores empregados no processo, a Eliminação de Gauss produz os fatores L, U de A, onde L é triangular inferior com diagonal unitária e U é triangular superior. Veja:

$$\begin{aligned} M_{n-1}M_{n-2}\dots M_2M_1[A|b] &= [U|y] \\ A &= M_1^{-1}M_2^{-1}\dots M_{n-2}^{-1}M_{n-1}^{-1}U \\ M_1^{-1}M_2^{-1}\dots M_{n-2}^{-1}M_{n-1}^{-1} &= L \\ A &= LU \end{aligned}$$

3.2.3 Visão por colunas da Fatoração $A = LU$

Uma maneira bastante conveniente de se formalizar as operações da Eliminação de Gauss é verificar que, a cada op, subtraímos de A uma matriz de rank 1, e depois operamos sobre a diferença, repetindo o processo, até a última op. Esta visão alternativa é chamada de visão coluna, ou visão de soma de matrizes de posto 1, para a fatoração.

Para deduzirmos esta visão alternativa, vamos partir da fatoração $A = LU$ que resulta da Eliminação de Gauss, escrevendo-a como uma soma de $n = posto(A)$ matrizes de posto 1.

$$\begin{aligned} A &= LU \\ A &= \begin{bmatrix} | & | & \dots & | \\ L_1 & L_2 & \dots & L_n \\ | & | & \dots & | \end{bmatrix} \begin{bmatrix} u_1^T \\ u_2^T \\ \dots \\ u_n^T \end{bmatrix} \\ &= \sum_{j=1}^{n} L_j u_j^T, \end{aligned}$$

onde L_j corresponde à $j-$ésima coluna de L e u_j^T é a $j-$ésima linha de U. O último termo da soma, $L_n u_n^T$ não foi explicitamente calculado na Eliminação de Gauss, pois não era necessário (já que a última coluna de A^{n-1} sempre é uma coluna de uma triangular superior de ordem n), sendo trivialmente dado por $e_n u_n^T$ (recorde-se de nossa notação que utiliza e_i para representar um vetor $n-$dimensional de zeros, exceto pela $i-$ésima entrada que é 1). Veja que para todo índice $j = 1, \dots, n$, as linhas u_j^T são as linhas pivotais: a linha j da matriz A^{j-1} obtida ao longo da Eliminação de Gauss. Já as colunas de L_j satisfazem: $l_{ij} = 0$ para $i < j$, $l_{jj} = 1$, $l_{ij} = -m_{ij}$.

Naturalmente, quando aplicarmos a Eliminação, não temos todos os termos $L_j u_j^T$ para $j = 1, \dots, n$. Estes termos são descobertos ao longo do processo. Porém, veja que para um determinado índice j de op, por exemplo, $j = 1$, temos

$$\begin{aligned}
A &= \sum_{k=1}^{n} L_k u_k^T \\
&= L_1 u_1^T + \sum_{k=2}^{n} L_k u_k^T \\
A - L_1 u_1^T &= \sum_{k=2}^{n} L_k u_k^T \\
&= A2
\end{aligned}$$

A matriz $A2 = \sum_{k=2}^{n} L_k u_k^T$ é uma soma de $n-2$ matrizes de posto 1. Novamente salientamos que, só ao final da Eliminação de Gauss é que dispomos dos demais termos $L_2 u_2^T, \ldots, L_n u_n^T$ na fatoração. Mas podemos aplicar a mesma ideia agora à matriz $A2$. Veja que a matriz $A2$ corresponde às últimas $n-1$ colunas e linhas da matriz A^1 que obtivemos ao final da primeira op na Eliminação de Gauss. Então, a linha u_2^T é a linha 2 de A^1 e a coluna L_2 é obtida calculando-se os multiplicadores pertinentes à segunda op. Incorporando mais uma op ao processo temos:

$$\begin{aligned}
A - L_1 u_1^T &= \sum_{k=2}^{n} L_k u_k^T \\
&= A2 \\
A - L_1 u_1^T - L_2 u_2^T &= \sum_{k=3}^{n} L_k u_k^T \\
&= A3
\end{aligned}$$

Então, a matriz Aj para $j = 1, \ldots, n$ é simplesmente $A - \sum_{k=1}^{j-1} L_k u_k^T$, onde $A1 = A$. Repetindo o processo por $n-1$ ops, temos que

$$A - \sum_{j=1}^{n-1} L_j u_j^T = e_n u_n^T,$$

o que nos permite escrever $A = \sum_{j=1}^{n-1} L_i u_i^T + e_n u_n^T$, sendo a fatoração final.

Exemplo 24 *Vamos ilustrar a visão de colunas da Fatoração, por meio do exemplo da seção anterior, interpretando agora cada op como a subtração de uma matriz de posto* 1 *que é dada pelo produto externo de uma coluna de* L *por uma linha de* U *(a linha pivotal de* A^{j-1}*).*

1. *primeira op*

   - $A = \begin{bmatrix} 1 & \textit{vez a linha pivot 1} \\ l_{21} & \textit{vez a linha pivot 1} \\ l_{31} & \textit{vez a linha pivot 1} \\ l_{41} & \textit{vez a linha pivot 1} \end{bmatrix} + A2 = \begin{bmatrix} 1 & \textit{vez a linha pivot 1} \\ l_{21} & \textit{vez a linha pivot 1} \\ l_{31} & \textit{vez a linha pivot 1} \\ l_{41} & \textit{vez a linha pivot 1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix}$

   - *Recorde-se dos valores que foram calculados para os multiplicadores que permitem escrever a primeira coluna de L com as entrada* $l_{21} = a_{21}/a_{11}, l_{31} = a_{31}/a_{11}, l_{41} = a_{41}/a_{11}$. *Recorde-se também que as entradas* l_{21}, l_{31}, l_{41} *são os simétricos de* m_{21}, m_{31}, m_{41}.

   - *Então temos* $\begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ 3 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 3 & 5 & 5 \\ 0 & 4 & 6 & 8 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 2 & 2 & 0 \\ 8 & 4 & 4 & 0 \\ 6 & 3 & 3 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 3 & 5 & 5 \\ 0 & 4 & 6 & 8 \end{bmatrix}$

2. *segunda op:*

   - $A2 = L_2 u_2^T + A3$.

   - $A2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 3 & 5 & 5 \\ 0 & 4 & 6 & 8 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 4 \end{bmatrix}$

   - $A2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 3 & 5 & 5 \\ 0 & 4 & 6 & 8 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 3 & 3 & 3 \\ 0 & 4 & 4 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 4 \end{bmatrix}$

3. *terceira op temos:*

   - $A3 = L_3 u_3^T + A4$.

- $A3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 2 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

4. o quarto termo, referente à uma quarta op não necessária na visão por linhas, mas necessária na visão colunas, corresponde a:

- $A4 = L_4 u_4^T = e_4 u_4^T = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 2 \end{bmatrix}$

3.2.4 Complexidade computacional de $A = LU$

Na Figura 3.5, apresentamos o Algoritmo que produz a fatoração $A = LU$, à partir da Eliminação de Gauss, sem ainda incorporar a troca de linhas. Vamos discutir sua complexidade computacional e mostrar que é um algoritmo na classe $O(n^3)$. O algoritmo usa a notação $B(p : k, l : m)$ para representar a submatriz de B que contém as linhas de índices p até k e colunas de índices l até m de B.

```
function [U,L] = EliminacaoGauss(A,n)
U = A
L = eye(n,n)
for j=1:n-1
  for i = j+1:n
    L(i,j) = U(i,j)/U(j,j)
    U(i,j:n) = U(i,j:n) - L(i,j)*U(j,j:n)
  end
end
endfunction
```

Figura 3.5: Algoritmo para Fatoração $A = LU$.

A instrução mais relevante para a complexidade computacional é

```
U(i,j:n) = U(i,j:n) - L(i,j)*U(j,j:n)
```

Observe que o termo

```
L(i,j)*U(j,j:n)
```

na instrução em estudo corresponde ao produto de um escalar por um vetor $n - j$ dimensional. Portanto, requer $n-j$ operações aritméticas de ponto flutuante. Vamos analisar apenas a complexidade adicionada por esta instrução, que é dada então pela soma $\sum_{j=1}^{n-1}\sum_{i=j+1}^{n}\sum_{k=j+1}^{n} 1$. Para o desenvolvimento que segue, recorde-se do valor das somas

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

e também

$$\sum_{i=1}^{n} i^2 = \frac{n}{6}(n+1)(2n+1).$$

Então, a função que determina o número de vezes que a instrução

```
U(i,j:n) = U(i,j:n) - L(i,j)*U(j,j:n)
```

será executada pode ser obtida desenvolvendo-se

$$\begin{aligned}
\sum_{j=1}^{n-1}\sum_{i=j+1}^{n}\sum_{k=j+1}^{n} 1 &= \\
\sum_{j=1}^{n-1}\sum_{i=j+1}^{n} (n-j) &= \\
\sum_{j=1}^{n-1} (n-j)^2 &= \\
\sum_{j=1}^{n-1} n^2 - 2n\sum_{j=1}^{n-1} j + \sum_{j=1}^{n-1} j^2 &= \\
\sum_{j=1}^{n-1} j^2 &= \\
\frac{n-1}{6} n(2n-1) &= \tfrac{2n^3-3n^2+n}{6}
\end{aligned}$$

Veja que a complexidade adicionada pela instrução é $\frac{2n^3-3n^2+n}{3}$, pois cada vez que a instrução for executada, serão realizadas uma subtração e uma soma. Portanto, a fatoração $A = LU$ custa $O(n^3)$ operações aritméticas, e a constante do termo cúbico na função de complexidade é $\frac{2}{3}$.

3.2.5 Introduzindo o pivoteamento de colunas: $PA = LU$

Problemas ao não se trocar as linhas da matriz

A Eliminação de Gauss na forma como apresentamos até aqui, sem incorporar a troca de linhas de A, não é prática e não funciona para a quase totalidade dos casos de interesse. Foi apresentada apenas por razões didáticas, visando ilustrar que, se forem armazenados os multiplicadores e se não houver divisão por zero em (3.4), a Eliminação produz os fatores L e U de A.

O problema da Eliminação de Gauss, no entanto, não se resume ao caso em que a_{jj}^{j-1} em (3.4) é identicamente nulo, caso em que $j-$ésima iteração não seria definida no algoritmo da Figura (3.5) que apresentamos. Quando o denominador em (3.4) não é zero, mas é muito pequeno comparado ao numerador, o módulo do multiplicador tende a ser muito grande. Esse multiplicador será utilizado nas transformações lineares (3.5) e, desta forma, erros numéricos muito representativos devem ser observados. Recorde-se das duas fontes principais de erros numéricos:

- Subtração de quantidades muito próximas.
- Soma de quantidades muito díspares.

Exemplo 25 *Para este exemplo, seja* $A = \begin{bmatrix} 10^{-17} & 1 \\ 1 & 1 \end{bmatrix}$ *a matriz a ser fatorada. Veja que de acordo com* (3.4) *e* (3.5), 10^{17} *vezes a primeira linha é subtraído da segunda linha. Assumindo que utilizemos aritmética de precisão infinita, o que não é o caso com o uso de computadores digitais, a Eliminação de Gauss produziria os seguintes fatores exatos (sem erros numéricos)* L, U *abaixo.* $L = \begin{bmatrix} 1 & 0 \\ 10^{17} & 1 \end{bmatrix}$, $U = \begin{bmatrix} 10^{-17} & 1 \\ 0 & 1 - 10^{17} \end{bmatrix}$

Considere agora uma condição realista, em que empregamos aritmética de precisão finita e a precisão da máquina é $\epsilon \approx 10^{-16}$. *A grandeza* $1 - 10^{17}$ *na entrada de* U *não seá representada de forma exata. Ao invés disso, obteremos o resultado* -10^{17}. *Desta forma, os fatores obtidos com* $\epsilon \approx 10^{-16}$ *são* $\tilde{L}, \tilde{U}$ *dados por* $\tilde{L} = \begin{bmatrix} 1 & 0 \\ 10^{17} & 1 \end{bmatrix}$, $\tilde{U} = \begin{bmatrix} 10^{-17} & 1 \\ 0 & -10^{17} \end{bmatrix}$, *cujo produto é* $\tilde{L}\tilde{U} = \begin{bmatrix} 10^{-17} & 1 \\ 1 & 0 \end{bmatrix}$, *uma matriz substancialmente distinta de* A. *Se agora desejarmos resolver o sistema linear* $Ax = b$ *para* $b = (1, 0)^T$ *via* $\tilde{L}\tilde{U}x = b$, *obteremos* $\tilde{x} = (0, 1)^T$. *Porém, a solução verdadeira do sistema linear é* $x = (-1, 1)^T$. *Ou seja, a solução numérica e a solução verdadeira são muito distantes. Verifique o resultado deste experimento numérico no seguinte código* `scilab`.

```
-->Ltilde = [1 0;1E17 1]
 Ltilde  =
   1.          0.
   1.000D+17   1.
-->Utilde = [1E-17 1;0 1-1E17]
```

```
 Utilde  =
    1.000D-17   1.
    0.         -1.000D+17
-->[y] = SubsSucessivas(Ltilde,b,2)
 y  =
    1.
  -1.000D+17
-->[x] = SubsRetroativas(Utilde,y,2)
 x  =
    0.
    1.
```

Entretanto, observe que implementarmos o pivoteamento parcial, teremos o resultado correto. Veja o resultado com o uso da função `lu` *do* `scilab`*.*

```
-->[x,L,U,P] = ResolveParTriangulares(A,b)
 x  =
  -1.
    1.
 L  =
    1.          0.
    1.000D-17   1.
 U  =
    1.   1.
    0.   1.
 P  =
    0.   1.
    1.   0.
```

A resposta do procedimento `ResolveParTriangulares(A,b)`*, ilustrado na Figura 3.4, é composta pela solução x, a matriz de permutação P e os fatores L, U, de forma que $PA = LU$. A matriz P mostra que, alterando a ordem das linhas da matriz A, o problema foi resolvido. Observe que, diante da troca de linhas, os fatores L, U foram calculados sem que os erros numéricos inerentes à computação digital produzissem uma resposta muito diferente da verdadeira.*

Pelas razões discutidas acima e ilustradas no exemplo, verificamos que os multiplicadores na Eliminação de Gauss não devem ser muito grandes. Por meio da operação de pivotemento de linhas, isto é, da troca de linhas de A, garantimos que os multiplicadores empregados tenham módulo não superior a 1. Como produzir esta fatoração é o assunto da próxima seção.

Fatoração $PA = LU$

O instrumento que usaremos para representar a troca de linhas de A visando reduzir os erros numéricos e divisão por zero é o de matriz de permutação. Dada uma permutação $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ dos inteiros $\{1, \ldots, n\}$, uma matriz P é uma permutação (de linhas, associado a π) da matriz identidade I_n se a i-ésima linha de P for a π_i-ésima linha de I_n. Duas propriedades importantes das matrizes de permutação devem ser recordadas aqui. Se P é matriz de permutação, sua inversa é sua transposta, isto é: $P^{-1} = P^T$. Uma matriz de permutação é um caso particular de uma ortogonal (ou unitária). A segunda propriedade é que o determinante de uma matriz de permutação é $(-1)^p$, onde p é o número de trocas de linhas necessárias realizadas em P, para que a matriz resultante destas trocas seja a identidade de mesma ordem.

Defina $\hat{A}$ como uma matriz que contém as mesmas linhas de A, apenas apresentadas em ordem diferente. Assuma que as linhas em que são apresentadas em $\hat{A}$ seja dada por uma permutação π dos índices das linhas de A. Então existe uma matriz de permutação P tal que $\hat{A} = PA$. Se A é não singular (logo, $\det(A) \neq 0$), então existe uma matriz de permutação P tal que podemos aplicar o Método de Eliminação de Gauss à matriz PA, sem que ocorra divisão por zero, no cálculo dos multiplicadores. Logo $PA = LU$. Assim, a permutação de linhas resolve o primeiro problema que identificamos, que é a divisão por zero. Resta-nos eleger algum bom critério para permutar as linhas de A, obtendo P e PA. Essa matriz P será *descoberta* ao longo do processo, ao longo da Eliminação de Gauss.

O critério para definir uma boa P é o seguinte. Desejamos uma P tal que a matriz L obtida ao se fatorar PA tenha entradas cujos módulos sejam no máximo iguais a 1. Para garantir esta propriedade, em cada op de índice j, a linha pivotal não será necessariamente a linha j de A^{j-1}. Vamos comparar as entradas na coluna j, nas linhas $j, j+1, \ldots, n$ de A^{j-1}, e eleger como linha pivotal p aquela que contiver o maior elemento em módulo naquela coluna. O pivoteamento é chamado de parcial pois envolve a comparação dos módulos apenas nas linhas e não na matriz toda, sem que haja troca de colunas de A também. O pivoteamento total é uma alternativa mais cara: escolhemos o elemento de maior módulo da submatriz quadrada de A^{j-1} que envolve as colunas e linhas de j até n. Isso implicaria em trocar a ordem de colunas e linhas e pesquisar o maior elemento dentre $O((n-j+1)^2)$ alternativas, elevando o custo total. Também seriam necessárias duas matrizes de permutação, uma para troca de linhas, P, e outra para troca de colunas, $\hat{P}$, de forma que $PA\hat{P} = LU$. A permutação total não será empregada aqui.

Resumindo então o que fazemos na permutação parcial, com troca de linhas: na

$j-$ésima op, escolhemos como elemento pivô o valor

$$a_{p,j}^{j-1} = \max\{|a_{k,j}^{j-1}| : k = j, \ldots, n\}$$

e a linha pivotal como a linha

$$p = \arg\max\{|a_{k,j}^{j-1}| : k = j, \ldots, n\}.$$

Caso $p \neq j$, trocamos as linhas p e j de A^{j-1} antes de fazer as operações. No exemplo e no algoritmo que apresentaremos na sequência, não armazenaremos explicitamente a matriz P, mas sim um vetor auxiliar, *pivot*, que guarda a ordem das linhas pivotais. O valor inteiro $pivot(i)$ indica o índice da linha do sistema original representado na linha i. Este vetor será inicializado como $pivot(i) = i, i = 1, \ldots, n$, assumindo que não haverá troca de linhas. Sempre que alguma troca ocorrer, trocamos o conteúdo armazenado em $pivot(j)$ por $pivot(p)$, e vice versa. Vamos ilustrar o processo por meio de um exemplo.

Exemplo 26 *Vamos resolver o sistema linear apresentado no Exemplo 21, usando a fatoração $PA = LU$. Vamos também usar a fatoração para o cálculo do determinante. Neste exemplo, indicamos as linhas pivotais pelo elemento em vermelho de maior módulo nas linhas que competem para serem as linhas pivotais.*

A ordem das linhas do sistema original é indicada à direita de $[A^j|b^j]$. Esta ordem deve refletir as informações armazenadas no vetor pivot. Agora, com o pivoteamento parcial, podemos ter multiplicadores do tipo m_{ii} (com índices de coluna e linha iguais), pois a linha i pode não ter sido a linha pivotal na op i. Diante desta abordagem, o primeiro índice i associado ao multiplicador m_{ij} não faz referência à posição física i da linha, mas sim qual linha i do sistema original é representada na posição considerada para o cálculo dos multiplicadores.

1. *Inicialização: pivot* $= (1, 2, 3, 4)^T$.

2. *(op1)* $j = 1$, $[A|b] = \left[\begin{array}{cccc|ccc} 2 & 1 & 1 & 0 & 1 & & (E1) \\ 4 & 3 & 3 & 1 & 3 & & (E2) \\ 8 & 7 & 9 & 5 & 7 & & (E3) \\ 6 & 7 & 9 & 8 & 3 & & (E4) \end{array}\right]$

   - $a_{p1} = 8$, $p = 3$.
   - *Atualizamos "$pivot(1) = pivot(3), pivot(3) = pivot(1)$", trocamos o conteúdo das linhas 1/3.*
   - *Pre-multiplicamos por P_1, que difere de I nas linhas 1 e 3 apenas.*

- $m_{11} = -\frac{1}{4}$, $m_{21} = -\frac{1}{2}, m_{41} = -\frac{3}{4}$

$$M_1P_1[A|b] = \left[\begin{array}{cccc|cc} 8 & 7 & 9 & 5 & 7 & (E3) \\ 0 & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} & -\frac{1}{2} & (E2') \\ 0 & -\frac{3}{4} & -\frac{5}{4} & -\frac{5}{4} & -\frac{3}{4} & (E1') \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} & -\frac{9}{4} & (E4') \end{array}\right]$$

3. *(op2)* $j = 2$, $M_1P_1[A|b] = \left[\begin{array}{cccc|cc} 8 & 7 & 9 & 5 & 7 & (E3) \\ 0 & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} & -\frac{1}{2} & (E2') \\ 0 & -\frac{3}{4} & -\frac{5}{4} & -\frac{5}{4} & -\frac{3}{4} & (E1') \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} & -\frac{9}{4} & (E4') \end{array}\right]$

- $a_{p2} = \frac{7}{4}$, $p = 4$.
- *Trocamos o conteúdo da linha* $j = 2$ *pela linha* $p = 4$, *que corresponde a fazer "pivot(2) = pivot(4), pivot(4) = pivot(2)".*
- *Pré-multplicamos por* P_2, *que difere de* I *nas linhas 2 e 4 apenas.*
- $m_{22} = \frac{2}{7}$, $m_{12} = \frac{3}{7}$

$$M_2P_2M_1P_1[A|b] = \left[\begin{array}{cccc|cc} 8 & 7 & 9 & 5 & 7 & (E3) \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} & -\frac{9}{4} & (E4') \\ 0 & 0 & -\frac{2}{7} & \frac{4}{7} & -\frac{12}{7} & (E1'') \\ 0 & 0 & -\frac{6}{7} & -\frac{2}{7} & -\frac{8}{7} & (E2'') \end{array}\right]$$

4. *(op3)* $j = 3$, $M_2P_2M_1P_1[A|b] = \left[\begin{array}{cccc|cc} 8 & 7 & 9 & 5 & 7 & (E3) \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} & -\frac{9}{4} & (E4') \\ 0 & 0 & -\frac{2}{7} & \frac{4}{7} & -\frac{12}{7} & (E1'') \\ 0 & 0 & -\frac{6}{7} & -\frac{2}{7} & -\frac{8}{7} & (E2'') \end{array}\right]$

- $a_{p3} = -\frac{6}{7}$, $p = 4$.
- *Trocamos o conteúdo da linha* $j = 3$ *pela linha* $p = 4$, *que corresponde a fazer "pivot(3) = pivot(4), pivot(4) = pivot(3)".*
- *Pré-multplicamos por* P_3, *que difere de* I *apenas nas linhas 3 e 4.*
- $m_{13} = -\frac{1}{3}$

$$M_3P_3M_2P_2M_1P_1[A|b] = \left[\begin{array}{cccc|cc} 8 & 7 & 9 & 5 & 7 & (E3) \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} & -\frac{9}{4} & (E4') \\ 0 & 0 & -\frac{6}{7} & -\frac{2}{7} & -\frac{8}{7} & (E2'') \\ 0 & 0 & 0 & \frac{2}{3} & -\frac{4}{3} & (E1''') \end{array}\right]$$

Agora, vamos resolver o sistema linear $Ax = b$. Observe que a solução de $Ly = Pb$ já é disponível, pois operamos sobre $[A|b]$ e não apenas sobre b. Bastaria, portanto, resolvermos $Ux = y$. Entretando, vamos explicitar todos os fatores obtidos com a fatoração e a resolução dos dois sistemas lineares triangulares, recalculando y.

Os fatores já disponíveis são P (pois dispomos de pivot) e U.

- $pivot = \begin{bmatrix} 3 & 4 & 2 & 1 \end{bmatrix} \rightarrow P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

- $U = \begin{bmatrix} 8 & 7 & 9 & 5 \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ 0 & 0 & -\frac{6}{7} & -\frac{2}{7} \\ 0 & 0 & 0 & \frac{2}{3} \end{bmatrix}$

Para termos a fatoração completa, o fator que nos resta determinar é L. Como definir L à partir dos multiplicadores ? Veja que o índice da linha de A que gera uma linha pivotal na op j é a linha $pivot(j)$. Portanto, para a linha $pivot(j)$ de L teremos $j - 1$ multiplicadores, calculados nas ops anteriores. Estes serão armazenados na linha j de L, na ordem em que foram gerados.

- $m_{11} = -\frac{1}{4}$, $m_{21} = -\frac{1}{2}, m_{41} = -\frac{3}{4}$, $m_{22} = \frac{2}{7}$, $m_{12} = \frac{3}{7}$, $m_{13} = -\frac{1}{3}$.

- $L = \begin{bmatrix} 1 & & & \\ \frac{3}{4} & 1 & & \\ \frac{1}{2} & -\frac{2}{7} & 1 & \\ \frac{1}{4} & -\frac{3}{7} & \frac{1}{3} & 1 \end{bmatrix}$

Para resolver o sistema, resolvemos dois sistemas lineares triangulares, já que $Ax = b \rightarrow PAx = Pb \rightarrow LUx = Pb$.

- $Ly = Pb : \begin{bmatrix} 1 & & & \\ \frac{3}{4} & 1 & & \\ \frac{1}{2} & -\frac{2}{7} & 1 & \\ \frac{1}{4} & -\frac{3}{7} & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 7 \\ 3 \\ 3 \\ 1 \end{bmatrix}, \; y = \begin{bmatrix} 7 \\ -\frac{9}{4} \\ -\frac{8}{7} \\ -\frac{4}{3} \end{bmatrix}$

- $Ux = y : \begin{bmatrix} 8 & 7 & 9 & 5 \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ 0 & 0 & -\frac{6}{7} & -\frac{2}{7} \\ 0 & 0 & 0 & \frac{2}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 7 \\ -\frac{9}{4} \\ -\frac{8}{7} \\ -\frac{4}{3} \end{bmatrix}, \; x = \begin{bmatrix} -1 \\ 1 \\ 2 \\ -2 \end{bmatrix}$

A fatoração também nos permite calcular o determinante da matriz A: $\det(PA) = \det(LU) = \det(U)$. Logo, $\det(A) = \det(U)(-1)^p$, onde p é o número de trocas de pares de linhas necessárias para transformar P em I. Com isso temos: $\det(A) = (8)(\frac{7}{4})(\frac{-6}{7})(\frac{2}{3})(-1)^3 = 8$.

Visão por colunas da fatoração $PA = LU$

Para ilustrar como podemos representar a Eliminação com pivoteamento parcial, através da visão de colunas, vamos escrever $A = \tilde{L}U$ onde $\tilde{L}$ não é triangular inferior como desejamos, mas poderá ser transformada em L, através da permutação de linhas pertinente. Ou seja, vamos escrever uma fatoração $A = \tilde{L}U$, onde não vamos nos preocupar com a forma de $\tilde{L}$. Esta matriz apenas deve representar as transformações lineares que desejamos fazer. As linhas pivotais que descobrirmos irão abastecer as linhas de U, na ordem em que forem descobertas. Com o processo, vamos descobrir uma matriz de permutação P tal que $PA = (P\tilde{L})U = LU$, onde $P\tilde{L}$ é a L que desejamos. Vamos ilustrar o processo com o exemplo seguinte.

Exemplo 27

- *$j = 1$, primeira op.*
  Seguindo a mesma estratégia de escolher o pivot de maior módulo, a primeira linha de U é a terceira linha de A. Os (simétricos dos) multiplicadores são calculados sem relação a esta linha, obtendo a primeira coluna de $\tilde{L}$. Assim, a primeira op deve produzir o resutlado: $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 3 & 7 \\ 2 & 4 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 3 \\ 2 & 4 & 8 \end{bmatrix}$. *Vamos guardar os vetores $\tilde{L}_i$ sem nos preocupar com a forma triangular (inferior) para eles.* $u_1^T = \begin{bmatrix} 2 & 4 & 8 \end{bmatrix}$ $\tilde{L}_1 = \begin{bmatrix} 0 \\ 0.5 \\ 1 \end{bmatrix}$ $A = \tilde{L}_1 u_1^T + A2 \rightarrow A2 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$

- *Note que a terceira linha de $A2$ (e não a primeira como na aplicação do método sem troca de linhas) é toda composta de zeros. $pivot(1) = 3$.*

- *$j = 2$, segunda op. Há empate para escolha do pivot. Adotamos $pivot(2) = 1$. A segunda operação de pivoteamento deve gerar o resultado* $A2 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$. *Então temos:* $u_2^T = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$, $\tilde{L}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. $A2 = \tilde{L}_2 u_2^T +$

$$A3 \to A3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

- *$j = 3$, terceira op. $pivot(3) = 2$. $u_3^T = \begin{bmatrix} 0 & 0 & 2 \end{bmatrix}$, $\tilde{L}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. $A3 =$*

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Veja que neste momento, já podemos escrever a A como soma de matrizes de rank-1 que acumulamos:

- $A = \begin{bmatrix} 0 & 1 & 0 \\ 0.5 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 & 8 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$

- $A = \tilde{L}_1 u_1^T + \tilde{L}_2 u_2^T + \tilde{L}_3 u_3^T = \tilde{L}U$

A forma acima ainda não é a desejada pois as colunas $\tilde{L}_i$ não são colunas de uma triangular inferior com diagonal unitária. Como dispomos de pivot, calculamos:

- *$pivot = (3, 1, 2) \to P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$*

- *Pré-multiplicando por P a fatoração acima, temos: $PA = (P\tilde{L})U = LU$*

- $PA = \begin{bmatrix} 2 & 4 & 8 \\ 0 & 1 & 1 \\ 1 & 3 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.5 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 8 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$

3.3 Fatoração de Cholesky

A fatoração de Cholesky $A = LL^T$ é uma forma particular da fatoração $A = LU$, na qual $L = U^T$ e as entradas na diagonal de L são positivas. Isto é, L é triangular inferior e $l_{ii} > 0$ para todo $i = 1, \ldots, n$. A fatoração é possível se e somente se a matriz $A \in \mathcal{S}_n^{++}$, isto é, se A é simétrica e positiva definida.

A caracterização da positividade de uma matriz simétrica pode ser realizada de diversas formas. Uma delas consiste em empregar o algoritmo de fatoração de Cholesky apresentado nesta seção. Se o algoritmo for bem sucedido, chegando ao final sem efetuar divisão por zero ou radiciação de argumento negativo, o fator L é produzido e a positividade é caracterizada. Caso contrário, conclui-se que a fatoração não é possível e que a matriz não é positiva.

Duas são as vantagens da fatoração de Cholesky em relação à Fatoração LU. A positividade da matriz garante não haver divisão por zero, ou pivot nulo, de forma que a fatoração é bastante estável numericamente. A segunda é que, em função de se explorar a simetria de A, o número de operações aritméticas envolvidas é aproximadamente a metade das necessárias na Eliminação de Gauss. De igual forma, há apenas um fator a ser armazenado, L, de forma que a complexidade de memória também é a metade da fatoração LU.

O algoritmo de Fatoração de Cholesky é baseado no seguinte Teorema.

Teorema 3.3.1 *Teorema de Cholesky. Uma matriz simétrica $A \in \mathbb{R}^{n \times n}$ é positiva definida ($A \in \mathcal{S}^n_{++}$) se e somente se possui uma fatoração (chamada fatoração de Cholesky) da forma:*

$$A = LL^T$$

onde $L \in \mathbb{R}^{n \times n}$ é uma matriz triangular inferior com diagonal positiva. Esta fatoração é única.

Uma interpretação do uso do algoritmo para a caracterização da positividade pode ser dada da seguinte forma. O algoritmo assume que A seja simétrica positiva definida e usa a definição $A = LL^T$ para se obter os fatores envolvidos. Em uma perspectiva otimista, se tudo der certo com a aplicação do algoritmo, obtemos os fatores e caracterizamos a positividade. Tudo de uma vez só.

Assim, a ideia é inicialmente assumir que A seja simétrica e positiva definida. Então, pelo teorema, existe L com a diagonal positiva tal que $A = LL^T$. Pela definição dos elementos de A, como o produto interno das linhas de L pelas colunas de L^T (ou pelas próprias linhas de L, já que $L = L^T$) temos que

$$a_{ij} = l_i^T l_j : i, j = 1, \ldots, n$$

onde l_i denota a i-ésima linha de A. A ideia é então *percorrer* os elementos de A por colunas, das de menor índice j para as de maior índice e, para um índice j de coluna fixo, percorrer os elementos das linhas de índices $j = i, i+1, \ldots, n$, calculando as entradas de L. Mais precisamente, para cada par i, j de linha e coluna de A, usamos a definição de a_{ij} para calcular a entrada l_{ij} de L. Para tanto, vamos empregar as definições de l_{jj} e de l_{ji}, obtidas por meio da definição $a_{ij} = l_i^T l_j$.

Para definir a expressão analítica de l_{jj}, l_{ji}, vamos nos recordar que L é triangular inferior $\iff$ $l_{ij} = 0, j > i$. Então:

$$\begin{aligned}
a_{jj} =& l_j^T l_j \\
=& \sum_{k=1}^{n} l_{jk}^2 \\
=& \sum_{k=1}^{j} l_{jk}^2 + \sum_{k=j+1}^{n} l_{jk}^2 \\
=& \sum_{k=1}^{j} l_{jk}^2 \\
=& l_{jj}^2 + \sum_{k=1}^{j-1} l_{jk}^2
\end{aligned}$$

Adotamos a raiz positiva e então obtemos a expressão:

$$l_{jj} = +\sqrt{a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2} \quad \text{para } j = 1, \ldots, n. \tag{3.8}$$

Veja que o método pode falhar (e com ele a hipótese de positividade) quando $l_{jj} \leq 0$. Se não for esse o caso, para um determinado índice $i > j$, temos:

$$\begin{aligned}
a_{ij} =& l_i^T l_j \\
=& \sum_{k=1}^{n} l_{ik} l_{jk} \\
=& \sum_{k=1}^{j} l_{ik} l_{jk} + \sum_{k=j+1}^{n} l_{ik} l_{jk} \\
=& \sum_{k=1}^{j} l_{ik} l_{jk} \\
=& l_{ij} l_{jj} + \sum_{k=1}^{j-1} l_{ik} l_{jk}
\end{aligned}$$

o que nos permite escrever:

$$l_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}}{l_{jj}} \quad \text{para } i = j+1, \ldots, n. \tag{3.9}$$

Assim sendo, a ideia pode ser sistematizada da seguinte forma. Para todo $j = 1, \ldots, n$, calculamos o elemento l_{jj} de acordo com (3.8). Se $l_{jj}^2 > 0$, calculamos os elementos l_{ij}, para todos os valores de $i = j+1, \ldots, n$, de acordo com

(3.9). Esta ideia é sistematizada na implementação em `scilab`, apresentada na Figura 3.6. Algumas implementações de Cholesky, por exemplo aquela disponível no pacote `scilab` retorna a matriz triangular superior. A implementação apresentada na Figura 3.6 retorna a triangular inferior.

A implementação apresentada na Figura 3.6 faz uso das expressões (3.8) e (3.9), que foram derivadas através da definição de $a_{ij} = l_i^T l_j$. Por esta razão é denominada implementação por linhas ou por produto interno da Fatoração de Cholesky.

```
function [L] = Cholesky(A)
    n = size(A,1)
    L = zeros(n,n)
    for j = 1:n
        soma = A(j,j)
        for k = 1:j-1
            soma = soma - L(j,k)*L(j,k)
        end
        if soma > 0.0
            L(j,j) = sqrt(soma)
        else
           print("Matriz nao e SPD \n")
        end
        for i = j+1:n
           soma = A(i,j)
           for k = 1:j-1
              soma  = soma - L(i,k)*L(j,k)
           end
           L(i,j) = soma / L(j,j)
        end
    end
endfunction
```

Figura 3.6: Algoritmo de Fatoração de Cholesky que explora a definição dos elementos a_{ij} de A para o cálculo dos fatores.

Verifique você mesmo a corretude da fatoração de Cholesky apresentada no exemplo abaixo.

Exemplo 28 *Vamos empregar o algoritmo para verificar a positividade e obter os fatores L na fatoração de Cholesky de A.*

```
A  =

   1.  -1.    3.   -4.
  -1.   5.  -1.     2.
   3.  -1.   14.  -9.
```

```
  -4.   2.  -9.    22.
-->L = Cholesky(A)
 L  =
   1.   0.   0.   0.
  -1.   2.   0.   0.
   3.   1.   2.   0.
  -4.  -1.   2.   1.
```

3.3.1 Complexidade da Fatoração de Cholesky

A instrução mais custosa do Algoritmo de Cholesky apresentado na Figura 3.6 é

```
soma  = soma - L(i,k)*L(j,k)
```

Cada vez que a instrução é executada, duas operações aritméticas de ponto flutuante são realizadas. Vamos calcular a contribuição desta instrução para a complexidade total do algoritmo e mostrar que o algoritmo também é $O(n^3)$, assim como a fatoração LU.

O número de vezes em que a instrução é chamada é dada pelo somatório abaixo avaliado.

$$
\begin{aligned}
\sum_{j=1}^{n}\sum_{i=j+1}^{n}\sum_{k=1}^{j-1} 1 &= \sum_{j=1}^{n}\sum_{i=j+1}^{n}(j-1) \\
&= \sum_{j=1}^{n}(n-j)(j-1) \\
&= \sum_{j=1}^{n}((n+1)j - n - j^2) \\
&= \frac{n(n+1)^2}{2} - n^2 - \frac{n}{6}(n+1)(2n+1) \\
&= \frac{n^3}{6} + O(n^2)
\end{aligned}
$$

Para a dedução acima, fizemos uso de $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$ e de $\sum_{i=1}^{n} i^2 = \frac{n}{6}(n+1)(2n+1)$. Veja que o total de operações aritméticas incorridas na instrução deve ser multiplicado por 2. Portanto, a constante do termo cúbico na complexidade adicionada pela instrução é $\frac{1}{3}$.

Quando comparada à complexidade da fatoração $A = LU$, discutida na Seção 3.2.4, verificamos que o termo cúbico na função de complexidade de Cholesky é $\frac{1}{3}$, enquanto que em LU é $\frac{2}{3}$. Assim sendo, embora sejam assintonticamente equivalentes (ambos são $O(n^3)$), para um dado valor de n, o custo computacional da fatoração de Cholesky é aproximadamente a metade do custo da Fatoração LU.

3.3.2 Visão por colunas ou *outer Cholesky*

Da mesma forma como apresentamos uma formulação mais abstrata para a fatoração LU, na forma de produtos externos, vamos proceder para a Fatoração de Cholesky.

Veja que se a matriz A a ser fatorada é positiva definida, $a_{11} > 0$ deve valer. Vamos então particionar A em dois blocos, o primeiro deles sendo a_{11}, uma matriz positiva definida de dimensão 1 e o bloco de dimensão $n-1$, $K = A(2:n, 2:n)$.

Feito o particionamento, escrevemos a matriz A simétrica como o produto de duas matrizes, triangulares em blocos (ainda não triangulares, apenas triangulares em blocos):

$$A = \begin{bmatrix} a_{11} & w^T \\ w & K \end{bmatrix} = \begin{bmatrix} r_{11} & 0 \\ s & \hat{R}^T \end{bmatrix} \begin{bmatrix} r_{11} & s^T \\ 0 & \hat{R} \end{bmatrix} \tag{3.10}$$

Veja que podemos calcular r_{11} e s facilmente. Na verdade, o conjunto de operações que apresentamos anteriormente (na versão produto escalar ou linha da Fatoração de Cholesky) para computar a primeira coluna de L ou (primeira linha de L^T) corresponde a calcular estas entradas. Veja:

$$\begin{aligned} r_{11} &= \sqrt{a_{11}} \\ s &= \frac{w}{r_{11}}. \end{aligned}$$

Além disso, sabemos que

$$\hat{R}^T \hat{R} = K - ss^T, \tag{3.11}$$

já que o segundo bloco K de A é o produto da segunda linha de blocos do primeiro fator pela segunda coluna de blocos no segundo fator em (3.10): $K = \hat{R}^T \hat{R} + ss^T$. Veja que, ao calcularmos $\hat{R}^T \hat{R} = K - ss^T$ estamos essencialmente subtraindo de A uma matriz de posto 1, $L_1 L_1^T$, o produto externo da primeira coluna de L pela primeira linha de L^T, para então fatorarmos o bloco não nulo da diferença.

Assim, para completar a fatoração, aplicamos a mesma ideia recursivamente, ao bloco de dimensão $(n-1)$ $\hat{R}^T \hat{R} = K - ss^T$ e assim por diante, até que o bloco a ser fatorado seja um escalar positivo.

Cabe destacar que podemos adotar esta abordagem pois

$$A \in \mathcal{S}_n^{++} \rightarrow K \in \mathcal{S}_{n-1}^{++} \iff \hat{R}^T \hat{R} \in \mathcal{S}_{n-1}^{++}.$$

Exemplo 29 *Computar a Fatoração de Cholesky $A = LL^T = R^T R$ de A dada abaixo por meio da visão de produtos externos, outer Cholesky.*

$$A = \begin{bmatrix} 1 & -1 & 3 & -4 \\ -1 & 5 & -1 & 2 \\ 3 & -1 & 14 & -9 \\ -4 & 2 & -9 & 22 \end{bmatrix}$$

1. *Primeira op, determinamos a primeira linha de $L^T = R$.*

$$r_{11} = \sqrt{a_{11}} = 1$$
$$s^T = \frac{w^T}{r_{11}} = (-1, 3, -4)$$

Com isso, calculamos o segundo bloco a ser fatorado $R^T R$:

$$\hat{R}^T \hat{R} = \begin{bmatrix} 5 & -1 & 2 \\ -1 & 14 & -9 \\ 2 & -9 & 22 \end{bmatrix} - \begin{bmatrix} -1 \\ 3 \\ -4 \end{bmatrix} \begin{bmatrix} -1 & 3 & -4 \end{bmatrix}$$
$$= \begin{bmatrix} 4 & 2 & -2 \\ 2 & 5 & 3 \\ -2 & 3 & 6 \end{bmatrix}$$

2. *Segunda op, desejamos a fatoração de Cholesky de $\begin{bmatrix} 4 & 2 & -2 \\ 2 & 5 & 3 \\ -2 & 3 & 6 \end{bmatrix}$, ou seja, desejamos a segunda linha de $L^T = R$.*

$$r_{22} = \sqrt{4} = 2$$
$$s^T = \frac{w^T}{2} = (1, -1)$$

O termo que sobra R^TR para ser fatorado é:

$$\begin{aligned}\hat{R}^T\hat{R} &= \begin{bmatrix} 5 & 3 \\ 3 & 6 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 4 \\ 4 & 5 \end{bmatrix}\end{aligned}$$

3. *Terceira op, desejamos a fatoração de Cholesky de* $\begin{bmatrix} 4 & 4 \\ 4 & 5 \end{bmatrix}$*, ou seja determinamos a terceira linha de* $L^T = R$

$$\begin{aligned} r_{33} &= \sqrt{4} = 2 \\ s^T &= \frac{w^T}{2} = (2) \end{aligned}$$

O termo que sobra R^TR:

$$\hat{R}^T\hat{R} = 5 - (2)(2)^T = 1$$

4. *Quarta op, determinamos a quarta linha de* $L^T = R$*, obtendo a fatoração de Cholesky da matriz* [1].

$$r_{44} = \sqrt{1} = 1$$

Compondo as linha que calculamos, podemos escrever o fator $L^T = R = \begin{bmatrix} 1 & -1 & 3 & -4 \\ & 2 & 1 & -1 \\ & & 2 & 2 \\ & & & 1 \end{bmatrix}$.

3.4 Sistemas lineares malcondicionados

Com as fatorações $PA = LU$ e de Cholesky $A = LL^T$ podemos resolver uma grande variedade de problemas em Álgebra Linear, por exempo, a resolução de sistemas lineares. Elas são úteis para se resolver boa parte dos sistemas lineares que usualmente encontramos em aplicações, sobretudo se estruturas de dados adequadas para representação de matrizes esparsas forem empregadas.

Entretanto, há limitações para o uso destas fatorações. Quando os sistemas lineares são malcondicionados, outras fatorações devem ser empregadas. Estas outras fatorações, seja para a resolução de sistemas lineares ou para outros propósitos, e a base matemática que as fundamenta é o foco do restante do curso de ALC.

Porém, antes de apresentarmos estas outras fatorações, devemos definir quais são estes tipos de sistemas lineares que não devemos esperar que as fatorações básicas sejam capazes de resolver. Devemos caracterizar os sistemas lineares $Ax = b$ malcondidicionados, que são definidos por matrizes de coeficientes A malcondicionadas. Na apresentação desta seção, salvo menção contrária, assumimos que A é não singular e que b é diferente de zero. Assim sendo, o sistema linear $Ax = b$ admite solução única não nula.

O bom ou o mau condicionamento de um sistem linear $Ax = b$ depende de uma grandeza associada à matriz de coeficientes: o número de condição da matriz A, $\kappa(A)$.

Dada uma norma matricial $\|\cdot\|$ induzida por uma norma vetorial p, definimos o número de condição na norma p como

$$\kappa_p(A) = \|A\|_p \|A^{-1}\|_p. \tag{3.12}$$

Quando o valor de p não for relevante para a análise ou quando for claro pelo contexto, usaremos $\kappa(A)$ para fazer menção ao número de condição da matriz. Lembre-se que a norma matricial espectral é induzida pela norma vetorial $p = 2$, ou norma Euclideana. Recorde-se também que a norma de Frobenius é subordinada à norma Euclideana, mas não é induzida por ela. Portanto, não podemos definir o número de condição de uma matriz, à partir da norma de Frobenius.

Veja que o número de condição de uma matriz depende da sua inversa, de forma que o seu cálculo é computacionalmente custoso (no mínimo $O(n^3)$). Observe também que, pela definição, $\kappa(A) = \kappa(A^{-1})$ vale. Além disso, se a norma matricial espectral for a norma escolhida, a avaliação de $\kappa(A)_2$ torna-se mais onerosa. Portanto, é comum utilizamos limites inferiores para $\kappa(A)$ para inferir propriedades da matriz. Em particular, sendo $A_1, \ldots, A_n$ as colunas de A, podemos derivar limites inferiores para o número de condição de A utilizando qualquer par de índices de colunas $i, j \in \{1, \ldots, n\}$ e computando o lado direito da seguinte desigualdade:

$$\kappa_p(A) \geq \frac{\|A_i\|_p}{\|A_j\|_p}, \;\; 1 \leq p \leq \infty. \tag{3.13}$$

Veja que para $i = j$, a desigualdade (3.13) indica que $\kappa_p(A) \geq 1$, para qualquer norma matricial induzida por norma vetorial.

Exemplo 30 *Duas matrizes notavelmente malcondicionadas são as matrizes de*

Hilbert H e de Vandermonde V, indicadas abaixo. O termo geral da entrada h_{ij} da matriz de Hibert é $h_{ij} = 1/(i+j-1)$. Assim, para $n=4$, temos

$$H_4 = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/4 & 1/5 & 1/6 \\ 1/4 & 1/5 & 1/6 & 1/7 \end{pmatrix}$$

Para definirmos uma matriz de Vandermonde, precisamos definir um vetor de pontos ou de dados $x = (x_1, \ldots, x_m)^T$, com entradas distintas par a par: $x_i \neq x_j$ para qualquer par $i \neq j$. Já a expressão do termo geral da matriz de V associada a este conjunto de dados é $v_{ij} = x_i^{j-1}$ para $i = 1, \ldots, m$ e $j = 1, \ldots, n$. Veja o caso quadrado:

$$V = \begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{pmatrix}$$

Note que para qualquer diferença de magnitude entre a mínima e a máxima das entradas em $\{x_1, \ldots, x_m\}$, a potenciação destas entradas aumenta a diferença de escala. Usando a expressão (3.13)*, fica evidente que tanto H quando V são malcondicionadas.*

A utilidade do número de condição, em qualquer norma que seja, consiste em dizer o quão sensível é a solução x do sistema linear $Ax = b$, quando as entradas do sistema linear, A, b ou ambos, são perturbados.

A matriz A é bem condicionada quando seu número de condição, não importa qual p você use, é pequeno. A definição precisa de *pequeno* depende de muitos fatores, por exemplo o número de bits sendo empregado na representação dos números de ponto flutuante, a quantidade de erros numéricos aceitáveis para uma dada aplicação e a confiança que temos na qualidade dos dados do problema (o quão precisos são A, b na modelagem da aplicação representada pelo sistema). Apesar destas observações, admite-se que um número de condição *bom* deva ser inferior a 100. Voltaremos a discutir este aspecto em breve.

O fato é que os dados A, b carregam erros. Seja porque os valores armazenados internamente na máquina são aproximações dos dados verdadeiros (armazenamos $fl(A), fl(b)$ e não A, b propriamente), ou porque são dados que vieram do laboratório, de cálculos computacionais anteriores ou de modelos matemáticos simplificados para se representar problemas muito complexos. Os dados carregam erros. Assim é razoável pensarmos em A, b como dados verdadeiros de um sistema linear hipotético, não sujeitos a qualquer tipo dos erros discutidos e em $\tilde{A}, \tilde{b}$ como os dados do sistema linear que efetivamente vamos resolver, pois são os dados que dispomos para aproximar A, b.

Diante desta perspectiva, informalmente, se $\tilde{b}$ é um vetor muito próximo de b, as soluções $\tilde{x}$ e x respectivamente de $\tilde{A}\tilde{x} = \tilde{b}$ e $Ax = b$ devem ser próximas, quando o sistema linear é bemcondidiconado. Quando o sistema é malcondicionado (isto é, $\kappa(A)$ é elevado), pequenas variações nos dados do sistema linear acarretam grandes variações na solução do sistema: $\|x - \tilde{x}\|$ é muito grande mesmo quando $\|b - \tilde{b}\|$ (ou $\|A - \tilde{A}\|$ ou ambos) é muito pequeno. Esta afirmativa precisa ser demonstrada matematicamente. É o que vamos fazer, mostrar que $\kappa(A)$ funciona como uma garantia de que erros muito grandes na resposta não sejam obtidos se os dados A, b e $\tilde{A}, \tilde{b}$ não diferem muito. Para isso, como a discussão até aqui já sugere, vamos usar normas matriciais e vetoriais para mensurar a magnitude das perturbações nos dados $\delta(b) = b - \tilde{b}$ e $\delta(A) = A - \tilde{A}$ e na solução do sistema linear $\delta(x) = x - \tilde{x}$.

De início, vamos considerar o caso em que apenas o vetor b é sujeito à erros, de forma que $A = \tilde{A}$. Então temos que $A\tilde{x} = \tilde{b}$ equivale a $A(x\delta(x)) = b + \delta(b)$. Como $Ax = b$, o sistema anterior pode ser escrito como $A\delta(x) = \delta(b)$ ou equivalentemente $\delta(x) = A^{-1}\delta(b)$. Recordamos que sempre usaremos normas matriciais induzidas por normais matriciais. Assim sendo, usando as propriedades de normas matriciais induzidas por normas vetoriais, temos que $\|\delta(x)\| \leq \|A^{-1}\|\|\delta(b)\|$. Aplicando a mesma relação ao sistema não perturbado por erros, $b = Ax$, temos $\|b\| \leq \|A\|\|x\| \rightarrow \frac{1}{\|x\|} \leq \|A\|\frac{1}{\|b\|}$. Combinando as duas desigualdades obtemos:

$$\begin{aligned} \frac{\|\delta(x)\|}{\|x\|} &\leq \|A\|\|A^{-1}\|\frac{\|\delta(b)\|}{\|b\|} \\ \frac{\|\delta(x)\|}{\|x\|} &\leq \kappa(A)\frac{\|\delta(b)\|}{\|b\|}. \end{aligned} \tag{3.14}$$

Veja que a expressão acima mostra que $\kappa(A)$ é uma grandeza que surge naturalmente quando tentamos relacionar a perturbação relativa nos dados, $\frac{\|\delta(b)\|}{\|b\|}$, em função da perturbação observada na solução do sistema linear, $\frac{\|\delta(x)\|}{\|x\|}$. Daí decorre a definição e utilidade do número de condição.

Observe agora que a desigualdade (3.14) indica que $\kappa(A)$ funciona como uma trava: se $\kappa(A)$ é pequeno e $\frac{\|\delta(b)\|}{\|b\|}$ também é pequeno, não há como esperar $\frac{\|\delta(x)\|}{\|x\|}$ grande. Por outro lado, mesmo que $\frac{\|\delta(b)\|}{\|b\|}$ seja pequeno, se $\kappa(A)$ é grande, a perturbação da resposta pode ser grande. E o fato de usarmos normas matriciais induzidas por normas vetoriais, sempre haverá, para uma matriz A e um b, uma perturbação $\delta(b)$ que faça a desigualdade (3.14) ser satisfeita de forma justa, na igualdade (o lado direito e esquerdo da desigualdade assumindo valores iguais).

Usando argumentos semelhantes, podemos demonstrar a validade da seguinte desigualdade, caso as perturbações ocorram apenas em A e o vetor b não seja perturbado:

$$\frac{\|\delta(x)\|}{\|x+\delta(x)\|} \leq \kappa(A)\frac{\|\delta(A)\|}{\|A\|}$$

Exemplo 31 *Neste exemplo, vamos aplicar aplicar uma perturbação pequena em b e verificar o que acontece com a resposta na solução do sistema linear definido por* $Ax = b$ *onde* $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$, $A^{-1} = \begin{pmatrix} -998 & -999 \\ 999 & -1000 \end{pmatrix}$ *e* $b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. *Vamos assumir que a perturbação é dada por* $\delta b = \begin{pmatrix} 10^{-3} \\ 0 \end{pmatrix}$.

Para este exemplo, vamos calcular explicitamente o número de condição nas normas $p = 1, \infty$. *Para tanto, usamos o fato de que* $\|A\|_1 = \|A\|_\infty = \|A^{-1}\|_1 = \|A^{-1}\|_\infty = 1999$. *Logo* $\kappa_1(A) = \kappa_\infty(A) = 3.992 \times 10^6$.

A solução de $Ax = b$ *é* $x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. *Por outro lado, a solução do sistema linear perturbado* $Ax = (b+\delta_b)$ *é* $\hat{x} = \begin{pmatrix} 2 \times 10^{-3} \\ -1 \times 10^{-3} \end{pmatrix}$. *Veja que* $\delta x = \begin{pmatrix} -0.998 \\ 0.999 \end{pmatrix}$. *Portanto* $\|\delta b\|_1 = 1.0 \times 10^{-3}$, $\|\delta x\|_1 = 1.997$, $\frac{\|\delta x\|_1}{\|\delta b\|_1} = 1.997 \times 10^{+3}$. *A perturbação na resposta é cerca de 2000 vezes maior que a perturbação nos dados.*

A matriz A do exemplo não é singular, pois $\det(A) = 1$. Porém, verificamos também que as duas restrições do sistema linear são praticamente linearmente dependentes. E esse fato nos motiva a produzir uma discussão adicional.

Observe que o número de condição $\kappa(A)$ é uma propriedade da matriz e nada tem a ver com a precisão da máquina onde eventualmente resolveremos a solução do sistema linear. É uma propriedade inexorável da matriz. É dela, não depende da máquina.

Então por que afirmamos que sistemas lineares malcondicionados são difíceis de serem resolvidos ou de serem fatorados ? Se usarmos aritmética de precisão infinita, um sistema linear malcondicionado (em que A é inversível) não é pior do que outro bem condicionado. Porém, no processo de fatoração das matrizes, inevitavelmente calculamos fatores sujeitos a erros numéricos. E estas perturbações levam a perturbações grandes nas respostas.

No caso de matrizes malcondicionadas, esses erros numéricos são grandes o suficiente para, muitas vezes, introduzir dependência linear onde não há, matematicamente, de forma exata, depedência linear. Para ser mais preciso, recorde-se do exemplo (25) que apresentamos. A matriz A daquele exemplo é bem condicionada,

$\kappa(A)_1 = 4$. Naquele exemplo, empregamos a fatoração LU sem troca de linhas e diante disso, o fator $\tilde{U}$ que foi obtido é uma matriz onde as duas linhas são praticamente linearmente dependentes, quando a matriz A original é uma matriz em que suas linhas são claramente linearmente independentes. E em função disso, sugerimos a troca de linhas, que para matrizes bem condicionadas é uma ideia numérica eficaz para produzir soluções com boa qualidade numérica.

Porém, o *potencial de redução de erros numéricos da permutação de linhas* quando a matriz é malcondicionada é limitado. Para uma matriz malcondicionada, mesmo a introdução de mecanismos de pivoteamento resulta na introdução de dependência linear no sistema equivalente $Ux = y$ a ser resolvido. Isso porque as entradas das matrizes $A^{j-1} : j = 1, \ldots, n-1$ que competem para definir a linha pivotal já são substancialmente diferentes daquelas que seriam obtidas se aritmética de precisão infinita tivesse sido empregada. Isso faz com que mecanismos distintos da Eliminação de Gauss sejam empregados. Os mecanismos empregados na Eliminação de Gauss para produzir um sistema triangular superior $Ux = y$ equivalente a $Ax = b$, qual seja, empregar combinações lineares das linhas do sistema e troca de linhas, são propensos a erros.

3.4.1 Resolvendo um sistema linear definido por uma matriz de Vandermonde usando $PA = LU$.

O nosso objetivo nesta seção é ilustrar o efeito do malcondicionamento da matriz na qualidade da solução numérica produzida pela fatoração $PA = LU$ para se resolver o sistema linear. Para tanto, construiremos um sistema linear malcondicionado, definido por uma matriz de Vandermonde. O sistema linear será construído partindo da solução desejada para o sistema linear, um vetor n-dimensional de 1's. Resolveremos o sistema linear para diversos valores de n e verificaremos a diferença entre a resposta numérica e a resposta que esperávamos obter.

Para criar este experimento, vamos considerar o polinômio de grau $n-1$ na variável t:

$$p(t) = \sum_{i=0}^{n-1} a_i t^i$$

onde $a_0 = a_1 = \cdots = a_{n-1}$ são os coeficientes do polinômio. No nosso experimento, vamos fixar todos os coeficientes em 1. Isto é, o polinômio que vamos considerar é:

$$p(t) = 1 + t + t^2 + \cdots + t^{n-1} \tag{3.15}$$

Agora, vamos arbitrariamente escolher um conjunto de n (o número de coeficientes do polinômio de grau $n-1$) abscissas t e, à partir delas, vamos calcular suas correspondentes ordenadas $p(t)$. Com as n tuplas $(t, p(t))$ construiremos um sistema linear no qual a matriz de coeficientes é uma matriz de Vandermonde, o vetor de termos independentes consiste no vetor formando pelos valores assumidos pelo polinômio para cada uma destas entradas e o vetor x que procuramos é o vetor dos coeficientes de um polinômio que interpola todos estes pontos. Veja que já sabemos a resposta correta para o sistema linear: o vetor n-dimensional $(1, \ldots, 1)^T$.

A criação destes dados seguirá os seguintes passos:

- Parametrizamos $t = i + 1$, para diversos valores distintos de i.

- Para cada valor de $t = i + 1$, $p(t)$ pode ser reescrito como o valor da soma de uma Progressão Geométrica, cuja expressão analítica é desenvolvida abaixo:

$$p(i+1) = (i+1)^0 + (i+1) + (i+1)^2 + \cdots + (i+1)^{n-1} = \sum_{j=0}^{n-1} (i+1)^j$$

  Ou seja, para um dado valor de i, $p(i + 1)$ corresponde à soma dos termos de uma Progressão Geométrica de n termos, com o primeiro termo igual a 1 e razão $(i + 1)$. Então podemos escrever:

$$p(i+1) = \sum_{j=0}^{n-1} (i+1)^j = \frac{(1+i)^n - 1}{i}$$

A Figura 3.7 apresenta o código da geração do sistema linear que desejamos resolver (função GeraVandermonde) e o procedimento GeraExperimento, que seá executado com o seguinte vetor de dados de entrada para valores de n: $(5, 8, 10, 15, 20)^T$. Observe que para cada $n \in valoresden$, o procedimento calcula $\|x\|_\infty$, a norma do vetor x, solução numérica do sistema linear. Veja na Figura 3.8, os valores das normas das soluções x encontradas. Para $n = 10$ em diante, em nada estas normas conferem com a norma infinito de um vetor de uns, que é 1. Em particular, para valores de $n = 15, 20$, os fatores encontrados na fatoração LU (usando a implementação profissional disponível no `scilab`) são absolutamente distintos dos valores corretos. Isso ocorre pois a matriz de coeficientes é extremamente mal condicionada.

O problema da fatoração LU para lidar com matrizes malcondicionadas está na ideia central do método, que é construir combinaçõesl lineares das linhas de $Ax = b$, visando triangularizar a matriz. Recorde-se das transformações lineares $M_{n-1}M_{n-2}\ldots M_1 A = U$. Estas são instáveis, pois as matrizes de multiplicadores M_j possuem entradas obtidas por divisões pelo elemento pivot, que pode ter magnitude muito pequena, ainda que usemos o pivoteamento de colunas. Assim, em boa

```
function [A,b] = GeraVandermonde(n)
 for i = 1:n
    for j = 0:n-1
      A(i,j+1) = (1 + i)^j;
    end
    b(i) = ((1 + i)^n - 1)/i;
 end
endfunction

function [normas] = GeraExperimento(valoresn)
    s = size(valoresn)
    for i = 1:s(1)
        [A,b] = GeraVandermonde(valoresn(i))
        [L,U,P] = lu(A)
        [m,n] = size(A)
        [y] = SubsSucessivas(L,P*b,n)
        [x] = SubsRetroativas(U,y,n)
        printf("n = %d %8.7E \n",valoresn(i),norm(x,'inf'))
        normas(i) = norm(x,%inf)
    end
endfunction
```

Figura 3.7: Procedimentos para o experimento numérico com a matriz de Vandermonde.

```
-->valoresden'
 ans  =
   5.   8.   10.   15.   20.
-->normas = GeraExperimento(valoresden)
n = 5 1.0000000E+00
n = 8 1.0000001E+00
n = 10 1.0003594E+00
n = 15 5.3951429E+05
n = 20 2.0111339E+18
 normas  =
   1.
   1.0000001
   1.0003594
   539514.29
   2.011D+18
```

Figura 3.8: Solução obtida para o experimento da Figura 3.7.

parte do restante deste curso, apresentaremos outras ideias para produzirmos outras fatoração matriciais, mais estáveis numericamente, no espírito que narramos aqui.

Exercícios Propostos

As questões 1 a 4 foram adaptadas de [2]. As questões 6 a 8 foram adaptadas de [5]
Questão 01: Qual a matriz M que transforma A em uma matriz triangular superior U $(MA = U)$? Multiplique por $M^{-1} = L$ para fatorar $A = LU$.

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 4 & 2 \\ 6 & 3 & 5 \end{bmatrix}.$$

Questão 02: Quais são as duas matrizes de multiplicação M_1 e M_2 que transformam a matriz A que uma matriz triangular superior U $(M_2M_1A = U)$? Multiplique a matriz U pelas inversas de $M_1 e M_2$ para fatorar A em $A = LU$.

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \\ 3 & 4 & 5 \end{bmatrix}.$$

Questão 03: Defina as matrizes L e U para a matriz simétrica A. Quais são as condições em a, b, c, d que definem os pivôs na diagonal da matriz U para que A seja fatorada em LU?

$$A = \begin{bmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{bmatrix}.$$

Questão 04: Considere as matrizes L, U e o vetor b. Resolva $Lc = b$. Então encontre a solução de $Ux = c$. Encontre a matriz A, do sistema original $Ax = b$.

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \text{ e } b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}.$$

Questão 05: Uma das aplicações da solução de sistemas lineares é no cálculo da inversa da matriz A. Considere a matriz $A = \begin{bmatrix} 1 & -3 & 2 \\ -2 & 8 & -1 \\ 4 & -6 & 5 \end{bmatrix}$ e a sua fatoração em $PA = LU$. Encontre a primeira coluna da inversa de A, através da solução de um sistema linear, usando explicitamente os fatores P, L, U. Lembre-se que $AA^{-1} = I$.

Questão 06: Utilize a Decomposição de Cholesky (baixa abstração) para determinar se a matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 10 \\ 3 & 10 & 16 \end{bmatrix}$ é positiva definida.

Questão 07: Utilize a Decomposição outer Cholesky para fatorar a matriz: $A = \begin{bmatrix} 4 & -2 & 4 & 2 \\ -2 & 10 & -2 & -7 \\ 4 & -2 & 8 & 4 \\ 2 & -7 & 4 & 7 \end{bmatrix}$.

Questão 08: Resolva o sistema $Ax = b$ com: $A = \begin{bmatrix} 0 & 4 & 1 \\ 1 & 1 & 3 \\ 2 & -2 & 1 \end{bmatrix}$ e $b = \begin{bmatrix} 9 \\ 6 \\ -1 \end{bmatrix}$, utilizando a fatoração $PA = LU$.

Questão 09: Suponha que $A \in \mathbb{R}^{n \times n}$ seja não singular e $B \in \mathbb{R}^{n \times p}$. Considere o problema de encontrar a matriz $X \in \mathbb{R}^{n \times p}$ tal que $AX = B$. Construa um algoritmo que encontre X em não mais que $O(\max\{pn^2, n^3\})$ operações aritméticas de ponto flutuante.

Questão 10: Deseja-se resolver o sistema linear $A^k x = b$ sem computar a matriz A^k (k é um inteiro qualquer). Sabe-se que a matriz A é não singular. Construa um algoritmo que resolva este sistema linear sem explicitamente avaliar A^k.

Questão 11: Suponha que dispomos de $A \in \mathbb{R}^{n \times n}$, $d \in \mathbb{R}^n$, $c \in \mathbb{R}^n$ e que desejemos encontrar $s = c^T A^{-1} d$. Uma abordagem seria computar A^{-1} conforme o exercício 1 acima sugere e depois calcular $s = cXd$. Entretanto, há uma forma mais econômica de se proceder. Identifique esta forma mais econômica.

Capítulo 4

Projetores e Ajuste de Curvas

Neste capítulo, apresentamos uma classe de matrizes quadradas especiais, chamadas de projetores ou matrizes de projeção. A propriedade central destas matrizes é que, se forem aplicadas em vetores que já pertencem ao seu espaço coluna, o resultado é o próprio vetor sobre o qual foram aplicadas. Quando além desta propriedade são também simétricas, estas matrizes são denonimadas projetores ortogonais. Estas propriedades são descritas na Seção 4.1 deste capítulo.

A imagem da transformação linear que uma matriz de projeção simétrica induz é o vetor em seu espaço coluna que dista o mínimo possível do ponto sobre o qual a matriz foi aplicada. Na Seção 4.2, apresentamos como podemos projetar um ponto em um subespaço vetorial e como são respresentadas as matrizes de projeção ortogonais associadas. O sistema de equações normais, que permite obter a projeção no espaço, é também desenvolvido naquela seção. Na Seção 4.3, mostramos como projetar em um conjunto afim, indicando que nada mais é que uma aplicação de projeção em subespaços vetoriais.

Tendo desenvolvido o conceito de matrizes de projeção ortogonal e de projeção em subsespaços vetoriais, apresentamos o Método dos Mínimos Quadrados para ajuste de curvas a um conjunto de dados na Seção 4.4. Salientamos que na Seção 4.4, o único ferramental matemático que empregamos para encontrar os coeficientes ótimos das funções de base empregadas no Método de Mínimos Quadrados é o conceito de projeção e de matrizes de projeção.

Por fim, na Seção 4.5, a última seção deste capítulo, complementamos a apresentação do Método de Mínimos Quadrados, desenvolvendo-o sem o uso dos conceitos de projeção que empregamos ao longo de todo o capítulo. Para tanto, usamos elementos do Cálculo Diferencial. Embora não haja resultado novo nesta última seção, entendemos que a metodologia de desenvolvimento das equações normais via Cálculo Diferencial complementa bem a abordagem que apresentamos antes, inteiramente centrada no conceito de projeção.

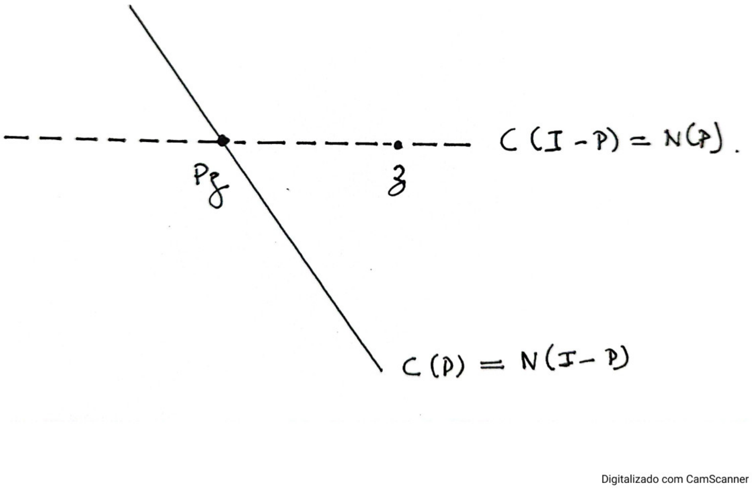

Figura 4.1: Projetor oblíquo.

4.1 Matrizes de projeção ou projetores

Uma matriz quadrada P é chamada de *projetor* se satisfaz a relação de idempotência:

$$P^2 = P. \tag{4.1}$$

Há dois tipos de projetores, os *projetores oblíquos* e os *projetores ortogonais*. Os projetores ortogonais são simétricos, isto é, $P^T = P$. Os oblíquos englobam os demais casos.

Observação 1 *É importante não confundir um projetor ortogonal com uma matriz ortogonal, pois $P^TP \neq I$, ou seja, um projetor ortogonal não é necessariamente uma matriz ortogonal.*

Vamos considerar que $P \in \mathbb{R}^{n\times n}$ seja um projetor e interpretar o efeito das transformações lineares Px, P^2x, P^3x e assim por diante. Veja que $z = Px \in C(P)$ e que $Pz = P^2x = Px = z$. Generalizando, se $z \in C(P)$, $P^kz = Px = z$ para qualquer inteiro $k \geq 1$.

As Figuras 4.1 e 4.2 ilustram o efeito das transformações lineares de dois projetores, oblíquos e ortogonais, respectivamente.

Exemplo 32 *Considere a matriz $P = \frac{uu^T}{u^Tu}$ para algum $u \in \mathbb{R}^n, u \neq 0$ e responda:*

- *P é projetor ?*

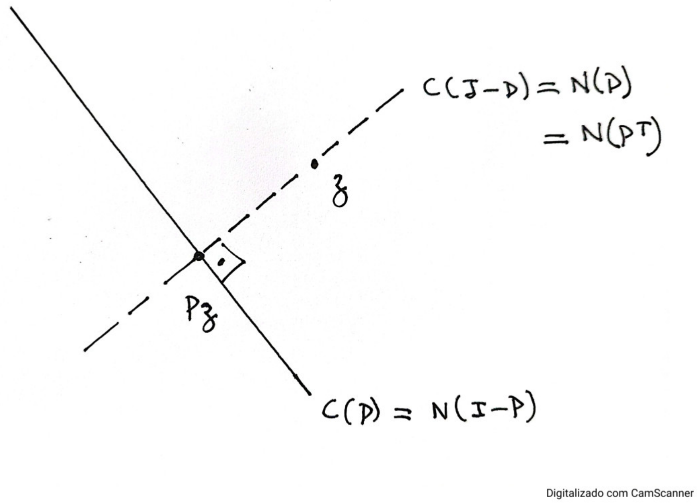

Figura 4.2: Projetor ortogonal.

- *P é projetor ortogonal ou oblíquo ?*
- *Qual é o posto de P ?*
- *Em qual espaço P projeta ? Qual a dimensão deste espaço ?*

Vamos verificar se a relação (4.1) se verifica: $P^2 = \frac{(uu^T)(uu^T)}{(u^Tu)^2} = \frac{u(u^Tu)u^T}{(u^Tu)^2} = \frac{uu^T}{u^Tu} = P$. *Portanto a matriz quadrada P é um projetor e como é simétrica ($P = P^T$), P é projetor ortogonal.*
Considere $y \in \mathbb{R}^n$ e veja que $Py = \left(\frac{uu^T}{u^Tu}\right) y = \frac{(u^Ty)}{u^Tu} u$. *Portanto Py é um vetor ao longo da linha u, cuja dimensão é 1.*

Observação 2 *Um projetor projeta em um espaço cuja dimensão é igual ao posto de P.*

A um projetor P, ortogonal ou oblíquo, associa-se um projetor complementar, $I - P$. Para mostrarmos que $I - P$ é de fato um projetor, devemos mostrar que vale a propriedade (4.1). Veja:

$$(I - P)^2 = I - 2P + P^2 = I - 2P + P = I - P,$$

comprovando que $I - P$ é de fato idempotente.

Em qual espaço $I - P$ projeta ? Naturalmente, $I - P$ projeta em $C(I - P)$. Mas como podemos relacionar $C(I - P)$ com alguns dos espaços fundamentais associados a P ? Para responder a esta questão, vamos examinar o resultado de aplicar P às colunas de $I - P$:

$$P(I-P) = P - P^2 = P - P = 0.$$

Uma vez que $P(I-P)$ é uma matriz nula, cada coluna de $I-P$ pertence ao subespaço $N(P)$. De forma análoga, $(I-P)P = 0$ e cada coluna de P pertence ao $N(I-P)$. Consequentemente, o vetor $z-Pz$ (veja a Figura 4.1) satisfaz $P(z-Pz) = P(I-P)z = 0$, de forma que $z - Pz \in N(P)$. Ou seja $(I-P)z$ projeta z em $N(P)$, pois $(I-P)z$ nada mais é do que uma combinação linear de vetores em $N(P)$.

Considerando os quatro espaços fundamentais associados a P, sabemos que $N(P^T) \perp C(P)$. Quando $P = P^T$, $N(P) = N(P^T)$ e portanto $I-P$ projeta z em $N(P)$ que, nesse caso, é um subespaço ortogonal a $C(P)$. Veja a Figura 4.2.

Complementando, temos também o seguinte resultado válido para qualquer projetor, oblíquo ou ortogonal.

Resultado 4.1.1 *$C(I-P) = N(P)$.*

Prova 4.1.1 *Veja que, por um lado temos:*

> *$C(I-P) \supseteq N(P)$, pois tomando algum $z \in N(P)$, temos que $Pz = 0$ e então $(I-P)z = z$, isto é $z \in C(I-P)$.*

Temos também:

- *$C(I-P) \subseteq N(P)$, pois para qualquer $y \in C(I-P)$ temos $(I-P)z = y$ para algum z, e então: $Py = P(I-P)z = Pz - P^2z = 0$. Logo $y = (I-P)z \in N(P)$.*

No caso de um projetor ortogonal, note que $C(I-P) = N(P)$ é coerente com o fato de que $I-P$ projeta em um espaço ortogonal a $C(P)$. A definição algébrica que apresentamos para um projetor ortogonal é que $P^T = P$. Por definição geométrica, um projetor ortogonal P deve ser tal que P e $I-P$ projetam em subespaços ortogonais: $C(I-P) \perp C(P)$. Estas duas definições são equivalentes, isto é, $C(I-P) \perp N(P) \iff P = P^T$.

Tomando o projetor $Z = I - P$ e seu complementar $I - Z$, por (4.1.1) deduzimos que $C(I-Z) = N(Z)$, de forma que obtemos o resultado complementar a (4.1.1): (4.1.1):

$$N(I-P) = C(P) \tag{4.2}$$

Exemplo 33 *Neste exercício desejamos mostrar, por meio de um exemplo numérico, que as transformações lineares associadas a P e $I-P$ levam a vetores ortogonais. Para tanto, considere o projetor ortogonal $P = \frac{uu^T}{u^Tu}$ para $u = (1,-1,1)^T$. Tome os vetores $z = (1,2,-1)^T$ e $y = (0,1,2)^T$, por exemplo, e verifique algebricamente que:*

- $(Py)^T(I-P)z = y^TP^T(I-P)z = y^T(P-P^2)z = 0$

Então temos que $P = \begin{bmatrix} 1/3 & -1/3 & 1/3 \\ -1/3 & 1/3 & -1/3 \\ 1/3 & -1/3 & 1/3 \end{bmatrix}$ *e* $I-P = \begin{bmatrix} 2/3 & 1/3 & -1/3 \\ 1/3 & 2/3 & 1/3 \\ -1/3 & 1/3 & 2/3 \end{bmatrix}$.
$Py = \begin{bmatrix} 1/3 \\ -1/3 \\ 1/3 \end{bmatrix}$ *e* $(I-P)z = \begin{bmatrix} 5/3 \\ 4/3 \\ -1/3 \end{bmatrix}$, *e o produto interno entre* Py *e* $(I-P)z$ *é* $5/9 - 4/9 - 1/9 = 0$.

Dois outros resultados relevantes são:

Resultado 4.1.2 *Para qualquer projetor, oblíquo ou ortogonal, valem as relações entre seus espaços:*

$$N(P) \cap N(I-P) = \{0\} \tag{4.3}$$

$$C(P) \cap C(I-P) = \{0\} \tag{4.4}$$

Prova 4.1.2 *Para demonstrar o primeiro deles,* (4.3), *veja que qualquer vetor* $z \in N(I-P) \cap N(P)$ *satisfaz*

$$\begin{aligned} 0 &= (I-P)z & &\textit{pois } z \in N(I-P) \\ &= z - Pz & & \\ &= z. & &\textit{pois } z \in N(P) \end{aligned}$$

A demonstração de (4.4) *segue raciocínio análogo.*

Uma consequência importante das relações (4.2)-(4.4) é que, a partir de um projetor $P \in \mathbb{R}^{n \times n}$ e os espaços ele e seu complemento geram, podemos escrever o espaço $\mathbb{R}^n$ como a soma de dois subespaços, $S_1 = N(P)$ e $S_2 = C(P)$, uma vez que $S_1 \cap S_2 = \{0\}$, isto é, $\mathbb{R}^n = N(P) \oplus C(P)$. Cabe recordar que, o $\mathbb{R}^n$ sendo uma soma direta de $C(P)$ e $N(P)$, qualquer vetor $z \in \mathbb{R}^n$ pode ser decomposto como $z = z^1 + z^2$, onde $z^1 \in C(P)$ e $z^2 \in N(P)$ são únicos.

Exercício 4.1.1 *Mostre que se* P *é projetor ortogonal, a matriz* $I - 2P$ *é unitária. Veja que* $(I-2P)^T(I-2P) = (I-2P)^2 = I^2 - 4P^2 + 4P^2 = I$.

4.2 Projetando em subespaços vetoriais

4.2.1 Motivação

Uma das grandes aplicações do conceito de projeção surge ao tentarmos *resolver* um sistema linear $Ax = b$, onde $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, para o qual $b \notin C(A)$. Claramente, a condição $b \notin C(A)$ implica que não há como resolver o sistema linear; estamos, na verdade, empregando um *abuso de linguagem* ao dizer que desejamos resolver o sistema.

Ao invés de resolver o sistema linear no sentido como fizemos ao empregar os métodos de Eliminação de Gauss, Fatorações $PA = LU$ ou mesmo Fatoração de Cholesky, vamos reformular o problema, permitindo que, associado à *solução* x do sistema, tenhamos um resíduo ou erro, $e = b - Ax$, diferente de zero.

Assim sendo, nosso objetivo agora ao *resolvermos* o sistema $Ax = b$, consiste em encontrar um vetor x que minimize este erro, para uma dada norma, a norma Euclideana. Mais formalmente, vamos resolver o problema de Quadrados Mínimos:

$$\min\|Ax - b\|_2. \tag{4.5}$$

Observação 3 *Associado a um problema de projeção, há três aspectos principais:*

1. *O que será projetado ?*

2. *Qual é o alvo ou onde será projetado ?*

3. *Como implementamos a projeção ?*

No caso do problema (4.5), desejamos projetar o vetor b no subespaço vetorial $C(A)$. Este problema de projeção consiste em resolver o problema de Otimização formulado em (4.5). O fato é que a solução desse problema de projeção, ou melhor dizendo desse problema de otimização particular, admite solução analítica. Nas seções seguintes vamos apresentar a forma analítica da solução ótima x que resolve (4.5), por meio de uma matriz de projeção P que projeta o ponto b em um ponto $p \in C(A)$, cuja distância $\|p - b\|_2$ é mínima.

Cabe destacar que, conceitualmente, poderíamos empregar outra norma, por exemplo, as normas $p = 1$ ou $p = \infty$. Porém, obteríamos uma solução x possivelmente distinta da que resolve (4.5) e, além disso, não teríamos uma expressão analítica para o ponto $p \in C(A)$ associado a esta solução, pois precisaríamos recorrer a algum algoritmo para resolver o problema de otimização equivalente na norma alternativa.

4.2.2 Projetando um vetor em subespaços vetoriais

Projetando em uma linha

O nosso primeiro caso de interesse consiste em projetar um vetor, digamos $b \in \mathbb{R}^m$, na linha *span*$\{u\}$ ou seja, no espaço vetorial associado ao vetor $u \in \mathbb{R}^m$. A partir de agora, vamos denominar o ponto em $span\{u\}$ mais próximo de b, na norma Euclideana, como p. O desenvolvimento que faremos nos permitirá determinar uma expressão analítica para a matriz de projeção P que sintetiza ou resume o processo de projetar um vetor qualquer em $span\{u\}$.

A geometria do processo de projeção é ilustrada na Figura 4.2. Veja que podemos decompor o vetor b a ser projetado como a soma de sua projeção p em $span\{u\}$ e de um vetor erro, ou diferença que pertence a $e \in span\{u\}^{\perp}$

$$b = p + e. \tag{4.6}$$

Sabemos que estas parcelas, p e e, são únicas pois $\mathbb{R}^m = span\{u\} \oplus span\{u\}^{\perp}$ uma vez que $span\{u\} \cap span\{u\}^{\perp} = \{0\}$.

Uma vez que $p \in span\{u\}$, escrevemos

$$p = \hat{x}u, \tag{4.7}$$

para um escalar $\hat{x} \in \mathbb{R}$ a ser determinado. Reescrevendo (4.6) a partir da expressão acima, temos $e = b - \hat{x}$ e como desejamos a projeção ortogonal, sabemos que $e \perp u$. Impondo esta condição temos:

$$\begin{aligned} u^T(b - \hat{x}u) &= 0 \\ \hat{x} &= \frac{u^T b}{u^T u}. \end{aligned} \tag{4.8}$$

Logo, a projeção de b pode ser determinada como

$$\begin{aligned} p &= \frac{u^T b}{u^T u} u \\ &= \frac{uu^T}{u^T u} b. \end{aligned}$$

Veja que a matriz de projeção

$$P = \frac{uu^T}{u^T u} \tag{4.9}$$

pode ser definida a partir da expressão acima e que a partir dela podemos sintetizar o processo de projeção como

$$p = Pb.$$

Veja que a matriz P definida em (4.9) é simétrica e de rank-1. Recorde-se do exercício (32) onde estudamos uma matriz de projeção P idêntica àquela que acabamos de deduzir. No exercício, mostramos que P é de fato um projetor, discutimos sua dimensão e a de seu projetor complementar, $I - P$. Por fim, veja também que

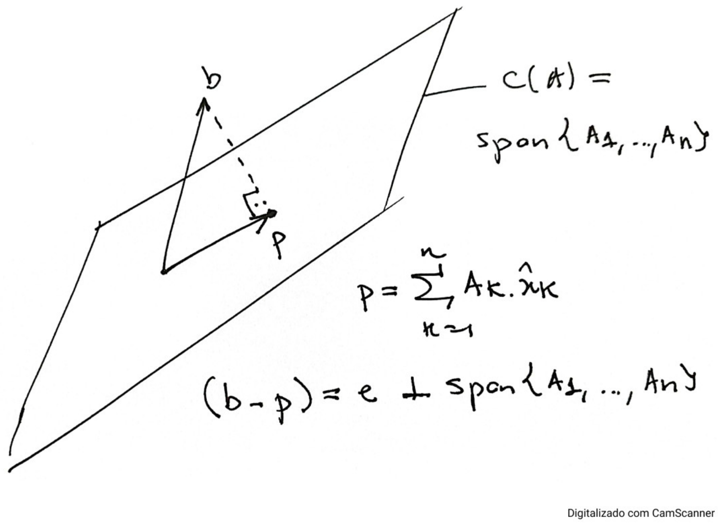

Figura 4.3: Projeção de um vetor no espaço coluna de A.

$(I-P)b = b-p = e$, de forma que o projetor $I-P$ projeta b em $N(P)$.

4.2.3 Projetando um vetor em um subespaço coluna

Vamos generalizar o resultado que desenvolvemos na seção anterior, no sentido de que agora desejamos projetar $b \in \mathbb{R}^m$ em um espaço vetorial para o qual conhecemos uma base, formada pelos n vetores linearmente independentes $A_1, A_2, \ldots, A_n$, onde $A_k \in \mathbb{R}^m : k = 1, \ldots, n$. Por conveniência, definimos a matriz $A \in \mathbb{R}^{m \times n} = [A_1, A_2, \ldots, A_n]$ de posto completo n, a partir da base para o espaço vetorial onde b deve projetado. Dessa forma, nosso problema passa a ser o de obter a expressão analítica para o ponto $p \in C(A)$ mais próximo de b na norma Euclideanda que, em outras palavaras, corresponde à projeção ortogonal de b em $C(A)$. Também desejamos a expressão para o projetor P associado a $C(A)$. Veja a Figura 4.3.

Como $p \in C(A)$, escrevemos $p = A\hat{x}$ para algum $\hat{x} \in \mathbb{R}^n$ a ser determinado. Mais uma vez, decompomos o vetor b segundo a expressão (4.6) e usamos o fato de que o erro $e = b - A\hat{x}$ é ortogonal a $C(A)$. Para garantir esta condição de ortogonalidade, impomos que e seja ortogonal a cada um dos elementos que definem a base para $C(A)$:

$$
\begin{aligned}
A_k^T(b - A\hat{x}) &= 0, & k &= 1, \ldots, n \\
A_k^T A\hat{x} &= A_k^T b, & k &= 1, \ldots, n \\
A^T A\hat{x} &= A^T b. & & \qquad (4.10)
\end{aligned}
$$

O sistema linear (4.10) é conhecido como *Sistema de Equações Normais*, pois garante que o vetor erro e seja ortogonal ou normal ao vetor $b - p$, onde p é a projeção de b em $C(A)$.

Observe que como a matriz A possui posto coluna completo, $(A^TA)^{-1}$ existe e a solução $\hat{x}$ existe e é única. A partir de (4.10), escrevemos:

$$\hat{x} = (A^TA)^{-1}A^Tb, \tag{4.11}$$

o que nos permite deduzir a expressão da projeção

$$p = A(A^TA)^{-1}A^Tb, \tag{4.12}$$

e também da matriz de projeção ortogonal associada

$$P = A(A^TA)^{-1}A^T. \tag{4.13}$$

Veja que A^TA é uma matriz simétrica e, consequentemente sua inversa, A^TA^{-1} também é. Assim, a matriz de projeção P dada por (4.13) é simétrica e portanto ortogonal. Mais uma vez, verificamos a equivalência entre as definições geométricas e algébricas de projetores ortogonais.

Por fim, veja que o vetor $e = b - A\hat{x}$ pertence a $C(A)^{\perp} = N(A^T)$. Então, o projetor complementar $I - P$ projeta em $N(A^T)$, isto é, $e = (I - P)b$.

Projeção com base ortonormal

Quando a base para o espaço onde desejamos projetar é ortonormal, isto é, possui elementos com norma unitária e ortogonais entre si, a expressão do projetor é simplificada. Como de costume, vamos designar uma matriz $A \in \mathbb{R}^{m \times n}$ com colunas ortonormais como $A = Q$. Como $Q^TQ = I$, podemos reescrever a expressão (4.13) como $P = QQ^T$, que é uma soma de matrizes de n rank-1.

Há várias vantagens em representarmos um espaço vetorial por meio de uma base ortonormal. A principal delas é que transformações lineares do tipo Qx associadas a matrizes Q satisfazendo $Q^TQ = I$ propagam menos erros numéricos. Estas matrizes preservam a norma Euclideana e os ângulos entre os vetores.

Além das vantagens acima identificadas, podemos transformar o problema de projetar em um espaço n dimensional, como uma sequência de problemas de projeção independentes, em subespaços ortogonais. Para explicar esta observação, considere que as colunas de Q sejam representadas pelos vetores $q_k \in \mathbb{R}^m : k = 1, \dots, n$, isto é, $Q = [q_1, q_2, \dots, q_n]$. Dessa forma, podemos escrever P como a soma

$$P = \sum_{i=1}^{n} q_i q_i^T. \tag{4.14}$$

de forma que a projeção p de b em $C(Q)$ pode ser escrita como

$$Pb = \sum_{i=1}^{n} q_i(q_i^T b). \tag{4.15}$$

Veja que as entradas da solução $\hat{x}$ de (4.11) que combina as colunas de Q de forma a escrever o ponto $p \in C(Q)$ estão explícitas na equação (4.15), isto é, $\hat{x}_i = q_i^T b$. Não por acaso, a expressão de $\hat{x}$ dada por (4.11) é simplificada para

$$\hat{x} = Q^T b, \tag{4.16}$$

quando $A = Q$, uma matriz com colunas ortonormais.

A expressão (4.15) nos permite interpretar o processo de projeção de b em $C(Q)$ como um processo mais simples, o de projetar b de forma independente em cada um dos subespaços $span\{q_1\}$, $span\{q_2\}, \ldots,$ $span\{q_n\}$, uma projeção a cada vez, e então somar os pontos correspondentes a cada uma destas projeções independentes para se obter o vetor p desejado.

Matematicamente, há uma equivalência entre o resultado que obteremos ao projetar de forma independente e o resultado dado por (4.15), pois a matriz Q é ortogonal. Defina o erro e^k como $e^k = b - \sum_{i=1}^{k}(q_i^T b)q_i$ para todo $k = 1, \ldots, n$ e veja que $e^n = b - Pb$.

Exemplo 34 *Considere a matriz* $A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 1 & 2 \end{bmatrix}$ *e sua fatoração* $A = QR = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \\ \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 2 \end{bmatrix}$. *Determine a projeção de* $b = (3, 2, 1)^T$ *em* $C(A)$*, usando a base fornecida pelas colunas de* Q *(de forma acoplada) e de forma independente, sequencialmente.*

Resolução 34.1 *Como* $C(A) = C(Q)$ *e* $Q^TQ = I$*, podemos calcular* $P = QQ^T = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$, $p = Pb = (2, 2, 2)^T$. *Associado a este* p *temos o erro* $e = b - p = (3, 2, 1)^T - (2, 2, 2)^T = (1, 0, -1)^T$.

Agora, vamos fazer a projeção de forma independente, que é possível pois dispomos de um base ortonormal para $C(A)$. *Projetamos* b *em* $span\{q_1\}$ *para obter* $e^1 = b - (b^Tq_1)q_1 = (3, 2, 1)^T - \frac{4}{\sqrt{2}}(1/\sqrt{2}, 0, 1/\sqrt{2})^T = (1, 2, -1)^T$. *Na sequência, projetamos* b *em* $span\{q_2\}$ *e descontamos de* e^1*, de forma que* $e^2 = e^1 - (b^Tq_2)q_2 = (1, 2, -1)^T - 2(0, 1, 0) = (1, 0, -1)^T$.

4.3 Projeção de um vetor em um conjunto afim

Um conjunto afim é um conjunto do tipo

$$Y = \{x \in \mathbb{R}^n : Ax = b\}, \tag{4.17}$$

onde $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. Veja que um conjunto afim é portanto o conjunto de soluções de um sistema linear. Para os resultados desta seção vamos assumir que:

- $posto(A) = r$.
- $b \in C(A)$, de forma que o sistema linear $Ax = b$ é consistente, isto é, admite solução: $Y \neq \emptyset$.

O nosso objetivo nesta seção é encontrar a projeção de um vetor $z \in \mathbb{R}^n$ no conjunto $Y \subset \mathbb{R}^n$. Isto é, desejamos encontrar o ponto $u \in Y$ que resolve

$$\min_{u \in Y} \|z - u\|_2^2. \tag{4.18}$$

Note que se $b = 0_m$, as soluções do sistema linear homogêneo $Ax = 0$ caracterizam um subsespaço linear, $N(A)$. É de se esperar que haja relação entre o problema de se projetar em Y, que desejamos resolver aqui, e o de se projetar em $N(A)$, que já sabemos resolver.

A Figura B.3 ilustra o conjunto afim $Y = \{x \in \mathbb{R}^2 : -x_1 + x_2 = 1\}$. Além de Y, a figura também mostra o subespaço linear $span\{(1,1)^T\}$. Observe que o conjunto Y corresponde a uma translação do subespaço linear ali indicado. A figura ilustra ainda: um ponto qualquer $x^0 = (0,1)^T \in Y$, o ponto $z = (1/2, 2)^T$ que desejamos projetar, o ponto $u \in Y$ que corresponde à projeção de z em Y e o ponto $p = (5/4, 5/4)$, que corresponde à projeção de z em $span\{(1,1)^T\}$.

Para o caso ilustrado na Figura B.3, a matriz A é $A = [-1, \; 1] \in \mathbb{R}^{1 \times 2}$, de forma que $N(A) = span\{(1,1)^T\}$. Observe através da figura que $z - u = \alpha(z - p)$ para algum $\alpha \in \mathbb{R}$ de forma que o ponto u, projeção de z em Y, e z se relacionam de acordo com a relação de ortogonalidade:

$$(z - u) \perp N(A). \tag{4.19}$$

Tanto quanto $(z - p) \perp N(A)$ garante que p é a projeção de z em $N(A)$ quanto $(z - u) \perp N(A)$ garante que $u \in Y$ é a projeção de z em Y. A propriedade de ortogonalide estabelecida em (4.19) sugere a nossa abordagem para se resolver o problema 4.18: o problema (4.18) se reduz ao problema de se projetar um ponto em um subespaço linear, assunto que estudamos em detalhes nas seções anteriores.

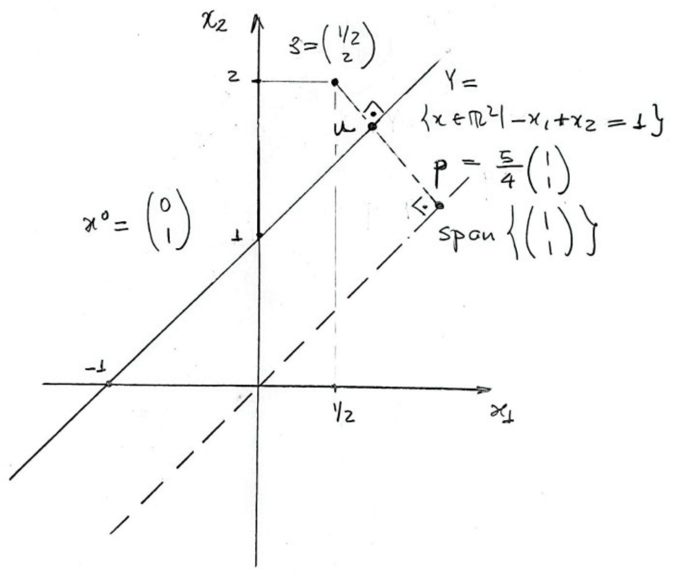

Figura 4.4: Projeção em conjunto afim.

Para o caso do exemplo ilustrado na Figura B.3, já podemos calcular o ponto u desejado, bastando para isso usarmos a propriedade (4.19), impondo a ortogonalidade. Ou seja, temos que $u = (0,1)^T + x(1,1)^T$ para algum $x \in \mathbb{R}$ que garanta $((0,1) + x(1,1) - (1/2,2))(1,1)^T = 0$. Resolvendo em x temos $x = 3/4$ e $u = (3/4, 7/4)^T$.

Nosso primeiro passo para formalizar a ideia que apresentamos para um conjunto Y mais geral que o ilustrado na figura consiste em reformular o conjunto Y adequadamente. Uma vez que $Y \neq \emptyset$, vamos tomar $x^0 \in Y$ como uma solução particular qualquer do sistema linear $Ax = b$. Recorde-se que um vetor $d \in N(A)$ satisfaz $Ad = 0$. Logo, para qualquer $x^0 \in Y$ e $d \in N(A)$ temos

$$A(x^0 + d) = b.$$

Desta forma, qualquer ponto $u \in Y$ pode ser escrito como

$$u = x^0 + \sum_{i=1}^{k} v^i x_i \tag{4.20}$$

onde $N(A) = span\{v^1, v^2, \ldots, v^k\}$, $k = dim(N(A)) = n - r$, sendo $\{v^1, v^2, \ldots, v^k\}$ uma base para $N(A)$.

Diante desta obervação, a definição do conjunto afim dada por (4.17) pode ser reformulada como:

$$Y = x^0 + span\{v^1, v^2, \ldots, v^k\} \tag{4.21}$$

No caso ilustrado na Figura B.3, temos que

$$Y = (0,1)^T + span\{(1,1)^T\}.$$

A expressão acima deve ser lida assim: o conjunto afim Y é caracterizado por um ponto mais um subsespaço linear.

Substituindo (4.20) em (4.18), reformulamos o problema em termos das variáveis $x_i : i = 1, \ldots, k$. Definindo como $V = [v^1, v^2, \ldots, v^k]$ a matriz $n \times k$ de posto k, e procedendo a uma mudança de variáveis (de $y \in Y$ para $x \in \mathbb{R}^k$), o problema (4.18) equivale a:

$$\min_{x \in \mathbb{R}^k} \|z - (x^0 + Vx)\|_2^2 \tag{4.22}$$

Observe que o problema (4.22) pode ser entendido como o problema de se projetar $z - x^0$ em $N(A) = C(V)$. Portanto, o sistema de equações normais que nos permite obter os pesos x é

$$V^T V x = V^T (z - x^0). \tag{4.23}$$

Mais especificamente, temos que $x = (V^T V)^{-1} V^T (z - x^0)$ é a solução do sistema de equações normais. Com os pesos x, obtemos a solução $p = x^0 + Vx$ como a solução de (4.17), o problema de projeção que desejávamos resolver.

O procedimento que explicamos aqui pode ser resumido nos seguintes passos:

1. Encontre uma solução particular x^0 para $Ax = b$.
2. Caracterize $N(A)$ por meio de uma base $v^1, v^2, \ldots, v^k$ para este subespaço e defina $V = [v^1, \ldots, v^k]$ uma matriz tal que $C(V) = N(A)$.
3. Resolva o sistema de equações normais $V^T V x = V^T (z - x^0)$ e obtenha a projeção de z em Y por meio de $p = x^0 + Vx$.

Exemplo 35 *Vamos resolver o problema de projetar o vetor z em Y, definidos pelos dados abaixo:*

```
A  =

   1.  -2.  -1.
  -3.   6.   3.
   1.  -2.  -1.
b  =
  -1.
   3.
```

```
  -1.
z  =
   4.
   0.
   1.
```

Seguindo os passos do procedimento que apresentamos acima temos:

1. *Resolvendo* $Ax^0 = b$*, temos que* $x^0 = (-1, 0, 0)^T$*.*

2. *Caracterizando* $N(A)$*. Fazendo uma fatoração de* A *que revele seu posto temos:* $A = \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & -1 \end{bmatrix}$*. Com isso caracterizamos* $N(A)$ *via* $N(A) \perp C(A^T)$ *e então* $N(A) = span\{v^1, v^2\}$*, onde* $v^1 = (-2, -1, 0)^T, v^2 = (-1, 0, -1)^T$*.*

```
 V  =
  -2.  -1.
  -1.   0.
   0.  -1.
```

3. *Resolvemos o sistema de equações normais* $V^T V x = V^T(z - x^0)$

```
VTV  =
   5.   2.
   2.   2.
z-x0 =
   5.
   0.
   1.
V'*(z-x0) =
  -10.
  -6.
x = inv(VTV)*V'*(z-x0);
x  =
  -1.3333333
  -1.6666667
```

4. *Calculamos o ponto* $u = x^0 + Vx$*.*

```
u = x0+V*x;
u  =
    3.3333333
    1.3333333
    1.6666667
```

Verifique que $V^T(z-u) = 0$*:*

```
-->V'*(z-u)
 ans  =
    1.332D-15
    4.441D-16
```

4.4 Ajuste de curvas e o método de Mínimos Quadrados via Projeção

Retornamos agora à motivação descrita na Seção 4.2.1, descrevendo uma aplicação para a resolução do problema (4.5). A aplicação surge no contexto de descrever o comportamento de uma função ou sistema em termos de variáveis explicativas ou independentes.

Em muitos casos, um sistema complexo pode ser descrito como uma *caixa-preta* na qual uma variável dependente é função de uma ou mais variáveis independentes. O valor exato da função pode ser conhecido para diferentes valores das variáveis independentes, porém, o comportamento exato deste sistema em função destas variáveis, isto é, a função que relaciona a variável dependente com as independentes ou não é analiticamente conhecida ou é muito complexa ou cara de ser avaliada. Neste contexto, o especialista no problema em questão pode propor uma função de base que aproxima ou substitui a caixa preta.

Mais formalmente, assuma que tenhamos uma função $y(x) : \mathbb{R} \to \mathbb{R}$ que não é conhecida analiticamente. Ao invés disso, dispomos de um conjunto de pares de pontos $\{(x_i, y_i) : i = 1, \ldots, m\}$, que representa a dependência de y em relação à variável x, restrita às abscissas $x_i : i = 1, \ldots, m$. Assumimos que as m abscissas x_i são distintas. Desejamos dispor de um modelo de uma função $g(x)$, em uma determinada classe de funções escolhida pelo especialista no problema, que aproxime os m pontos dados. Com esta função $g(x)$ podemos estimar o valor de $y(\overline{x})$ por meio de $g(\overline{x})$, para algum valor de $\overline{x} \notin \{x_i : i = 1, \ldots, m\}$.

Uma classe de função frequentemente usada como função de aproximação $g(x)$ é a classe dos polinômios. Polinômios são funções contínuas, fáceis de serem avaliados, derivados e integrados.

Suponha então que $g(x) = \sum_{k=0}^{n-1} \alpha_k x^k$ represente um polinômio de grau não superior a $n-1$. O polinômio é definido a partir dos coeficientes $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ que, de início, são desconhecidos. O objetivo é encontrar o polinômio (isto é, seus coeficientes) que melhor represente os pontos, de acordo com algum conceito previamente escolhido do que seja o melhor.

No caso do Método dos Mínimos Quadrados, objeto de estudo desta seção, o melhor polinômio é aquele cujos coeficientes $\alpha_i : i = 0, 1, \ldots, n-1$ minimiza a função soma de quadrados de desvios:

$$D(\alpha_0, \ldots, \alpha_{n-1}) = \sum_{i=1}^{m} \left(y_i - \sum_{k=0}^{n-1} \alpha_k x_i^k \right)^2. \tag{4.24}$$

Veja que a função 4.24 envolve um termo $(y_i - \sum_{k=0}^{n-1} \alpha_k x_i^k)^2$ para cada um dos m pontos. Dificilmente, o melhor polinômio escolhido permitirá que a soma dos quadrados dos desvios $\sum_{i=1}^{n} \left(y_i - \sum_{k=0}^{n-1} \alpha_k x_i^k\right)^2$ seja nula. Isso só ocorreria caso a função $y(x)$ fosse um polinômio de grau $n-1$ ou se o grau do polinômio fosse de grau $m-1$, o que seria altamente não recomendável. Isso jamais deve ser feito.

Veja também que, a partir da função soma de quadrados acima, podemos definir uma matriz $A \in \mathbb{R}^{m \times n}$ e um vetor $b \in \mathbb{R}^m$ $A = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & x_m^2 & \ldots & x_m^{n-1} \end{bmatrix}$

$b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$ de forma que minimizar a função (4.24) equivale a encontrar o vetor $\alpha = (\alpha_0, \ldots, \alpha_{n-1})^T$ que minimiza:

$$\|A\alpha - y\|_2 \tag{4.25}$$

Para um conjunto de coeficientes α qualquer, $\|A\alpha - y\|_2^2$ é exatamente o valor da função soma de quadrados de desvios formulada para aquele vetor α de coeficientes. Como a norma Euclideana $\|\cdot\|_2$ é uma função não negativa, o vetor α que minimiza $\|A\alpha - y\|_2^2$ também minimiza $\|A\alpha - y\|_2$.

Em outras palavras, resolver o problema $\min\|A\alpha - y\|_2$ é exatamente o problema de projeção em $C(A)$ que estudamos nas seções anteriores. O vetor y deve ser projetado em $C(A)$. A solução analítica deste problema de projeção é:

$$\alpha = (A^T A)^{-1} A^T y \tag{4.26}$$

$$p = A\alpha \tag{4.27}$$

onde p é o ponto em $C(A)$ mais próximo do vetor y.

O desenvolvimento acima considerou o caso em que a função $g(x)$ empregada para explicar o conjunto de pontos é um polinômio. Na verdade, qualquer função que seja linear nos parâmetros a serem estimados pode ser empregada. No caso do polinômio considerado, os parâmetros estimados pelo Método dos Mínimos Quadrados são os coeficientes do polinômio. A título de ilustração, poderíamos ter empregado uma função de aproximação do tipo $g(x) = \alpha_0 + \alpha_1 ln(x)$, por exemplo, caso $x_i \geq 0 : i = 1, \ldots, n$. Esta função não é linear em x mas permanece linear nos parâmetros. Nesse caso, a matriz A a ser empregada no Método dos Mínimos Quadrados seria $A = \begin{bmatrix} 1 & \ln(x_1) \\ 1 & \ln(x_2) \\ \vdots & \vdots \\ 1 & \ln(x_m) \end{bmatrix}$, uma vez que a função de quadrados mínimos seria substituída por

$$D(\alpha_0, \alpha_1) = \sum_{i=1}^{m} \left(y_i - (\alpha_0 + \alpha_1 \ln(x_i))\right)^2 .$$

Quando a função $g(x)$ escolhida para representar os pontos não for linear nos parâmetros, é necessário implementar uma linearização do modelo, para que o Método de Mínimos Quadrados seja empregado.

A título de ilustração, considere que a função $g(x)$ escolhida seja $g(x) = \alpha_0 e^{-\alpha_1 x}$, que não é linear em α_0, α_1. Veja que podemos linearizar esta função da seguinte forma:

$$\begin{aligned} g(x) &= \alpha_0 e^{-\alpha_1 x} \\ \ln(g(x)) &= \ln(\alpha_0 e^{-\alpha_1 x}) \\ &= \ln(\alpha_0) + \ln(e^{-\alpha_1 x}) \\ &= \hat{\alpha}_0 - \alpha_1 x \\ &= \hat{\alpha}_0 + \hat{\alpha}_1 x \end{aligned}$$

Ou seja, neste caso, devemos projetar o vetor $\begin{bmatrix} \ln(y_1) \\ \ln(y_2) \\ \vdots \\ \ln(y_m) \end{bmatrix}$ no espaço coluna da matriz $A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix}$, para obter os parâmetros ótimos $\hat{\alpha}_0$ e $\hat{\alpha}_1$, com os quais os coeficientes $\alpha_0 = e^{\hat{\alpha}_0}$ e $\alpha_1 = -\hat{\alpha}_1$ desejados podem ser obtidos.

Exemplo 36 *Considere os dados tabelados abaixo que devem ser ajustados por uma função do tipo $g(x) = \alpha_0 x^{\alpha_1}$.*

x	10	20	30	40	50	60	70	80
y	25	70	380	550	610	1220	830	1450

Note que o modelo não é linear em α_0, α_1. Procedendo à linearização, ajustamos:

$$\ln(y) = \ln(\alpha_0) + \alpha_1 \ln(x.)$$

Desta forma, o sistema de equações normais que determinará os parâmetros ótimos $\ln(\alpha_0)$ e α_1 a ser resolvido é definido pela matriz A cuja primeira coluna é um vetor m dimensional de 1's e segunda coluna é $(\ln(x_1), \ldots, \ln(x_m))^T$. Por sua vez, temos o vetor $b = (\ln(y_1), \ldots, \ln(y_m))^T$. Veja o detalhamento dos cálculos empregado o `scilab`*, lembrando que a base empregada na função* `log` *do* `scilab` *é e. No exemplo abaixo, para ilustração, calculamos explicitamente a solução via $(A^TA)^{-1}$. Cabe a ressalva, mais uma vez, que ao invés disso, a matriz A^TA deve ser fatorada (via Cholesky, ou idealmente via QR) e a solução dos sistema recuperada sem a inversão explícita de A^TA.*

```
-->x = [10;20;30;40;50;60;70;80];
-->y = [25;70;380;550;610;1220;830;1450];
-->n = 2;
-->m = 8;
-->A = zeros(m,n);
-->b = log(y)
 b  =
   3.2188758
   4.2484952
   5.9401713
   6.3099183
```

```
   6.413459
   7.1066061
   6.7214257
   7.2793188
-->A(:,1) = ones(m,1);
-->A(:,2) = log(x)
 A  =
   1.   2.3025851
   1.   2.9957323
   1.   3.4011974
   1.   3.6888795
   1.   3.912023
   1.   4.0943446
   1.   4.2484952
   1.   4.3820266
-->ATA = A'*A
 ATA  =
   8.           29.025284
   29.025284   108.77174
-->ATb = A'*b
 ATb  =
   47.23827
   178.25992
-->sol = inv(ATA)*ATb
 sol  =
  -1.294126
   1.9841763
```

Desta forma, $\ln(\alpha_0) = -1.294126 \rightarrow \alpha_0 = e^{-1.294126} = 0.2741374$ *e* $\alpha_1 = 1.9841763$.

4.5 Desenvolvimento do Método dos Mínimos Quadrados Via Cálculo Diferencial

Retomamos ao problema de encontrar os coeficientes $\alpha_k : k = 0, \ldots, n-1$, visando minimizar a função (4.24), que por conveniência reescrevemos abaixo.

$$D(\alpha_0, \ldots, \alpha_{n-1}) = \sum_{i=1}^{m} \left(y_i - \sum_{k=0}^{n-1} \alpha_k x_i^k \right)^2 .$$

A abordagem que desenvolvemos nesta seção é a de encontrar os coeficientes α_k impondo que os coeficientes desejados devem resolver o problema de otimização não linear irrestrito abaixo:

$$\begin{aligned} \min D(\alpha_0, \ldots, \alpha_{n-1}) \\ \alpha_k \in \mathbb{R} : k = 0, \ldots, n-1 \end{aligned} \tag{4.28}$$

Uma condição necessária para que que um vetor $\alpha \in \mathbb{R}^n$ seja o minimizador de (4.28), o vetor gradiente de D em $\hat{\alpha}$ deve ser nulo. O problema de otimização (4.28) é estritamente convexo e portanto esta condição necessária também é suficiente para otimalidade de $\hat{\alpha}$. Impondo tais condições temos:

$$\frac{\partial D}{\partial \alpha_k} |_{\hat{\alpha}} = 0, \text{ para todo } k = 0, 1, \ldots, n-1. \tag{4.29}$$

Para um problema de otimização mais geral que (4.28), o sistema de equações de primeira ordem (4.29), que impõe que $\nabla D = 0$, é usualmente um sistema de equações não lineares. Entretanto, a formulação do Problema de Mínimos Quadrados é linear nos parâmetros $\alpha_k : k = 0, 1, \ldots, n-1$ que desejamos estimar. Por esta razão, o sistema que obtemos quando calculamos as derivadas de D explicitamente em (4.29) é um sistema linear. Mostraremos a seguir que o sistema (4.29) nada mais é do que o sistema de equações normais que apresentamos anteriormente.

Aplicando a regra da cadeia, calculamos as derivadas parciais de D em relação a cada um dos parâmetros $\alpha_j : j = 1, \ldots, n-1$ que desejamos estimar:

$$\begin{aligned} \frac{\partial D}{\partial \alpha_j} &= \frac{\partial}{\partial \alpha_j} \sum_{i=1}^{m} \left(y_i - \sum_{k=0}^{n-1} \alpha_k x_i^k \right)^2 \\ &= \sum_{i=1}^{m} \frac{\partial}{\partial \alpha_j} \left(y_i - \sum_{k=0}^{n-1} \alpha_k x_i^k \right)^2 \\ &= 2 \sum_{i=1}^{m} \left(y_i - \sum_{k=0}^{n-1} \alpha_k x_i^k \right) \frac{\partial}{\partial \alpha_j} \left(y_i - \sum_{k=0}^{n-1} \alpha_k x_i^k \right) \\ &= 2 \sum_{i=1}^{m} \left(y_i - \sum_{k=0}^{n-1} \alpha_k x_i^k \right) \left(-x_i^j \right) \qquad\qquad j = 0, 1, \ldots, n-1 \end{aligned}$$

Impondo a condição de gradiente nulo no ponto α que desejamos encontrar:

$$
\begin{aligned}
0 &= \sum_{i=1}^{m} \left(y_i - \sum_{k=0}^{n-1} \alpha_k x_i^k \right) x_i^j && j = 0, 1, \ldots, n-1 \\
\sum_{i=1}^{m} y_i x_i^j &= \sum_{k=0}^{n-1} \left(\sum_{i=1}^{m} x_i^j x_i^k \right) \alpha_k && j = 0, 1, \ldots, n-1
\end{aligned} \tag{4.30}
$$

Salientamos que o sistema de equações (4.30) é exatamente o sistema de Equações Normais (4.10) que obtivemos quando projetamos b em $C(A)$. Quando escrito em termos de y e de α o sistema de equações normais toma a forma:

$$A^T y = (A^T A)\alpha$$

Para verificar a equivalência, recorde-se da definição da matrix $A = (x_i^{k-1}) : i = 1, \ldots, m, k = 1, \ldots, n$, $A \in \mathbb{R}^{m \times n}$:

$$
A = \begin{bmatrix}
1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\
1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
1 & x_m & x_m^2 & \ldots & x_m^{n-1}
\end{bmatrix},
$$

Veja que o lado direito em (4.30) corresponde a exatamente $A^T y$, pois para cada linha j do sistema de equações normais, o termo independente corresponde ao produto interno $\langle A_j, y \rangle = \sum_{i=1}^{m} y_i x_i^{j-1}$, onde A_j é a $j-$ésima coluna de A. De forma análoga, os coeficientes que multiplicam α_k na linha $j \in \{0, \ldots, n-1\}$ no lado direito da igualdade correspondem a $\langle A_{j+1}, A_k \rangle = \sum_{i=1}^{m} x_i^j x_i^k$.

Exercícios Propostos

As questões de 1 a 3, 7 e 8 foram adaptadas de [2]. As questões 5 e 6 foram adaptadas de [4].

Questão 01: Encontre a matriz de projeção P_C no espaço coluna da matriz e a matriz de projeção P_R no espaço linha de A. O que podemos dizer da matriz $B = P_C A P_R$?

$$A = \begin{bmatrix} 3 & 6 & 6 \\ 4 & 8 & 8 \end{bmatrix}.$$

Questão 02: Considere vetor b e o vetor p que é a combinação de $A_1, \cdots, A_n$ pertencentes à $\mathbb{R}^m$. Como podemos verificar se p é uma projeção de b no subespaço gerado pelos vetores de A_i?

Questão 03: Considere o vetor b. Suponha que P_1 seja a matriz de projeção no subespaço $\mathbb{R}^1$ gerado pela primeira coluna da matriz $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$. Suponha que P_2 seja a matriz de projeção no espaço coluna de A. Qual é o resultado do produto $P_2 P_1$?

Questão 04: Se A é uma matriz quadrada e inversível, qual é matriz de projeção P no espaço gerado pelas colunas de A?

Questão 05: Seja E uma matriz $m \times m$, com $Ex = \frac{x+Fx}{2}$ onde F é uma matriz $m \times m$ que transforma $[x_1, \cdots, x_m]$ em $[x_m, \cdots, x_1]$. A matriz E é um projetor ortogonal, um projetor oblíquo ou não é um projetor?

Questão 06: Se P é um projetor ortogonal, então $I - 2P$ é uma matriz unitária.

Questão 07: Suponha que as colunas de A não sejam independentes. Como podemos definir uma matriz B tal que $P = B(B^T B)^{-1} B^T$ seja a matriz de projeção no espaço coluna de A?

Questão 08: Considere um conjunto de valores t_i, deslocado da média $\hat{t} = (t_1 + \cdots + t_m)/m$ para obter $T_i = t_i - \hat{t}$, sabendo que $\sum T_i = 0$. A partir desta transformação, qual a estrutura da nova matriz A? Qual a relação entre as novas colunas de A, que representa a equação de ajuste da equação $C + DT$? Quais os valores dos parâmetros C e D?

Capítulo 5

Fatoração QR

Neste capítulo, discutimos uma das mais importantes fatorações matriciais, a chamada fatoração QR, isto é, $A = QR$, onde Q possui colunas ortonormais e R é triangular superior. Inicialmente, apresentamos as duas formas desta fatoração, reduzida e completa. Na sequência, revisamos os motivos que fazem da fatoração QR tão importante em Computação Científica. Apresentamos alguns algoritmos para se calcular a fatoração QR reduzida de uma matriz: o algoritmo de Gram-Schmidt (GS), Gram-Schmidt Revisado (GSR), Gram-Schmid Revisado com Permutação de Colunas e, finalmente, o mais estável dos métodos aqui discutidos, um algoritmo de triangularização ortogonal que emprega refletores de Householder.

5.1 Fatoração QR reduzida e completa

Toda matriz $A \in \mathbb{R}^{m\times n}$ de posto n pode ser fatorada na forma

$$A = QR$$

onde a matriz $Q \in \mathbb{R}^{m\times n}$ possui colunas ortonormais ($Q^T Q = I_n$ - matriz identidade de ordem n) e a matriz $R \in \mathbb{R}^{n\times n}$ satisfaz $r_{ii} \neq 0$ para todo $i = 1, \ldots, n$ e também $r_{ij} = 0$ para todo par de índices j de coluna e i de linha que satisfazem $j < i$. Isto é, R é uma triangular superior, com diagonal principal não nula. Esta fatoração, em que Q é uma matriz retangular com número de colunas igual ao número de colunas de A é chamada de fatoração QR reduzida. Há uma outra fatoração, denominada fatoração QR completa, satisfazendo $Q \in \mathbb{R}^{m\times m}, Q^T Q = I_m$ e $R \in \mathbb{R}^{m\times n}$, em que as últimas $m - n$ colunas de Q fornecem uma base (ortonormal) para $N(A^T)$ e as últimas $m - n$ linhas de R são vetores de zeros, introduzidos em R para garantir conformabilidade.

A não ser que mencionemos o contrário, assumimos que A possui posto completo, ou seja, $posto(A) = n \leq m$. Também assumimos que as colunas de Q são definidas como $q_i \in \mathbb{R}^m : i = 1, \ldots, n$.

5.2 Importância da fatoração $A = QR$

Possivelmente, a fatoração SVD, *Sigular Value Decomposition* ou Decomposição em Valores Singulares ($A = U\Sigma V^T$), seja a mais relevante fatoração matricial conhecida, se forem considerados aspectos como as informações que revela sobre a matriz A e o uso que se faz destas informações. A fatoração SVD não apenas fornece bases ortonormais para os espaços fundamentais associados à matriz A, mas também revela, por meio da magnitude dos valores singulares σ_i, uma hierarquia sobre quais matrizes de rank-1 na forma $\sigma_i u_i v_i^T$ são as mais importantes para se aproximar a matriz fatorada.

Se do ponto de vista das informações que revela a fatoração SVD reina soberana, do ponto de vista algorítmico, pelo amplo uso que adquire em Computação Científica, inclusive como ingrediente essencial para se computar as fatorações SVD e espectral, a fatoração $A = QR$ é possivelmente a mais importante fatoração matricial. Esta fatoração é de importância crucial em Computação Científica.

Em capítulos anteriores, discutimos que para se resolver um Problema de Projeção em $C(A)$ ou para fazermos Ajuste de Curvas pelo método de Mínimos Quadrados, ou seja, para se resolver $\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$, precisamos resolver um sistema de equações normais

$$A^T A\hat{x} = A^T b.$$

O sistema linear acima é tipicamente mal-condicionado, isto é, o número de condição $\kappa(A^T A)$ de A tende a ser elevado pois é o quadrado do número de condição de A. Para uma matriz A quadrada, sabemos que o número de condição é definido como

$$\kappa_p(A) = \|A\|_p \|A^{-1}\|_p,$$

para qualquer norma matricial p induzida por norma vetorial p. No caso de uma matriz retangular, o número de condição pode ser equivalentemente redefinido como

$$\kappa_p(A) = \frac{\max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}}{\min_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}}. \tag{5.1}$$

Intuitivamente, $\kappa_p(A)$ será grande quando as colunas de A forem aproximadamente linearmente dependentes. Veja que, para o caso da norma matricial espectral (induzida pela norma 2 vetorial, ou norma Euclideana) a definição (5.1) acima equivale ao que conhecemos de nossos estudos anteriores: Se $A \neq A^T$, $\kappa_2(A) = \frac{\sigma_1(A)}{\sigma_n(A)} = \frac{\sqrt{\lambda_1(A^TA)}}{\sqrt{\lambda_n(A^TA)}}$, onde λ_1 e λ_n são, respectivamente, o maior e o menor autovalor de A^TA. Observe então que $\kappa_2(A^TA) = \frac{\lambda_1(A^TA)}{\lambda_n(A^TA)} = \kappa_2(A)^2$. Ou seja, um eventual mal condicionamento de A é agravado em A^TA.

Em resumo, a matriz de coeficientes A^TA do sistema de equações normais satisfaz $\kappa(A^TA) = \kappa(A)^2$, de forma que tende a ser mal condicionada. A fatoração de Cholesky, embora possa ser empregada dada a simetria e positividade de A^TA, na prática é muitas vezes desaconselhada. Ao invés dela, o sistema deve ser resolvido via Fatoração QR:

$$\begin{aligned} A^TA\hat{x} &= A^Tb \\ R^TQ^TQR\hat{x} &= R^TQ^Tb \\ R\hat{x} &= Q^Tb. \end{aligned} \tag{5.2}$$

O sistema linear triangular superior (5.2) é o sistema final que deve ser resolvido para se encontrar o vetor de pesos $\hat{x}$ desejado no ajuste de curvas. Se, ao invés do vetor de pesos, se desejar apenas o ponto $p \in C(A)$ correspondente à projeção de b em $C(A) = C(Q)$, basta calcularmos $p = \sum_{i=1}^{n} q_i(q_i^Tb)$.

As fatorações $PA = LU$ e de Cholesky são de grande importância para a resolução de sistemas lineares em que as matrizes de coeficientes (quadradas) são bem condicionadas, isto é, possuem número de condição pequeno. Estas duas fatorações básicas resolvem grande parte dos problemas de sistemas lineares com os quais podemos nos deparar. Entretanto, se a matriz de coeficientes do sistema for mal condicionada, como muitas vezes observado no caso das Equações Normais, fatorações numericamente mais estáveis devem ser empregadas. A fatoração $A = QR$ é a principal alternativa. Veja que em (5.2), suprimimos a necessidade de se resolver um sistema linear intermediário do tipo $Ly = Pb$, quando A^TA é fatorada via $PA = LU$, ou $Ly = b$, quando $A^TA = LL^T$ é fatorada via Cholesky. Quando dispomos da fatoração $A = QR$, a resolução do sistema linear intermediário pode ser substituído pela simples resolução de $Rx = Q^Tb$, mais estável numericamente.

Na próxima seção, apresentamos uma ideia elegante que permite, em teoria, a produção de uma fatoração QR para A. Trata-se do algoritmo de ortogonalização de bases de Gram-Schmidt (GS). Na prática, na sua versão clássica, GS não é um algoritmo recomendável pois é fortemente sujeito a erros numéricos e há perda de ortogonalidade entre as colunas de Q, sobretudo para as últimas colunas calculadas, aquela cujos índices são mais próximos de n. Estes aspectos serão discutidos e refinados na sequência.

5.3 Algoritmos para fatoração $A = QR$

5.3.1 Algoritmos baseados na ortogonalização de Gram-Schmidt

GS assim como outros algoritmos que produzem uma fatoração $A = QR$ exploram o fato de que, para A com posto completo, as primeiras k colunas de A e de Q produzem os mesmos subespaços vetoriais.

Para ilustrar a ideia central do procedimento, considere que A_k, q_k representam as $k-$ésimas colunas de A e de Q respectivamente. Então temos que:

$$\begin{aligned} span\{A_1\} &= span\{q_1\} \\ span\{A_1, A_2\} &= span\{q_1, q_2\} \\ \vdots \quad & \quad \vdots \\ span\{A_1, \ldots, A_n\} &= span\{q_1, q_2, \ldots, q_n\} \end{aligned}$$

Assim sendo, desejamos encontrar por meio da fatoração QR uma base mais conveniente para $C(A)$. E por mais conveniente queremos dizer, *base ortonormal.* Preservamos, entretanto, os subespaços gerados pela sequência de colunas de A. Veja que a forma abaixo

$$\left[\begin{array}{c|c|c|c} A_1 & A_2 & \cdots & A_n \end{array}\right] = \left[\begin{array}{c|c|c|c} q_1 & q_2 & \cdots & q_n \end{array}\right] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix} \tag{5.3}$$

fornece uma visão conveniente para a fatoração, ao demonstrar, por meio das entradas de R, como as bases para cada um destes n subespaços se relacionam. Verifique que o sistema (5.3) corresponde a:

- $A_1 = r_{11}q_1$
- $A_2 = r_{12}q_1 + r_{22}q_2$
- $\cdots$
- $A_i = \sum_{k=1}^{i} r_{ki}q_k$
- $A_n = r_{1n}q_1 + r_{2n}q_2 + \cdots + r_{nn}q_n$

onde:

- $q_i^T q_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$

- $r_{ii} \neq 0 : i = 1, \ldots, n$ (como consequência de $posto(A) = n$).

Além disso, também temos que:

$$span\{A_1\} \subseteq span\{A_1, A_2\} \subseteq \cdots \subseteq span\{A_1, A_2, \ldots, A_n\}.$$

Se as colunas de A são linearmente independentes, ao adicionarmos uma coluna nova, estamos gerando um espaço que contém os espaços anteriores, de forma estrita. Isto é, sempre teremos uma *informação nova*, ao inserirmos a coluna A_i em $span\{A_1, \ldots, A_{i-1}\}$. Na medida em que $A_i \notin span\{A_1, \ldots, A_{i-1}\} = span\{q_1, \ldots, q_{i-1}\}$, a condição $r_{ii} \neq 0$ deve ser observada. O quão mais próximo de zero for a entrada r_{ii}, mais próximo de linearmente dependente A_i será das primeiras $i - 1$ colunas de Q. Se $r_{ii} = 0$, A_i é linearmente dependente de $A_1, \ldots, A_{i-1}$.

Neste momento, estamos em condição de estabelecer a invariante do algoritmo de GS, isto é o conjunto de propriedades que se observam no início da $j-$ésima iteração típica do algoritmo. São elas:

1. As $j - 1$ colunas $q_i : i = 1, \ldots, j - 1$ são ortonormais;

2. Estas $j - 1$ colunas satisfazem:

$$span\{A_1, \ldots, A_{j-1}\} = span\{q_1, \ldots, q_{j-1}\}.$$

Satisfeitas estas propriedades, o objetivo ao longo da $j-$ésima iteração é determinar uma nova coluna q_j tal que:

1. $span\{A_1, \ldots, A_j\} = span\{q_1, \ldots, q_j\}$

2. $q_j \perp span\{q_1, \ldots, q_{j-1}\}$

3. $\|q_j\|_2 = 1$.

Essencialmente, o processo empregado para se determinar q_j como desejado acima é que merece o nome de *Ortonogalização de Gram-Schmidt*.

A ideia do processo de ortogonalização é simples e elegante. Para produzirmos a primeira coluna de Q, basta fazermos:

$$\begin{aligned} r_{11} &= \|A_1\|_2 \\ q_1 &= \frac{A_1}{r_{11}}. \end{aligned}$$

Para as demais colunas de Q, fazemos uso da invariante. Dispomos de $j-1$ colunas $q_1,\dots,q_{j-1}$ ortonormais satisfazendo $span\{A_1,\dots,A_{j-1}\} = span\{q_1,\dots,q_{j-1}\}$. Seguimos os seguintes passos:

1. Calculamos a projeção ortogonal p_j de A_j em $span\{q_1,\dots,q_{j-1}\}$. Sabemos que esta projeção ortogonal é dada por:

$$p_j = (q_1^T A_j)q_1 + (q_2^T A_j)q_2 + \cdots + (q_{j-1}^T A_j)q_{j-1}. \tag{5.4}$$

   Veja que se definirmos a matriz $Q_{j-1} \in \mathbb{R}^{m\times j-1}$ como a matriz formada pelas $j-1$ colunas de Q já calculadas, temos a expressão equivalente para a projeção p_j:

$$p_j = (Q_{j-1}Q_{j-1}^T)A_j. \tag{5.5}$$

2. Calculamos a diferença v_j (ou erro) entre A_j e sua projeção p_j e normalizamos a diferença. Este erro nada mais é do que a aplicação do projetor ortogonal complementar $I - Q_{j-1}Q_{j-1}^T$ ao vetor A_j:

$$v_j = (I - Q_{j-1}Q_{j-1}^T)A_j. \tag{5.6}$$

   A coluna q_j é este vetor erro após a normalização:

$$v_j = A_j - p_j \tag{5.7}$$

$$q_j = \frac{v_j}{\|v_j\|_2}. \tag{5.8}$$

Veja que na iteração típica essencialmente decompomos o vetor A_j em seus constituintes principais, isto é, os elementos $\{q_1,\dots,q_j\}$ da nova base para o espaço vetorial formado pelas primeiras j colunas de A:

$$A_j = \sum_{i=1}^{j} r_{ij}q_i$$

.

Por fim, cabe mencionar que todas as informações relativas à $j-$ésima coluna de R foram obtidas com o procedimento acima. Veja que $r_{jj} = \|v_j\|_2$ e que $r_{ij} = (q_i^T A_j) : i = 1,\dots,j-1$. Note que no procedimento que descrevemos acima preenchemos as informações (as entradas) da matriz R ao longo de suas colunas. Isto é, para um índice j de coluna a ser ortogonalizada, calculamos $r_{ij} : i = 1,2,\dots,j$, nesta ordem.

O processo acima é repetido até que a última coluna, A_n, tenha sido ortogonalizada.

Exemplo 37 *Vamos utilizar o procedimento de GS para produzir a fatoração QR de* $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & -3 \\ 0 & -2 & 3 \end{bmatrix}$. *Observe que as colunas* $A_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$, $A_3 = \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix}$ *são linearmente independentes.*

Primeira iteração - $j = 1$:

- *Dado A_1 desejamos $span\{q_1\} = span\{A_1\}$, $\|q_1\| = 1$.*
- *Então fazemos $q_1 = \frac{A_1}{\|A_1\|}$*
- *Ou seja, $r_{11} = \|A_1\|$*

Logo: $A_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $q_1 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $r_{11} = \sqrt{2}$.

Segunda iteração - $j = 2$:

- *Dados:* $q_1 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$
- *Projeção:* $p_2 = (q_1^T A_2)q_1 = \sqrt{2}q_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $r_{12} = \sqrt{2}$.
- *Erro:* $v_2 = A_2 - p_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\|v_2\| = r_{22} = \sqrt{6}$.
- $q_2 = \frac{\sqrt{6}}{6}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$

Terceira iteração - $j = 3$:

- *Dados:* $q_1 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $q_2 = \frac{\sqrt{6}}{6}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $A_3 = \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix}$

- *Projeção:* $p_3 = (q_1^T A_3)q_1 + (q_2^T A_3)q_2 = 3\sqrt{2}q_1 - \sqrt{6}q_2 = \begin{bmatrix} 2 \\ -4 \\ 2 \end{bmatrix}$, $r_{13} = 3\sqrt{2}, r_{23} = -\sqrt{6}$.

- *Erro:* $v_3 = A_3 - p_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

- $q_3 = \frac{\sqrt{3}}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $r_{33} = \sqrt{3}$.

Em resumo, obtemos:

$$\begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & -3 \\ 0 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \sqrt{2} & 3\sqrt{2} \\ & \sqrt{6} & -\sqrt{6} \\ & & \sqrt{3} \end{bmatrix}$$

O algoritmo de ortogonalização de GS é apresentado na Figura 5.1.

```
function [Q,R] = GS_Classico(A)
   [m,n] = size(A)
   R = zeros(n,n)
   Q = zeros(m,n)
   for j = 1:n
       V = A(:,j)
       for i = 1:j-1
           R(i,j) = Q(:,i)'*A(:,j)
           V = V - R(i,j)*Q(:,i)
       end
       R(j,j) = norm(V,2)
       Q(:,j) = 1.0/R(j,j) * V
   end
endfunction
```

Figura 5.1: Algoritmo clássico para ortogonalização de bases de GS.

Observe que a complexidade computacional de GS é de $O(mn^2)$ operações aritméticas. Este algoritmo assume que a matriz A possui posto coluna completo, uma vez que na atribuição

$$q_j = \frac{1}{r_{jj}} v_j$$

não é verificado se $r_{jj} \neq 0$. Na próxima seção, discutimos o que ocorre quando há dependência linear entre as colunas de A.

GS quando o posto é incompleto

Nesta subseção, vamos analisar o caso em que A seja uma matriz com $posto(A) = r < n = \min\{m, n\}$. Consideramos um exemplo que ilustra o resultado do algoritmo GS implementado de forma exata, isto é, sem a presença de erros numéricos. Posteriormente, discutimos o efeito dos erros numéricos.

Exemplo 38 *O objetivo deste exemplo é ilustrar o comportamento do processo de ortogonalização de de GS na ausência de erros numéricos. Para tanto, considere a matriz* $A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ -1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 1 \end{bmatrix}$ *que deve ser fatorada na forma* $A = QR$.

Primeira iteração - $j = 1$:

$$A_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}, \; q_1 = \tfrac{\sqrt{6}}{6}\begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}, \; r_{11} = \sqrt{6}.$$

Segunda iteração - $j = 2$:

$$A_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix}, \; q_1 = \tfrac{\sqrt{6}}{6}\begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}$$

Projeção: $p_2 = (q_1^T)A_2 q_1 = 0q_1$ *(ou seja: $q_1 \perp A_2$),* $r_{12} = 0$

$$q_2 = \tfrac{\sqrt{6}}{6}\begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix}, \; r_{22} = \sqrt{6}.$$

Terceira iteração - $j = 3$:

$$A_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \; q_1 = \tfrac{\sqrt{6}}{6}\begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}, \; q_2 = \tfrac{\sqrt{6}}{6}\begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix}$$

<u>*Projeção:*</u> $p_3 = (q_1^T)A_3q_1 + (q_2^T)A_3q_2$
$r_{13} = q_1^T A_3 = \sqrt{6}$, $r_{23} = q_2^T A_3 = 2\sqrt{6}$, $\Rightarrow \ \ p_3 = A_3$, $v_3 = 0, r_{33} = 0$.

Observe que neste momento detectamos que $A_3 \in span\{q_1, q_2\}$ (A_3 e sua projeção em $span\{q_1, q_2\}$ são o mesmo vetor) e a matriz A não possui posto completo. Isso significa que a coluna A_3 não é necessária para caraterizar $C(A)$. Desta forma, não é necessário que Q possua 4 colunas, mas no máximo 3 colunas. Similarmente, não é necessário que R possua 4 linhas, mas sim 3. *Assim sendo, na posição da terceira coluna de Q vamos armazenar alguma eventual coluna adicional, linearmente independente com q_1, q_2, necessária para descrever $C(A)$.*

<u>*Quarta iteração - $j = 4$:*</u>

$$A_4 = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \ q_1 = \frac{\sqrt{6}}{6} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}, \ q_2 = \frac{\sqrt{6}}{6} \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix}$$

<u>*Projeção:*</u> $p_4 = (q_1^T)A_4q_1 + (q_2^T)A_4q_2 + 0q_3$
$r_{14} = q_1^T A_4 = \frac{\sqrt{6}}{2}$, $r_{24} = q_2^T A_4 = -\frac{\sqrt{6}}{6}$, $r_{34} = 0$.

$$p_4 = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{6} \\ -\frac{5}{6} \\ \frac{5}{6} \end{bmatrix}$$

<u>*Erro:*</u> $v_4 = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{6} \\ -\frac{1}{6} \\ \frac{1}{6} \end{bmatrix}, \ r_{44} = \frac{\sqrt{3}}{3}, \ q_4 = \sqrt{3} \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{6} \\ -\frac{1}{6} \\ \frac{1}{6} \end{bmatrix}$

Neste momento, dispomos de uma fatoração $A = CR$ (isto mesmo, CR e não QR), onde $C \in \mathbb{R}^{4\times3}$ contempla apenas as colunas q_1, q_2, q_4, e $R \in \mathbb{R}^{3\times4}$. A coluna q_3 não foi empregada e uma linha a menos em R foi necessária, pois posto$(A) = 3$.

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ -1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 1 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}}{6} & 0 & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} \\ -1 & \frac{2\sqrt{6}}{6} & -\frac{\sqrt{3}}{6} \\ 2 & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 & \sqrt{6} & \frac{\sqrt{6}}{2} \\ & \sqrt{6} & 2\sqrt{6} & -\frac{\sqrt{6}}{6} \\ & & & \frac{\sqrt{3}}{3} \end{bmatrix}$$

Alguns textos denominam a fatoração acima de QR generalizada. De forma estrita, a fatoração acima não é uma QR (embora $Q^TQ = I_3$) pois R não é triangular superior. Veja que esta fatoração corresponde à soma de três matrizes de rank-1: $A = q_1r_1^T + q_2r_2^T + q_4r_4^T$, onde r_1^T, r_2^T, r_4^T são as linhas de coeficientes que obtivemos na ortogonalização de GS. O termo faltante, $q_3r_3^T$, que não adiciona posto à matriz

A, corresponde ao produto externo de uma coluna de zeros q_3 por uma linha de zeros r_3^T. Estas foram removidas na apresentação da fatoração acima.

Para obtermos a fatoração QR, neste caso completa para uma matriz de posto incompleto, pós multiplicaremos A por uma matriz de permutação P, que troca a posição das colunas 3 e 4 de A. Assim sendo, $C(AP)$ será representado pelo subespaço associado às primeiras três colunas de Q. A última coluna de Q será reservada par um vetor que caracteriza a base de $N(A^T)$. Vamos calcular esta coluna para produzir a fatoração QR completa de AP.

Vamos seguir os seguintes passos:

- *Apenas para facilitar as contas no exemplo, multiplicamos q_1, q_2, q_4 respectivamente por $\frac{6}{\sqrt{6}}, \frac{6}{\sqrt{6}}, \frac{6}{\sqrt{3}}$, e resolvemos o sistema linear:*
$$\begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 2 & 1 \\ -3 & 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

- *Após eliminação temos* $\begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & -6 & 6 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, *que nos permite obter* $y = \begin{bmatrix} 1 & 3 & -1 & -1 \end{bmatrix}^T$, *que normalizado dará origem a uma coluna de Q como* $\frac{1}{\|y\|} y = \frac{\sqrt{3}}{6} \begin{bmatrix} 1 & 3 & -1 & -1 \end{bmatrix}^T$.

- *Para termos a estrutura que desejamos, com as três primeiras colunas de Q produzindo uma base ortonormal para $C(A)$ e a última para $N(A^T)$ fazemos um pivoteamento de colunas, trocando a 4a. com a 3a., de forma que as três primeiras entradas na diagonal de R sejam não nulas, uma vez que posto$(A) =$ 3.*

Então temos a seguinte fatoração completa para AP:

$$AP = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ -1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} \frac{\sqrt{6}}{6} & 0 & -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{6} \\ 0 & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} & \frac{3\sqrt{3}}{6} \\ -\frac{\sqrt{6}}{6} & \frac{2\sqrt{6}}{6} & -\frac{\sqrt{3}}{6} & -\frac{\sqrt{3}}{6} \\ \frac{2\sqrt{6}}{6} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} & -\frac{\sqrt{3}}{6} \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 & \frac{\sqrt{6}}{2} & \sqrt{6} \\ & \sqrt{6} & -\frac{\sqrt{6}}{6} & 2\sqrt{6} \\ & & \frac{\sqrt{3}}{3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = QR$$

Observe que na fatoração QR acima indicada, a terceira coluna de Q representa a ortogonalização da projeção de A_4 em span$\{q_1, q_2\}$ e que a quarta coluna de Q é uma base ortonormal para $N(A^T)$. Note também que a quarta linha de R é identicamente nula já que a quarta coluna de Q na representação acima não deve ser empregada na combinação linear que descreve as colunas de A.

No próximo exemplo, apresentamos o resultado do algoritmo GS na presença de erros numéricos, considerando uma representação de ponto flutuante padronizada que emprega 64 bits, dos quais 52 são dedicados à mantissa. Como tal, a precisão desta máquina é $\epsilon \approx 10^{-16}$.

Antes de apresentarmos o exemplo, vamos supor que A_j seja a primeira coluna linearmente dependente das demais colunas $q_1, \ldots, q_{j-1}$. Como vimos no exemplo anterior, na ausência de erros numéricos, o valor correto de r_{jj} seria zero. Entretanto, na presença de erros numéricos, este valor de r_{jj} não será exatamente zero, mas um valor muito pequeno, de magnitude muito menor que a magnitude das demais entradas de R. Esse aspecto e suas consequências são os pontos centrais a serem examinados no exemplo que segue.

Exemplo 39 *O objetivo deste exemplo é ilustrar o comportamento do algoritmo GS clássico, sem nenhuma alteração, quando a matriz de entrada é deficiente em posto. Considere a matriz A de posto incompleto* $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}$*, e verifique que* $A_3 = 2A_2 - A_1$*,* $A_4 = 2A_3 - A_2$*, de forma que posto*$(A) = 2$*. Esta matriz possui deficiência de posto (em duas unidades).*

```
-->[q,r] = GS_Classico(A)
 q  =
   0.09245      0.5392919  -0.5465825  -0.5465825
   0.4622502    0.2927584  -0.370154    -0.370154
   0.8320503    0.046225   -0.2118872  -0.2118872
   0.09245     -0.0616334   0.0397829   0.0397829
   0.2773501   -0.7858253   0.7195517   0.7195517
 r  =
   10.816654    11.926054   13.035455   14.144855
   0.           1.6641006   3.3282012   4.9923018
   0.           0.          3.209D-14  -7.650426
   0.           0.          0.          7.650426
```

Veja que a entrada $r_{33} \approx 10^{-14}$ indica que a terceira coluna de A é linearmente dependente de A_1, A_2 e, portanto de q_1, q_2. Apesar disso, o procedimento prossegue e produz as colunas q_3 e q_4 que não tem significado numérico algum. Em particular, o resultado correto esperado para as grandezas computadas no passo em que q_3 é avaliada é um vetor de zeros, em função da dependência linear de A_3 com q_1, q_2. Observação idêntica vale para a coluna q_4 obtida na sequência.

Vamos analisar o que aconteceu. O resultado do erro v_3, antes do passo de normalização, é um vetor bastante próximo de zero. Sua norma é r_{33}. Porém, como este vetor foi normalizado para se produzir q_3, esta terceira coluna de Q deveria ser nula e não é. Após a normalização, esta coluna é linearmente independente de q_1, q_2, simplesmente porque v_3 foi dividido por sua norma. Portanto, não tem nenhum significado. Na sequência, esta coluna q_3 é usada na projeção de A_4 em span$\{q_1, q_2, q_3\}$. Naturalmente, a coluna q_4 não tem significado algum, pois emprega uma coluna q_3 inexistente para caracterizar $C(A)$. O resultado acima sugere que a matriz fatorada possui posto 3 *(3 entradas significativamente distintas de zero na diagonal de R), quando esse não é o caso.*

Ortogonalização de GS revisada

O algoritmo clássico de ortogonalização de GS visto na seção anterior pode ser aprimorado, visando ganhar mais estabilidade numérica. Essencialmente, o algoritmo desconsidera a magnitude das projeções das colunas nas bases calculadas. Além disso, o erro numérico associado às colunas de Q de índices maiores (mais próximos de n) aumenta, de forma que, na prática, ao final do processo de GS podemos ter $\|Q^T Q - I\|$ muito grande.

Os erros na ortogonalização clássica de GS tendem a ser maiores em função de dois motivos:

1. O primeiro deles, inerente ao desenho ou à concepção do algoritmo, é que estamos usando um procedimento de triangularização para obtermos uma matriz Q.

   Em outras palavras, o processo de obtenção de Q pode ser reinterpretado como o de aplicar uma sequência de matrizes triangulares superiores R_i convenientemente escolhidas em A de forma que $(((AR_1)R_2)\dots)R_n = Q$, ou seja, que ao final obtenhamos a matriz Q desejada.

   Quando aplicamos R_1 em A, isto é, quando calculamos AR_1, obtemos uma matriz cuja primeira coluna é a q_1 de GS. Veja na transformação linear abaixo o formato da triangular R_1 e o resultado de AR_1.

$$\left[\begin{array}{c|c|c|c} A_1 & A_2 & \cdots & A_n \end{array}\right]\begin{bmatrix} \frac{1}{r_{11}} & \frac{-r_{12}}{r_{11}} & \frac{-r_{13}}{r_{11}} & \cdots & \frac{-r_{1n}}{r_{11}} \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} = \left[\begin{array}{c|c|c|c} q_1 & A_2R_1 & \cdots & A_nR_1 \end{array}\right] \tag{5.9}$$

Ao aplicarmos R_2 à direita de AR_1, isto é com a transformação linear $(AR_1)R_2$, preservamos a primeira coluna q_1 e construímos a segunda, q_2. Esse processo se repete até que $AR_1R_2 \dots R_n$ seja a matriz Q que o procedimento de GS calcula. Veja a segunda iteração de GS detalhada abaixo.

$$\left[\begin{array}{c|c|c|c|c} q_1 & A_2R_1 & A_3R_1 & \cdots & A_nR_1 \end{array}\right]\begin{bmatrix} 1 & & & & \\ & \frac{1}{r_{22}} & \frac{-r_{23}}{r_{22}} & \cdots & \frac{-r_{2n}}{r_{22}} \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} = \left[\begin{array}{c|c|c|c|c} q_1 & q_2 & A_3R_1R_2 & \cdots & A_nR_1R_2 \end{array}\right] \tag{5.10}$$

Observe que cada uma das matrizes R_i é uma matriz triangular superior, com diagonal unitária, que difere da matriz identidade apenas para a linha i e as colunas de índice $j \geq i$ naquela linha. Esta reinterpretação será examinada em detalhes, mais tarde neste capítulo, quando apresentarmos a fatoração QR por refletores de Householder. O fato objetivo e relevante neste momento da exposição é que este processo de triangularização é pouco (ou menos) estável numericamente. Veja que se a magnitude de r_{ii} é muito pequena, as grandezas ao longo da linha i de R_i tendem a ser muito grandes. O resultado é que, neste caso, a transformação linear $(AR_1 \dots R_{i-1})R_i$ produz erros numéricos grandes.

Recorde-se que um problema análogo ao descrito acima surge na Eliminação de Gauss, quando os pivôs são muito pequenos. No caso da Eliminação de Gauss, lidamos com este aspecto implementando pivoteamento de linhas, visando encontrar o elemento pivô de maior magnitude e, com isso, reduzindo erros numéricos. Aqui, no contexto da ortogonalização de GS, aplicaremos ideia análoga, que é o pivoteamento de colunas. A implementação de pivoteamento de colunas em GS será apresentada na seção 5.3.2.

2. O segundo aspecto é que calculamos o erro v_j de uma única vez, após todas as colunas $q_i : i = 1, \dots, j-1$ terem sido calculadas, segundo a expressão:

$$v_j = (I - Q_{j-1}Q_{j-1}^T)A_j$$

ou, equivalentemente,

$$v_j = A_j - \sum_{i=1}^{j-1} (q_i^T A_j) q_i$$

No algoritmo, esta expressão de erro é calculada no laço:

```
V = A(:,j)
for i = 1:j-1
    R(i,j) = Q(:,i)'*A(:,j)
    V = V - R(i,j)*Q(:,i)
end
```

Figura 5.2: Trecho de interesse do procedimento de ortogonalização de GS.

Para tratar este segundo aspecto, e conferir maior estabilidade numérica ao procedimento de ortogonalização de GS, faremos uma alteração na ordem em que o desconto do termo $(q_i^T A_j)q_i$ de A_j ocorre. Esta alteração dá origem à versão revisada de GS. Na versão revisada, o desconto ocorrerá o mais cedo possível, assim que a coluna $q_i : i < j$ esteja disponível. Esta antecipação do desconto modifica o comportamento numérico do algoritmo mas não muda a equivalência matemática entre os dois procedimentos, GS Clássico e GS Revisado, na ausência de erros numéricos.

O algoritmo revisado (GSR) é apresentado na Figura 5.3. Veja no algoritmo revisado que a matriz V recebe uma cópia da matriz A a ser fatorada e que cada vez que uma coluna q_i de Q é calculada, as colunas de V de índice $j > i$ são descontadas de sua projeção em q_i.

O resultado seguinte demonstra que os dois algoritmos são matematicamente equivalentes, isto é, computam as mesmas grandezas na ausência de erros numéricos.

Resultado 5.3.1 *Vamos mostrar que, ao final da iteração $i = k$, a expressão armazenada para a coluna v_k de V no algoritmo revisado é idêntica ao vetor v_j do algoritmo clássico. Na sequência, mostramos o efeito do desconto das colunas $q_i : i = 1, \ldots, k-1$ até o momento em que q_k é calculada no algoritmo revisado:*

```
function [Q,R] = GS_Revisado(A)
   [m,n] = size(A)
   R = zeros(n,n)
   Q = zeros(m,n)
   V = A
   for i = 1:n
       R(i,i) = norm(V(:,i),2)
       Q(:,i) = V(:,i)/R(i,i)
       for j = (i+1):n
           R(i,j) = Q(:,i)'*V(:,j)
           V(:,j) = V(:,j) - R(i,j)*Q(:,i)
       end
   end
endfunction
```

Figura 5.3: Algoritmo de ortogonalização de GS revisado. Algoritmo assume que posto coluna é completo.

$\Rightarrow$*Primeira iteração*, $i = 1$

$$\begin{aligned} v_k &= A_k - (q_1^T A_k) q_1 \\ &= A_k - r_{1k} q_1 \end{aligned}$$

$\Rightarrow$*Segunda iteração*, $i = 2$

$$\begin{aligned} v_k &= v_k - (q_2^T v_k) q_2 \\ &= A_k - r_{1k} q_1 - (q_2^T (A_k - r_{1k} q_1)) q_2 \\ &= A_k - r_{1k} q_1 - (q_2^T A_k) q_2 - r_{1k} (q_2^T q_1) q_2 \\ &= A_k - r_{1k} q_1 - r_{2k} q_2 \qquad \textit{uma vez que } (q_2^T q_1) = 0 \\ &= A_k - \sum_{j=1}^{i} (q_j^T A_k) q_j \\ &= A_k - q_1 q_1^T A_k - q_2 q_2^T A_k \\ &= A_k - Q_2 Q_2^T A_k \qquad \textit{lembrando que } Q_2 \textit{ contém colunas } q_1, q_2 \\ &= (I - Q_2 Q_2^T) A_k \end{aligned}$$

$\vdots\vdots$

$\Rightarrow$$i = (k-1) -$ *ésima iteração*

$$\begin{aligned} v_k &= A_k - r_{1k} q_1 - r_{2k} q_2 - \cdots - r_{k-1} q_{k-1} \qquad \textit{uma vez que } (q_i^T q_j) = 0, i, j < k, i \neq j \\ &= (I - Q_{k-1} Q_{k-1}^T) A_k \end{aligned}$$

$\Rightarrow$$k -$ *ésima iteração*, $i = k$

$$r_{kk} = \|(I - Q_{k-1} Q_{k-1}^T) A_k\|_2$$

$$v_k = \frac{v_k}{r_{kk}} \quad (= q_k)$$

Veja que a expressão $v_k = (I - Q_{k-1}Q_{k-1}^T)A_k$ produzida pelo algoritmo GSR é a mesma expressão (5.6) produzida pelo GS para o erro. Porém, na presença de erros numéricos, os dois algoritmos produzem resultados distintos. Para ilustrar este ponto, considere o exemplo computacional seguinte.

Exemplo 40 *Vamos empregar o algoritmo GS revisado para fatorar a matriz de posto* 2 *empregada no exemplo* (39). *Veja o resultado da fatoração:*

```
-->[Q,R] = GS_Revisado(A)
 Q  =
   0.09245      0.5392919  -0.1740118   0.2700975
   0.4622502    0.2927584  -0.4524306  -0.8757263
   0.8320503    0.046225   -0.843957    0.3894845
   0.09245     -0.0616334  -0.0957065   0.0170553
   0.2773501   -0.7858253  -0.2088141   0.0903422
 R  =
   10.816654    11.926054   13.035455   14.144855
   0.           1.6641006   3.3282012   4.9923018
   0.           0.          6.380D-15   8.836D-15
   0.           0.          0.          7.599D-16
```

Embora as colunas q_3, q_4 de Q oferecidas pelo algoritmo não representem nada, pois o algoritmo assume que o posto da matriz é completo, os valores de r_{33} e r_{44} indicam que o posto de A é 2.

5.3.2 Ortogonalização de GS com pivoteamento de colunas

Nesta seção, apresentamos como implementar o pivoteamento de colunas em GS. Antes de apresentarmos o algoritmo propriamente, vamos ilustrar a ideia da permutação, sem considerar os erros numéricos, considerando para isso o método de ortogonalização de GS clássico. Na sequência, combinaremos a permutação de colunas com a versão revisada de GS para obter um algoritmo bastante mais estável. Veja que ao assim procedermos, vamos tratar os dois problemas de GS que descrevemos na seção anterior.

Essencialmente, as premissas para o pivoteamento das colunas são as seguintes:

- Como nehuma coluna de Q foi computada, na primeira iteração ($j = 1$), a coluna de A de maior norma Euclideana é a escolhida para gerar a primeira coluna de Q, q_1.

- Nas demais iterações ($j \geq 2$), consideramos o subsepaço $span\{q_1, \ldots, q_{j-1}\}$ gerado pelas $j-1$ colunas já calculadas de Q até aquela iteração e fazemos o seguinte:
  - Seja $\mathcal{K}^j \subset \{1, 2, \ldots, n\}$ o conjunto de índices de colunas de A que não foram empregadas para gerar colunas de Q até a iteração j.
  - Para cada coluna $A_k : k \in \mathcal{K}^j$, calculamos p_k, a projeção da coluna A_k em $span\{q_1, \ldots, q_{j-1}\}$ e, na sequência, calculamos a diferença $v_k = A_k - p_k$.
  - A coluna $A_k : k \in \mathcal{K}^j$ que gera a coluna q_j de Q é aquela que tem a maior diferença, medida pela norma Euclideana, ou seja:

    $$k = \arg\max \|v_j\|_2 : j \in \mathcal{K}^j$$

    Veja que quando escolhemos a coluna com base em $\arg\max \|v_j\|_2$ geramos uma matriz R cujas entradas na diagonal principal são não crescentes, favorecendo a estabilidade numérica e a detecção do posto da matriz. Quando um determinado valor $|r_{jj}|$, escolhido mediante este critério, for inferior a um determinado valor de tolerância numérica, podemos estabelecer que $posto(A) = j - 1$.

Para ilustrar a ideia do pivoteamento de colunas em GS clássico, considere o seguinte exemplo, implementado na ausência de erros numéricos.

Exemplo 41 *Vamos fatorar a matriz A abaixo usando a ideia do pivoteamento de colunas.* $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}$.

Primeira coluna de Q, j = 1:

- *A coluna de maior norma Euclideana de A é A_4, com $\|A_4\| = 15$.*
- $r_{11} = 15$, $pivot(1) = 4, pivot(4) = 1$
- $q_1 = \frac{1}{15} \begin{bmatrix} 4 \\ 8 \\ 12 \\ 1 \\ 0 \end{bmatrix}$

Segunda coluna de Q*,* $j = 2$:

- *Colunas candidatas:* $\begin{bmatrix} A_1 & A_2 & A_3 \\ \hline 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \\ 1 & 1 & 1 \\ 3 & 2 & 1 \end{bmatrix}$, $q_1 = \frac{1}{15}\begin{bmatrix} 4 \\ 8 \\ 12 \\ 1 \\ 0 \end{bmatrix}$

- *Produtos internos:* $q_1^T A_1 = \frac{153}{15}$, $q_1^T A_2 = \frac{177}{15}$, $q_1^T A_3 = \frac{201}{15}$.

- *Projeções:* $p_1^1 = \frac{153}{15} q_1$, $p_2^1 = \frac{177}{15} q_1$, $p_3^1 = \frac{201}{15} q_1$

- *Diferenças:* $v_1^1 = A_1 - \frac{153}{15} q_1$, $v_2^1 = A_2 - \frac{177}{15} q_1$, $v_3^1 = A_3 - \frac{201}{15} q_1$

Ou seja:

- $v_1^1 = \frac{1}{225}\begin{bmatrix} -387 \\ -99 \\ 189 \\ 72 \\ 675 \end{bmatrix}$, $v_2^1 = \frac{1}{225}\begin{bmatrix} -162 \\ 126 \\ 414 \\ 72 \\ 450 \end{bmatrix}$, $v_3^1 = \frac{1}{225}\begin{bmatrix} 63 \\ 351 \\ 639 \\ 72 \\ 225 \end{bmatrix}$

- $\|v_1^1\| = \frac{810}{225}$, $\|v_2^1\| < \frac{649}{225}$, $\|v_3^1\| < \frac{769}{225}$

- *Logo,* q_2 *vem de* v_1^1, $r_{22} = \frac{810}{225} = 3.6$ *(pivot*$(2) = 1$*)*

- $q_2 = v_1^1 : \left(\frac{810}{225}\right) = \frac{1}{810}\begin{bmatrix} -387 \\ -99 \\ 189 \\ 72 \\ 675 \end{bmatrix}$

- *Como* q_2 *veio de* A_1, $r_{12} = q_1^T A_1 = \frac{153}{15} = 10.2$

Terceira coluna de Q*,* $j = 3$:

- *Candidatas:* $\begin{bmatrix} A_2 & A_3 \\ \hline 2 & 3 \\ 6 & 7 \\ 10 & 11 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}$, $q_1 = \frac{1}{15}\begin{bmatrix} 4 \\ 8 \\ 12 \\ 1 \\ 0 \end{bmatrix}$, $q_2 = \frac{1}{810}\begin{bmatrix} -387 \\ -99 \\ 189 \\ 72 \\ 675 \end{bmatrix}$

- *Produtos internos já calculados:* $q_1^T A_2 = \frac{177}{15}$, $q_1^T A_3 = \frac{201}{15}$.

- *Novos produtos internos:* $q_2^T A_2 = \frac{12}{5}$, $q_2^T A_3 = \frac{6}{5}$

- *Projeções:* $p_2^2 = (q_1^T A_2)q_1 + (q_2^T A_2)q_2 = \frac{177}{15}q_1 + \frac{12}{5}q_2$,
  $p_3^2 = (q_1^T A_3)q_1 + (q_2^T A_3)q_2 = \frac{201}{15}q_1 + \frac{6}{5}q_2$

- *Diferenças:* $v_2^2 = A_2 - p_2^2$ *e* $v_3^2 = A_3 - p_3^2$ *Ou seja:*

- *Projeções:* $p_2^2 = (q_1^T A_2)q_1 + (q_2^T A_2)q_2 = \frac{177}{15}q_1 + \frac{12}{5}q_2$,
  $p_3^2 = (q_1^T A_3)q_1 + (q_2^T A_3)q_2 = \frac{201}{15}q_1 + \frac{6}{5}q_2$

- $p_2^2 = \begin{bmatrix} 2 \\ 6 \\ 10 \\ 1 \\ 2 \end{bmatrix} = A_2 \rightarrow v_2^2 = 0$

- $p_3^2 = \begin{bmatrix} 3 \\ 7 \\ 11 \\ 1 \\ 1 \end{bmatrix} = A_3 \rightarrow v_3^2 = 0$

- *Neste momento, sabemos que posto*$(A) = 2$ *e concluímos a fatoração reduzida.*

- *As demais colunas na fatoração completa, caso seja desejada, devem ser resolvidas via eliminação.*

- *Com o vetor pivô, construímos a matriz* P *em* $AP = QR$*, onde* Q *neste caso (fatoração reduzida) é uma matriz* 5×2 *e a matriz* R *é* 2×4.

Veja que a fatoração $AP = QR$ reduzida ficou assim:

$$
\begin{aligned}
AP &= \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{4}{15} & -\frac{387}{810} \\ \frac{8}{15} & -\frac{99}{810} \\ \frac{12}{15} & \frac{189}{810} \\ \frac{1}{15} & \frac{72}{810} \\ 0 & \frac{675}{810} \end{bmatrix} \begin{bmatrix} 15 & \frac{153}{15} & \frac{177}{15} & \frac{201}{15} \\ 0 & 3.6 & \frac{12}{5} & \frac{6}{5} \end{bmatrix} \\
&= QR
\end{aligned}
$$

Retornamos agora à descrição do algoritmo que combina tanto a versão revisada da ortogonalização de GS quanto a permutação de colunas. Este algoritmo é apresentado na Figura 5.4. As funções auxiliares para o algoritmo são apresentadas na Figura 5.5. O algoritmo computa a fatoração $AP = QR$ e também identifica o posto numérico da matriz de entrada A.

Assim como na versão revisada, uma cópia de A é feita na matriz V, de forma que ao longo das iterações do algoritmo, as colunas de V armazenam as diferenças entre as colunas de A e suas projeções nas colunas de Q já determinados. A lógica da troca de colunas é a seguinte. No início da $i-$ésima iteração, verifica-se qual coluna armazenada nas posições $V_i, V_{i+1}, \ldots, V_n$ tem a maior norma Euclideana, e portanto dista mais de $span\{q_1, \ldots, q_{i-1}\}$, onde $q_1, \ldots, q_{i-1}$ são as colunas de Q geradas até a iteração $i - 1$. Se esta coluna de maior norma Euclideana não estiver armazenada na coluna V_i, isto é, se estiver armazenada na coluna V_p, trocamos o conteúdo da coluna V_i e V_p de V, atualizando também as entradas do vetor *pivot*.

A detecção do posto numérico de A fica então facilitada. Sempre que a norma Euclideana máxima das colunas $V_i, V_{i+1}, \ldots, V_n$ for inferior a um valor de controle, que depende de alguma norma de A e da precisão da máquina, o posto de A é o valor de $i - 1$. Assim sendo, quando $|r_{ii}|$ é inferior à tolerância, o algoritmo interrompe a ortogonalização das colunas restantes e retorna a fatoração de colunas, na qual Q tem como número de colunas o posto numérico detectado.

Exemplo 42 *Vamos ilustrar o resultado do algoritmo de ortogonalização de GS revisado com permutação de colunas para a matriz A do exemplo 41.*

```
-->[Q,R] = GS_Revisada_PermutaColunas(A)
```

```
function [Q,R,pivot] = GS_Revisada_PermutaColunas(A)
   [m,n] = size(A)
   nainf = norm(A,'inf'); eps = 1.0E-14
   pivot = zeros(n); R = zeros(n,n); Q = zeros(m,n)
   posto = n;
   for i=1:n
       pivot(i) = i
   end
   V = A
   for i = 1:n
       [p,nmax] = DeterminaNormaMaxima(V,i,n)
       if (p <> i) then
           [R,V,pivot] = TrocaConteudoColunas(i,p,R,V,pivot)
       end
       R(i,i) = nmax
       if ((nmax < eps * nainf) & (posto == n)) then
           posto = i-1
           break
       end
       Q(:,i) = V(:,i)/R(i,i)
       for j = (i+1):n
           R(i,j) = Q(:,i)'*V(:,j)
           V(:,j) = V(:,j) - R(i,j)*Q(:,i)
       end
   end
   if posto <> n
       Q = Q(:,1:posto)
       R = R(1:posto,:)
       printf("Rank numérico detectado: %d \n",posto)
   end
endfunction
```

Figura 5.4: Algoritmo de ortogonalização de Gram-Schmidt revisado com permutação de colunas.

```
Rank numérico detectado: 2
 Q  =
   0.2666667  -0.4777778
   0.5333333  -0.1222222
   0.8         0.2333333
   0.0666667   0.0888889
   0.          0.8333333
 R  =
   15.   10.2   11.8   13.4
```

```
function [R,V,pivot] = TrocaConteudoColunas(i,p,R,V,pivot)
    ColAux = V(:,i)
    V(:,i) = V(:,p)
    V(:,p) = ColAux
    ColAux2 = R(1:i-1,i)
    R(1:i-1,i) = R(1:i-1,p)
    R(1:i-1,p) = ColAux2
    t = pivot(i)
    pivot(i) = pivot(p)
    pivot(p) = t
endfunction
function [p,nmax] = DeterminaNormaMaxima(V,ini,n)
    p = ini
    nmax = norm(V(:,ini),2)
    for j = ini+1:n
        nor = norm(V(:,j),2)
        if (nor > nmax) then
            nmax = nor
            p = j
        end
    end
endfunction
```

Figura 5.5: Funções auxiliares para o algorimto de ortogonalização de Gram-Schmidt revisado com permutação de colunas.

```
   0.    3.6    2.4    1.2
 pivot  =
   4.
   1.
   2.
   3.
```

5.4 Análise de erros de arrendondamento e reortogonalização

A qualidade da fatoração $A = QR$ (ou idealmente, $AP = QR$ quando há incorporação de pivoteamento de colunas no processo), pode ser avaliada por meio da grandeza $\|I - Q^T Q\|$. Quanto melhor a fatoração, mais próximo de zero $\|I - Q^T Q\|$ deve ser, indicando que as colunas de A (ou de AP) foram bem ortogonalizadas.

Nesta seção, vamos apresentar um experimento computacional comparando as fatorações QR que já estudamos até o momento e também uma outra, bastante mais

m	n	$\kappa(V)$	$\|I - Q^TQ\|$ para diversos algoritmos GS Clássico	GS Rev	GS Rev + Permuta	qr Scilab
6	4	1.066D+02	5.243D-15	4.696D-15	7.226D-15	9.174D-16
9	6	2.752D+03	8.253D-11	1.283D-13	1.363D-13	6.753D-16
12	8	7.280D+04	0.0000026	3.274D-12	2.468D-12	9.491D-16
15	10	1.952D+06	0.1265365	8.151D-11	3.466D-11	6.636D-16
18	12	5.280D+07	2.6409387	1.037D-09	1.669D-09	8.429D-16
25	20	3.243D+14	11.387547	0.0079543	0.0083214	1.314D-15

Tabela 5.1: Resultado de diversas fatorações QR aplicados a matrizes de Vandermonde de diferentes dimensões e números de condição.

estável, disponível em pacotes de Álgebra Linear Numérica, como `Scilab, MATLAB, NumPy`, que emprega Refletores de Householder. O algoritmo que produz essa fatoração será nosso objeto de estudo detalhado nas últimas seções deste capítulo.

Para nosso experimento computacional, vamos assumir que a matriz a ser fatorada possui posto completo, $posto(A) = n$. Idealmente, ao final do processo de fatoração, devemos ter $I - Q^TQ = 0$. Como trabalhamos com precisão finita, a fatoração será considerada boa se $\|I - Q^TQ\|_2 \approx c\epsilon$, onde c é uma constante pequena e $\epsilon \approx 1.11 \times 10^{-16}$ é a precisão da máquina,

Vamos comparar os algoritmos que estudamos até o momento para fatorar uma matriz de Vandermonde V, retangular $m \times n$, para diversos valores de m, n. As entradas da matriz de Vandermonde considerada são

$$v_{ij} = \left(\frac{j}{n}\right)^{i-1}.$$

Para cada algoritmo, avaliamos o indicador $\|I - Q^TQ\|_2$ obtido com a matriz Q produzida pelo algoritmo. Os resultados obtidos por cada algoritmo, para dimensão m, n da matriz de Vandermonde, são apresentados na Tabela 5.1.

Na tabela, $\kappa_p(V)$ é o número de condição da matriz de Vandermonde (empregamos $p = 2$). Observe a diferença de ordem de grandeza das quantidades $\|I - Q^TQ\|$ computadas pelos diversos algoritmos. Veja que para a condição mais extrema testada, isto é, quando $\kappa(V) \approx 3.243D+14$, apenas o algoritmo que emprega refletores de Householder apresentou resultados satisfatórios. Todos os demais produziram colunas pouco ortogonais.

Se as colunas de Q não forem suficientemente ortogonais como no caso acima ilustrado, podemos reortogonalizar Q. Basta fatorarmos a matriz Q em sua QR algumas vezes. Com isso, devemos melhorar a qualidade dos fatores Q e R produzidos. Na Figura 5.6, ilustramos o processo de *reortogonalização automática*, considerando como algoritmo base, o GS Revisado. A ideia consiste em aplicar o mesmo algoritmo de fatoração, GS revisado, tantas quantas vezes forem necessárias, até que a grandeza $\|I - Q^T Q\|$ avaliada para a matriz Q produzida seja suficientemente boa. Na primeira aplicação de GS Revisado, a matriz de entrada é a matriz A a ser fatorada. Nas seguintes, sempre passamos o último fator Q encontrado. Veja que durante o processo, precisamos armazenar o produto das matrizes R's encontradas.

Cabe destacar que no algoritmo de reortogonalização automática apresentado na Figura 5.6, a reortogonalização é realizada *a posteriori*, isto é, quando todas as colunas de Q foram calculadas. Na prática, melhores resultados são obtidos quando a segunda (ou as múltiplas) ortogonalização(ões) é (são) feita(s) internamente, ao longo da ortogonalização promovida pelo algoritmo.

```
function [Q,R] = Fatoracao_Revisada_Reortogonalizacao(A)
   [m,n] = size(A)
   [Q,R] = GS_Revisado(A)
   n2 = norm(eye(n,n)-Q'*Q,2)
   printf("%4.3E \n",n2)
   while (n2 > 100*%eps)
     [Qn,Rn] = GS_Revisado(Q)
      n2 = norm(eye(n,n)-Qn'*Qn,2)
      printf("%4.3E \n",n2)
      R = Rn*R
   end
endfunction
```

Figura 5.6: Procedimento de ortogonalização automática.

Veja o resultado da reortogonalização para a matriz V de ordem 25×20. Com uma única reortogonalização adicional, conseguimos reduzir $\|I - Q^T Q\|$ de $7.954E-03$ para $4.572E-16$, algo próximo da precisão da máquina.

```
-->V = GeraVandermondeMod(25,20);
-->[Q,R] = Fatoracao_Revisada_Reortogonalizacao(V);
7.954E-03
4.572E-16
-->norm(V-Q*R,'inf')
 ans  =

   1.634D-12
```

Na próxima seção, vamos estudar os refletores de Householder e a fatoração QR que faz uso destes refletores.

5.5 Triangularização de Householder

Nesta seção, voltamos a assumir que $A \in \mathbb{R}^{m\times n}$ tem posto completo, isto é, $posto(A) = n$. Quando relaxarmos esta hipótese, faremos menção explícita ao fato.

Em nossos estudos anteriores, mais precisamente na Seção 5.3.1, mencionamos que podemos interpretar o procedimento de ortogonalização de bases de GS como o processo de escolher matrizes triangulares superiores convenientes, aplicando-as à direita de A e com isso ortogonalizar a matriz. Veja: $AR_1R_2\dots R_n = Q$. Recorde-se da estrutura das matrizes triangulares superiores discutida naquela seção e veja que estas matrizes $R_i : i = 1,\dots,n$ são quadradas com diagonal não nula, portanto são não singulares, de forma que admitem inversa. Por esta razão podemos escrever que

$$\begin{gathered} AR_1R_2\dots R_n = Q \\ AR_1R_2\dots R_n(R_n)^{-1}\dots(R_2^{-1})(R_1^{-1}) = Q(R_n)^{-1}\dots(R_2^{-1})(R_1^{-1}) \\ \Rightarrow \\ A = QR \end{gathered}$$

onde $R = (R_n)^{-1}\dots(R_2^{-1})(R_1^{-1})$. Em outras palavras, GS ortogonaliza por meio de transformações lineares triangulares.

Uma ideia complementar a esta consiste em, por meio de transformações lineares ortogonais (mais estáveis), triangularizar a matriz A. Veja que se conseguirmos fazer uma sequência de transformações ortogonais que produza uma triangular superior R com a diagonal não nula, isto é,

$$Q_n\dots Q_2Q_1A = R,$$

temos uma fatoração QR para A.

Este é exatamente o artifício que usaremos nesta seção para fatorar $A = QR$. Em particular, usaremos a triangularização de Householder, que recebe este nome uma vez que as matrizes ortogonais empregadas, $Q_i : i = 1,\dots,n$, são *refletores de Householder*. Para a finalidade de fatorar $A = QR$, a alternativa de triangularizar por meio de matrizes ortogonais deve nos parecer mais atrativa do que a ortogonalização por matrizes triangulares, uma vez que as transformações lineares produzidas por matrizes ortogonais são mais estáveis. Essa é a motivação para construirmos uma nova classe de algoritmos para fatoração $A = QR$.

Cabe destacar que os refletores de Householder adquirem enorme importância em Computação Científica, não apenas porque dão origem a esta fatoração QR estável. Também são um ingrediente fundamental no cálculo de autovalores e de autovetores

$$\begin{bmatrix} x & x & x \\ x & x & x \\ x & x & x \\ x & x & x \\ x & x & x \end{bmatrix} \xrightarrow{Q_1} \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & x & x \\ 0 & x & x \\ 0 & x & x \end{bmatrix} \xrightarrow{Q_2} \begin{bmatrix} x & x & x \\ & x & x \\ & 0 & x \\ & 0 & x \\ & 0 & x \end{bmatrix} \xrightarrow{Q_3} \begin{bmatrix} x & x & x \\ & x & x \\ & & x \\ & & 0 \\ & & 0 \end{bmatrix}$$

Figura 5.7: Tranformações lineares desejadas para construirmos um algoritmo que faz uma triangularização ortogonal. Nesta figura, as entradas das matrizes que preservam sua cor a cada tranformação, preservam seus valores a cada transformação.

de matrizes diagonalizáveis, no cálculo de fatorações SVD e também de Schur.

Antes de discutirmos tais refletores, vamos detalhar um pouco mais as transformações que levam à fatoração. Para as observações que fazemos na sequência, considere a Figura 5.7, que ilustra as transformações que desejamos. Destacamos os seguintes aspectos:

1. O procedimento calcula matrizes ortogonais $Q_k : k = 1, \ldots, n$ capazes de garantir $Q_n \ldots Q_2 Q_1 A = R$, onde R é triangular superior. Recorde-se que o produto de matrizes ortogonais é uma matriz ortogonal, portanto $A = Q_1^T \ldots Q_n^T R$ é uma fatoração QR de A (produto de matriz ortogonal por triangular superior). Isto é, $Q_1^T \ldots Q_n^T = Q$ é uma matriz ortogonal.

2. Veja o efeito desejado de Q_1, Q_2, $Q_3, \ldots$ na matriz A. O efeito da matriz Q_k sobre a coluna k de $Q_{k-1} Q_{k-2} \ldots Q_1 A$ é de zerar as últimas $m - k$ entradas daquela coluna. A matriz Q_k não tem efeito sobre as primeiras $k - 1$ linhas de $Q_{k-1} Q_{k-2} \ldots Q_1 A$, isto é, elas são preservadas. Na figura, as cores preservadas indicam entradas das matrizes que são inalteradas com as transformações indicadas.

3. Para que as transformações lineares sejam ortogonais e, ao mesmo tempo preservem as primeiras $k-1$ linhas de $Q_{k-1} \ldots Q_1 A$, a norma Euclideana do vetor $m - (k - 1)$ dimensional, correspondente às últimas $m - (k - 1)$ entradas na coluna k de $Q_{k-1} \ldots Q_1 A$, deve ser o mesmo antes e após a transformação linear. Como todas, exceto a entrada na linha k daquele vetor, serão nulas após a transformação, toda a norma Euclideana daquele vetor *deve ser transferida para a posição na diagonal* daquela coluna. Atenção aqui: Este é o aspecto que inspira o uso do refletor de Householder.

5.5.1 Refletores de Householder

Vamos apresentar o conceito dos refletores de Householder de uma maneira genérica e, na sequência, vamos indicar como podem ser especializados para produzir as transformações lineares indicadas na Figura 5.7.

Definição 5.6 *Dado um vetor $v \in \mathbb{R}^n$ e $u = \frac{v}{\|v\|_2}$ sua normalização, a matriz*

$$F = I - \frac{2}{v^T v} v v^T = I - 2uu^T \tag{5.11}$$

é chamada de refletor de Householder.

Veja que o refletor de Householder consiste em uma atualização de posto 1 na matriz identidade. Examine a Figura 5.8 e identique as seguintes entidades relevantes na mesma, pertinentes para a interpretação do refletor:

- o ponto a (a ser refletido);
- o ponto $r = Fa$ (a reflexão de a, gerada pela aplicação de F em a). Note que $\|r\|_2 = \|a\|_2$, pois ambos situam-se sob o círculo de raio $\|a\|_2$;
- o ponto p, ponto médio do segmento de reta que liga a até r,
- o vetor $v = a - r$. Note que este foi o vetor empregado para se construir o refletor de Householder F. A normalização de v é o vetor u, que recebe o nome de *vetor de Householder;*
- os subespaços $span\{v\}$ e $span\{v\}^\perp$. Note que $span\{v\}^\perp$ é o plano de simetria da reflexão, de forma que o ponto refletido e sua reflexão distam a mesma quantidade do plano. Note também que este ponto de distância mínima entre a e $span\{v\}^\perp$ e entre r e $span\{v\}^\perp$ é o ponto p;
- os dois triângulos retângulos que têm a origem e o ponto p como vértices comuns e que diferem pelo outro vértice, a saber, o ponto a e sua reflexão r.

Agora que já identificamos os elementos essenciais na figura, vamos analisar como a reflexão de a foi construída. Tome o ponto a e considere a transformação linear $(I - \frac{vv^T}{v^T v})a = p$. Veja que o projetor $P = (I - \frac{vv^T}{v^T v})$ projeta em $span\{v\}^\perp$ e que $\frac{vv^T}{v^T v}$, como sabemos, projeta em $span\{v\}$. O segmento $a - p$ corresponde à diferença ou erro entre o projetado a e sua projeção em $span\{v\}^\perp$. Portanto, esta diferença está em $(span\{v\}^\perp)^\perp = span\{v\}$.

Veja que $a = p + \frac{1}{2}v \rightarrow p = a - \frac{1}{2}(a - r) = \frac{1}{2}(a + r)$. Ou seja, se subtrairmos de a a quantidade $\frac{1}{2}v$ obtemos a projeção de a em torno do plano simétrico à reflexão. Continuando o processo, se subtrairmos $\frac{1}{2}v$ de p, obtemos a reflexão r de a.

Outro aspecto importante que merece ser destacado é que se tormarmos $\hat{v} = r - a = -v$, os refletores $I - 2\frac{\hat{v}\hat{v}^T}{\hat{v}^T\hat{v}}$ e $I - 2\frac{vv^T}{v^T v}$ são idênticos. Por fim, veja que o refletor preserva a norma Euclideana, pois r,a distam a mesma quantidade da origem.

Nesse momento em que a geometria da transformação foi bastante discutida, já podemos enunciar os dois principais resultados dos refletores de Householder.

Resultado 5.6.1 *O refletor de Householder F é uma matriz simétrica, ortogonal e não singular.*

Prova 5.6.1

- *Simetria ($F = F^T$): F é a diferença entre duas matrizes simétricas, a identidade e $2uu^T$.*
- *Ortogonal ($FF^T = F^TF = I$)*

$$\begin{aligned}(I - 2uu^T)^T(I - 2uu^T) &= I - 4uu^T + (2uu^T)^T(2uu^T) \\ &= I - 4uu^T + 4u(u^Tu)u^T \\ &= I\end{aligned}$$

- *É não singular, admite inversa, pois $F^TF = I$ e então $F^{-1} = F^T$.*

Resultado 5.6.2 *Dados dois vetores $a, r \in \mathbb{R}^n$ tais que $\|a\|_2 = \|r\|_2$ e $v = a - r$, a matriz $F = I - 2\frac{vv^T}{v^Tv}$ aplicada em a satisfaz:*

$$Fa = r.$$

Prova 5.6.2

$$\begin{aligned}Fa &= \left(I - 2\frac{(a-r)(a-r)^T}{(a-r)^T(a-r)}\right)a \\ &= a - 2\frac{(a-r)(a^Ta - r^Ta)}{(a-r)^T(a-r)} \\ &= a - 2\frac{(a-r)(a^Ta - r^Ta)}{a^Ta - 2r^Ta + r^Tr} \\ &= a - 2\frac{(a-r)(a^Ta - r^Ta)}{2a^Ta - 2r^Ta} \\ &= a - 2\frac{(a-r)(a^Ta - r^Ta)}{2(a^Ta - r^Ta)} \\ &= r\end{aligned}$$

Na sequência, apresentamos alguns exemplos que ilustram como se construir refletores de Householder de forma a garantir algum resultado particular para as transformaçõse lineares desejadas. Iniciamos com a mais simples delas, na qual o refletor é aplicado em um vetor, tendo como restrição a reflexão desejada.

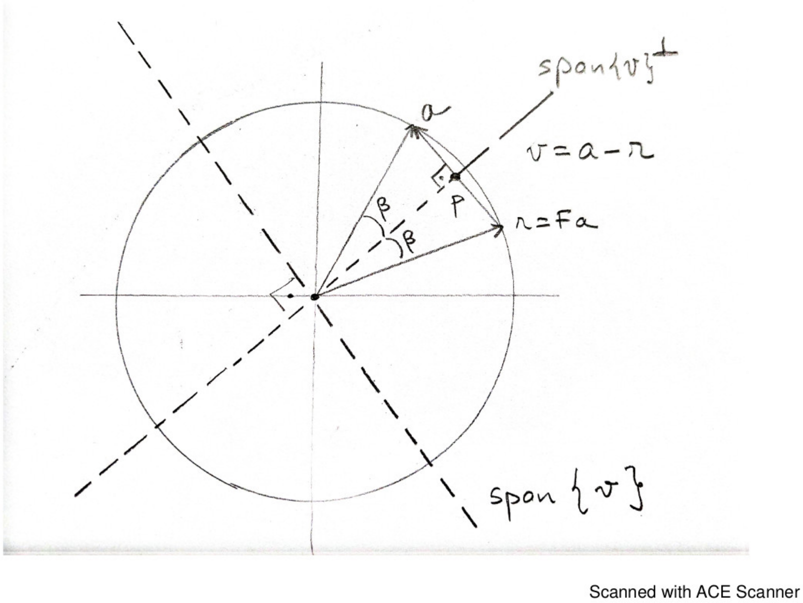

Figura 5.8: Ilustração da geometria do Refletor de Householder no plano.

Exemplo 43 *Construir o refletor de Householder que reflita o vetor* $a = (1, 0, 1, 2)^T$ *no sentido positivo da linha* $e_4 = (0, 0, 0, 1)^T$. *Na sequência, calculamos o vetor* r, *a reflexão desejada para* a. *Note que* a *e* r *precisam ter a mesma norma Euclideana. Então, calculamos o vetor de Householder* u, *o refletor* $F = I_4 - 2uu^T$ *e verificamos que* $Fa = u$, *como desejado.*

```
-->a = [1;0;1;2]
-->r = [0;0;0;1]*norm(a,2)
 r  =
   0.
   0.
   0.
   2.4494897
-->v = a - r
 v  =
   1.
   0.
   1.
  -0.4494897
-->u = v/norm(v,2)
 u  =
   0.6738873
   0.
   0.6738873
```

```
  -0.3029054
-->F = eye(4,4)-2*u*u'
 F  =
   0.0917517   0.  -0.9082483   0.4082483
   0.          1.   0.          0.
  -0.9082483   0.   0.0917517   0.4082483
   0.4082483   0.   0.4082483   0.8164966
-->F*a
 ans  =
  -7.772D-16
   0.
  -7.772D-16
   2.4494897
```

No próximo exemplo, nosso objetivo é construir um refletor de Householder que reflita especificamente um subvetor de uma coluna de uma matriz, em um ponto desejado.

Exemplo 44 *Construir um refletor de Householder que reflita o vetor* $B(2:5,3)$ *(vetor da 3 coluna de B, da segunda até a quinta linha de B) de uma matriz* $B \in \mathbb{R}^{7\times 4}$ *de entradas aleatórias no vetor* $(0,0,0,1)\|B(2:5,3)\|_2$*. O refletor deve manter inalteradas as demais linhas de B, isto é, as linhas* $1,6,7$ *de B não devem ser afetadas pela transformação linear* FB*. Na resolução abaixo apresentada, utilizamos uma matriz* $F \in \mathbb{R}^{7\times 7}$ *ortogonal cujo bloco* $F(2:5,2:5) \in \mathbb{R}^{4\times 4}$ *é o refletor de Householder propriamente dito. As demais colunas e linhas de F são necessárias para manter inalteradas as linhas* $1,6,7$ *de B. Veja os cálculos:*

```
->B = rand(7,4)
 B  =
   0.2113249   0.685731    0.5442573   0.9329616
   0.7560439   0.8782165   0.2320748   0.2146008
   0.0002211   0.068374    0.2312237   0.312642
   0.3303271   0.5608486   0.2164633   0.3616361
   0.6653811   0.6623569   0.8833888   0.2922267
   0.6283918   0.7263507   0.6525135   0.5664249
   0.8497452   0.1985144   0.3076091   0.4826472
-->a = B(2:5,3)
 a  =
   0.2320748
   0.2312237
```

```
   0.2164633
   0.8833888
-->r = [0;0;0;1]*norm(a,2)
 r  =
   0.
   0.
   0.
   0.966724
-->F = eye(7,7)
 F  =
   1.   0.   0.   0.   0.   0.   0.
   0.   1.   0.   0.   0.   0.   0.
   0.   0.   1.   0.   0.   0.   0.
   0.   0.   0.   1.   0.   0.   0.
   0.   0.   0.   0.   1.   0.   0.
   0.   0.   0.   0.   0.   1.   0.
   0.   0.   0.   0.   0.   0.   1.
-->v = a-r
 v  =
   0.2320748
   0.2312237
   0.2164633
  -0.0833352
-->u = v/norm(v,2)
 u  =
   0.5781594
   0.5760391
   0.5392669
  -0.2076098
-->F(2:5,2:5) = eye(4,4)-2*u*u'
 F  =
   1.   0.           0.           0.           0.           0.   0.
   0.   0.3314635  -0.6660848  -0.6235645   0.2400631   0.   0.
   0.  -0.6660848   0.3363579  -0.6212777   0.2391828   0.   0.
   0.  -0.6235645  -0.6212777   0.4183823   0.2239142   0.   0.
   0.   0.2400631   0.2391828   0.2239142   0.9137963   0.   0.
   0.   0.           0.           0.           0.           1.   0.
   0.   0.           0.           0.           0.           0.   1.
```

```
-->F*B
 ans  =
   0.2113249   0.685731    0.5442573   0.9329616
   0.2042069   0.054836   -2.776D-16  -0.2924643
  -0.5495921  -0.7519869  -2.220D-16  -0.1925636
  -0.1843881  -0.2071436  -1.943D-16  -0.1113191
   0.8635389   0.9580226   0.966724    0.4743074
   0.6283918   0.7263507   0.6525135   0.5664249
   0.8497452   0.1985144   0.3076091   0.4826472
-->norm(a,2)
 ans  =
   0.9667240
```

Neste terceiro exemplo a seguir, nosso objetivo é ilustrar a construção de um refletor de Householder para produzir um resultado particular em uma linha da matriz alvo da transformação linear. Ou seja, ao invés de aplicarmos FA, aplicamos AF, ou equvivalentemente, FA^T, uma vez que F é uma matriz simétrica.

Exemplo 45 *Construir um refletor de Householder que reflita o vetor $B(2, 2:4)$ de uma matriz $B \in \mathbb{R}^{4\times5}$ em com entradas aleatoriamente escolhidas no interalo $(0, 1)$ em $e_2\|B(2, 2:4)\|$. As demais colunas de B devem ser preservadas.*
Veja que neste caso, queremos aplicar uma transformação ortogonal em uma linha de B. Portanto, a matriz ortogonal que iremos construir deve ser aplicada à direita de B.

```
-->B = rand(4,5)
 B  =
   0.3321719   0.2693125   0.0437334   0.2806498   0.1121355
   0.5935095   0.6325745   0.4818509   0.1280058   0.6856896
   0.5015342   0.4051954   0.2639556   0.7783129   0.1531217
   0.4368588   0.9184708   0.4148104   0.211903    0.6970851
-->a = B(2,2:4)'
 a  =
   0.6325745
   0.4818509
   0.1280058
-->F = eye(5,5)
 F  =
   1.   0.   0.   0.   0.
   0.   1.   0.   0.   0.
```

```
   0.   0.   1.   0.   0.
   0.   0.   0.   1.   0.
   0.   0.   0.   0.   1.
-->r = [0;1;0]*norm(a,2)
 r  =
   0.
   0.8054292
   0.
-->v = a - r
 v  =
   0.6325745
  -0.3235783
   0.1280058
-->u = v/norm(v,2)
 u  =
   0.8761798
  -0.4481888
   0.1773011

-->F(2:4,2:4) = eye(3,3)-2*u*u'
 F  =
   1.   0.           0.           0.           0.
   0.  -0.5353822    0.785388    -0.3106953    0.
   0.   0.785388     0.5982535    0.1589287    0.
   0.  -0.3106953    0.1589287    0.9371287    0.
   0.   0.           0.           0.           1.
-->B*F
 ans  =
   0.3321719  -0.197034    0.2822818   0.1862814   0.1121355
   0.5935095   1.180D-16   0.8054292   0.          0.6856896
   0.5015342  -0.251445    0.5998443   0.6454371   0.1531217
   0.4368588  -0.2317831   1.0031952  -0.0208588   0.6970851
```

5.6.1 Fatoração $A = QR$ via refletores de Householder

Nesse momento, já dispomos de todos os elementos para mostrar como a triangularização de A poderá ser feita, por meio de matrizes ortogonais, construídas com o auxílio de refletores de Householder. Em linhas gerais, o algoritmo implementa uma transformação ortogonal a cada iteração, uma iteração por coluna de

A, até que a matriz transformada tenha a forma de uma triangular superior, isto é, toda entrada r_{ij} da matriz transformada final R é nula para índices satisfazendo $i > j : j = 1, \ldots, n, i = j + 1, \ldots, n$. Recorde-se da Figura 5.7, que ilustra os resultados desejados com as operações ortogonais.

A invariante do algoritmo é a seguinte:

- No início da iteração k, as colunas de índice menor que k estão prontas, pois já tem a forma de uma triangular superior.

Na iteração k dispomos da matriz transformada pelas $k-1$ iterações anteriores. Esta matriz é $Q_{k-1} \ldots Q_1 A$. O objetivo na iteração k consiste em transformar a coluna k de $Q_{k-1} \ldots Q_1 A$ em uma coluna de uma triangular superior. Para tanto, deve-se manter inalteradas as colunas e linhas de $Q_{k-1} \ldots Q_1 A$, de índices iguais ou inferiores a $k-1$. Para se garantir esta propriedade, escolhemos a matriz Q_k ortogonal, cujo particionamento em blocos é:

$$Q_k = \left[\begin{array}{c|c} I & 0 \\ \hline 0 & F \end{array}\right],$$

onde I é a matriz identidade de ordem $k-1$ e F é uma matriz ortogonal de ordem $m-(k-1)$. O refletor F é o elemento responsável por produzir zeros nas posições corretas da coluna k enquanto as demais entradas de Q_k preservam a estrutura desejada, criada nas iterações anteriores. Veja que, como $F^T F = I_{m-(k-1)}$, $Q_k^T Q_k = I_m$ e Q_k é, de fato, uma matriz ortogonal.

Nossa tarefa agora é usar o que sabemos sobre os refletores de Householder para construir a matriz F necessária. Para sistematizar o procedimento, assumimos que no início da iteração k, as entradas nas linhas $k, k+1, \ldots, m$ da coluna k de $Q_{k-1} \ldots Q_1 A$ sejam respresentadas pelo vetor $x \in \mathbb{R}^{m-(k-1)}$. Após a aplicação de F em x, o resultado Fx deve ser um vetor que tenha a norma de x na entrada de posição 1 do vetor. Ou seja, o refletor F aplicado em x deve produzir a transformação linear: $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{m-(k-1)} \end{bmatrix}$, $Fx = \begin{bmatrix} \|x\| \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \|x\| e_1$, onde e_1 é um vetor $m-(k-1)$ dimensional de zeros, exceto pela primeira posição que é 1.

Observe que, pela forma como definimos, F reflete o espaço $\mathbb{R}^{m-(k-1)}$ em torno do hiperplano H_1, indicado na Figura 5.9, de forma que o vetor $v = \|x\| e_1 - x$ é perpendicular ao hiperplano H. Observe o subespaço H_2 indicado na figura, e note também que dispomos de mais de uma escolha para o ponto de reflexão e, consequentemente, para F. Ao invés de refletir o vetor x em $+\|x\|e_1$, podemos refletir em $-\|x\|e_1$. Naturalmente, o refletor F obtido muda dependendo do ponto de reflexão. Em resumo, temos duas opções para o vetor v que será usado para a construção do refletor F:

- Opção 1: $v = x - \|x\|e_1$. Na figura, este vetor v é perpendicular ao subespaço H_1, sendo este o subespaço em torno do qual se dá a reflexão.
- Opção 2: $v = x + \|x\|e_1$. Na figura, este vetor v é perpendicular a H_2. De forma análoga, caso este vetor v dê origem ao refletor, refletido e reflexão são simétricos a H_2.

Para aumentar a estabilidade numérica, devemos escolher o ponto de reflexão que promove o máximo deslocamento, isto é, o ponto em que a distância entre x e sua reflexão é máxima. Assim sendo, se a entrada x_1 for negativa, devemos refletir x em $\|x\|e_1$ e, se for positiva, devemos refletir em $-\|x\|e_1$.

Na implementação do algoritmo (que será discutida mais à frente), vamos calcular a reflexão da seguinte forma. Dado o vetor x a ser refletido, definimos

$$sinal(x_1) = \begin{cases} 1 & \text{se } x_1 \geq 0 \\ -1 & \text{caso contrário} \end{cases} . \tag{5.12}$$

Visando evitar cancelamentos numéricos indesejados, responsáveis por ampliar erros numéricos, adotamos o ponto de reflexão como

$$r = -sinal(x_1)\|x\|_2 e_1 \tag{5.13}$$

e o vetor de Householder (antes da normalização) como

$$v = x + sinal(x_1)\|x\|_2 e_1. \tag{5.14}$$

Vamos analisar a expressão (5.14). Veja que quando $x_1 \geq 0$, devemos adotar a opção 1 acima indicada. Nesse caso, o ponto de reflexão mais distante de x seria $-\|x\|_2 e_1$ e o correspondente $v = x - (-\|x\|_2 e_1)$, conforme dado por (5.14). Se por outro lado $x_1 < 0$, a opção 2 é a indicada. O ponto de reflexão mais distante seria $\|x\|_2 e_1$ e o vetor $v = x - \|x\|_2 e_1 = x + sinal(x_1)\|x\|_2 e_1$, comprovando a corretude da expressão (5.14).

Exemplo 46 *Neste exemplo, nosso objetivo é triangularizar a matriz A dada. Apenas para simplificar este primeiro exemplo, vamos assumir que o ponto de reflexão sempre será $+\|x\|e_1$.*

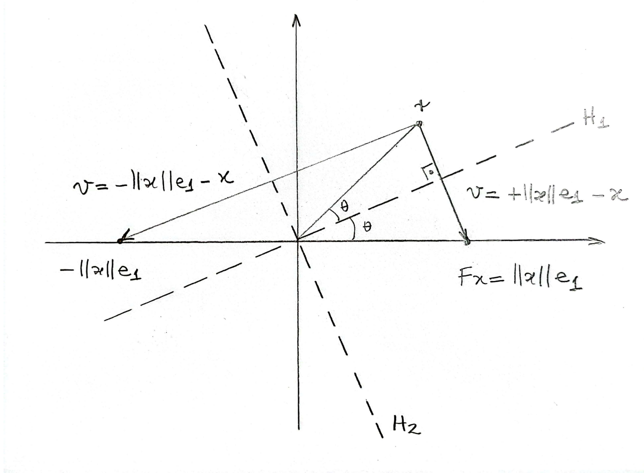

Figura 5.9: Alternativas para o vetor e para o refletor de Householder, que devem ser exploradas visando incrementar a estabilidade numérica do processo de triangularização.

$$A = \begin{bmatrix} -4 & 1 & 1 \\ 2 & 1 & -1 \\ 4 & 1 & 1 \end{bmatrix}.$$

Primeira iteração:

- $x = \begin{bmatrix} -4 & 2 & 4 \end{bmatrix}^T$, $\|x\| = 6$.
- $v^T = \begin{bmatrix} 6 & 0 & 0 \end{bmatrix}^T - \begin{bmatrix} -4 & 2 & 4 \end{bmatrix}^T = \begin{bmatrix} 10 & -2 & -4 \end{bmatrix}^T$, $v^T v = 120$.
- $F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{2}{120} \begin{bmatrix} 10 \\ -2 \\ -4 \end{bmatrix} \begin{bmatrix} 10 & -2 & -4 \end{bmatrix}$

$$F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{2}{120} \begin{bmatrix} 100 & -20 & -40 \\ -20 & 4 & 8 \\ -40 & 8 & 16 \end{bmatrix}$$

$$F = \begin{bmatrix} -2/3 & 1/3 & 2/3 \\ 1/3 & 14/15 & -2/15 \\ 2/3 & -2/15 & 11/15 \end{bmatrix}$$

- $FA = \begin{bmatrix} -2/3 & 1/3 & 2/3 \\ 1/3 & 14/15 & -2/15 \\ 2/3 & -2/15 & 11/15 \end{bmatrix} \begin{bmatrix} -4 & 1 & 1 \\ 2 & 1 & -1 \\ 4 & 1 & 1 \end{bmatrix} =$

$$\begin{bmatrix} 6 & 1/3 & -1/3 \\ 0 & 17/15 & -11/15 \\ 0 & 19/15 & 23/15 \end{bmatrix}$$

- *Para a primeira iteração, temos* $Q_1 = F$.

Segunda iteração:

- $Q_1 A = \begin{bmatrix} 6 & 1/3 & -1/3 \\ 0 & 17/15 & -11/15 \\ 0 & 19/15 & 23/15 \end{bmatrix}$

- Q_2 *tem a forma* $\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & F_2 \end{array}\right]$, *onde* $F_2 \in \mathbb{R}^{2\times 2}$ *é o segundo refletor.*

- F_2 *será contruído a partir de* $x = \begin{bmatrix} 17/15 & 19/15 \end{bmatrix}^T$, $\|x\| = \frac{\sqrt{26}}{3}$.

- $v = \frac{\sqrt{26}}{3}\begin{bmatrix} 1 & 0 \end{bmatrix}^T - \begin{bmatrix} 17/15 & 19/15 \end{bmatrix}^T = \begin{bmatrix} \frac{5\sqrt{26}-17}{15} & -\frac{19}{15} \end{bmatrix}^T$, $v^T v \approx 1.9251853$.

- $F_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 2/1.9251853 \begin{bmatrix} \frac{5\sqrt{26}-17}{15} \\ -\frac{19}{15} \end{bmatrix} \begin{bmatrix} \frac{5\sqrt{26}-17}{15} & -\frac{19}{15} \end{bmatrix}$

  $F_2 = \begin{bmatrix} 0.6667949 & 0.7452413 \\ 0.7452413 & -0.6667949 \end{bmatrix}$

- $Q_2 Q_1 A = R = \begin{bmatrix} 6 & 1/3 & -1/3 \\ 0 & 1.6996732 & 0.6537205 \\ 0 & 0 & -1.5689291 \end{bmatrix}$

- $Q_2 Q_1 = \begin{bmatrix} -0.6666667 & 0.3333333 & 0.6666667 \\ 0.7190925 & 0.5229764 & 0.4576043 \\ -0.1961161 & 0.7844645 & -0.5883484 \end{bmatrix}$

- *Fatoração QR resultante:* $A = (Q_2 Q_1)^T R$

- *Veja que não dispomos explicitamente do fator* $(Q_2 Q_1)^T$.

Observe também que como $m \leq n$, $n-1$ *matrizes ortogonais* Q_k *foram suficientes para a triangularização. Caso* $m > n$, *precisaríamos de uma transformação adicional.*

O algoritmo que produz a fatoração $A = QR$ utilizando os refletores de Householder é apresentado na Figura 5.10. A matriz ortogonal utilizada a cada iteração é denominada Qa. Veja que o produto das matrizes ortogonais Qa é armazenado na matriz Q, a cada iteração do algoritmo (laço em k). Ao final, as matrizes Q^T e R são retornadas como os fatores desejados. A implementação apresentada aqui visa ser mais didática e do que computacionalmente eficiente, pois é possível reduzir em uma ordem (em m) a sua complexidade computacional.

O algoritmo na Figura 5.10 não implementa pivoteamento de colunas, dado que o vetor a ser refletido, definindo na intrução

```
x = R(k:m,k)
```

não é escolhido com base em sua norma, dentre as colunas restantes de R. Na sequência, apresentaremos uma implementação do mesmo algoritmo na qual, durante a iteração k, percorremos as colunas restantes da matriz, de índices $j = k, \ldots, n$, e verificamos qual vetor

```
x = R(k:m,j)
```

possui a maior norma Euclideana. Caso o índice p da coluna $R(k : m, p)$ de maior norma Euclideana seja distinto de k, o conteúdo da coluna k e p são trocados. O algoritmo apresentado na Figura 5.11 introduz o pivoteamento de colunas na fatoração QR via refletores de Householder. As funções auxiliares para o algoritmo da Figura 5.11 são apresentadas na Figura 5.12.

Exemplo 47 *Vamos empregar o algoritmo de fatoração de Householder para fator a matriz A com deficiência de posto do exemplo 41.*

```
A  =
   1.   2.    3.    4.
   5.   6.    7.    8.
   9.   10.   11.   12.
   1.   1.    1.    1.
   3.   2.    1.    0.
-->[Q,R,pivot,P] = APQR_Householder(A)
Posto detectado: 2
 Q  =
  -0.2666667   0.4777778
  -0.5333333   0.1222222
  -0.8        -0.2333333
  -0.0666667  -0.0888889
```

```
function [Q,R] = QR_Householder(A)
    [m,n] = size(A)
    R = A
    Q = eye(m,m)
    for k = 1:n
        x = R(k:m,k)

        vk = sign(x(1))*norm(x,2) * eye(m-k+1,1) + x
        vk = 1.0 / norm(vk,2) * vk

        R(k:m,k:n)   = R(k:m,k:n)   - 2.0 * vk*(vk'*R(k:m,k:n))

        Qa = eye(m,m)
        Qa(k:m,k:m) = eye(m-k+1,m-k+1) - 2.0 * vk*(vk'*Qa(k:m,k:m))
        Q = Qa*Q

    end
    Q = Q'
endfunction
```

Figura 5.10: Algoritmo para Fatoração QR via refletores de Householder, sem pivoteamento de colunas.

```
   0.          -0.8333333
 R  =
  -15.         -10.2  -11.8  -13.4
  -1.776D-15  -3.6   -2.4   -1.2
 pivot  =  4. 1. 2. 3.
 P  =
   0.   1.   0.   0.
   0.   0.   1.   0.
   0.   0.   0.   1.
   1.   0.   0.   0.
-->norm(A*P-Q*R,1)
 ans  =
   1.725D-14
```

```
function [Q,R,pivot,P,posto] = APQR_Householder(A)
    ninfA = norm(A,'inf')
    eps = 1.0E-14
    [m,n] = size(A)
    posto = min(m,n)
    R = A
    Q = eye(m,m)
    pivot = zeros(n,1)
    for j = 1:n
        pivot(j) = j;
    end
    nitermax = min(m,n)
    for k = 1:n
        [p,maxn] = DeterminaMelhorRefletor(R,m,n,k)
        if (p <> k) // troca conteudo de R
           [R,pivot] = TrocaConteudoColunas(R,pivot,k,p)
        end
        if (maxn < (eps * ninfA))
            posto = k - 1
            printf('Posto detectado: %d \n',posto)
            break
        end
        x = R(k:m,k)

        vk = sign(x(1))*norm(x,2) * eye(m-k+1,1) + x
        vk = 1.0 / norm(vk,2) * vk

        R(k:m,k:n)   = R(k:m,k:n)   - 2.0 * vk*(vk'*R(k:m,k:n))

        Qa = eye(m,m)
        Qa(k:m,k:m) = eye(m-k+1,m-k+1) - 2.0 * vk*(vk'*Qa(k:m,k:m))
        Q = Qa*Q
    end
    Q = Q'
    P = zeros(n,n)
    for j = 1:n
        P(pivot(j),j) = 1
    end
    Q = Q(:,1:posto)
    R = R(1:posto,:)
endfunction
```

Figura 5.11: Algoritmo para Fatoração QR via refletores de Householder, incorporando pivoteamento de colunas e detecção de posto incompleto.

```
function [p,maxn] = DeterminaMelhorRefletor(R,m,n,k)
    p = k
    maxn = norm(R(k:m,k),2)
    for j = k+1:n
        normaj = norm(R(k:m,j),2)
        if normaj > maxn
           p = j
           maxn = normaj
        end
    end
endfunction
function [R,pivot] = TrocaConteudoColunas(R,pivot,k,p)
    t = R(1:m,k)
    R(1:m,k) = R(1:m,p)
    R(1:m,p) = t
    t1 = pivot(k)
    pivot(k) = pivot(p)
    pivot(p) = t1
endfunction
```

Figura 5.12: Funções auxiliares para a fatoração QR via refletores de Householder com pivoteamento de colunas e detecção de posto incompleto.

Exercícios Propostos

As questões de 1 a 9 foram adaptadas de [2].

Questão 01: Aplicando o algoritmo de Gram-Schmidt, encontre a fatoração completa q_1, q_2, q_3 da matriz $A = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{bmatrix}$ tal que q_1, q_2 sejam uma base para $C(A)$.

Questão 02: Aplicando o algoritmo de Gram-Schmidt, encontre uma base ortonormal para o espaço coluna da matriz $A = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}$. Calcule a projeção do vetor $b = \begin{bmatrix} -4 \\ -3 \\ 3 \\ 0 \end{bmatrix}$ neste subespaço.

Questão 03: Aplicando o algoritmo de Gram-Schmidt, encontre uma base ortonormal para o espaço coluna da matriz $A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix}$. Escreva $A = QR$.

Questão 04: Se Q tem colunas ortonormais, qual é a solução $\hat{x}$ para o ajuste linear $Qx = b$?

Questão 05: Calcule a matriz de projeção $P = QQ^T$ quando $q_1 = \begin{bmatrix} 0.8 \\ 0.6 \\ 0 \end{bmatrix}$ e $q_2 = \begin{bmatrix} -0.6 \\ 0.8 \\ 0 \end{bmatrix}$.

Questão 06: Se A é uma matriz $m \times n$, com $r(A) = n$ e após a sua fatoração em QR é produzida uma matriz $Q = [Q_1\, Q_2]$ quadrada de ordem m e uma matriz $R = [R\, 0]^T$ $m \times n$, com 0 uma matriz nula, responda:

1. As n colunas de Q_1 formam uma base ortonormal para qual subespaço fundamental?

2. As $m-n$ colunas de Q_2 formam uma base ortonormal para qual subespaço fundamental?

3. Como as colunas de Q_2 devem ser obtidas.

Questão 07: A matrix $P = QQ^T$ é a matriz de projeção no espaço coluna de $Q^{m\times n}$. Agora adicione uma nova coluna a, fazendo $A = [Q\, a]$. A coluna a é substituída por qual nova coluna q, após a aplicação do algoritmo de Gram-Schmidt?

Questão 08:

1. Encontre os vetores ortonormais q_1, q_2, q_3 tais que q_1 e q_2 gerem o espaço coluna de $A = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{bmatrix}$.

2. Qual dos 4 espaços fundamentais contém q_3?

3. "Resolva"$Ax = [1\,2\,7]^T$ usando mínimos quadrados (ou ajuste).

Questão 09: Qual o múltiplo α de $a = [4\,5\,2\,2]^T$ tal que αa é o vetor mais próximo de $b = [1\,2\,0\,0]$? Encontre os vetores ortornormais q_1 e q_2 no plano gerado por a e b.

Questão 10: Considere que $A \in \mathbb{R}^{m\times n}$ possui posto completo igual a n. Considere a fatoração $A = QR$ (reduzida) onde $Q \in \mathbb{R}^{m\times n}$ é ortonormal e $R \in \mathbb{R}^{n\times n}$ é triangular superior, com a diagonal positiva. Mostrar que

1. A fatoração é única

2. A matriz R é o fator triangular superior da fatoração de Cholesky de A^TA.

Questão 11: A matriz $A \in \mathbb{R}^{m\times n}$ de posto completo n foi fatorada $A = QR$ (reduzida). Deseja-se "resolver"o sistema linear $Ax = b$, isto é, encontrar o ponto $p \in C(A)$ que minimiza $\|p-b\|_2$. Conhecendo-se o vetor $\hat{x}$ que combina as colunas de A e obtém o ponto p, seria possível determinar algum vetor $\hat{y}$ que combina as colunas de Q e leva ao mesmo ponto p ? Em caso positivo, justifique sua resposta e apresente o vetor y. Em caso negativo, indique a razão pela qual não se pode obter tal vetor y.

Capítulo 6

Fatorações Espectral, de Schur e SVD

Neste capítulo, discutimos as fatorações espectral $A = Q\Lambda Q^T$ de matrizes reais simétricas e a fatoração SVD de uma matriz $A \in \mathbb{R}^{m\times n}$, isto é $A = U\Sigma V^T$. O objetivo do capítulo ainda não é apresentar algoritmos que computam estas fatorações; os algoritmos serão discutidos mais à frente em nosso curso. Os objetivos aqui são mostrar a existência das fatorações, assim como apresentar as informações que revelam.

Iniciamos nosso estudo caracterizando o conjunto das matrizes reais diagonalizáveis e mostramos que as simétricas não apenas são similares a matrizes diagonais, mas são ortogonalmente similares a matrizes diagonais. Durante o processo, mostramos que toda matriz quadrada, embora nem sempre seja diagonalizável, sempre é similar a uma matriz triangular superior. Esta observação que dá origem à Fatoração de Schur, também de relevante importância em Computação Científica.

Na sequência, apresentamos a fatoração SVD e como ela generaliza a fatoração espectral. Concluímos o capítulo discutindo algumas aplicações importantes desta fatoração, sobretudo na redução de dimensionalidade de matrizes de dados, na aproximação de matrizes por outras, de posto baixo, e também apresentando uma pseudo-inversa para matrizes singulares e para matrizes retangulares.

6.1 Introdução

O primeiro tema a ser estudado neste capítulo é a caracterização das matrizes que admitem uma fatoração espectral. Para tanto, recordamos alguns conceitos que utilizamos ao longo do curso. Essencialmente nos dedicaremos à fatorar matrizes reais. Porém, precisaremos eventualmente trabalhar em aritmética complexa pois algumas matrizes reais não admitem autovalores reais.

Um *autopar* de uma matriz $A \in \mathbb{R}^{n\times n}$, é uma tupla (λ, x) formada por um escalar $\lambda \in \mathbb{C}$ (possivelmente complexo), denominado *autovalor* de A, e por um vetor $x \in \mathbb{C}^n$, também possivelmente complexo, denominado *autovetor* de A, que satisfaz a relação

$$Ax = \lambda x.$$

O conjunto de todos os autovalores de uma matriz é denominado *espectro* da matriz.

Os autovalores de A são as raízes do *polinômio característico* de A, obtido quando reescrevemos $Ax = \lambda x$ como $(A - \lambda I)x = 0$ e impomos que para que exista solução não trivial (distinta de zero) para este sistema linear homogêneo, a matriz $(A - \lambda I)$ precisa ter determinante igual a zero. Quando impomos que $\det(A - \lambda I) = 0$, obtemos o polinômio $p_A(\lambda)$, denominado polinômio característico de A. A existência de autovalores complexos para uma matriz A real (ou de $p_A(\lambda)$ admitir raízes complexas) é relacionada ao fato de que alguns sistemas dinâmicos, cuja dinâmica é representada pela matriz em questão, podem ter regime transiente que oscila, decae ou cresce.

Objetivamente, o autovalor faz com que o espaço nulo da matriz $(A - \lambda I)$, isto é $\Phi = N(A - \lambda I)$, seja distinto do vetor nulo. Este subespaço Φ é denominado *autoespaço* associado a λ. Claramente, um autovetor x (ou qualquer múltiplo dele) satisfaz $x \in \Phi$. Associado a um autovalor λ de A há duas grandezas inteiras relevantes:

- μ_λ, ou simplesmente μ, chamado de *multiplicidade algébrica* de λ. A multiplicidade algébrica representa o número de vezes que λ é raiz do polinômio característico de A. Pelo Teoreoma Fundamental da Álgebra, um polinômio com coeficientes reais (como é o caso do polinômio característico de uma matriz real n dimensional) posui n raízes, entre reais e complexas, contando sua multiplicidade. As raízes complexas, caso existam, aparecem aos pares conjugados.

  Por exemplo, suponha que uma matriz A possua 3 autovalores distintos, $\lambda_1, \lambda_2, \lambda_3$ e que seu polinômio característico tenha sido fatorado da seguinte forma:

  $$p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)^3(\lambda - \lambda_3)^2.$$

  As mulitiplicadades algébricas de $\lambda_1, \lambda_2, \lambda_3$ são, respectivamente, $\mu_1 = 1, \mu_2 = 3, \mu_3 = 2$.

- β_λ, ou siplesmente β, chamado de *multiplicidade geométrica* de λ. Ele é a dimensão do autoespaço Φ associado a λ e portanto, indica o número de vetores em qualquer base para Φ.

Um autovalor λ de A é denominado *defectivo* se sua multiplicidade algébrica excede sua multiplicidade geométrica. Uma matriz é defectiva se possui algum autovalor defectivo.

Exemplo 48 *Considere a matriz* $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ *e* $B = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$. *Tanto A quanto B possuem* $p(\lambda) = (2-\lambda)^3$ *como polinômio característico. Desta forma, em ambos os casos, a multiplicidade algébrica de* $\lambda = 2$ *é* $\mu = 3$. *Entretanto, para a matriz A,* $\Phi = span\{e_1, e_2, e_3) = \mathbb{R}^3$ *e, assim sua multiplicidade geométrica é* $\beta = 3$. *A matriz A não é defectiva e, por esta razão, o conjunto de seus n autovetores linearmente independentes geram o* $\mathbb{R}^n$. *Por outro lado, associado ao autovalor* $\lambda = 3$ *de B, há apenas o autovetor* e_1 *(ou múltiplos dele). A multiplicidade geométrica de B é* $\beta = 1$. *Como* $\beta = 1 < \mu = 3$, *B é defectiva, pois possui um autovalor defectivo. Não é possível gerar o* $\mathbb{R}^3$ *apenas com os autovetores de B. Por esta razão, dizemos que há falta de autovetores.*

Dizer que uma matriz é diagonalizável significa dizer que é similar a uma matriz diagonal. Veremos que as matrizes diagonalizáveis, isto é, que podem ser escritas como $A = X\Lambda X^{-1}$, são as matrizes não defectivas. Observe a força de escrevemos a similaridade de A com uma matriz diagonal. Se é diagonaliazável, possui n autovetores linearmente independentes.

Nunca é demais enfatizar o ponto seguinte. Quando estabelecemos a relação de similaridade $A = X\Lambda X^{-1}$ para uma Λ diagonal, estamos estabelecendo que:

- os n autovalores de A são as n entradas na diagonal de Λ. A e Λ são similares e como tal possuem o mesmo espectro.
- Como $A = X\Lambda X^{-1} \to AX = X\Lambda$, temos que as n colunas linearmente independentes de X fornecem os n autovetores li de A. X admite inversa !

Desta forma, nem toda matriz quadrada é diagonalizável. Porém, qualquer matriz $A \in \mathbb{R}^{n\times n}$, inclusive as defectivas, admite uma *fatoração de Schur*

$$A = QTQ^*$$

em que Q é unitária ($Q^*Q = I$) e T é triangular superior. Claramente A é similar a T, de forma que possuem o mesmo conjunto de autovalores, situados na diagonal de T. Porém, diferentemente do caso em que a matriz considerda é similar a uma diagonal, (seja ortogonalmente similar, $A = Q\Lambda Q^T$, ou simplesmente similar, $A = X\Lambda X^{-1}$) não podemos dizer que $AQ = QT$ forneça os autovetores de A nas colunas de Q. T é triangular e não diagonal. A pode não ter n autovetores linearmente independentes, nesse caso não há como ser similar a uma diagonal.

Provaremos a existência da fatoração de Schur, em uma etapa intermediária da próxima seção, dedicada à caracterização das matrizes diagonalizáveis. Durante toda a seção, empregamos B^* para designar a matriz transposta conjugada de B, sempre que for necessário operar nos complexos.

6.2 A fatoração de Schur de matrizes quadradas e a diagonalização de matrizes simétricas

Ao longo do nosso curso, mencionamos que toda matriz real simétrica é ortogonalmente similar a uma matriz diagonal. Isto é, qualquer matriz A simétrica pode ser escrita como $A = Q\Lambda Q^T$ onde os autovetores de A são as colunas de Q, seus autovalores são as entradas reais na diagonal de Λ e Q é ortogonal, $Q^TQ = I$. Dada sua importância teórica e prática, sobretudo para a demonstração dos resultados particulares à fatoração SVD, demonstramos este resultado, denominado Teorema Espectral.

A demonstração é fracionada na apresentação de três resultados que, encadeados, levam ao resultado que desejamos mostrar. Os primeiros resultados não requerem que a matriz A seja simétrica e tratam da independência linear de autovetores associados a autovalores que são distintos. Ao longo do processo, demonstraremos um resultado de vital importância, a existência da Fatoração de Schur, pertinente para qualquer matriz quadrada. Como de costume, ao longo deste capítulo $N(A), C(A), N(A^T), C(A^T)$ representam os quatro espaços fundamentais associados a uma matriz A.

Resultado 6.2.1 *Sejam $\lambda_i : i = 1, \ldots, k$ os autovalores distintos de $A \in \mathbb{R}^{n\times n}$ (não necessariamente simétrica). Seja $\Phi_i = N(\lambda_i I_n - A)$ o autoespaço associado a λ_i e $u^{(i)}$ qualquer vetor não nulo tal que $u^{(i)} \in \Phi_i : i \in 1, \ldots, k$. Então, os $u^{(i)}$ são linearmente independentes.*

Prova 6.2.1 *O resultado é demonstrado em duas partes. Na primeira parte demonstramos que $u^{(i)} \notin \Phi_j$ para $j \neq i$. Na sequência, usamos este resultado para mostrar que dois autovetores $u^{(j)}$,$u^{(i)}$ de autoespaços distintos não podem ser linearmente dependentes.*

1. *Primeira parte. Suponha que $u^{(i)} \in \Phi_j$ para $j \neq i$. Sendo verdade temos:*

$$
\begin{aligned}
Au^{(i)} &= \lambda_j u^{(i)} \\
Au^{(i)} &= \lambda_i u^{(i)} \\
0 &= (\lambda_i - \lambda_j)u^{(i)} \\
&\Rightarrow \\
\lambda_j &= \lambda_i
\end{aligned}
$$

Temos assim uma contradição, pois por hipótese os autovalores λ_i e λ_j são distintos, para $i \neq j$. Logo, $i \neq j$ implica em $u^{(i)} \notin \Phi_j$.

2. *Segunda parte. Vamos supor, por absurdo, que exista algum $u^{(i)}$, por exemplo $u^{(1)}$, que possa ser escrito como combinação linear dos demais. Então temos $u^{(1)} = \sum_{i=2}^{k} \alpha_i u^{(i)}$, que implica nas seguintes identidades:*

$$\begin{aligned} \lambda_1 u^{(1)} &= \sum_{i=2}^{k} \alpha_i \lambda_1 u^{(i)} \\ &= Au^{(1)} \\ &= \sum_{i=2}^{k} \alpha_i A u^{(i)} \\ &= \sum_{i=2}^{k} \alpha_i \lambda_i u^{(i)} \end{aligned}$$

Subtraindo a primeira e a última equação, obtemos:

$$\sum_{i=2}^{k} \alpha_i (\lambda_i - \lambda_1) u^{(i)} = 0$$

Como para todo $i \geq 2$ temos $\lambda_1 - \lambda_i \neq 0$ por hipótese, a identidade acima implica que $u^{(2)}, \ldots, u^{(k)}$ são linearmente dependentes. Assim sendo, algum destes vetores, digamos $u^{(2)}$ pode ser escrito como combinação linear dos demais $u^{(3)}, \ldots, u^{(k)}$. Repetindo o mesmo racicínio, chegaríamos ao ponto de mostrar que $u^{(3)}, \ldots, u^{(k)}$ são linearmente dependentes. Repetindo o processo, chegaríamos a conclusão, no último passo, que $u^{(k-1)}$ e $u^{(k)}$ são linearmente dependentes, o que implicaria em $u^{(k-1)} \in \Phi_k$, o que é uma contradição ao resultado que mostramos na primeira parte.

O próximo resultado também não requer que a matriz A seja simétrica e permitirá que caracterizemos as matrizes que são diagonalizáveis.

Resultado 6.2.2 *Toda matriz real $A \in \mathbb{R}^{n \times n}$ (simétrica ou não) é similar a uma matriz bloco triangular, onde um dos blocos é $\lambda_i I_{\beta_i}$ onde λ_i é um dos autovalores distintos de A, β_i é a dimensão do autoespaço $\Phi_i = N(\lambda_i I_n - A)$.*

Prova 6.2.2 *Sem perda de generalidade, podemos assumir que $U^{(i)}$ é uma matriz $n \times \beta_i$ cujas colunas formam uma base ortonormal para Φ_i (por exemplo, usamos o procedimento de Gram-Schmidt para ortogonalizar qualquer base para Φ_i). Observe que os autovalores λ_i podem ser complexos, assim como as colunas em $U^{(i)}$, elementos da base para Φ_i. Vamos considerar agora uma matriz $Q^{(i)}, n \times (n - \beta_i)$, cujas* *colunas formam uma base ortonormal para* $C(U^{(i)})^{\perp}$. *Veja que qualquer coluna de* $U^{(i)}$ *é ortogonal a qualquer coluna de* $Q^{(i)}$ *e assim sendo, a matriz* $P^{(i)}, n \times n$ *definida como* $P^{(i)} = [U^{(i)}, Q^{(i)}]$ *é unitária, de forma que* $(P^{(i)})^* P^{(i)} = I_n$ *e* $(P^{(i)})^{-1} = (P^{(i)})^*$. *Então temos:*

$$\begin{aligned} AU^{(i)} &= \lambda_i U^{(i)} \\ (U^{(i)})^* AU^{(i)} &= \lambda_i (U^{(i)})^* U^{(i)} \\ &= \lambda_i I_{\beta_i} \end{aligned}$$

e também

$$\begin{aligned} (Q^{(i)})^* AU^{(i)} &= \lambda_i (Q^{(i)})^* U^{(i)} \\ &= 0_{n-\beta_i,\beta_i}. \end{aligned}$$

Portanto, podemos escrever que

$$(P^{(i)})^{-1} AP^{(i)} = (P^{(i)})^* AP^{(i)} = \begin{bmatrix} \lambda_i I_{\beta_i} & (U^{(i)})^* AQ^{(i)} \\ 0_{n-\beta_i,\beta_i} & (Q^{(i)})^* AQ^{(i)} \end{bmatrix}, \tag{6.1}$$

provando o resultado.

Neste momento, à partir de (6.1), podemos estabelecer um corolário e demonstrar que toda matriz admite uma fatoração de Schur.

Resultado 6.2.3 Fatoração de Schur
Toda matriz $A \in \mathbb{R}^{n\times n}$ *(simétrica ou não) é unitariamente similar a uma matriz triangular superior, isto é,* $A = QTQ^*$ *onde* $Q^*Q = I$ *e* T *é uma matriz triangular superior.*

Prova 6.2.3 *Veja que podemos aplicar o mesmo raciocínio empregado na demosntração do resultado 6.2.2, para o bloco* $n - \beta_i \times n - \beta_i$ *correspondente à matriz* $(Q^{(i)})^* AQ^{(i)}$, *que é uma matriz com entradas reais. Há um autovetor* λ_j *para o bloco, que também é autovalor para* A. *Associado a* λ_j, *há uma base para* $\hat{\Phi}_j \subset \mathbb{C}^{n-\beta_i}$, *o autoespaço associado ao autovalor* λ_j *de* $(Q^{(i)})^* AQ^{(i)}$. *Identificamos uma base para* $\hat{\Phi}_j^{\perp}$, *construímos uma matriz unitária de ordem* $n - \beta_i$ *com a base para* $\hat{\Phi}_j$ *e* $\hat{\Phi}_j^{\perp}$, *e repetimos a análise acima, recursivamente. Se a matriz resultante no bloco distinto de* $\lambda_j I_{\beta_j}$ *for triangular superior* $(\beta_i + \beta_j = n)$, *concluímos a prova. Caso contrário, repetimos o processo, com o novo bloco que tem dimensão pelo menos uma unidade menor que* $n - \beta_i$.

Observação 6.2.1 *Cabe chamar atenção para um elemento na demonstração acima. Observe que* λ_j *é autovalor tanto para* A *quanto para o bloco* $(Q^{(i)})^* A Q^{(i)}$. *Porém os autoespaços de* A *e de* $(Q^{(i)})^* A Q^{(i)}$ *associados a* λ_j *são diferentes,* $\Phi_j \neq \hat{\Phi}_j$*; veja que a própria dimensão destes autoespaços é diferente. Quando a matriz* A *for real simétrica, não teremos dificuldade em construir um autovetor para* A *a partir de um autovetor para* $(Q^{(i)})^* A Q^{(i)}$*, associado ao autovalor* λ_j *comum entre estas duas matrizes. Para o caso em que* A *não é simétrica, isso nem sempre é possível. Outro ponto importante é que a dimensão de* Φ_j *não é* β_j *necessariamente.* β_j *é a dimensão de* $\hat{\Phi}_j$*. Na verdade,* β_j *será a multiplicidade algébrica de* λ_j *em* $p_A(\lambda)$*.*

Além da existência da Fatoração de Schur, o Resultado 6.2.2 acima nos permite também estabelecer que a dimensão do autoespaço associado a λ_i, β_i, é limitada superiormente pela multiplicidade algébrica μ_i de λ_i. Sabemos que matrizes similares possuem os mesmos autovalores, contando suas multiplicidades. Sabemos também que os autovalores de uma matriz bloco triangular são a união dos autovalores de cada bloco. Desta forma, o bloco $(Q^{(i)})^* A Q^{(i)}$ em (6.1) não pode conter λ_i como autovalor. Portanto, $\beta_i \leq \mu_i$.

O próximo resultado caracteriza o conjunto das matrizes diagonalizáveis.

Resultado 6.2.4 *Seja* $\{\lambda_i : i = 1, \ldots, k\}$ *o conjunto de autovalores distintos de uma matriz* $A \in \mathbb{R}^{n \times n}$ *(não necessariamente simétrica),* $\mu_i : i = \{1, \ldots, k\}$ *suas multiplicidades geométricas,* $\Phi_i = N(\lambda_i I_n - A)$*. Além disso, considere que* $U^{(i)} = [u_1(i), \ldots, u_{\beta_i}(i)]$ *uma matriz contendo uma base ortonormal para* Φ_i*, onde* β_i *é a dimensão do autoespaço* Φ_i*. Então:*

1. $\beta_i \leq \mu_i : i = 1, \ldots, k$

2. *e, se* $\beta_i = \mu_i : i = 1, \ldots, k$ *então a matriz*

$$U = [U^{(1)}, \ldots, U^{(k)}]$$

admite inversa e

$$A = U \Lambda U^{-1},$$

onde

$$\Lambda = \begin{bmatrix} \lambda_1 I_{\mu_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{\mu_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_k I_{\mu_k} \end{bmatrix}$$

Prova 6.2.4 *1. A discussão precedente à apresentação do resultado demonstrou que* $\beta_i \leq \mu_i$ *para qualquer autovalor* λ_i *de* A*.*

2. *Assumimos então que* $\beta_i = \mu_i$ *para todo* $i = 1, \ldots, k$. *Então, os vetores* $u_1^{(i)}, \ldots, u_{\mu_i}^{(i)}$ *são linearmente independentes pois formam uma base para* Φ_i. *Pelo Resultado* (6.2.1), *sabemos que autovetores associados a autovalores distintos são linearmente independentes. Desta forma, o conjunto de autovetores* $\{u_j^i : i = 1, \ldots, k, j = 1, \ldots, \beta_i\}$ *são linearmente independentes. Como* $\mu_i = \beta_i$ *para todo* i, *temos que* $\sum_{i=1}^k \beta_i = \sum_{i=1}^k \mu_i = n$. *Portanto, a matriz* U *possui posto completo,* n, *admitindo inversa.*
   Como temos $AU^{(i)} = \lambda_i U^{(i)}, i = 1, \ldots, k$, *podemos escrever*

$$AU = U\Lambda$$
$$A = U\Lambda U^{-1}$$

O Resultado (6.2.4) mostrou que as matrizes não defectivas, isto é, que não possuem autovalores defectivos ($\beta_i = \mu_i$ para todo autovalor λ_i) são similares a matrizes diagonais e possuem n autovetores linearmente independentes, de forma que os autovetores formam uma base para o $\mathbb{C}^n$ (recorde-se que os autoespaços, até o momento, podem não ser espaços complexos). Na verdade, a classe das matrizes similares à matrizes diagonais é exatamente a classe das matrizes não defectivas. Esta é, portanto, uma condição necessária e suficiente para que a matriz seja diagonalizável.

Agora, vamos particularizar nosso estudo para as matrizes reais simétricas. Mostraremos que tais matrizes possuem autovalores e autovetores reais, que não são defectivas e que não apenas admitem uma diagonalização $A = X\Lambda X^{-1}$, mas que admitem uma diagonalização ortogonal $A = Q\Lambda Q^T$, onde $Q^TQ = I$ se verifica.

Resultado 6.2.5 Decomposição espectral de uma matriz real simétrica
Seja uma matriz $A \in \mathbb{R}^{n\times n}$ *real simétrica e* $\lambda_i : i = 1, \ldots, k$ *seus* k *autovalores distintos. Seja* μ_i *a multiplicidade algébrica do autovalor* λ_i, *isto é, o número de vezes que* λ_i *é raiz do polinômio característico de* A. *Além disso, denote por* $\Phi_i = N(\lambda_i I_n - A)$ *o auto-espaço associado ao autovalor* λ_i *e* β_i *a dimensão de* Φ_i. *Então, para todo* $i = 1, \ldots, k$ *valem os resultados abaixo:*

1. $\lambda_i \in \mathbb{R}$ *e os autovetores associados a* λ_i *sempre podem ser reais.*

2. $\Phi_i \perp \Phi_j$ *para* $i \neq j$. *Isto é, os autoespaços associados a autovalores distintos são subespaços vetoriais ortogonais.*

3. *A dimensão de* Φ_i *é* μ_i, *isto é,* $\beta_i = \mu_i$.

Prova 6.2.5 Demonstração do Teorema de Decomposição Espectral para matrizes reais simétricas

1. *Tome um autopar* (λ, u) *de* A*, de forma que temos* $Au = \lambda u$*. Então, tomando o conjugado complexo temos*

$$(Au)^* = (\lambda u)^* \rightarrow u^* A^* = \lambda^* u^*.$$

Pré-multiplicando $Au = \lambda u$ *por* u^* *e pós-multiplicando* $u^* A^* = \lambda^* u^*$ *por* u*, temos:*

$$u^* A u = \lambda^* u^* u$$
$$u^* A^* u = \lambda^* u^* u$$

Subtraindo a segunda da primeira equação temos

$$u^*(A - A^*)u = (\lambda - \lambda^*)u^* u.$$

Como $u^* u = \|u\|_2^2 \neq 0$ *e lembrando que* $A = A^T$ *e que* $A^T = A^*$ *pois* A *é real e simétrica, concluímos que* $\lambda - \lambda^* = 0$*. Logo, os autovalores são reais.*
Uma consequência deste resultado é que o autovetor u *associado a* λ *sempre pode ser real. Veja que se* u *é complexo e satisfaz* $Au = \lambda u$ *com* λ *real, temos que a parte real* $Re(Au)$ *de* Au *e a parte real* $Re(\lambda u)$ *de* λu *devem ser iguais. Isto é,* $Re(Au) = A(Re(u)) = Re(\lambda u) = \lambda(Re(u))$*. Portanto,* $Re(u)$ *é um autovetor real de* A *associado a* λ*.*

2. *Vamos tomar* $v_i \in \Phi_i$*,* $v_j \in \Phi_j$ *para* $i \neq j$*, isto é, os autovalores associados* $\mu_i \neq \mu_j$*. Então temos:*

$$Av_i = \lambda_i v_i$$
$$Av_j = \lambda_j v_j$$
$$v_j^T A v_i = \lambda_i v_j^T v_i$$

Além disso

$$\begin{aligned} v_j^T A v_i &= v_i^T A^T v_j \\ &= v_i^T A v_j \\ &= \lambda_j v_i^T v_j \\ &= \lambda_j v_j^T v_j, \end{aligned}$$

o que nos permite estabelecer, ao subtrair as duas equações finais acima, que:

$$(\lambda_i - \lambda_j)v_j^T v_i = 0.$$

Como $\lambda_i \neq \lambda_j$, $v_i^T v_j = 0$ *e os dois autovalores, de autoespaços distintos são ortogonais.*

3. *Seja* λ *um autovalor qualquer de* A *e* μ, β *suas multiplicidades algébrica e geométrica, respectivamente, e* Φ *o autoespaço associado a* λ*. Sabemos que* $\beta \leq \mu$ *e agora mostramos que para matrizes simétrias sempre temos* $\mu = \beta$*. Para isso, vamos construir uma base ortonormal para* Φ *composta por* μ *vetores. O primeiro passo neste sentido é usar uma construção similar àquela que nos permitiu escrever a equação (6.1), ao final da demonstração do Resultado 6.2.2. Assim, especializamos aquele argumento e, na sequência, concluímos a demonstração da parte 3 deste resultado.*

   ** *(resultado intermediário) Para qualquer matriz* B *quadrada de ordem* m *e simétrica, com autovalor* λ*, existe uma matriz ortogonal* $U = [u, Q] \in \mathbb{R}^{m \times m}$, $Q \in \mathbb{R}^{m \times m-1}$*, tal que* $Bu = \lambda u, \|u\|_2 = 1$*,* Q *possui colunas ortonormais, que formam uma base para o espaço* $m-1$ *dimensional span*$\{u\}^\perp$*. Como a matriz* B *agora é simétrica, já demonstramos que seus autavalores são reais e que também possuem autovetores reais, podemos substituir a operação de transposição conjugada por simples transposição. Veja então que* $Q^T u = 0$*, e portanto:*

$$U^T B U = \begin{bmatrix} \lambda & 0_{1,m-1} \\ 0_{m-1,1} & Q^T B Q \end{bmatrix},$$

   onde o bloco $Q^T BQ$ *é uma matriz simnétrica de ordem* $m-1$*. Veja então que* B *simétrica é similar a uma matriz bloco diagonal (e não apenas bloco triangular).*

Agora aplicamos o resultado acima para a matriz A*, visando concluir a prova. Para o autovalor* λ *de* A*, temos* $\mu \geq 1$*. Como* A *é simétrica* $n \times n$*, existe uma matriz* $U_1 \in \mathbb{R}^{n \times n}$*,* $U_1 = [u_1, Q_1]$ *onde* $Au_1 = \lambda u_1, \|u_1\|_2 = 1$ *e* $Q_1 \in \mathbb{R}^{n \times n-1}$ *possui colunas ortonormais formando uma base para span*$\{u_1\}^\perp$*, de forma que*

$$U_1^T A U_1 = \begin{bmatrix} \lambda & 0_{1,n-1} \\ 0_{n-1,1} & A_1 \end{bmatrix},$$

onde $A_1 = Q_1^T A Q_1$ *é simétrica. Se* $\mu = 1$*, concluímos a prova, uma vez que* *encontramos uma base* (u_1) *para* ϕ *que contém* $\mu = 1$ *vetores linearmente independentes (ortonormais). Se, ao invés disso,* $\mu > 1$*, em função da estrutura bloco diagonal de* $U_1^T A U_1$*,* λ *é autovalor de* A_1 *com multiplicidade* $\mu - 1$*. Neste caso, aplicamos o mesmo raciocínio à matriz* A_1*: existe* U_2 *ortogonal, de ordem* $n-1$*, tal que* $U_2 = [\tilde{u}_2, Q_2] \in \mathbb{R}n - 1 \times n - 1$ *onde* $A_1\tilde{u}_2 = \lambda\tilde{u}_2, \|\tilde{u}\|_2 = 1$ *e*

$$U_2^T A_1 U_2 = \begin{bmatrix} \lambda & 0_{1,n-1} \\ 0_{n-1,1} & A_2 \end{bmatrix},$$

onde $A_2 = Q_2^T A_1 Q_2$ *é simétrica, de ordem* $n-2$*.*

Veja que o vetor $u_2 = U_1 \begin{bmatrix} 0 \\ \tilde{u}_2 \end{bmatrix}$ *é um autovetor de norma Euclideana unitária de A:*

$$\begin{aligned} A &= U_1 \begin{bmatrix} \lambda & 0 \\ 0 & A_1 \end{bmatrix} U_1^T \\ Au_2 &= U_1 \begin{bmatrix} \lambda & 0 \\ 0 & A_1 \end{bmatrix} U_1^T U_1 \begin{bmatrix} 0 \\ \tilde{u}_2 \end{bmatrix} \\ &= U_1 \begin{bmatrix} 0 \\ A\tilde{u}_2 \end{bmatrix} \\ &= \lambda u_2 \end{aligned}$$

Além disso, a norma de u_2 *é unitária*

$$\begin{aligned} \|u_2\|_2 &= u_2^T u_2 \\ &= \begin{bmatrix} 0 \\ \tilde{u}_2 \end{bmatrix}^T U_1^T U_1 \begin{bmatrix} 0 \\ \tilde{u}_2 \end{bmatrix} \\ &= \tilde{u}_2^T \tilde{u}_2 \\ &= 1 \end{aligned}$$

e u_2 *é ortogonal a* u_1*:*

$$\begin{aligned} u_1^T u_2 &= u_1^T [u_1, Q_1] \begin{bmatrix} 0 \\ \tilde{u}_2 \end{bmatrix} \\ &= (u_1^T Q_1)\tilde{u}_2 \\ &= 0^T \tilde{u}_2 \\ &= 0 \end{aligned}$$

De forma análoga, se $\mu = 2$*, concluímos a prova, pois* u_1, u_2 *formam uma base* *ortonormal de dimensão $\mu = 2$ para Φ. Caso contrŕio, se $\mu > 2$, repetimos o processo na matriz A_2, e encontramos um autovetor u_3 ortogonal a u_1, u_2. O processo se repete até que disponhamos de μ autovetores ortogonais.*

Observação 6.2.2 *Veja que no caso de uma matriz simétrica, não foi difícil construir um autovetor u_2 para A a partir do autovetor $\tilde{u}_2$ para o bloco A_1. Veja que a simetria da matriz $U_1^T A U_1$, que possui uma linha (e uma coluna) de zeros à direita (abaixo) de λ foi fundamental nesta construção. No caso da fatoração de Schur que apresentamos no resultado 6.2.3 isso é bastante mais difícil e nem sempre é possível ser feito. Ou seja, a matriz não simétrica pode de fato ter deficiência de autovetores.*

Finalmente, podemos enunciar o principal resultado desta seção, que é um corolário dos Resultados 6.2.4 e 6.2.5.

Resultado 6.2.6 Teorema Espectral
Seja $A \in \mathbb{R}^{n\times n}$ simétrica e sejam $\lambda_i \in \mathbb{R} : i = 1, \ldots, n$ seus autovalores, contando as multiplicidades. Então, existe um conjunto de n autovetores $u_i \in \mathbb{R}^n$ ortonormais. Equivalentemente, existe U, $n \times n$, ortogonal tal que

$$A = U\Lambda U^T = \sum_{i=1}^{n} \lambda_i u_i u_i^T$$

onde Λ é uma matriz diagonal de ordem n, onde sua diagonal é formada pelos elementos $\lambda_1, \lambda_2, \ldots, \lambda_n$, nesta ordem.

6.3 Fatoração SVD reduzida e completa

Assim como no caso da fatoração QR, vamos apresentar a fatoração SVD reduzida e a completa de A. De início, assumimos que a matriz $A \in \mathbb{R}^{m\times n}$ a ser fatorada satisfaz $posto(A) = r$, podendo ser completo ou não.

A fatoração SVD de A pode ser escrita como

$$A = U\Sigma V^T \qquad \text{ou equivalentemente} \qquad (6.2)$$

$$AV = U\Sigma \qquad (6.3)$$

$$Av_i = \sigma_i u_i \qquad i = 1, \ldots, r \qquad (6.4)$$

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T \qquad (6.5)$$

onde:

- $U \in \mathbb{R}^{m\times r}$ é uma matriz com r colunas u_i linearmente independentes e ortonormais, isto é, $U^TU = I_r$. As colunas de U são denominadas *vetores singulares à esquerda* de A.

- $\Sigma \in \mathbb{R}^{r\times r}$ é uma matriz diagonal, contendo em sua diagonal os valores singulares de A, de forma que, por convenção, $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_r > 0$,

- $V \in \mathbb{R}^{n\times r}$ é uma matriz com n colunas v_i linearmente independentes, ortonormais: $V^TV = I_r$. As colunas de V são denominados de *vetores singulares à direita* de A.

A forma $A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$ mostra que A é a soma de r matrizes de posto-1, sendo o $i-$ésimo termo representado por $\sigma_i u_i v_i^T$. Diferentemente das outras fatorações que estudamos até o momento, que também permitiam escrever a matriz fatorada como uma soma de matrizes de posto-1, a fatoração SVD revela uma *hierarquia* para cada um destes elementos na soma. Isso ocorre pois os valores singulares, ao satisfazerem $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, conferem uma importância maior aos primeiros termos na soma (termos com índices i menores). Essencialmente, o valor singular σ_i reflete a importância do termo $\sigma_i u_i v_i^T$ na representação de A. Assim sendo, são de relevante importância para se produzir aproximações da matriz fatorada. A Figura 6.2 ilustra a forma dos termos na fatoração.

Veja que a fatoração SVD revela o posto r da matriz A. Na fatoração reduzida, as r colunas de U fornecem uma base já ortogonalizada para a $C(A)$ e as colunas de V uma base (também já ortogonalizada) para $C(A^T)$. Como as colunas de U e de V têm norma Euclideana unitária, a matriz Σ, diagonal com os valores singulares, é responsável por fazer a mudança de escala necessária para escrever as colunas de A na transformação, uma vez que estas colunas via de regra não serão unitárias.

A interpretação geométrica da fatoração reduzida é a seguinte. Vamos considerar a forma $Av_i = \sigma_i u_i$. Veja que, para algum $i \in \{1, \ldots, r\}$, o vetor $v_i \in \mathbb{R}^n$ possui $\|v_i\|_2 = 1$ e a imagem $Av_i = \sigma_i u_i$ satisfaz $\|\sigma_i u_i\|_2 = \sigma_i$. Assim sendo, ao aplicarmos A em v_i obtemos um vetor em $C(A)$ na direção de u_i. A menos que $\sigma_i = 1$, esta imagem não possui norma unitária. A Figura 6.1 ilustra o caso particular em que $r = posto(A) = 2$ e $A \in \mathbb{R}^{2\times 2}$.

No caso da transformação ilustrada na Figura 6.1, a matriz possui posto completo, igual ao número de colunas da matriz, 2. Assim, a não ser que o vetor v seja nulo, a imagem Av é distinta do vetor zero. Agora, considerando todos os vetores $v \in \mathbb{R}^2$ de norma Euclideana untária, há dois vetores (e seus simétricos) relevantes, identificados na parte à esquerda da figura. São os dois vetores singulares. Quando a matriz A é aplicada em v_1, o vetor singular à direita associado ao maior valor

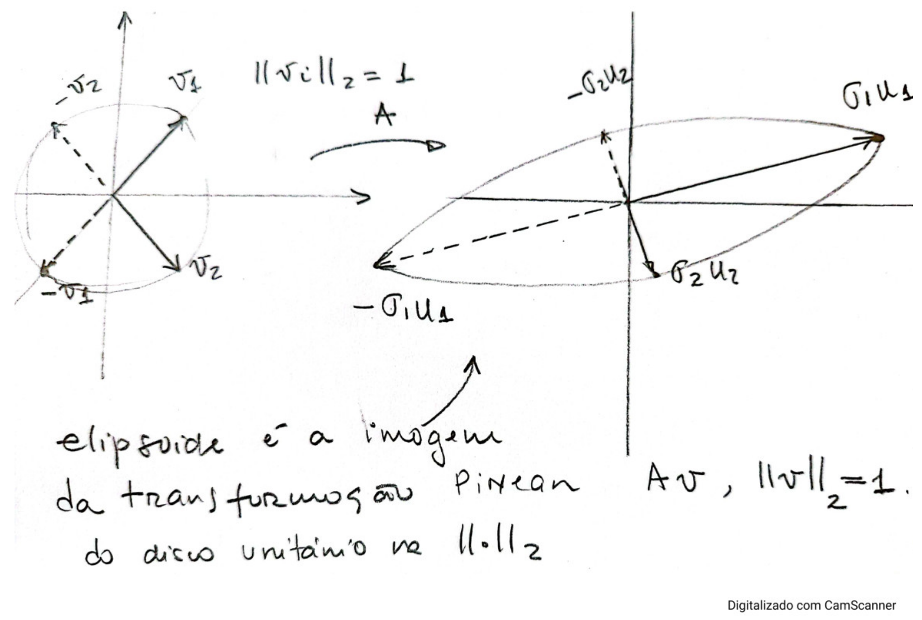

Figura 6.1: Ilustração no plano da geometria da tranformação linear Av para $\|v\|_2 = 1$, à luz da fatoração SVD.

singular σ_1, obtemos a imagem de maior norma Euclideana, $\sigma_1 u_1$. Quando A é aplicada em v_2, associado a σ_2, obtemos a imagem de menor norma Euclideana. No caso ilustrado na figura, $\sigma_1 \geq 1$ e $\sigma_2 < 1$, uma vez que Av_1 produziu um vetor de norma superior à norma do vetor de entrada, e que Av_2 produziu um vetor com norma inferior à norma do vetor de entrada. O elipsóide à direita na figura é a imagem de A no disco unitário. Como o posto de A é a dimensão de seu espaço coluna, e no caso tratado na figura, o posto é completo, este elipsóide possui exatamente $posto(A) = 2$ eixos principais, definidos pelos vetores $u_i : i = 1, 2$. Caso houvesse deficiência de posto, $N(A) \neq \{0\}$, teríamos $Av_i = 0$ para algum vetor $v_i \neq 0$, e um dos eixos principais deste elipsóide seria degenerado.

Vamos agora discutir a forma da fatoração SVD completa. Assim como na forma reduzida, as primeiras r colunas da matriz U na fatoração completa contém uma base ortonormal para $C(A)$. Porém, a matriz U passa a ser quadrada de ordem m, de forma que suas últimas $m - r$ colunas são preenchidas com uma base ortonormal para $N(A^T)$. De forma similar, as últimas $n - r$ colunas de V recebem uma base ortonormal para $N(A)$, preservando em suas primeiras r colunas a base ortonormal para $C(A^T)$. Com a expansão das matrizes U e V, precisamos garantir que a matriz Σ tenha dimensão $m \times n$. Assim sendo, para que a matriz A de fato seja o produto de $U \Sigma V^T$, as últimas $m - r$ linhas e as últimas $n - r$ colunas da matriz Σ na fatoração completa precisam ser linhas e colunas de zeros. A Figura 6.3 ilustra a forma dos termos na fatoração completa considerando-se que $posto(A) = r < \min\{n, m\}$. Veja

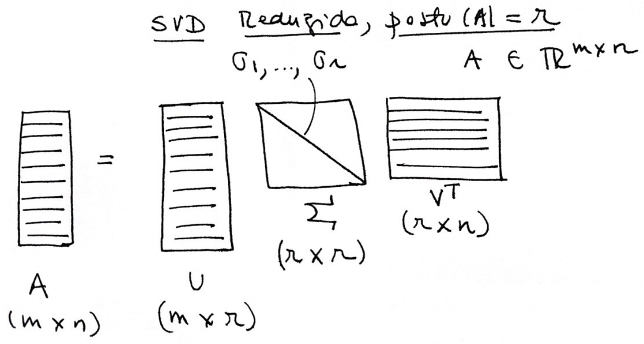

Figura 6.2: Formato da fatoração SVD reduzida de uma matriz A, com $posto(A) = r$.

que completamos as linhas e colunas de U e V com zeros, assim como na matriz Σ.

6.4 A fatoração SVD como uma generalização da fatoração espectral

Sabemos que as matrizes normais (matrizes para as quais se observa que $AA^T = A^TA$) são unitariamente diagonalizáveis, isto é, admitem uma fatoração espectral

$$A = Q\Lambda Q^T, \tag{6.6}$$

onde Λ é uma matriz diagonal com os autovalores de A, $QQ^T = Q^TQ = I$ onde Q é uma matriz quadrada com os autovetores de A. Os autovetores de A formam uma base para $\mathbb{R}^n$ e podem ser ortogonalizados. Nesse caso, A e Λ são não apenas matrizes similares (possuindo portanto os mesmos autovalores), mas são similares por meio de uma Q ortogonal.

Veja que no caso de uma matriz não diagonalizável, seja por ser quadrada e defectiva, ou por ser retangular, não é possível produzir a fatoração espectral. No caso de uma matriz não quadrada, $C(A)$ e $C(A^T)$ são subespaços de espaços vetoriais de dimensões distintas. A fatoração espectral assim sequer faz sentido. Em qualquer um destes dois casos em que a fatoração espectral não pode ser produzida, podemos obter a fatoração SVD de A.

A fatoração SVD sempre pode ser produzida e, de certa forma, generaliza a fatoração espectral. A fatoração SVD de A relaciona-se com a fatoração espectral de A^TA e de AA^T. Veremos que os vetores singulares $v_i : i = 1, \ldots, n$ de A são

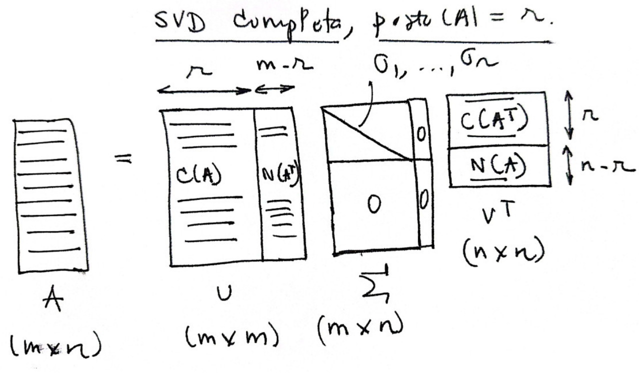

Figura 6.3: Formato da fatoração SVD completa de uma matriz A, com $posto(A) = r < \min\{n, m\}$.

os autovetores de A^TA. Por sua vez, os vetores singulares $u_i : i = 1, \ldots, m$ são os autovetores de AA^T. Vamos mostrar estes dois resultados, isto é, como estes vetores singulares podem ser obtidos.

Resultado 6.4.1 *Os vetores singulares v_i à direita de A são os autovetores de A^TA.*

Prova 6.4.1 *Veja que $A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^TU^TU\Sigma V^T = V\Sigma^T\Sigma V^T$. Logo $(A^TA)V = V\Sigma$, que mostra que $(v_i, \sigma_i^2) : i = 1, \ldots, n$ formam um autopar para A^TA. Ou seja, cada autovalor λ_i da matriz simétrica semipositiva definida A^TA é σ_i^2. Seus autovetores são $v_i : i = 1, \ldots, n$. A^TA possui r autovalores não nulos e $n - r$ nulos.*

Resultado 6.4.2 *Os vetores singulares u_i à esquerda de A são os autovetores de AA^T.*

Prova 6.4.2 *De forma análoga, $AA^T = (U\Sigma V^T)(U\Sigma V^T)^T = U\Sigma V^TV\Sigma^TU^T = U\Sigma\Sigma^TU^T$. Logo, $(AA^T)U = U(\Sigma\Sigma^T)$. AA^T possui r autovalores maiores que zero, $\lambda_i = \sigma_i^2 : i = 1, \ldots, r$ e $m - r$ autovalores nulos. Os autovetores de AA^T são os vetores singulares à esquerda de A. Os autovalores não nulos de A^TA e de AA^T são os mesmos, inclusive em suas multiplicidades algébricas.*

Um outro ponto importante diz respeito à orgonalidade das transformações lineares Av_i e Av_j, para vetores singulares v_i e v_j distintos, isto é, $i \neq j$. Veja o resultado a seguir.

Resultado 6.4.3 *Para $i \neq j$, temos $Av_i \perp Av_j$.*

Prova 6.4.3 *Veja que $Av_i = \sigma_i u_i$ e $Av_j = \sigma_j u_j$. Portanto:*

$$\begin{aligned} u_i^T u_j &= \left(\frac{Av_i}{\sigma_i}\right)^T \frac{Av_j}{\sigma_j} \\ &= \frac{v_i^T(A^T A v_j)}{\sigma_i \sigma_j} \\ &= \frac{\sigma_j^2}{\sigma_i \sigma_j} v_i^T v_j \\ &= 0 && \textit{pois } v_i \perp v_j \end{aligned}$$

Veja que o passo final do resultado acima demonstrado depende do fato de que as matrizes simétricas, $A^T A$ no caso aqui tratados, são unitariamente similares a uma matriz diagonal, possuindo autovetores ortogonais.

Vamos investigar o caso em que a matriz A fatorada na forma SVD é normal. Considere então os seguintes resultados.

Resultado 6.4.4 *Se A é uma matriz normal, isto é, $AA^T = A^T A$, os autopares (λ_i, q_i) de $A^T A$ são autopares de AA^T.*

Prova 6.4.4 *Considere a fatoração SVD completa $A = U\Sigma V^T$ de A. Então temos:*

$$\begin{aligned} AA^T &= (U\Sigma V^T)^T (U\Sigma V^T) \\ &= U\Sigma\Sigma^T U^T \\ &= U\Sigma^2 U^T \\ (AA^T)U &= U\Sigma^2 \end{aligned}$$

e também

$$\begin{aligned} A^T A &= (U\Sigma V^T)^T (U\Sigma V^T) \\ &= V\Sigma^T \Sigma V^T \\ &= V\Sigma^2 V^T \\ (A^T A)V &= V\Sigma^2 \end{aligned}$$

Portanto, $AA^T = A^T A$ implica em

$$\begin{aligned} A^T A &= V\Sigma^2 V^T \\ &= AA^T \\ &= U\Sigma^2 U^T \end{aligned}$$

Portanto $(A^TA)U = U\Sigma^2$ *e* $(AA^T)V = V\Sigma^2$*, comprovando que* AA^T *e* A^TA *possuem os mesmos autopares. Chame estes autopares de* (σ_i^2, q_i) *e veja que* $A^TA = Q\Lambda Q^T = AA^T$*, onde* Λ *é diagonal com entradas* σ_i^2 *na diagonal.*

Uma consequência do resultado acima é o seguinte.

Resultado 6.4.5 *Quando* A *é normal, os vetores singulares à direita* v_i *e à esquerda* u_i *de* A *são iguais e correspondem aos autovetores de* A^TA *(ou de* AA^T*).*

Recorde-se que uma matriz simétrica é um caso particular de matrizes normais, que como acima indicado, admitem fatorações espectrais. Observe que o preço que pagamos por não podermos produzir uma diagonalização espectral (6.6) (quando A não é normal) é termos dois conjuntos de vetores singulares: os u_i's e os v_i's. No caso de matrizes diagonalizáveis, precisamos apenas de um conjunto deles: os autovetores da matriz.

Considere a fatoração SVD completa e veja que podemos escrever $\Sigma = U^TAV$, para $U^TU = I$ e $V^TV = I$, onde $U^T = U^{-1}$ e $V^T = V^{-1}$. Esta observação sugere uma relação análoga à relação de similaridade entre matrizes. Veremos mais à frente, que se escrevemos

$$A = UBV$$

para U e V matrizes ortogonais, B e A são denominadas *ortogonalmente equivalantes* ou *unitariamente equivalantes*. Matrizes ortogonalmente equivalentes possuem os mesmos valores singulares. Assim sendo, podemos pensar que a fatoração SVD produziu uma *diagonalização* para a matriz A, considerando que a matriz Σ (ainda que retangular) é uma matriz diagonal, com os r valores singulares de A em sua diagonal (as eventuais demais $n - r$ entradas da diagonal de Σ sendo nulas).

Exemplo 49 *Vamos empregar os resultados acima para produzir uma fatoração* SVD *para a matriz* A *dada abaixo. Salientamos que, por razões que discutiremos em breve, os algoritmos que iremos empregar para esse propósito não computam as matrizes* A^TA *e* AA^T*.*

- $A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$ *com* $posto(A) = 2$

- $A^TA = \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix}$, $AA^T = \begin{bmatrix} 9 & 12 \\ 12 & 41 \end{bmatrix}$.

- *Autovalores de* $A^TA : \lambda_1 = \sigma_1^2 = 45, \lambda_2 = \sigma_2^2 = 5$

- *Autovetores de* $A^TA :$

- *Associado a* $\lambda_1 = 45$*:* $v_1 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
- *Associado a* $\lambda_2 = 5$*:* $v_2 = \frac{\sqrt{2}}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

- *Vetores* u_1, u_2 *calculados a partir de* v_1, v_2*:*
  - $u_1 = \frac{1}{\sqrt{45}} A v_1 = \frac{\sqrt{10}}{10}\begin{bmatrix} 1 & 3 \end{bmatrix}^T$
  - $u_2 = \frac{1}{\sqrt{5}} A v_2 = -\frac{\sqrt{10}}{10}\begin{bmatrix} -3 & 1 \end{bmatrix}^T$

- *A não é simétrica: os u's e os v's são diferentes.*

- *Verifique que* $u_1 = \frac{\sqrt{10}}{10}\begin{bmatrix} 1 & 3 \end{bmatrix}^T$ *e* $u_2 = -\frac{\sqrt{10}}{10}\begin{bmatrix} -3 & 1 \end{bmatrix}^T$ *são autovelores de* $AA^T = \begin{bmatrix} 9 & 12 \\ 12 & 41 \end{bmatrix}$*, para* $\lambda_1 = 45$ *e* $\lambda_2 = 5$*.*

- *Termos na fatoração:*
  $U = \frac{\sqrt{10}}{10}\begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}$,
  $\Sigma = \begin{bmatrix} \sqrt{45} & 0 \\ 0 & \sqrt{5} \end{bmatrix}$ *e*
  $V = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

6.5 Aplicações da fatoração SVD

Nesta seção tratamos de duas aplicações importantes da fatoração SVD, a saber a aproximação de matrizes por outras de posto baixo e a análise de compomenentes principais, temas conhecidos na literatura como *Low-rank approximation* e *Principal Component Analisys*, que dá origem ao acrônimo PCA.

6.5.1 Avaliação de potências de matrizes

Considere uma matriz A quadrada. Em algumas aplicações em Otimização e em Aprendizado de Máquinas é necessário avaliar A^k, para alguns valores positivos de k. Na maioria dos casos, calcular explicitamente a $k-$ésima potência de A é uma atividade muito cara. Isso sem contar o fato de que quando $k \to \infty$, não há como avaliar a matriz por este meio. Assim sendo, se a matriz A for diagonalizada, podemos computar

$$A^k = X\Lambda^k X^{-1},$$

bastando implementar a exponenciação nas entradas da matriz diagonal Λ. Quando $k \to \infty$, a potência A^k ou vai para zero (quando seu maior autovalor em módulo possui módulo inferir à unidade) ou explode, caso seu maior autovalor tiver módulo superior a 1. Independentemente do caso, podemos determinar uma função polinômial em A por meio de uma função polinomial nas entradas de Λ, isto é, nos autovalores de A.

Uma aplicação típica de potências de matrizes relevantes para a Ciência da Computação aparece na exponenciação de matrizes de adjacência de grafos.

6.5.2 Aproximação de posto baixo

Uma aplicação importante da fatoração SVD é a chamada *Low Rank Approximation* (LRA). Dada uma matriz, $A \in \mathbb{R}^{m \times n}$, de posto r, deseja-se encontrar uma outra matriz A_k de mesma ordem de A e de posto $k : k \leq r$ que resolva o seguinte problema de otimização:

$$f_0 = \min_{A_k \in \mathbb{R}^{m \times n} | \text{ posto}(A_k) = k} \|A - A_k\|_F^2. \tag{6.7}$$

Devemos ler o problema de otimização (6.7) da seguinte forma: Dentre todas as matrizes reais A_k com m linhas e n colunas, e com posto exatamente $k \leq r$, qual é a que melhor aproxima A, na norma de Frobenius ? Esta é a formulação matemática do LRA da matriz A.

Para resolver LRA, vamos assumir que dispomos da fatoração SVD reduzida de A, isto é $A = U\Sigma V^T = \sum_{i=1}^r \sigma_i u_i v_i^T$. Recorde-se que tanto a norma matricial espectral quanto a norma de Frobenius são invariantes às tranformações unitárias, isto é, se Q é ortogonal (ou unitária) $\|QA\|_F = \|A\|_F$ e $\|QA\|_2 = \|A\|_2$. Por esta razão podemos reescrever a função objetivo do problema (6.7) da seguinte forma:

$$\begin{aligned} \|A - A_k\|_F^2 &= \|U^T(A - A_k)V\|_F^2 \\ &= \|\Sigma - Z\|_F^2 \end{aligned}$$

onde $Z \in \mathbb{R}^{m \times n}$ é uma matriz de posto exatamente k. Diante disso, o problema (6.7) pode ser reescrito na seguinte forma mais conveniente:

$$f_0 = \min_{Z\in\mathbb{R}^{m\times n}|\ \text{posto}(Z)=k} \left\| \begin{bmatrix} \text{diag}\{\sigma_1,\ldots,\sigma_r\} & 0_{r,n-r} \\ 0_{m-r\times r} & 0_{m-r,n-r} \end{bmatrix} - Z \right\|_F^2 . \tag{6.8}$$

Na expressão acima, diag$\{\sigma_1,\ldots,\sigma_r\}$ representa uma matriz diagonal que tem o vetor $[\sigma_1,\ldots,\sigma_r]$ na diagonal principal. Note que, visando reduzir o valor da função objetivo (do quadrado da norma de Frobenius), não faz sentido que a matriz Z possua entradas não nulas fora da diagonal (elementos $z_{ii} : i = 1,\ldots,\min\{m,n\}$). Introduzir elementos z_{ij} não nulos fora da diagonal de Z apenas criará uma contribuição adicional positiva z_{ij}^2 para a função que desejamos minimizar. Desta forma, podemos assumir que Z é também uma matriz diagonal, assim como Σ. Então, a função objetivo em (6.8) pode ser escrita como

$$\|\text{diag}\{\sigma_1,\ldots,\sigma_r\} - \text{diag}\{z_{11},\ldots,z_{rr}\}\|_F^2 = \sum_{i=1}^{r}(\sigma_i - z_{ii})^2.$$

Claramente, a matriz Z ótima, que resolve portanto o problema de otimização (6.8), é uma matriz de zeros, exceto pelas entradas $z_{ii} = \sigma_i, i = 1,\ldots,k$. Para esta escolha de Z, a função objetivo é $f_0 = \sum_{i=k+1}^{r}\sigma_i^2$.

Então, a matriz A_k ótima que desejávamos determinar pode ser agora recuperada a partir de $Z = U^T A_k V$, que leva a

$$A_k = UZV^T = \sum_{i=1}^{k}\sigma_i u_i v_i^T.$$

A razão

$$\eta_k = \frac{\|A_k\|_F^2}{\|A\|_F^2} = \frac{\sigma_1^2+\sigma_2^2+\cdots+\sigma_k^2}{\sigma_1^2+\sigma_2^2+\cdots+\sigma_r^2}$$

indica a proporção da *variância* total dos dados representados pela matriz A que é explicada pela aproximação A_k de posto k de A. O erro ϵ_k de aproximação, portanto, é dado por

$$\epsilon_k = 1-\eta_k = \frac{\sigma_{k+1}^2+\sigma_{k+2}^2+\cdots+\sigma_r^2}{\sigma_1^2+\sigma_2^2+\cdots+\sigma_r^2}$$

Exemplo 50 *Aproximar a matrix A por A_2 na norma espectral e na norma de Frobenius.* $A = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

A fatoração SVD desta matriz é simples de ser verificada:

- *Valores singulares de A: 4, 3, 2, 1.*

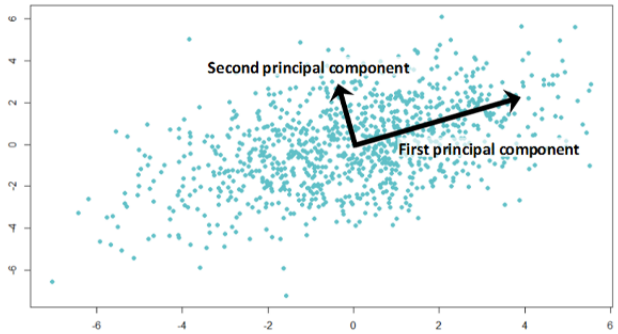

Figura 6.4: Figura extraída de `https://www.analyticsvidhya.com/`

- *Vetores singulares são colunas da matriz I.*

Veja que $A_2 = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ *e portanto,* $\|A - A_2\|_2 = 2$ *e* $\|A - A_2\|_F = 2 + 1 = 3.$

6.5.3 Análise de componentes principais

O PCA é uma técnica de Aprendizado de Máquina não supervisionada, no sentido de que os dados a serem utilizados na fase de treinamento não são rotulados. No PCA, o objetivo é encontrar as informações mais relevantes em um conjunto de dados, ou seja, encontrar as direções em um conjunto de dados ao longo das quais os dados variam mais. Considere a Figura 6.4 que ilustra um conjunto de dados bidimensionais: cada ponto representado na figura é um indivíduo com duas propriedades, representadas no eixo horizontal e vertical da figura. Verifique que há uma direção, indicada como *first principal component*, fazendo um ângulo de aproximadamente $\frac{\pi}{6}$ com o eixo horizontal, que é aquela em que há mais variação dos dados (em abuso de linguagem, *é uma direção ao longo da qual os dados se espalham mais*). Na segunda direção, ortogonal à primeira, há menor variação dos dados.

Para formalizar a ideia, vamos considerar que dispomos de m indivíduos $x_i \in \mathbb{R}^n$ e que cada indivíduo possua n características ou *features*. O indivíduo médio é representado pelo vetor $\overline{x} := \frac{\sum_{i=1}^m x_i}{m}$; este representando portanto o vetor com as médias das n características observadas em cada indivíduo. À partir dos vetores x_i

e do vetor médio $\overline{x}$, construímos a *matriz de dados centralizados*:

$$\tilde{X} = \left[\begin{array}{cccc} \tilde{x}_1 & \tilde{x}_2 & \cdots & \tilde{x}_m \end{array} \right],$$

onde cada coluna $\tilde{x}_i$ é um vetor n-dimensional dado por $x_i - \overline{x}$. Veja que nesta seção, a matriz de interesse possui ordem $n \times m$, diferentemente do normalmente empregamos até aqui. Tipicamente, temos mais indivíduos do que características associadas aos indivíduos, de forma que $m \geq n$ usualmente se aplica.

No PCA, matematicamente procuramos uma direção $z \in \mathbb{R}^n$ (uma *coluna sintética* de $\tilde{X}$, ou um *indivíduo sintético*) de norma Euclideana unitária, $\|z\|_2 = 1$, tal que a variância das projeções dos indivíduos centralizados ao longo de $span\{z\}$ seja a máxima. Este é chamado de primeiro vetor principal. Escolhemos a norma Euclideana para a definição da direção uma vez que esta norma não favorece nenhuma direção particular.

Veja a Figura 6.4 novamente e observe que os dados estão centralizados na origem. Observe também que nossa percepção geométrica concorda com a definição matemática dada.

Vamos agora apresentar a formulação matemática do problema que nos permite identificar o primeiro vetor principal z. Não conhecemos z, mas sabemos que a projeção do indivíduo centralizado $\tilde{x}_i$ em $span\{z\}$ é $\alpha_i z$, onde $\alpha_i = \tilde{x}_i^T z$ para todo $i = 1, \ldots, m$. Se somarmos os valores de α_i^2 estamos então somando os quadrados dos coeficientes das projeções de $\tilde{x}_i$ em z. Como desejamos maximizar a média desta soma, desejamos maximizar a quantidade

$$\frac{1}{m}\sum_{i=1}^{m}\alpha_i^2.$$

Então temos:

$$\begin{aligned}
\frac{1}{m}\sum_{i=1}^{m}\alpha_i^2 &= \frac{1}{m}\sum_{i=1}^{m}\sum_{i=1}^{m}(\tilde{x}_i^T z)^2 \\
&= \frac{1}{m}\sum_{i=1}^{m}(z^T\tilde{x}_i)(\tilde{x}_i^T z) \\
&= \frac{1}{m}\sum_{i=1}^{m} z^T(\tilde{x}_i\tilde{x}_i^T)z \\
&= z^T\left(\frac{1}{m}\sum_{i=1}^{m}\tilde{x}_i\tilde{x}_i^T\right)z \\
&= z^T\left(\frac{\tilde{X}\tilde{X}^T}{m}\right)z.
\end{aligned}$$

Normalmente, a matriz $S := \frac{\tilde{X}\tilde{X}^T}{m}$ (simétrica semipostiva definida) é denominada de matriz de covariância amostral. Os elementos da diagonal principal de S são o valor da variância da variável i:

$$S_{ii} = \frac{1}{m}\sum_{k=1}^{m}(\tilde{X}_{ik})^2.$$

Por outro lado, os elementos fora da diagonal principal da matriz S são o valor da covariância para cada par de variáveis:

$$S_{ij} = \frac{1}{m}\sum_{k=1}^{m}\tilde{X}_{ik}\tilde{X}_{jk}, i \neq j.$$

A matriz de covariância explica o comportamento das variáveis usadas para representar os dados, no seguinte sentido:

- Quando $cov(x, y) > 0$, os valores das variáveis x, y mudam na mesma direção (crescem ou decrescem);
- Quando $cov(x, y) < 0$, os valores das variáveis x, y mudam em direções opostas;
- Quando $cov(x, y) = 0$, as variáveis são independentes.

Neste momento, já podemos formular o problema de encontrar o primeiro componente principal dos dados como o seguinte problema de otimização:

$$\max_{z \in \mathbb{R}^n} \quad z^T\left(\frac{\tilde{X}\tilde{X}^T}{m}\right)z \tag{6.9}$$

$$\text{sujeito à} \quad \|z\|_2 = 1 \tag{6.10}$$

Para resolver o problema (6.13), vamos assumir que dispomos da fatoração SVD reduzida da matriz $\frac{\tilde{X}}{\sqrt{m}}$, isto é $\frac{\tilde{X}}{\sqrt{m}} = \sum_{k=1}^{r}\sigma_k u_k v_k^T$. Veja que $S = \frac{\tilde{X}\tilde{X}^T}{m}$ é simétrica (semipositiva definida), de ordem n, admitindo a fatoração espectral

$$S = U\Sigma^2U^T$$

que fornece os vetores singulares à esquerda $u_i : i = 1, \dots, r$ de $\frac{\tilde{X}}{\sqrt{m}}$. No caso em questão, a matriz S possui posto r. Portanto, possui r autovalores não nulos $\lambda_i = \sigma_i^2 > 0 : i = 1, \dots, r$, sendo os $u_i : i = 1, \dots, r$ os autovetores associados.

Com a fatoração SVD de $\tilde{X}/m$ em mãos, reformulamos o problema (6.13) em termos da matriz $U \in \mathbb{R}^{n\times r}$, usando o fato de que $\frac{\tilde{X}\tilde{X}^T}{m} = U\Sigma\Sigma^T U^T = U\Sigma^2 U^T$.

$$\max_{z\in\mathbb{R}^n} \quad z^T\left(U\Sigma^2U^T\right)z \tag{6.11}$$

$$\text{sujeito à} \quad \|z\|_2 = 1 \tag{6.12}$$

Veja que podemos escrever a função objetivo como $z^T\left(U\Sigma^2U^T\right)z = \|(\Sigma U^T)z\|_2^2$. Diante disso, fica claro que buscamos o vetor z, de norma Euclideana unitária que maximiza a norma Euclideana da transformação linear $(\Sigma U^T)z$. Veja que quando $z = u_1$, temos que $U^T z = e_1$ e $(\Sigma U^T)z = \Sigma(U^T z) = \Sigma e_1 = \sigma_1$. Logo, o primeiro vetor principal é o vetor singular u_1 associado a σ_1, pois $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$. A variância explicada por este componente principal é σ_1^2, que é o resultado da função objetivo (6.11) quando $z = u_1$.

Para calcular os demais componentes principais, procedemos da seguinte forma:

1. Subtraímos as projeções dos dados centralizados na direção do primeiro vetor pricippal $z = u_1$ e calculamos os novos dados *descontados*:

$$\tilde{x}_i^{(1)} = \tilde{x}_i - u_1(u_1^T\tilde{x}_i), i = 1, \cdots, m.$$

2. Calculamos a nova matriz $\tilde{X}^{(1)}$ que tem como colunas os vetores $\tilde{x}_i^{(1)}$:

$$\tilde{X}^{(1)} = \begin{bmatrix} \tilde{x}_1^{(1)} & \cdots & \tilde{x}_m^{(1)} \end{bmatrix} = (I_n - u_1u_1^T)\tilde{X}$$

3. Desta forma, podemos escrever a fatoração SVD de $\frac{\tilde{X}^{(1)}}{\sqrt{m}}$ à partir da fatoração de $\frac{\tilde{X}}{\sqrt{m}}$:

$$\begin{aligned}
\frac{\tilde{X}^{(1)}}{\sqrt{m}} &= \frac{\tilde{X}}{\sqrt{m}} - u_1u_1^T\frac{\tilde{X}}{\sqrt{m}} \\
&= \sum_{k=1}^{r}\sigma_k u_k v_k^T - u_1u_1^T(\sum_{k=1}^{r}\sigma_k u_k v_k^T) \\
&= \sum_{k=1}^{r}\sigma_k u_k v_k^T - (\sigma_1(u_1u_1^T)u_1v_1^T + \sum_{k=2}^{r}\sigma_k u_1(u_1^Tu_k)v_k^T) \\
&= \sum_{k=2}^{r}\sigma_k u_k v_k^T
\end{aligned}$$

Veja que a matriz $\tilde{X}^{(1)}$ é uma matriz de posto $r-1$, com valores singulares $\sigma_2, \ldots, \sigma_r$.

Portanto, como $v_i^T v_j = 0$ para $i \neq j$ e $v_k^T v_k = 0$, o problema de se encontrar o segundo componente principal pode ser formulado como

$$\max_{z \in \mathbb{R}^n} \quad z^T \left(\frac{\sum_{k=2}^{r} \sigma_k^2 u_k u_k^T}{m} \right) z \tag{6.13}$$

$$\text{sujeito à} \quad \|z\|_2 = 1, \tag{6.14}$$

cuja solução é $z = u_2$, por analogia ao raciocínio que apresentamos no cáculo do primeiro componente. Podemos repetir este processo até que as $k = r$ direções principais tenham sido identificadas.

Em resumo, as direções principais são os vetores singulares a esquerda de $\tilde{X}$: $u_1, u_2, \ldots, u_r$, nesta ordem.

6.5.4 Pseudo-inversa

Considere uma matriz quadrada $A \in \mathbb{R}^{n\times n}$ não singular e a transformação linear $Ax = b$. Como $C(A) = \mathbb{R}^n$, podemos escrever a transformação linear inversa, partindo da imagem b e obtendo o certificado x de que $b \in C(A)$: $x = A^{-1}b$. Para tanto, usamos o conceito da inversa A^{-1} de A que existe quando $C(A) = \mathbb{R}^n$ (ou $\det(A) \neq 0$). Esta inversa satisfaz $A^{-1}A = AA^{-1} = I_n$. A^{-1} é inversa à direita e à esquerda de A.

Para uma matriz A singular a inversa não existe, pois $C(A) \neq \mathbb{R}^n$. Para uma retangular, não é possível esperar existir uma inversa A^{-1} satisfazendo $A^{-1}A = AA^{-1}$, pois o número de linhas e colunas de A pode diferir. Essa ideia de inversa precisa ser adaptada, preservando a capacidade de representar a transformação linear inversa: levar a imagem em $C(A) \subseteq \mathbb{R}^m$ para o domínio $\mathbb{R}^n$.

Portanto, vamos de uma certa forma generalizar a ideia de inversa para matrizes quadradas singulares ou mesmo para matrizes retangulares. De agora em diante, vamos considerar $A \in \mathbb{R}^{m\times n}$ possui posto r e definir uma *pseudo-inversa* para A, uma matriz representada como $A^+ \in \mathbb{R}^{n\times m}$.

Esta pseudo-inversa deve reproduzir algumas operações da inversa de uma matriz quadrada não singular, para matrizes retangulares. Ela será construída (ou definida) de forma a satisfazer algumas propriedades:

- Para qualquer vetor $y_c \in C(A)$, ou seja, algum vetor y_c para o qual existe x tal que onde $Ax = y_c$, queremos garantir que $A^+y_c = x$ se verifique. Veja que estamos fazendo o mapeamento inverso de $y_c \in C(A)$ para $x \in \mathbb{R}^n$.

- Além disso, para qualquer vetor $y_n \in N(A^T)$ devemos satisfazer $A^+y_n = 0$.

Satisfeitas as condições acima, vamos verificar o que ocorre quando A^+ é aplicada em um $y \in \mathbb{R}^m$ qualquer. Sabemos que y pode ser decomposto de forma única na forma $y = y_n + y_c$, onde $y_n \in N(A^T), y_c \in C(A)$. Então temos:

$$\begin{aligned} A^+ y &= A^+(y_n + y_c) \\ &= A^+ y_c. \end{aligned}$$

Desta forma, apenas a parcela relativa a $C(A)$ de y de fato tem impacto na transformação linear inversa, uma vez que $A^+ y_n = 0$.

Vamos agora construir a A^+. Sabemos que os vetores $\{u_i : i = 1, \ldots, r\}$ que compõem as primeiras r colunas de U na fatoração SVD de A fornecem uma base para $C(A)$. Recordando: $A = U\Sigma V^T = \sum_{i=1}^r \sigma_i u_i v_i^T$ é a fatoração SVD de A. Sabemos também que com a fatoração SVD completa, as últimas $m - r$ colunas de U na fatoração fornecem uma base para $N(A^T)$.

Para algum $i \in \{1, \ldots, r\}$, devemos esperar que $A^+ u_i$ produza algum vetor $z \in \mathbb{R}^n$ que certifique que $u_i \in C(A)$. Sabemos também que $Av_i = \sigma_i u_i$, de forma que um possível valor para z é $z = \frac{1}{\sigma_i} v_i$. Então, propomos:

$$A^+ = \sum_{i=1}^{r} \frac{1}{\sigma_i} v_i u_i^T \tag{6.15}$$

como a fatoração SVD reduzida de A^+. Veja que a definição de A^+ dada por (6.15) satisfaz aquilo que estabelecemos

- para $u_k \in C(A)$, elemento da base para $C(A)$, temos

$$A^+ u_k = \left(\sum_{i=1}^{r} \frac{1}{\sigma_i} v_i u_i^T \right) u_k = \frac{1}{\sigma_k} v_k$$

  uma vez que $u_i \perp u_k$ para $i \neq k$.

- para $u_k \in N(A^T)$, temos que $A^+ u_k$ pois $N(A^T) \perp C(A)$.

Além disso, a expressão (6.15) indica que o posto de A^+ é r, o mesmo de A como era esperado. Também indica que se σ_i é valor singular de A, associado a vetores singulares à esquerda u_i e à direita v_i, então $\frac{1}{\sigma_i}$ é valor singular de A^+, associado a vetores singulares v_i à esquerda e u_i à direita. Relação similar existe entre os autovalores de uma matriz quadrada e sua inversa, caso exista. Embora a inversa de uma matriz quadrada possa não existir, sempre haverá sua pseudo-inversa, definida como (6.15).

Cabe ainda uma observação adicional sobre o resultado de $A^+ y$ para $y \in \mathbb{R}^m, y = y_c + y_n$, $y_c \in C(A), y_n \in N(A^T)$. Sabemos que apenas a parcela $y_c \in C(A)$ é determinante para o resultado de $A^+y = A^+y_c$. Este vetor y_c por estar em $C(A)$ pode ser escrito como $y_c = \sum_{i=1}^r \alpha_i u_i$. Pela definição de A^+ dada por (6.15), A^+u_i sempre resulta em $\frac{1}{\sigma_i}v_i$ onde v_i é um elemento da base para $C(A^T)$. Portanto, a transformação que produzimos com a pseudo-inversa definida como $A^+ = \sum_{i=1}^r \frac{1}{\sigma_i}v_i u_i^T$ sempre mapeia $y \in \mathbb{R}^m$ para algum vetor em $C(A^T)$. Isso é importante de ser dito pois: $Ax = y$ não admite solução se $y \notin C(A)$. Porém, se A possui posto incompleto, o sistema linear $Ax = y_c$ possui infinitas soluções. A pseudo-inversa fornece uma solução em $C(A^T)$.

Podemos escrever a definição de A^+, expressa por (6.15), por meio da fatoração SVD reduzida

$$A^+ = V\Sigma^{-1}U^T, \tag{6.16}$$

onde $\Sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_r} \end{pmatrix}$.

Observe a consistência dimensional da expressão (6.16): V é $n \times r$, Σ^{-1} é $r \times r$ e U^T é $r \times m$. Completando as colunas de V e as linhas de U^T com os elementos $\{v_{r+1}, \ldots, v_n\}$ e $\{u_{r+1}, \ldots, u_m\}$ das bases para $N(A)$ e $N(A^T)$ temos a definição da pseudo-inversa de A como

$$A^+ = V\Sigma^+U^T, \tag{6.17}$$

onde agora V é $n \times n$, U é $m \times m$ e Σ^+ (a pseudo-inversa de Σ) é $n \times m$ e tem as últimas $n - r$ linhas e $m - r$ colunas de zeros, assumindo a forma abaixo:

$$\Sigma^+ = \left(\begin{array}{ccc|c} \frac{1}{\sigma_1} & & & \\ & \ddots & & \\ & & \frac{1}{\sigma_r} & \\ \hline & & & \\ & & & \end{array}\right).$$

Veja que pela fatoração SVD reduzida de A^+ (dada por (6.15) ou (6.16)) e de A, o produto A^+A é uma matriz $n \times n$, satisfazendo

$$A^+A = \left(\sum_{i=1}^r \frac{1}{\sigma_i}v_i u_i^T\right)\left(\sum_{i=1}^r \sigma_i u_i v_i^T\right) = \left(\begin{array}{c|c} I_r & \\ \hline & \end{array}\right), \tag{6.18}$$

que possui as últimas $n-r$ linhas e colunas de zeros. Veja também que $AA^+ \neq A^+A$, pois AA^+ é uma matriz $m \times m$, com um bloco I_r e $m - r$ linhas e colunas de zeros adicionais.

Verifique que, pela expressão (6.18), A^+A é idempotente, simétrica e, portanto, é projetor ortogonal. Por uma expressão análoga a (6.18) para AA^+, as mesmas observações são válidas. Veja a Figura 6.5 para verificar em quais espaços projetam. AA^* projeta qualquer vetor em $C(A)$ enquanto que A^+A projeta em $C(A^T)$. Veja que para um $y = \sum_{i=1}^m \alpha_i u_i \in \mathbb{R}^m$ qualquer temos:

$$\begin{aligned}
AA^+\left(\sum_{i=1}^m \alpha_i u_i\right) &= A\left(\sum_{i=1}^m \alpha_i (A^+ u_i)\right) \\
&= A\left(\sum_{i=1}^m \alpha_i \left(\sum_{k=1}^r \frac{1}{\sigma_k} v_k (u_k^T u_i)\right)\right) \\
&= A\left(\sum_{i=1}^r \frac{\alpha_i}{\sigma_i} v_i\right) \\
&= \left(\sum_{i=1}^r \frac{\alpha_i}{\sigma_i} A v_i\right) \\
&= \left(\sum_{i=1}^r \frac{\alpha_i}{\sigma_i} \sigma_i u_i\right) \\
&= \sum_{i=1}^r \alpha_i u_i
\end{aligned}$$

Pela expressão acima $AA^+y \in C(A)$, uma vez que $\{u_i : i = 1, \ldots, r\}$ fornece uma base para $C(A)$. Desta forma, se $y \in C(A)$, $AA^+y = y$. Usando a mesma abordagem, tomando um vetor $x \in \mathbb{R}^n$ qualquer, $x = \sum_{i=1}^n \beta_i v_i$, mostramos que A^+A projeta em $C(A^T)$.

Exemplo 51 *Encontre a pseudo-inversa de* $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. *Claramente a matriz A possui posto um, e possui um valor singular* $\sigma_1 = 2$ *e os vetores singulares* $u_1 = \frac{1}{\sqrt{2}}(1,1)^T$ *e* $v_1 = \frac{1}{\sqrt{2}}(1,1)^T$. *Caracterizando* $N(A) = N(A^T)$ *(a matriz é simétrica), temos que* $u_2 = \frac{1}{\sqrt{2}}(1,-1)^T, v_2 = \frac{1}{\sqrt{2}}(1,-1^T)$. *Então temos:*

$$A^+ = V\Sigma^+U^T = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1/2 & 0 \\ 0 & 0 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Observe que apenas os r termos na fatoração SVD reduzida de A bastam para caracterizar A^+. *Porém, a forma completa é necessária para fazer uso de* Σ^+ *ao invés de* Σ^{-1}, *conforme a nossa definição. Neste exemplo apresentamos a forma* (6.17) *para poder explicitar a forma de* Σ^+.

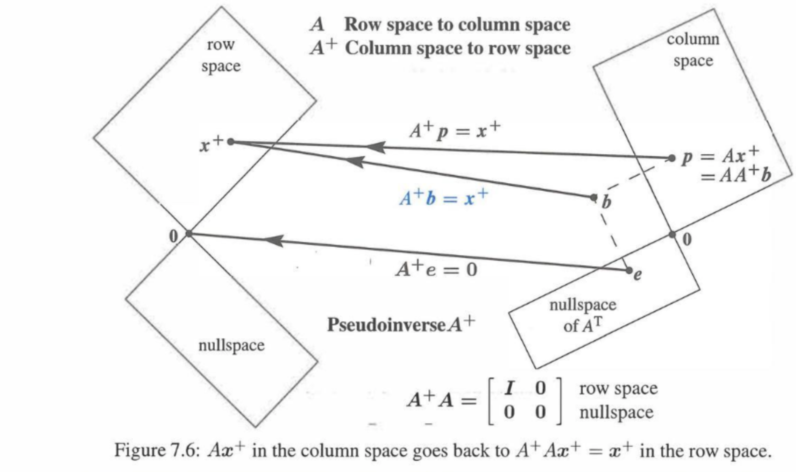

Figura 6.5: A pseudo-inversa de A e suas transformações lineares. Figura extraída de Introduction to Linear Algebra - Gilbert Strang.

Pseudo-invesa e o Problema de Mínimos Quadrados com deficiência de posto

Quando a matriz A possui posto $r < \min\{m, n\} = n$, o Problema de Mínimos Quadrados $\min_x \|Ax - b\|_2^2$ admite mais de uma solução, isto é, o sistema de equações normais $A^TAx = A^Tb$ é factível e admite infinitas soluções. Vamos usar a pseudo-inversa de A para obter uma destas possíveis soluções.

Desejamos encontrar algum x tal que $p = Ax$ onde p é a projeção ortogonal de b em $C(A)$. Sabemos que AA^+ é um projetor e que projeta em $C(A)$. Portanto $p = AA^+b$ é a projeção ortogonal de b em $C(A)$. Como $p \in C(A)$, $p = Ax$ para algum x. Vamos investigar a opção que surge naturalmente destas observações que é $p = A(A^+b) = Ax^+$, para $x^+ = A^+b$. Este vetor x^+ satisfaz as equações normais $A^TAx = A^Tb$. Para mostrar isso, vamos empregar a fatoração SVD de A, A^T, A^+ completas: $A = U\Sigma V^T, A^T = V\Sigma^T U^T, A^+ = V\Sigma^+ U^T$. Então temos:

$$\begin{aligned}
A^TAx^+ &= A^TAA^+b \\
&= (V\Sigma^TU^T)(U\Sigma V^T)(V\Sigma^+U^T)b \\
&= (V(\Sigma^T\Sigma\Sigma^+)U^T)b \\
&= (V\Sigma^TU^T)b \\
&= A^Tb
\end{aligned}$$

Exemplo 52 *Encontrar uma solução para o Problema de Mínimos Quadrados, de-*

finido por $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ *e* $b = (3,1)^T$. *Veja que* $b \notin C(A)$ *pois* $Ax = b$ *não admite solução. Usando pseudo-inversa* A^+ *de* A *obtida no Exemplo 51, encontramos* $x^+ = A^+b = (1,1)^T$. *Veja que tanto* x^+ *quanto* $x = (1+c, 1-c)$ *para qualquer* $c \in \mathbb{R}$ *resolvem o sistema de equações normais:*

$$\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

A solução $x^+ = A^+b$ fornecida pela pseudo-inversa é apenas uma solução. Porém, é a solução de menor norma Euclideana $\sqrt{(1+c)^2 + (1-c)^2} = \sqrt{2(1+c^2)}$. Verique que, como discutimos anteriormente, $x^+ \in C(A^T)$.

Exercícios Propostos

As questões foram adaptadas de [2].

Questão 01: Para as matrizes a seguir, encontre os valores singulares e os vetores singulares à esquerda e à direita: $A_1 \begin{bmatrix} 0 & 4 \\ 0 & 0 \end{bmatrix}, A_2 \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix}$. Verifique: $A_1 = U\Sigma V^T, A_2 = U\Sigma V^T$.

Questão 02: O espaço linha de $A = \begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix}$ tem dimensão igual a 1. Encontre v_1 no espaço linha de A e u_1 no espaço coluna. Qual o valor de σ_1? Escreva a matriz A como $A = U\Sigma V^T$.

Questão 03: Encontre bases ortonormais para os quatro subespaços fundamentais da matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$.

Questão 04: Seja a seguinte matriz construída a partir de dados coletados $X_0 = \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix}$. Encontre a média de cada uma das variáveis e encontre a matriz centralizada $\tilde{X}$. Calcule a matriz de covariaância amostral S e encontre os autovalores λ_1, λ_2. Qual a linha que passa pela origem e que é mais próxima das 5 amostras da matriz X_0?

Questão 05: Considere a matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$.

- Calcule a matriz A^TA e seus autovalores e autovetores. Encontre os valores singulares de A.
- Calcule a matriz AA^T e seus autovalores e autovetores.
- Verifique que $Av_1 = \sigma_1 u_1$. Fatore a matriz A usando a fatoração SVD reduzida e completa.
- Calcule a pseudoinversa de A.

Questão 06: Suponha que a matriz A tenha colunas ortogonais $w_1, w_2, \cdots, w_n$ com normas $\sigma_1, \sigma_2, \cdots, \sigma_n$, respectivamente. Descreva as matrizes U, Σ, V da fatoração

$A = U\Sigma V^T$.

Questão 07: Mostre que se v é um autovetor de $A^T A$, então Av é um autovetor de AA^T.

Questão 08: Aplicando a fatoração SVD, mostre que as matrizes $A^T A$ e AA^T possuem os mesmos autovalores diferentes de zero.

Questão 09: Suponha que $u_1, \cdots, u_n$ e $v_1, \cdots, v_n$ formam bases ortonormais para $\mathbb{R}^n$. Defina a matriz $A = U\Sigma V^T$ que transforme cada v_j em u_j tal que $Av_1 = u_1, \cdots, Av_n = u_n$.

Questão 10: Suponha que A seja uma matriz simétrica 2×2 com autovetores unitários u_1 e u_2 e autovalores $\lambda_1 = 3$ e $\lambda_2 = -2$. Quais são as matrizes U, Σ, V^T da fatoração $A = U\Sigma V^T$?

Capítulo 7

Cálculo de Autovalores, Autovetores e Vetores Singulares

Dada um matriz quadrada A, real de ordem n, nosso objetivo neste capítulo é construir algoritmos que determinem os autopares (λ, x) tais que $Ax = \lambda x$. Daremos ênfase ao caso em que A é real, mas mencionaremos também o caso em que a matriz é complexa hermitiana, isto é, $A = A^*$. A nossa abordagem para se obter os autovalores de A se baseia em fatorá-la, de forma a revelar seu espectro. Como as fatorações devem revelar o espectro, recorreremos ao uso de transformações similares como técnica de projeto de algoritmos. Também é nosso objetivo produzir algoritmos que computem a fatoração SVD $A = U\Sigma V^T$ de matrizes $A \in \mathbb{R}^{m\times n}$.

Essencialmente, encontrar as raízes do polinômio característico $p_A(\lambda)$ (obtido ao se impor $det(A-\lambda I) = 0$) não é uma opção, pois é um problema bastante mal condicionado. Na verdade, o problema de se encontrar as raízes de um polinômio qualquer (não necessariamente o polinômio característico de uma matriz) é normalmente formulado como o problema de se encontrar o espectro de uma matriz associada ao polinômio, chamada *matriz companheira* do polinômio. Ou seja, normalmente encontrar raízes de poliômios é resolvido como um problema de fatoração de uma matriz, que revele seu espectro. Como veremos, estes dois problemas, encontrar raízes de polinômios e descobrir o espectro de matrizes, são muito relacionados.

Concentramos nossos estudos em algoritmos para se produzir as seguintes fatorações matriciais:

1. Fatoração espectral de matrizes hermitianas (ou reais simétricas);

2. Fatoração de Schur para matrizes quadradas, não diagonalizáveis;

3. Fatoração SVD.

O objetivo das duas primeiras fatorações é revelar o espectro das matrizes quadradas e, se possível, também identificar seus autovetores. No caso das matrizes hermitianas, nosso algoritmo produzirá simultaneamente os autovalores e autovetores da matriz fatorada. No caso das matrizes defectivas, a fatoração de Schur nos dará o espectro da matriz e um de seus autovetores. Para o cálculo dos demais, será necessário trabalho computacional complementar. No caso da fatoração SVD, mostraremos como pode ser obtida sem recorrermos explicitamente à fatoração espectral de A^TA ou de AA^T, evitando problemas de condicionamento numérico que já discutimos anteriormente.

Logo abaixo, apresentamos um resumo sobre a existência das fatorações que revelam o espectro de matrizes quadradas:

1. A diagonalização de $A = X\Lambda X^{-1}$ existe se e somente se A é não defectiva. A e Λ são similares e portanto possuem os mesmos autovalores.

2. A diagonalização unitária de $A = Q\Lambda Q^*$ (com Q unitária, $QQ^* = Q^*Q = I$) existe se e somente se A é normal. Matrizes normais satisfazem $AA^* = A^*A$, sendo matrizes hermitianas (e reais simétricas) casos particulares de matrizes normais.

3. Toda matriz quadrada admite uma triangularização unitária, ou uma Decomposição de Schur, na forma $A = QTQ^*$, onde T é triangular superior.

Os algoritmos que iremos construir para produzirmos diagonalizações unitárias $A = Q\Lambda Q^*$ serão capazes de produzir uma fatoração de Schur, caso A não seja normal. Ou seja, a fatoração de Schur será obtida com o mesmo algoritmo. Apenas a resposta do algoritmo será distinta, dependendo da matriz de entrada. No caso geral, quando A não for normal, a fatoração produzida será uma fatoração de Schur. Os algoritmos que serão investigados nesta seção baseiam-se em tranformações lineares unitárias (ou ortogonais), numericamente desejáveis por serem estáveis.

7.1 Dificuldades no cálculo de autovalores

Sabemos que os autovalores de $A \in \mathbb{R}^{n \times n}$ são as raízes de seu polinômio característico $p_A(\lambda) = det(A - \lambda I) = 0$. Por outro lado, mostraremos aqui que qualquer problema de encontrar as raízes de um polinômio em coeficientes reais pode ser formulado como o problema de encontrar os autovalores de uma matriz associada ao polinômio, dita *matriz companheira do polinômio.*

Para verificar este resultado, considere o polinômio de grau m em z

$$p(z) = z^m + a_{m-1}z^{m-1} + \cdots + a_1 z + a_0 \tag{7.1}$$

e sua matriz companheira de ordem m

$$A = \begin{bmatrix} 0 & & & & & -a_0 \\ 1 & 0 & & & & -a_1 \\ & 1 & 0 & & & -a_2 \\ & & 1 & \ddots & & \vdots \\ & & & \ddots & 0 & -a_{m-2} \\ & & & & 1 & -a_{m-1} \end{bmatrix}. \tag{7.2}$$

Veja que a linha $(1, z, z^2, \ldots, z^{m-1})$ é um autovetor à esquerda de A, com autovalor z, quando z é uma raiz do polinômio $p(z)$. Isto é, verifique que $(1, z, z^2, \ldots, z^{m-1})A = (1, z, z^2, \ldots, z^{m-1})z$, quando $p(z) = 0$, observando os seguintes passos:

$$\begin{bmatrix} 1, z, z^2, \ldots, z^{m-1} \end{bmatrix} \begin{bmatrix} 0 & & & & & -a_0 \\ 1 & 0 & & & & -a_1 \\ & 1 & 0 & & & -a_2 \\ & & 1 & \ddots & & \vdots \\ & & & \ddots & 0 & -a_{m-2} \\ & & & & 1 & -a_{m-1} \end{bmatrix} =$$

$$\begin{bmatrix} z & z^2 & z^3 & \ldots & z^{m-1} & (-a_0 - a_1 z - a_2 z^2 + \cdots - a_{m-1}z^{m-1}) \end{bmatrix} =$$

$$\begin{bmatrix} z & z^2 & z^3 & \ldots & z^{m-1} & z^m \end{bmatrix} =$$

$$z \begin{bmatrix} 1 & z & z^2 & \ldots & z^{m-2} & z^{m-1} \end{bmatrix}$$

Note que na penúltima etapa do desenvolvimento acima, usamos o fato de que, quando z é raíz de $p(z)$, $z^m = -a_0 - a_1 z - a_2 z^2 + \cdots - a_{m-1}z^{m-1}$. Além da relação demonstrada entre as raízes de $p(z)$ e os autovalores da matriz companheira, cabe destacar que $p(z)$ é igual a $(-1)^m$ vezes o determinante da matriz (7.2). A consequência destes desenvolvilmentos é resumida no resultado abaixo.

Resultado 7.1.1 *O espectro da matriz companheira* (7.2) *fornece as raízes do polinômio* (7.1).

Em resumo, tanto quanto o problema de se encontrar autovalores de uma matriz se reduz ao problema de se encontraz raízes de um polinômio, tanto quanto o inverso também é verdadeiro: encontrar as raízes de qualquer polinômio com coeficientes reais equivale ao problema de se encontrar os autovalores de uma matriz convenientemente escolhida.

Da relação acima explicitada resulta a natureza dos algoritmos que iremos desenvolver para se calcular o espectro de uma matriz. Para tanto, considere o resultado abaixo enunciado, demonstrado no século XIX. Essencialmente, o teorema abaixo resumido indica que não há expressão analítica fechada, análoga por exemplo, à expressão de Báscara, para se expressar as raízes de um polinômio de grau igual ou superior a 5.

Resultado 7.1.2 (Abel, 1824) *Para qualquer $m \geq 5$, existe um polinômio de grau m com coeficientes racionais que possui uma raiz real r, com a propriedade de que r não pode ser escrita por uma expressão fechada envolvendo números racionais, adições, subtrações, multiplicações, divisões e radiciação.*

As consequências algorítmicas do resultado acima são muito importantes. Ainda que utilizássemos aritmética exata, não haveria algoritmo que produziria as raízes de um polinômio arbitrário em um número finito de passos. Naturalmente, a conclusão se aplica para o problema de se encontrar os autovalores de matrizes. E então, temos o mais importante resultado desta seção: Qualquer algoritmo para o cálculo de autovalores deve ser iterativo e não baseado em algum método direto, como os que vimos para a solução de sistemas lineares. Isso não significa que não sejamos capazes de produzir um bom algoritmo para se determinar o espectro de matrizes. Seremos e este é o tema da próxima seção.

7.2 Algoritmos para fatoração de Schur e diagonalização unitária

À luz dos resultados da seção anterior, não há um método direto para determinar autovalores de A. O objetivo dos algoritmos que discutiremos é *produzir uma sequência de matrizes que rapidamente convirjam para uma forma que revele os autovalores de A.* Assim sendo, o cálculo de autovalores é computacionalmente mais custoso do que a resolução de outras tarefas numéricas com as quais nos deparamos no curso, por exemplo, é mais cara que a solução de sistemas lineares. Apesar disso, em muitos casos, é possível produzir algoritmos que gerem sequências em que o número de digítos de precisão dobra ou triplica a cada iteração.

A ideia central dos algoritmos que vamos construir é resumida da seguinte forma. Vamos aplicar uma sequência de transformações unitárias

$$Q_j^* \cdots Q_2^* Q_1^* A Q_1 Q_2 \cdots Q_j,$$

de forma que

$$\lim_{j\to\infty} Q_j^* \cdots Q_2^* Q_1^* A Q_1 Q_2 \cdots Q_j = T,$$

onde T é uma matriz quadrada triangular superior. Isto é, no limite, A é similar a uma triangular superior, de forma que, no limite, obtemos uma fatoração de Schur de A.

Cabe destacar que, se a matriz de entrada A for hermitiana (ou real simétrica), $Q_j^* \cdots Q_2^* Q_1^* A Q_1 Q_2 \cdots Q_j$ será hermitiana (ou real simétrica). Portanto T será triangular e hermitiana, ou seja, uma matriz diagonal Λ. Em qualquer um dos casos apontados, os autovalores de A serão encontrados na diagonal de T (ou de Λ). Quando a matriz resultante das transformações for diagonal, apenas neste caso, as colunas do produto dos fatores $Q_1 Q_2 \dots Q_j$ nos dará os autovetores de A. Quando a matriz resultante for uma triangular superior T, apenas a primeira coluna de $Q_1 Q_2 \dots Q_j$ será um autovetor de A, associado ao autovalor λ_1, localizado na primeira posição T_{11} da diagonal de T. Por fim, destacamos que se A é real não simétrica, seus autovalores podem ser complexos (em conjugados). Portanto os algoritmos que produzem a fatoração de Schur devem admitir aritmética complexa.

7.2.1 Algoritmos

Independentemente de A ser Hermitiana (ou real simétrica) ou não, os métodos para o cálculo de autovalores se baseiam em duas fases. A segunda fase é imprescindível, existindo sem a aplicação da primeira fase. Esta, por sua vez, é importante para reduzir o custo por iteração e o número de iterações necessárias para a convergência da segunda fase. Ou seja, a segunda existe sem a primeira. A primeira ajuda a segunda.

De forma bastante resumida, as duas fases podem ser descritas da seguinte forma:

- Fase I: Trata-se de um <u>método direto</u> que visa transformar A em uma *matriz Hessenberg superior*, isto é, uma matriz com zeros abaixo da primeira subdiagonal (quase uma triangular superior, exceto pela primeira subdiagonal). Esta fase tem custo de $O(n^3)$. Seu único objetivo é melhorar a convergência e o custo por iteração da Fase II, descrita a seguir.

- Fase II: <u>método iterativo</u> para, assintoticamente, transformar a Hessenberg superior (ou a matriz A original caso a Fase I não tenha sido chamada) em uma triangular superior. Em princípio, *esta fase não termina nunca.* Porém, com $O(n)$ iterações, a norma da matriz subdiagonal inferior é reduzida para precisão da máquina. Sem a Fase I, o custo por iteração seria $O(n^3)$, pois a matriz seria densa e o número de iterações necessárias para convergência (pelo mesmo critério), bastante maior. Com a aplicação da Fase I, é possível reduzir a complexidade por iteração para $O(n^2)$, pois a matriz de entrada possui estrutura esparsa.

As Figuras abaixo ilustram a estrutura das matrizes obtidas ao longo das duas fases do algoritmo. A Figura 7.2 ilustra a topologia das matrizes quando $A = A^*$, enquanto que a Figura 7.1 trata do caso mais geral, não simétrico.

$$\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix} \underset{\text{Fase 1}}{\rightarrow} \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ & x & x & x & x \\ & & x & x & x \\ & & & x & x \end{bmatrix} \underset{\text{Fase 2}}{\rightarrow} \begin{bmatrix} x & x & x & x & x \\ & x & x & x & x \\ & & x & x & x \\ & & & x & x \\ & & & & x \end{bmatrix}$$

Figura 7.1: Estruturas das matrizes ao longo das duas fases, quando a matriz de entrada não é Hermitiana, $A \neq A^*$.

$$\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix} \underset{\text{Fase 1}}{\rightarrow} \begin{bmatrix} x & x & & & \\ x & x & x & & \\ & x & x & x & \\ & & x & x & x \\ & & & x & x \end{bmatrix} \underset{\text{Fase 2}}{\rightarrow} \begin{bmatrix} x & & & & \\ & x & & & \\ & & x & & \\ & & & x & \\ & & & & x \end{bmatrix}$$

Figura 7.2: Estruturas das matrizes ao longo das duas fases, quando a matriz de entrada é hermitiana, isto é, $A = A^*$.

Independentemente do caso (Hermitiana ou não), a premissa para o desenvolvimento de qualquer ideia aqui é usar transformações ortogonais, de forma a obter uma transformação similar.

Fase I. Vamos primeiro discutir uma ideia que seríamos tentados a empregar para a diagonalização (ou triangularização) de A mas que, sozinha, por um número finito de passos, não funciona. Apesar disso, esta primeira ideia, desde que convenientemente adaptada, será de valiosa importância mais tarde, na Fase II.

- Primeira ideia: constuir um refletor de Householder Q_1^*, aplicar Q_1^* à esquerda de A, gerando zeros nas linhas $2, \ldots, m$ na primeira coluna e, depois, Q_1 à direita de Q_1^*A. Veja que a transformação gera uma matriz $Q_1^*AQ_1$ similar a A.

  $\Rightarrow$ O problema desta ideia é que, ao fazer a segunda operação, à direita de Q_1^*A, combinaríamos as colunas de Q_1^*A, destruindo a estrutura de zeros criada na primeira coluna. Confira na Figura 7.3 que de fato, ao aplicar Q_1 à direita de Q_1^*A destruímos a estrutura de trinagular superior da primeira coluna de Q_1^*A. Não é surpresa à luz do resultado de Abel que não possamos por meio

$$\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix} \overset{Q_1^*A}{\longrightarrow} \begin{bmatrix} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{bmatrix} \overset{Q_1^*AQ_1}{\longrightarrow} \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix}$$

Figura 7.3: Efeito das transformações lineares $Q_1^T A Q_1$ caso $Q_1^T A$ gerasse uma coluna de $m-1$ zeros abaixo da diagonal principal da primeira coluna de A.

> um número finito destas transformações lineares triangularizar a matriz A. Curiosamente, esta ideia simplista que no momento descartamos, tem o efeito de tipicamente reduzir a magnitude das entradas abaixo da diagonal principal, apesar de não torná-las zero. Esta propriedade será explorada futuramente, na Fase II do algoritmo que computa autovalores.

Não temos como ser tão ambiciosos, para um algoritmo direto. Para resolver a questão, a da destruição da estrutura da primeira coluna, temos que ser menos ambiciosos ao empregar um método direto. Usaremos uma matriz ortogonal Q_1^*, construída a partir de um refletor de Householder, idealizado para, na primeira iteração manter inalterada a primeira linha de A, criando zeros apenas a partir da terceira linha daquela conluna em diante. Ao construirmos esta *matriz Q_1 mais modesta*, não perderemos a estrutura da primeira coluna, quando fizermos a transformação à direita de Q_1^*A. Veja o resultado que desejamos com a primeira iteração na Figura 7.4.

$$\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix} \overset{Q_1^*A}{\longrightarrow} \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{bmatrix} \overset{Q_1^*AQ_1}{\longrightarrow} \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ & x & x & x & x \\ & x & x & x & x \\ & x & x & x & x \end{bmatrix}$$

Figura 7.4: Efeito do primeiro par de transformações ortogonais que desejamos constuir, na Fase I. A figura ilustra um caso em que A não é Hermitiana.

Assim sendo, o algoritmo da primeira fase irá construir refletores de Householder para obter uma Hessenberg superior e não uma triangular superior.

A ideia que descrevemos para a primeira coluna será repetida para as demais colunas, exceto pela última, de forma que, no total, faremos $(n-2)$ transformações de Householder simétricas, à esquerda e à direita. Na segunda iteração, a matriz Q_2^* preserva as duas primeiras linhas de $Q_1^*AQ_1$, criando zeros a partir da quarta linha. Veja o resultado na Figura 7.5.

Veja que se procedermos como indicado, ao final de $n-2$ iterações a matriz resultante terá zeros abaixo da segunda subdiagonal, sendo portanto uma Hessenberg

$$\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ & x & x & x & x \\ & x & x & x & x \\ & x & x & x & x \end{bmatrix} \xrightarrow{Q_2^* Q_1^* A Q_1} \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ & x & x & x & x \\ & 0 & x & x & x \\ & 0 & x & x & x \end{bmatrix} \xrightarrow{Q_2^* Q_1^* A Q_1 Q_2} \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ & x & x & x & x \\ & & x & x & x \\ & & x & x & x \end{bmatrix}$$

Figura 7.5: Efeito do segundo par de transformações ortogonais na Fase I.

superior. Caso a matriz de entrada seja Hermitiana (ou real simétrica), a Hessenberg superior será uma matriz diagonal (possuindo elementos não nulos apenas na diagonal principal e nas duas diagonais acima e abaixo da principal).

Uma vez que a estratégia de abordagem da primeira fase foi delineada, podemos agora detalhar o cálculo dos refletores de Householder necessários em cada uma das $n-2$ iterações daquela fase. Para a explicação que segue, assuma que a matriz H receba uma cópia da matriz A.

- Na primeira iteração: Zeramos os elementos nas linhas $3,\ldots,m$, da primeira coluna de H. Portanto, para a construção de Q_1^*, empregamos:
  - $x \in \mathbb{R}^{m-1}$ como o vetor $H(2:m,1)$;
  - $v = signal(x_1)\|x\|e_1 + x$;
  - A matriz $Q_1^* \in \mathbb{R}^{m\times m}$ será $Q_1^* = \left[\begin{array}{c|c} 1 & 0_{1,m-1} \\ \hline 0_{m-1,1} & F_1 \end{array}\right]$, obtida a partir do refletor $F_1 = I - \frac{2}{\|v\|^2}vv^T \in \mathbb{R}^{(m-1)\times(m-1)}$.
- Para qualquer outra coluna, de índice $k \in \{2,\ldots,m-2\}$, empregamos:
  - $x \in \mathbb{R}^{m-k}$ é o vetor $H(k+1:m,k)$;
  - $v = signal(x_1)\|x\|e_1 + x$;
  - A matriz $Q_k^* \in \mathbb{R}^{m\times m}$ será $Q_k^* = \left[\begin{array}{c|c} I_k & 0_{k,m-k} \\ \hline 0_{m-k,k} & F_{m-k} \end{array}\right]$, obtida a partir do refletor $F_{m-k} = I_{m-k} - \frac{2}{\|v\|^2}vv^T \in \mathbb{R}^{(m-k)\times(m-k)}$.

A Fase I é detalhada no algoritmo identificado na Figura 7.6.

Exemplo 53 *Veja o resultado da aplicação do algoritmo da Fase I na matriz A abaixo identificada. Observe que, como a matriz de entrada é simétrica, o resultado é uma matriz tridiagonal simétrica.*

```
A1  =

   85.      102.    70.      129.    137.
   102.     167.    85.      157.    189.
```

```
function [H] = Householder_To_Hessenberg(A)
    [m,n] = size(A)
    H = A
    for k = 1:m-2
        x = H(k+1:m,k)

        vk = sign(x(1))*norm(x,2) * eye(m-k,1) + x
        vk = 1.0 / norm(vk,2) * vk

        // multiplicacao a esquerda, operacoes em linhas de H
        H(k+1:m,k:m)   = H(k+1:m,k:m)   - 2.0 * vk*(vk'*H(k+1:m,k:m))

        // multiplicacao a direita, operacoes em colunas de H
        H(1:m,k+1:m)   = H(1:m,k+1:m)   - 2.0 * (H(1:m,k+1:m) * vk) * vk'
    end
endfunction
```

Figura 7.6: Fase I para a fatoração de Schur de uma matriz quadrada. O algoritmo assume que A é quadrada, fazendo $n-2$ pares de transformações ortogonais simétricas em A.

```
   70.    85.    110.   91.    151.
   129.   157.   91.    272.   218.
   137.   189.   151.   218.   267.
->H =  Householder_To_Hessenberg(A1)
 H  =
   85.         -225.19769   1.701D-14   3.283D-14   2.122D-14
  -225.19769    683.96883   7.3768307  -6.085D-14  -8.860D-14
  -1.421D-14    7.3768307   72.467708  -28.312681   3.553D-14
  -2.842D-14    8.882D-16  -28.312681   30.104996   23.589342
  -2.842D-14   -1.776D-15   1.066D-14   23.589342   29.458471
```

Exemplo 54 *Neste segundo exemplo, a matriz de entrada não é simétrica mas admite autovalores reais. Veja que o resultado da Fase I é uma Hessenberg Superior.*

```
A2  =
   10.   4.   9.   8.   2.
   8.    1.   8.   7.   7.
   5.    3.   6.   1.   9.
   7.    4.   4.   3.   5.
   0.    2.   9.   7.   3.
->H =  Householder_To_Hessenberg(A2)
 H  =
```

```
  10.        -11.321712  -1.8717159   4.2299349  -3.9272345
 -11.74734    11.536232   11.539886  -5.3539833   3.8128417
  0.          9.3998608   3.0516512  -5.2453526   4.851592
  0.         -1.110D-16  -4.6624397   1.7493343   1.8270586
  0.          1.776D-15  -8.882D-16   1.2442563  -3.3372174
```

A Fase II. Podemos agora começar a discutir o algoritmo iterativo que corresponde à Fase II, salientando que o mesmo pode ser executado sem a prévia aplicação da Fase I. Nesse caso, serão necessárias mais iterações para sua convergência. Dependendo da estrutura da matriz de entrada A recebida, o algoritmo retornará fatorações diferentes para a mesma, de acordo com o seguinte:

- Se a Fase I foi aplicada e a matriz recebida é Hessenberg (triangular superior + subdiagonal), o algoritmo da Fase II irá produzir uma fatoração de Schur de A. O espectro de A estará representado na diagonal.
- Se a Fase I foi aplicada e A é tridiagonal, o algoritmo irá produzir uma fatoração espectral para A (ou uma diagonalização unitária de A), isto é, teremos a matriz de autovetores e uma diagonal com seus autovalores.
- As mesmas observações são válidas para o caso em que a Fase I não foi aplicada. A fatoração de Schur será retornada caso a matriz A não seja unitariamente diagonalizável, ou uma decomposição espectral unitária para A será apresentada, caso contrário.

A Fase II corresponde a uma das ideias mais brilhantes já concebidas em Computação Científica. Essencialmente, o ingrediente principal da Fase II é o algoritmo de fatoração QR, que será aplicado sequencialmente, por meio dos seguintes passos:

1. Fatoramos $A = QR$.
2. Veja que ao multiplicarmos A pelo fator Q^T, à esquerda, e Q à direita, obtemos uma transformação similar:

$$\begin{aligned} A &= QR \\ Q^T AQ &= Q^T QRQ \\ &= RQ. \end{aligned}$$

O desenvolvimento acima mostra que A e RQ são similares a $Q^T AQ$ e possuem os mesmos autovalores.

3. Por esta razão, atualizamos A como RQ e repetimos os passos 1 a 2, *até convergir*.

Discutiremos a caracterização da convergência do procedimento acima mais tarde. Para o momento, apenas afirmamos que o algoritmo converge para uma A triangular superior, com os autovalores na diagonal. Sobre o procedimento acima, cabem algumas observações adicionais importantes:

1. Este algoritmo essencialmente explora a primeira ideia que descartamos para a triangularização de A na Fase I. A ideia foi descartada pois era incapaz, em número finito de aplicações, de triangularizar ou diagonalizar A. Porém, a ideia é excelente para assintoticamente produzir uma matriz similar à matriz A, que seja triangular ou diagonal.

2. A ideia que descartamos para abordar o problema na Fase I é ruim para transformar A em uma triangular superior em um único passo, mas é bastante eficiente como estrutura de um processo iterativo, que gera uma forma de Schur para A, principalmente se a Fase I tiver sido chamada previamente.

Veja que a ideia do algoritmo é empregar, repetidamente a fatoração QR de uma matriz A que é substuída pelo produto de seus fatores Q, R, em ordem inversa. É fundamental que a implementação da fatoração QR empregada em cada iteração do método seja a mais estável possível. Por esta razão, usamos a fatoração QR via Refletores de Householder.

O algoritmo correspondente à Fase II, conhecido como Algoritmo QR *sem deslocamento* é apresentado na Figura 7.7. Sua implementação em `scilab` é apresentada na Figura 7.8 seguinte.

A esta altura, o leitor deve estar perplexo pois, o Passo 3 do algoritmo acima, que corresponde à multiplicar $R^{(k)}$ à direita por $Q^{(k)}$, destrói a estrutura triangular de $R^{(k)}$. O fato é que, apesar disso, no limite, para valores de k suficientemente grandes, estas matrizes $R^kQ^{(k)}$ serão triangulares superiores.

Para ilustrar o procedimento, vamos considerar dois exemplos.

Exemplo 55 `A1 =`

```
85.    102.   70.    129.   137.
102.   167.   85.    157.   189.
70.    85.    110.   91.    151.
129.   157.   91.    272.   218.
137.   189.   151.   218.   267.
[Qa1,Lambda1,T1] = QR_Para_Autovalores(A1,1.0E-15)
```

(Inicialização:) $k \leftarrow 1$, $A^{(k)} = A$.
Repita até convergir:

1. Fatoramos $A^{(k)} = Q^{(k)} R^{(k)}$ (Householder faz a triangularização de $A^{(k)}$), isto é, $R^{(k)} = (Q^{(k)})^T A^{(k)}$).
2. Como $Q^{(k)}$ é ortogonal, $(Q^{(k)})^T A^k = R^{(k)}$.
3. Multiplicando à direita: $(Q^{(k)})^T A^k Q^{(k)} = R^{(k)} Q^{(k)}$ (temos uma transformação similar, que preserva o espectro).
4. $A^{(k+1)} \leftarrow R^{(k)} Q^{(k)}$, $k \leftarrow k+1$ e repetimos o processo até que, após a atualização, $A^{(k+1)}$ *seja suficientemente triangular superior.*

Figura 7.7: Algoritmo QR sem deslocamento para fatoração de Schur de A. Na descrição do algoritmo, $A^{(k)}$ não é a k-ésima potência de A, mas sim a matriz disponível na $k-$ésima iteração do procedimento.

```
Lambda1  = 759.26225 88.740522 43.220379 9.7447831 0.0320717
T1  =
   759.26225  -2.324D-14   6.974D-15  -5.490D-14   9.601D-14
   0.          88.740522   1.030D-14   1.285D-14   1.225D-14
```

```
function [Qa,Lambda,H] = QR_Para_Autovalores(A,tol)
    [m,n] = size(A)
    H = A
    [H] = Householder_To_Hessenberg(A)
    CONVERGIU = 0
    k = 0
    Qa = eye(m,m)
    while (CONVERGIU == 0)
       [Q,R] = QR_Householder(H)
       Qa = Qa * Q
       H = R*Q
       ninf = norm(tril(H)-diag(diag(H,0)),1)
       printf("Iter: %d ninf: %7.6E \n",k,ninf)
       k = k + 1
       if (ninf < tol)
         CONVERGIU = 1
       end
       k = k+1
    end
    Lambda = diag(H)
endfunction
```

Figura 7.8: Algoritmo QR sem deslocamento para Fatoração de Schur.

```
0.          0.          43.220379   8.313D-15  -1.465D-14
0.          0.          0.          9.7447831   3.233D-15
0.          0.          0.          0.          0.0320717
```

Exemplo 56 `A2  =`

```
10.   4.   9.   8.   2.
8.    1.   8.   7.   7.
5.    3.   6.   1.   9.
7.    4.   4.   3.   5.
0.    2.   9.   7.   3.
[Qa2,Lambda2,T2] = QR_Para_Autovalores(A2,1.0E-15)
Lambda2  = 26.024819 -7.5953848 6.7069158 -3.3870332 1.2506831
 T2  =

26.024819   0.8270619   4.8318138   8.2921898  -2.0623582
0.         -7.5953848   0.7415849  -2.272945   -1.8566947
0.          0.          6.7069158   0.6733713   1.4334871
0.          0.          0.         -3.3870332  -0.5328745
0.          0.          0.          0.          1.2506831
```

7.3 Fatoração SVD

Nesta seção, apresentamos algoritmos para obtermos a fatoração SVD de $A \in \mathbb{R}^{m\times n}$: $A = U\Sigma V^T$. Sabemos que os vetores singulares de A relacionam-se aos autovetores de A^TA e AA^T, assim como os valores singulares $\sigma_i > 0$ de A são as raízes quadradas dos autovalores não nulos de AA^T e de A^TA. Assim sendo, poderíamos (mas não devemos) fazer a fatoração SVD de por meio da fatoração espectral de A^TA, através dos passos seguintes:

1. Calculamos explicitamente A^TA.
2. Fatoramos $A^TA = Q^T\Lambda Q$.
3. Os valores singulares σ_i de A são as raízes dos autovalores não nulos λ_i de A^TA, armazenados na diagonal de Λ.
4. Os vetores singulares à direita v_i de A são as colunas q_i de Q, associadas às entradas $\lambda_i > 0$.
5. Os vetores singulares à esquerda de A são obtidos via $u_i = \frac{1}{\sigma_i}Aq_i$, para todo $\sigma_i > 0$.

Enfatizamos que o procedimento acima não deve ser empregado, tendo em vista que devemos evitar o cálculo explícito de A^TA, pois a matriz é mais mal condicionada do que A. Além disso, o problema sofre da *perda de informação devido ao quadrado*, isto é, se A possui valores singulares distintos de zero mas muito pequenos, estes valores serão avaliados sem precisão ao se calcular A^TA. As abordagens que apresentaremos na sequência não fazem uso explícito de A^TA ou de AA^T.

7.3.1 Fatoração SVD sem o cálculo explícito de A^TA

Em linhas gerais, os algoritmos que vamos apresentar são divididos em duas fases:

- Fase I: Transformamos A em uma matriz bidiagonal $\hat{B}$ por meio de um método direto (baseado em *transformações ortogonalmente* equivalentes, via Refletores de Householder).
- Fase II: Extraímos de $\hat{B}$ sua submatriz B bidiagonal quadrada.
  - Construímos uma matriz auxiliar quadrada $(2n \times 2n)$

$$H = \begin{bmatrix} 0 & B^* \\ B & 0 \end{bmatrix}.$$

  - Fazemos a fatoração espectral de $H = Z\Sigma Z^T$. Com os autovetores de H (nas colunas de Z) recuperamos os vetores singulares de $B, \hat{B}$ e de A.

A estrutura que descrevemos acima faz uso de um conceito novo: *transformações ortogonalmente equivalentes.*

Definição 7.4 *Duas matrizes $\hat{B}, A \in \mathbb{R}^{m\times n}$ são ortogonalmente equivalentes (OE) se e somente se existem matrizes ortogonais $Q \in \mathbb{R}^{m\times m}, P \in \mathbb{R}^{n\times n}$ tais que*

$$A = Q\hat{B}P.$$

Veja que se A e $\hat{B}$ são OE podemos relacionar seus valores e vetores singulares de uma forma simples e conveniente. Seja $\hat{B} = \hat{U}\Sigma\hat{V}^T$ a fatoração SVD de $\hat{B}$. Então para para $A, \hat{B}$ OE, temos:

$$\begin{aligned} A &= Q\hat{B}P \\ &= Q(\hat{U}\Sigma\hat{V}^T)P \\ &= (Q\hat{U})\Sigma(\hat{V}^TP) \end{aligned}$$

Portanto, $A = U\Sigma V^T$, de forma que $A, \hat{B}$ possuem os mesmos valores singulares e $U = (Q\hat{U})$, $V^T = \hat{V}^T P$ são os fatores desejados na fatoração de A.

Veja que o conceito de matrizes OE não é muito diferente do conceito de matrizes similares que empregamos para obter diagonalizações ou triangularizações unitárias de matrizes quadradas. No caso da fatoração SVD, não precisamos de transformações similares, mas sim de tranformações OE. Agora, no contexto da fatoração SVD, vamos transformar A em uma $\hat{B}$ conveniente, ortogonalmente equivalente a A, e calcular a fatoração SVD de $\hat{B}$. No contexto de fatoração SVD, a forma *conveniente* da matriz $\hat{B}$ é bidiagonal, por razões que ficarão mais claras brevemente.

Os algoritmos que iremos apresentar são mais facilmente explicados quando a matriz A é quadrada. Como normalmente isso não é observado, em uma etapa de pré-processamento fazemos a fatoração QR reduzida de A e, na Fase I, bidiagonalizamos R ao invés de A. Na composição final dos vetores singulares de A a partir dos de R e de B, usamos o fator Q obtido na fase de pré-processamento. Além disso, podemos assumir que o número m de linhas de A seja pelo menos igual ao número de colunas n, pois caso contrário, podemos fazer a fatoração SVD de A^T. Assim sendo, vamos assumir que $m \geq n$. Se $m = n$, não precisamos fazer a etapa de pré-processamento (QR de A). Se $m > n$, fazemos a QR de A e bidiagonalizamos R.

Podemos bidiagonalizar $A \in \mathbb{R}^{m\times n}$, por meio de n matrizes ortogonais $E_i \in \mathbb{R}^{m\times m} : i = 1, \ldots, n$ e $n-2$ matrizes ortogonais $D_i \in \mathbb{R}^{n\times n} : i = 1, \ldots, n-2$, aplicados sequencialmente à esquerda e à direita de A, respectivamente. Isto é, existem $O(n)$ matrizes E_i, D_i ortogonais tais que:

$$\hat{B} = E_n E_{n-1} \ldots E_1 A D_1 D_2 \ldots D_{n-2}$$

onde $\hat{B}$ é bidiagonal, ortogonalmente equivalente a A. As matrizes E_i e D_i serão construídas por meio de refletores de Householder. Veja que se A é uma triangular superior, caso a etapa de pré-processamento tenha sido aplicada, não precisamos das transformações à esquerda de A (ou de R), isto é, $E_i = I$.

Visando recordar o uso dos refletores de Householder considere o seguinte exemplo.

Exemplo 57 *Vamos supor que a transformação $E_1 A$ na primeira coluna de A tenha sido realizada e que* $E_1 A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 2 & 4 & 1 \\ 0 & -3 & -2 & 5 \end{bmatrix}$ *Então temos que a primeira operação à direita pode ser realizada da seguinte forma:*

```
function [B] = Bidiagonaliza(A)
    [m,n] = size(A)
    if (n > m) then B = A'; [m,n] = size(B);
              else B = A;
    end
    for k = 1:n
        // operacoes a esquerda
        x = B(k:m,k)
        vk = sign(x(1))*norm(x,2) * eye(m-k+1,1) + x
        vk = 1.0 / norm(vk,2) * vk
        B(k:m,k:n)   = B(k:m,k:n)   - 2.0 * vk*(vk'*B(k:m,k:n))
        if (k <= n-2)
          // Operacoes a direita
          x  = B(k,k+1:n)'
          [n1,m1] = size(x)
          vk = sign(x(1))*norm(x,2) * eye(n-k,1) + x
          vk = 1.0 / norm(vk,2) * vk
          D = eye(n,n)
          F = eye(n-k,n-k)
          F = F - 2 * vk * vk'
          D(k+1:n,k+1:n) = F
          B = B * D
        end
    end
endfunction
```

Figura 7.9: Algoritmo para bidiagonalizacão de uma matriz $A \in \mathbb{R}^{m\times n}$.

- $D_1 = \begin{bmatrix} 1 & 0 \\ 0 & F_1 \end{bmatrix}$ *onde* $F_1 = I_{3\times 3} - 2\frac{vv^T}{v^Tv}$ *para* $x = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^T$, $v = \sqrt{2}\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T - \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^T = \begin{bmatrix} \sqrt{2}-1 & -1 & 0 \end{bmatrix}^T$

- $D_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

- $E_1AD_1 = \begin{bmatrix} 1 & 1.4142136 & 0 & 0 \\ 0 & 2.8284271 & -1.4142136 & -2 \\ 0 & 4.2426407 & -1.4142136 & 1 \\ 0 & -3.5355339 & -0.7071068 & 5 \end{bmatrix}$

O algoritmo apresentado na Figura 7.9 implementa a bidiagonalização de uma matriz com m linhas e n colunas, para $m \geq n$.

Veja o resultado da bidiagonalização das matrizes A_1, A_2 que já aprentamos neste capítulo.

Exemplo 58 *Resultado do algoritmo apresentado na Figura para a bidiagonalização das matrizes A_1, A_2 anteriormente definidas.*

```
-->Bidiagonaliza(A1)
 ans  =
  -240.70521   719.46083   0.          0.          0.
   0.          31.902753   22.076149   0.          0.
   0.          0.          83.604679  -17.802383   0.
   0.          0.          0.          43.573255   6.6092123
   0.          0.          0.          0.         -0.0325336
-->Bidiagonaliza(A2)
 ans  =
  -15.427249   20.884335   0.          0.          0.
   0.          12.153633   2.4112356   0.          0.
   0.          0.         -4.5019751  -5.7067634   0.
   0.          0.          0.          3.7099871   2.1224413
   0.          0.          0.          0.          1.7933167
```

Veja também o resultado da bidiagonalização de uma matriz retangular, A_3.

Exemplo 59 `->A`

```
 A3  =


   2.   4.   5.   8.
   6.   9.   3.   0.
   6.   9.   5.   7.
   7.   3.   5.   2.
   9.   4.   1.   4.
   5.   7.   2.   8.
   3.   3.   6.   6.

-->[B3] = Bidiagonaliza(A3)
 B3  =
  -15.491933   19.530425   0.          0.
   0.         -12.286121  -3.5853615   0.
   0.          0.         -6.8825697  -0.431447
   0.          0.          0.          5.1188868
   0.          0.          0.          0.
   0.          0.          0.          0.
   0.          0.          0.          0.
```

Fase II. Vamos agora iniciar a apresentação da segunda fase do algoritmo, assumindo que A é quadrada ou que, em caso contrário, procedemos à bidiagonalização de R (obtido via fatoração reduzida $QR = A$) e não de A.

Uma hipótese adicional importante aqui é que assumimos que a matriz bidiagonal B $(n \times n)$ obtida seja *propriamente bidiagonal*, ou seja, não há elementos nulos na diagonal principal e na primeira super diagonal. Se o contrário ocorresse, por exemplo, se $B_{k,k+1} = 0$ para algum k, poderíamos particionar $B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}$ em blocos, onde $B_1 \in \mathbb{R}^{k \times k}, B_2 \in \mathbb{R}^{m-k \times m-k}$. As fatorações SVD de B_1,B_2 podem ser feitas separadamente e, então, combinadas para a SVD de B. De forma similar, se $B_{k,k} = 0$ também podemos decompor o problema em dois subproblemas independentes. Omitimos estes detalhes nesta apresentação introdutória.

Para B propriamente bidiagonal, construímos a matriz auxiliar Hermitiana de ordem $2n$, $H = \begin{bmatrix} 0 & B^* \\ B & 0 \end{bmatrix}$, que é fatorada como

$$\begin{bmatrix} 0 & B^* \\ B & 0 \end{bmatrix} \begin{bmatrix} V & V \\ U & -U \end{bmatrix} = \begin{bmatrix} V & V \\ U & -U \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{bmatrix}.$$

Observe que:

1. Como B é propriamente bidiagonal, H é não singular, tendo todos seus autovalores reais não nulos.

2. A forma bidiagonal B ortogonalmente equivalente a R ou a A acelera o cálculo da fatoração espectral de H. Em princípio qualquer outra matriz quadrada não singular ortogonalmente equivalente a R poderia ter sido usada no lugar de B. Porém, com uma bidiagonal, aceleramos a Fase II.

3. Os autovalores de H aparecem aos pares $\sigma_i, -\sigma_i$. Os módulos destes valores fornecem os valores sigulares de B.

4. A fatoração espectral de H revela a fatoração espectral de B, pois temos:

$$BV = U\Sigma$$
$$B^*U = V\Sigma$$

5. Se $\begin{bmatrix} v & u \end{bmatrix}^T \in \mathbb{R}^{2n}$ é autovetor de H, $\begin{bmatrix} v & -u \end{bmatrix}^T$ também é. u, v são vetores singulares à direita e à esquerda de B. Os vetores v_i, após terem sido normalizados, são vetores singulares de B.

6. Os vetores singulares de A (e/ou de seu fator R) podem ser computados a partir de v, u.

Apresentamos a seguir um resumo dos passos que devemos proceder para obter a fatoração SVD de A e um exemlo completo, lembrando que nosso objetivo é obter os fatores U, Σ, V em $A = U\Sigma V^T$.

P. 1 Dada $A \in \mathbb{R}^{m\times n}$, fazemos a fatoração $A = QR$ reduzida de A. Assumimos que A possui posto coluna completo.

P. 2 Fatoramos $R \in \mathbb{R}^{n\times n}$ na forma SVD obtendo:

$$R = U_R \Sigma_R V_R^T$$

O passo P. 2 é detalhado da seguinte forma:

P. 2.1 Bidiagonalização de R por meio de transformações unitárias:

$$ERD = B \iff R = E^T B D^T$$

Recorde-se que para R triangular superior, $E = I$.

P. 2.2 Fatoração espectral de $H = \begin{bmatrix} 0 & B^* \\ B & 0 \end{bmatrix}$.

$$\begin{bmatrix} 0 & B^* \\ B & 0 \end{bmatrix} \begin{bmatrix} V_B & V_B \\ U_B & -U_B \end{bmatrix} = \begin{bmatrix} V_B & V_B \\ U_B & -U_B \end{bmatrix} \begin{bmatrix} \Sigma_B & 0 \\ 0 & -\Sigma_B \end{bmatrix}$$

Ou seja,

$$BV_B = U_B \Sigma_B,$$

e portanto, a fatoração espectral de H nos fornece a SVD de B (e vice versa).

Aqui cabe uma nota de atenção: Para uso na fatoração de A, precisamos normalizar V_B, U_B. Isso porque extraímos V_B, U_B de um autovetor $z = [V_B \quad U_B]^T$ tem $\|z\|_2 = 1$. Porém, seus subvetores V_B, U_B não.

P.3 Compomos o resultado:

$$A = (QU_R)\Sigma_R V_R^T$$

$\Rightarrow$ Ou seja: $V = V_R, U = QU_R, \Sigma = \Sigma_R$.

Desta forma, temos a composição final dos fatores de A pode ser construída da seguinte forma:

1. Fatoração SVD de B:

$$B = U_B \Sigma_B V_B^T$$

2. Fatoração SVD de R:

$$\begin{aligned} R &= E^T B D^T \\ &= (E^T U_B) \Sigma_B V_B^T D^T \end{aligned}$$

3. Fatoração QR de A:

$$\begin{aligned} A &= QR \\ &= (QE^T U_B) \Sigma_B (V_B^T D^T) \end{aligned}$$

4. Fatores obtidos em $A = U\Sigma V^T$: $U = QE^T U_B$, $\Sigma = \Sigma_R$, $V^T = V_B^T D^T$

Vamos ilustrar todo este algoritmo por meio de um exemplo completo.

Exemplo 60 *Vamos fazer a fatoração SVD da matriz retangular A indicada abaixo. Primeiro, vamos usar a implementação da fatoração disponível no* `Scilab` *e na sequência, usando o procedimento que acabamos de detalhar, comparando os resultados obtidos. Primeira parte: usando a função SVD do* `Scilab`*.*

```
A3 = [2.   4.   5.   8.;
      6.   9.   3.   0.;
      6.   9.   5.   7.;
      7.   3.   5.   2.;
      9.   4.   1.   4.;
      5.   7.   2.   8.;
      3.   3.   6.   6.];
-->[Us,Ss,Vs] = svd(A3);
-->Vs,Ss
 Vs  =
  -0.528988   -0.568151   -0.5957387   0.2060861
  -0.5710213  -0.2856276   0.7690584  -0.0300126
  -0.3702612   0.2547129  -0.214162   -0.8672731
  -0.5069645   0.7285209  -0.0881998   0.4521782
 Ss  =
   26.913271   0.          0.          0.
   0.          8.8292206   0.          0.
   0.          0.          5.5837141   0.
```

```
   0.           0.           0.           5.0539917
   0.           0.           0.           0.
   0.           0.           0.           0.
   0.           0.           0.           0.
```

Segunda parte: procedendo de acordo com o algoritmo que apresentamos. De início apresentamos nossa formulação para resolver o problema como a sequência dos seguintes passos, estruturados em linguagem `Scilab`*.*

```
[m,n] = size(A3);
[Q,R] = qr(A3);
[E,B,D] = BidiagonalizaExp(R(1:n,1:n))
H = zeros(2*n,2*n);
BTrans = B';
H(1:n,n+1:2*n) = BTrans;
H(n+1:2*n,1:n) = B;
[AutovetoresH,S] = spec(H)
VB = AutovetoresH(1:n,n+1:2*n)
UB = AutovetoresH(n+1:2*n,n+1:2*n)
for i = 1:n
    VB(:,i) = VB(:,i)/norm(VB(:,i),2)
    UB(:,i) = UB(:,i)/norm(UB(:,i),2)
end
U = Q(:,1:n)*E'*UB
V = D*VB
Sigma = diag(S(n+1:2*n,n+1:2*n))
```

Os passos acima produzem os seguintes resultados:

```
-->V
 V  =
   0.2060861  -0.5957387  -0.568151   -0.528988
  -0.0300126   0.7690584  -0.2856276  -0.5710213
  -0.8672731  -0.214162    0.2547129  -0.3702612
   0.4521782  -0.0881998   0.7285209  -0.5069645
-->Sigma
 Sigma  =
   5.0539917
   5.5837141
   8.8292206
   26.913271

-->U
```

```
U  =

  -0.0844516   0.0194042   0.5462452  -0.3436625
  -0.3235889   0.4843742  -0.5906995  -0.3501582
  -0.0405055   0.2970933   0.0445856  -0.5095322
  -0.4114458  -0.5570138  -0.2382241  -0.3077001
   0.5295151  -0.5108385  -0.3496427  -0.3508714
   0.5348687   0.2275891   0.169601   -0.425007
  -0.388277   -0.2317834   0.3780704  -0.3181844
```

Exercícios Propostos

As questões de 9 a 12 foram adaptadas de [3].

Questão 01: Falso ou Verdadeiro ou Não é possível dizer? Justifique. Se os autovalores de uma matriz A são iguais a 2,2,5 então a matriz é: (a) inversível; (b) diagonalizável; (c) não diagonalizável.

Questão 02: Falso ou Verdadeiro ou Não é possível dizer? Justifique: Se os autovetores de uma matriz A são múltiplos do vetor $[1\ \ 4]^T$, então A: (a) não tem inversa; (b) tem um autovalor repetido; (c) não é diagonalizável.

Questão 03: Seja $A^k = X\Lambda^k X^{-1}$. Em quais casos $A^k = X\Lambda^k X^{-1}$ se aproxima de uma matriz nula, ou seja, $A^k \to 0$?

Questão 04: Qual o valor de b na matriz $A = \begin{bmatrix} 2 & b \\ 1 & 0 \end{bmatrix}$ tal que: (a) $A = Q\Lambda Q^T$ exista? (b) A é não diagonalizável? (c) A é singular?

Questão 05: Se a matriz A é uma matriz ortogonal, quais são as matrizes Q e R da fatoração QR? Neste caso, o algoritmo QR para o cálculo dos autovalores de A irá convergir?

Questão 06: Para a matriz $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ aplique o método da potência com $x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Para qual autovetor os vetores x_k estão se aproximando? E para qual autovalor? Divida x_k por $||x_k||$.

Questão 07: Aplicando o algoritmo QR, calcule os autovalores da matriz:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 9 & 2 \\ 0 & -1 & 2 \end{bmatrix}$$

Questão 08: Quais são os valores singulares, diferentes de zero e em ordem decrescente, de $A - A_k$?

Questão 09: Encontre a melhor aproximação de rank-1 da matriz:

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Questão 10: Quais matrizes de rank igual a 3 possuem $||A - A_1||_2 = ||A - A_2||_2$?

Questão 11: Por que as matrizes A e A^+ possuem o mesmo rank? Se A é uma matriz quadrada, A e A^+ possuem os mesmos autovalores? Caso não possuam, quais são os autovalores de A^+?

Questão 12: Suponha que a matriz A tenha colunas independentes ($r(A) = r = n$ e o espaço nula de A possui somente o vetor nulo). Descreva a matriz $\Sigma^{m \times n}$ de $A = U\Sigma V^T$.

Questão 13: Mostre que A^T tem os mesmos valores singulares que a matriz A (diferentes de zero).

Questão 14: Quais são os valores singulares de AA^TA?

Capítulo A

Resolução dos Exercícios Propostos

A.1 Capítulo 2

Questão 01: Considere que B seja uma matriz 4×4 sobre a qual aplicamos as seguintes operações:

Solução: As operações serão representadas por pré (E) e pós (D) multiplicações de B por matrizes de dimensões compatíveis, que representem tais operações. Considerando as visões de alto nível para multiplicações de matrizes, se o objetivo é a modificação das colunas da matriz B, a matriz definida está à direita de B. Para modificações das linhas de B, a matriz definida está à esquerda de B.

1. dobrar os valores da coluna 1: $BD_1 = B \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

2. dividir os valores da linha 3: $E_1B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$B.

3. adicionar linha 3 à linha 1: $E_2B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$B.

4. trocar as linhas 1 e 4: $E_3B = PB = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$B

5. subtrair a linha 2 de cada uma das outras linhas: $E_4B = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$ B

6. substituir a coluna 4 pela coluna 3: $BD_2 = B \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

7. eliminar a coluna 1, de forma que a dimensão da matriz resultante seja uma coluna a menos. D_3 possui dimensão $4x3$
$BD_3 = B \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Questão 02: Considere a matriz em blocos $\begin{bmatrix} A & I \\ I & C \end{bmatrix}$, onde I é uma matriz identidade e A possui dimensões $p \times q$. Quais as dimensões de C ?

Solução: A partição de matrizes em blocos deve ser realizada de tal maneira que blocos vizinhos (considerando as linhas e colunas) possuam dimensões compatíveis. Como a matriz $A \in \mathbb{R}^{p \times q}$, a matriz I à direita de A possui dimensão $p \times p$, já que deve ter o total de linhas compatível com a matriz A. Por outro lado, a matriz I abaixo de A possui dimensão $q \times q$, já que deve ter o total de colunas compatível com a matriz A. A partir das dimensões dos blocos destas matrizes, C terá q linhas e p colunas.

Questão 03: Considere a matriz em blocos $K = \begin{bmatrix} I & A^T \\ A & 0 \end{bmatrix}$. Quais das seguintes afirmativas são necessariamente verdadeiras (necessariamente verdadeiras significa que são verdadeiras sem nenhuma consideração adicional).

a) K é simétrica.

Solução: Para verificar se a matriz K é simétrica, iremos aplicar a operação de transposição em K. Como a matriz K está dividida em blocos, a operação de transposição poderá ser realizada em cada partição. Assim: $K^T = \begin{bmatrix} I^T & A^T \\ (A^T)^T & 0^T \end{bmatrix} = \begin{bmatrix} I & A^T \\ A & 0 \end{bmatrix} = K$. *Então, K é simétrica.*

b) A é quadrada ou larga (a matriz é larga quando não é alta, isto é, número de colunas maior que o número de linhas).

Solução: Como a matriz K é simétrica, ela é quadrada. Vamos considerar $K \in \mathbb{R}^{n \times n}$. É possível definir a seguinte matriz $K = \begin{bmatrix} I & e_1^T \\ e_1 & 0^T \end{bmatrix}$, com $I \in \mathbb{R}^{1 \times 1}$ e e_1 um vetor coluna de $n-1$ linhas, ou seja, neste contra exemplo, $A = e_1$ é uma matriz alta e fina.

c) A submatriz identidade e a matriz de zeros em K possuem as mesmas dimensões.

Solução: Não, pelo contra exemplo acima.

d) A submatriz de zeros é quadrada.

Solução: Sim. A matriz de zeros possui o mesmo número de linhas de A e de colunas igual às colunas de A^T.

Questão 04: Seja A uma matriz $m \times n$ e considere a matriz *empilhada* $S = \begin{bmatrix} A \\ \hline I \end{bmatrix}$, onde I é a matriz identidade.

a) Quando as colunas de S são linearmente independentes ? *Solução: Sempre, pois há uma submatriz quadrada de mesma ordem de A (mesmo número de colunas de A) que admite inversa, que é I.*

b) Quando as linhas de S são linearmente independentes ? Obs: sua resposta pode depender de m, n ou do fato de A ter ou não linhas ou colunas linearmente independentes.

Solução: Nunca. As linhas de A são combinações lineares de I.

Questão 05: Considere que você necessite avaliar $z = (A + B)(x + y)$ onde A, B são matrizes conformáveis com os vetores x, y. Considere as seguintes alternativas e determine o número de operações de ponto flutuante de cada uma, indicando qual é a mais econômica ao final. Considere que A é $m \times n$ e que x, y são vetores n dimensionais.

a) Primeiro somar $A+B$, então somar $x+y$, e depois aplicar a soma $(A+B)$ na soma $(x+y)$.

Solução: O total de elementos em cada matriz é igual a mn. Assim, a soma $(A+B)$ tem o custo $O(mn)$. Para os vetores, a soma $(x+y)$ tem o custo $O(n)$. Esta primeira etapa tem um custo total da ordem de $mn+n$. A segunda operação $(A+B)(x+y)$, envolve a mutiplicação de uma matriz $C = (A+B)$ por um vetor coluna $v = x+y$, ou seja, temos o cálculo de m produtos escalares de vetores n dimensionais e m somas de $n-1$ termos, resultando em um custo de $m(n-1)$ operações. O custo total é de $3mn + n - m$.

b) Distribuir, avaliar cada termo e então somar: $z = Ax + Ay + Bx + By$.

Solução: Cada termo (mutliplicação de matriz por vetor) possui o custo computacional de $mn + m(n-1)$ *operações. A soma entre os quatro vetores resultantes é realizada com* $3m$ *operações. Custo total é de* $4mn + 4m(n-1) + 3m = 8mn - m$.

Questão 06: Escolha uma única matriz B (3×3) tal que para toda matriz A:

a) BA = 4A.

Solução: Todos os elementos da matriz A devem ser multiplicados por 4. Com B á esquerda, temos que definir uma matriz onde cada linha é multiplicada por 4. Assim $B = 4I$.

b) BA = 4B. *Solução: Como temos que definir a matriz B, neste caso B é a matriz nula.*

c) BA possui as linhas 1 e 3 de A trocadas, preservando a linha 2 .

Solução: B é uma matriz de permutação, obtida através da mudança dos elementos diferentes de zero de uma matriz I. *Ou seja* $B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

Questão 07: Descreva o espaço coluna (em termos de linhas ou planos) das seguintes matrizes:

a) $A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Solução: $C(A) \in \mathbb{R}^3$. *Como* $A_2 = 2A_1$, *uma base para* $C(A)$ *é o vetor* $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ *e a dimensão de* $C(A)$ *é igual a 1. Ou seja,* $C(A)$ *é uma linha que, neste caso, é o eixo x.*

b) $B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}$.

Solução: $C(A) \in \mathbb{R}^3$. *Como as colunas de A são LI, a base para* $C(A)$ *é dada pelos vetores* $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$ *e a dimensão de* $C(A)$ *é igual a 2. Ou seja,* $C(A)$ *é um plano formado pelas coordenadas x,y.*

c) $C = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 0 \end{bmatrix}$.

Solução: $C(A) \in \mathbb{R}^3$. *Como* $A_2 = 0A_1$, *uma base para* $C(A)$ *é o vetor* $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ *e a dimensão de* $C(A)$ *é igual a 1. Ou seja,* $C(A)$ *é uma linha.*

Questão 08: Considere os vetores $v_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ e $v_2 = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$. Responda às questões abaixo:

a) Estes vetores são linearmente independentes?

Solução: Sim. Para obter o vetor nulo 0_3, $\alpha_1 = \alpha_2 = 0$.

b) Eles formam uma base para um espaço $\mathcal{V}$? Qual espaço eles geram?

Solução: Sim. Formam um plano em $\mathbb{R}^3$

c) Qual a dimensão do espaço gerado?

Solução: Dimensão igual a 2, pois a cardinalidade da base é igual a 2.

d) Quais matrizes A possuem $\mathcal{V}$ como espaço coluna?

Solução: Todas as matrizes $A^{3\times n}$, *com rank(A) = 2.*

d) Descreva todos os vetores v_3 que completam a base para $\mathbb{R}^3$.

Solução: Todo vetor $v_3 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, *com* $c \neq 0$ *é um vetor que completa a base para* $\mathbb{R}^3$.

Questão 09: As colunas de A são n vetores pertencentes à $\mathbb{R}^m$. Se estes vetores são linearmente independentes, qual é o rank de A? Se estes vetores geram $\mathbb{R}^m$ qual o rank de A? Se estes vetores que geram $\mathbb{R}^m$ são base para $\mathbb{R}^m$, qual a relação entre m, n e rank de A?

Solução: Rank(A) = n. Para gerar $\mathbb{R}^m$, *rank(A) tem que ser igual a m. Dada as relações descritas, A é uma matriz quadrada, onde* $m = n = rank(A)$

Questão 10: Encontre as bases e as dimensões para cada um dos quatro espaços fundamentais associados às matrizes A e B:

a) $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix}$.

Solução: Temos que $A_2 = 2A_1$ *e* $A_3 = 4A_1$. *Uma base para* $C(A)$ *é o vetor* $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. *Assim, a dimensão de* $C(A)$ *é igual a 1. Sabemos que a dimensão de* $C(A^T)$ *é igual a dimensão de* $C(A)$, *portanto é igual a 1. Uma base para* $C(A^T)$ *é o vetor* $\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$ *(veja que* $a_2^T = 2a_1^T$*). Para a definição das dimensões de* $N(A)$ *e* $N(A^T)$, *temos que verificar em quais espaços* $C(A^T)$ *e* $C(A)$ *estão inseridos.* $C(A) \in \mathbb{R}^2$. *Portanto,* $N(A^T)$ *tem dimensão igual a 1. Como* $C(A^T) \in \mathbb{R}^3$, $N(A)$ *tem dimensão igual a 2. Para obtermos as bases dos subespaços nulos, temos que resolver os sistemas de equações* $Ax = 0$ *e* $A^T y = 0$. *Assim, uma base para* $N(A)$ *é dada pelos vetores* $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ *e* $\begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}$ *e uma base para* $N(A^T)$ *é dada pelo vetor* $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

b) $B = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 5 & 8 \end{bmatrix}$.

Solução: Os mesmos conceitos ressaltados na letra (a) devem ser empregados na solução de (b). Assim, apresentaremos as dimensões de cada subespaço e suas respectivas bases. A dimensão de $C(A)$ *é igual a 2, e uma base é dada pelos vetores* $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ *e* $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$. $N(A^T)$ *possui somente o vetor nulo.* $C(A^T)$ *possui dimensão 2 e uma base é dada pelos vetores* $\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$ *e* $\begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}$. *A dimensão* $N(A)$ *é igual a 1. Após a solução do sistema* $Ax = 0$, *uma base para* $N(A)$ *é dada pelo vetor* $\begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}$.

Questão 11: Se $Ax = b$ tem solução e $A^T y = 0$, $(y^T x = 0)$ ou $(y^T b = 0)$? Justifique.

Solução: Nas condições apresentadas no enunciado, $b \in C(A)$ *e* $y \in N(A^T)$. $C(A)$ *é complemento ortogonal de* $N(A^T)$, *ou seja, seus elementos são perpendiculares. Assim, temos que* $y^T b = 0$.

Questão 12: Suponha que A seja uma matriz simétrica ($A^T = A$). O espaço coluna de A é perpendicular ao espaço nulo de A? Justifique.

Solução: Como A é uma matriz simétrica, $C(A) = C(A^T)$. Assim, $C(A)$ é ortogonal a $N(A)$.

Questão 13: Considere a matriz $A = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$ e o vetor $b = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$. Uma solução para o sistema $Ax = b$ é o vetor $x = [1 \ \ 1 \ \ 2]^T$. Responda: Esta solução é única ? Em caso positivo, justifique. Em caso negativo, justifique e apresente uma solução alternativa.

Solução: 1) ***Considerando o posto da matriz.*** *A matriz é baixa ($r = m$) e larga ($r < n$), onde r é o rank da matriz e m e n seu número de linhas e colunas, respectivamente. Logo, por definição, $Ax = b$ possui infinitas soluções. Uma delas, por exemplo, é o vetor $x = [0 \ \ 0 \ \ 4]^T$.*

2) ***Considerando os subespaços fundamentais da matriz A.*** *Considerando $x \in R^3$ uma possível solução do sistema $Ax = b$, se $x \in C(A^T)$, o sistema poderá admitir uma ou várias soluções. Caso contrário, se $x \notin C(A^T)$ e $b \neq 0$, x é uma combinação linear de vetores pertencentes a $C(A^T)$ e a $N(A)$. No caso analisado, $\nexists y | A^T y = x$, implicando que $x \notin C(A^T)$. Assim, x é uma das possíveis soluções do sistema dado. Para encontrar outras soluções para o sistema, vamos analisar os vetores em $N(A)$. Para achar os vetores $z \in N(A)$, resolvemos $Ax = 0$. Fazendo-o, obtemos $N(A) = \{z \in R^3 | z_1 = z_2, z_3 = -2z_1\}$. Um vetor particular que satisfaz as condições para $N(A)$ é o vetor $z = [1, 1, -2]$. Assim para qualquer $z \in N(A)$, temos que $A(x + z) = Ax + Az = b$*

Questão 14: Uma matriz simétrica A possui os autovalores $3, -3$, com os respectivos autovetores $\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} \end{bmatrix}$ e $\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$. Qual é a matriz A ? Esta matriz é positiva definida, negativa definida ou indefinida ?

Solução: Seja $V = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$ a matriz composta pelos autovetores de A e $\Lambda = \begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix}$ a matriz diagonal composta pelos autovalores de A, temos $A = V\Lambda V^T = \begin{bmatrix} 0 & -3 \\ -3 & 0 \end{bmatrix}$. Como a matriz possui autovalores positivos e negativos, sabemos que a matriz é indefinida. Ou seja, existem $x \in R^2$ tal que $x^T Ax > 0$ e $x^T Ax < 0$. Note que, qualquer $x \in span(v_1)$, $x \neq 0$ satisfaz $x^T Ax > 0$ e qualquer $x \in span(v_2)$ satisfaz $x^T Ax < 0$.

Questão 15: Considere a matriz V formada pelos autovetores da matriz A acima identificada. O que você pode dizer sobre os quatro espaços fundamentais da matriz A ? Isto é, caracterize todos os quatro espaços fundamentais com suas dimensões.

Solução: Como A possui 2 autovalores diferentes, seus autovetores são ortogonais. Logo, $C(A) = span\{v_1, v_2\} = R^2$. Por consequência, $N(A) = \{0\}$. Pela simetria, temos que $C(A^T) = C(A)$ e $N(A^T) = N(A)$.

Questão 16: Suponha que a matriz A das duas questões anteriores tenha o seu autovalor -3 sustituído por 0, preservando os autovetores. O que você pode dizer sobre os quatro espaços fundamentais desta nova matriz A ? Isto é, caracterize todos os quatro espaços fundamentais com suas dimensões. Esta matriz é positiva definida, semi-positiva definida, negativa definida ou semi-negativa definida ?

Solução: A nova matriz $A = V\Lambda V^T = \begin{bmatrix} \frac{3}{2} & -\frac{3}{2} \\ -\frac{3}{2} & \frac{3}{2} \end{bmatrix}$. Note que $a_1 = -1 \times a_2$ e que $v_1 \perp v_2$. Logo, $C(A) = span\{v_1\}$, o rank de A é $r = 1$, e $v_2 \in N(A^T)$. Como a matriz ainda é simétrica, $C(A) = C(A^T)$, e $N(A) = N(A^T)$, e a dimensão do nulo é $m = n - r = 2 - 1 = 1$. Como A não possui autovalores negativos, mas possui um autovalor 0, e os determinantes de suas submatrizes principais são $\frac{3}{2}$ e 0, A é semi-positiva definida.

Questão 17: Responda verdadeiro ou falso e justifique.

1. $\{(x, y) : y = |x|, x \in \mathbb{R}\}$ é um subespaço do $\mathbb{R}^2$.

   Solução: Falso. Considere o seguinte contra-exemplo: $v = (-1, 1)^T$ e $u = (1, 1)^T$. A soma dos vetores $u + v = (0, 2)^T$ resulta em um vetor que não pertence ao conjunto definido.

2. $\{(x, y) : x^2 + y^2 = 0, x, y \in \mathbb{R}\}$ é um subespaço do $\mathbb{R}^2$.
   Verdadeiro. O conjunto acima, dos pares reais x, y tais que $x^2 = -y^2$, se resume a $\{(0, 0)\}$ que define um subespaço pois é fechado à soma e multiplicação por escalar.

3. $\{(x, y) : x^2 - y^2 = 0, x, y \in \mathbb{R}\}$ é um subespaço do $\mathbb{R}^2$.

   Solução:Falso Este conjunto contém os pontos em que $x = y$ ou $x = -y$. Novamente, tome $v = (-1, 1)^T$ e $u = (1, 1)^T$ e veja que $u + v = (0, 2)^T$ não pertence ao conjunto.

4. $\{(x, y) : x - y = 1, x, y \in \mathbb{R}\}$ é um subespaço do $\mathbb{R}^2$.

   Solução: Falso. O vetor nulo não pertence ao conjunto definido. Este conjunto é um *conjunto afim*, que é uma translação de um subespaço vetorial.

Questão 18: Sejam W_1, W_2 dois subespaços de um espaço vetorial V e seja

$$W_1 + W_2 = \{w_1 + w_2 : w_1 \in W_1, w_2 \in W_2\}$$

a soma de W_1 e W_2.

1. Mostre que $W_1 \cap W_2$ e $W_1 + W_2$ são subespaços.

   Solução: Para verificar se $W_1 \cap W_2$ é um subespaço vetorial, temos que: (i) verificar se o vetor nulo pertence a $W_1 \cap W_2$; (ii) verificar se o conjunto é fechado nas operações de multiplicação por escalar - na verdade o caso (i) é um caso particular de (ii), que optamos em deixar explícito; (iii) e soma.
   (i) Como W_1 e W_2 são subespaços vetoriais, o nulo pertence a ambos e, portanto à interseção.
   (ii) Seja $v \in W_1 \cap W_2$. Como W_1 e W_2 são subespaços vetoriais, $\alpha v \in W_1$ e $\alpha v \in W_2$ e, portanto, $\alpha v \in W_1 \cap W_2$.
   (iii) Sejam $v, u \in W_1 \cap W_2$. Como W_1 e W_2 são subespaços vetoriais, $\alpha v, \beta u \in W_1$ e $\alpha v, \beta u \in W_2$, $\alpha v + \beta u \in W_1$ e $\alpha v + \beta u \in W_2$ e, portanto $\alpha v + \beta u \in W_1 \cap W_2$.

   Para $W_1 + W_2$, temos:
   (i) $0 \in W_1$ e $0 \in W_2$, então $0 = 0 + 0$ e $0 \in W_1 + W_2$.
   (ii) Se $v \in W_1 + W_2$, temos que $v = w_1 + w_2 \to \alpha v = \alpha w_1 + \alpha w_2$, para $w_1 \in W_1$ e $w_2 \in W_2$. Como W_1 e W_2 são subespaços vetoriais $\alpha w_1 \in W_1$ e $\alpha w_2 \in W_2$ e, portanto $\alpha v = \alpha w_1 + \alpha w_2 \in W_1 + W_2$.
   (iii) Definir v, u como somas em $W_1 + W_2$, aplicar a multiplicação por escalares, cujos resultados estão em W_1 e W_2, e, portanto serão fatores da soma que define os elementos em $W_1 + W_2$.

2. Mostre que $W_1 \cap W_2 \subseteq W_1 \cup W_2 \subseteq W_1 + W_2$.

   Solução:

   - Primeira relação: $v \in W_1 \cap W_2 \to v \in W_1, v \in W_2$, logo $v \in W_1 \cup W_2$.

- Segunda relação: Tome $v \in W_1 \cup W_2$. Suponha que $v \in W_1$ e $v \notin W_2$. Então $v = v + 0$, onde $0 \in W_2$, portanto $v \in W_1 + W_2$. Agora suponha que $v \in W_1 \cap W_2$. Como W_1, W_2 são subespaços, para qualquer $\alpha \in \mathbb{R}$ temos que $\alpha v \in W_1, (1-\alpha)v \in W_2$, logo $v = \alpha v + (1-\alpha)v \in W_1 + W_2$.

3. $W_1 \cup W_2$ é um subespaço ? Justifique.
   Solução: Não. Como contra-exemplo, considere $W_1 = \{(x,y) \in \mathbb{R}^2 : x = y\} = span\{(1,1)^T\}$ e $W_2 = \{(x,y) \in \mathbb{R}^2 : x = -y\} = span\{(-1,1)^T\}$, que nada mais é que um particionamento do conjunto do exercício 19 desta lista. Claramente W_1, W_2 são subespaços mas $W_1 \cup W_2 = \{(x,y) : x^2 - y^2 = 0, x, y \in \mathbb{R}\}$ não é um subespaço.

4. Quando $W_1 \cup W_2$ é um subespaço ?
   Solução: Quando $W_1 \subseteq W_2$ ou $W_2 \subseteq W_1$. Sempre que existir a diferença simétrica, isto é, $W_1 \setminus W_2 \neq \emptyset$ e $W_2 \setminus W_1 \neq \emptyset$ podemos tomar $v \in W_1 \setminus W_2$ e $u \in W_2 \setminus W_1$ e a soma $u + v$ não pertence a $W_1 \cup W_2$.

5. Qual o menor subespaço de V contendo $W_1 \cup W_2$?
   Solução: $W_1 + W_2$. Pela questão 4 acima, para ser subespaço precisa ser pelo menos $W_1 + W_2$. Se o conjunto então envolver $v \notin W_1 + W_2$ e for subespaço, não é o menor subespaç contendo $W_1 \cup W_2$.

Questão 19: Sejam W_1 e W_2 subespaços vetoriais gerados respectivamente pelos v's e u's abaixo indicados.

- $v^1 = (1,2,-1,-2)^T$, $v^2 = (3,1,1,1)^T$ e $v^3 = (-1,0,1,-1)^T$
- $u^1 = (2,5,-6,-5)^T$, $u^2 = (-1,2,-7,3)^T$.

Encontre as dimensões e bases para $W_1 \cap W_2$ e $W_1 + W_2$.

Resolução: Precisamos investigar o sistema linear

$$\alpha_1 v^1 + \alpha_2 v^2 + \alpha_3 v^3 = \beta_1 u^1 + \beta_2 u^2.$$

Seja $W_1 = span(v^1, v^2, v^3)$ e $W_2 = span(u^1, u^2)$. As dimensões de $W_1, W_2, W_1 \cap W_2$ e $W_1 + W_2$ são $3, 2, 1, 4$, respectivamente. Basta calcular o posto das matrizes abaixo. u^1 é uma base para $W_1 \cap W_2$, $\{v^1, v^2, v^3, u^2\}$ formam uma base para $W_1 + W_2$.

```
-->VT = [1 2 -1 -2;3 1 1 1;-1 0 1 -1]
 VT  =

    1.    2.  -1.  -2.
```

```
   3.    1.    1.    1.
  -1.    0.    1.   -1.

-->V = VT'
 V  =

   1.    3.   -1.
   2.    1.    0.
  -1.    1.    1.
  -2.    1.   -1.

-->UT = [2 5 -6 -5;-1 2 -7 3]
 UT  =

   2.    5.   -6.   -5.
  -1.    2.   -7.   3.

-->U = UT'
 U  =

   2.   -1.
   5.    2.
  -6.   -7.
  -5.   3.

-->X = [V,U]
 X  =

   1.    3.   -1.    2.   -1.
   2.    1.    0.    5.    2.
  -1.    1.    1.   -6.   -7.
  -2.    1.   -1.   -5.    3.

-->rank(X)
 ans  =

   4.
```

```
-->rank(X(:,1:3))
 ans  =

   3.

-->rank(X(:,4:5))
 ans  =

   2.
```

Observação Suponha que $U_1, \ldots, U_m$ sejam subespaços de um espaço vetorial V. Cada elemento de $U_1 + U_2 + \cdots + U_m$ pode ser escrito como $u_1 + u_2 + \cdots + u_m$, onde $u_j \in U_j$. Estamos particularmente interessados em casos em que cada vetor em $U_1 + U_2 + \cdots + U_m$ pode ser representado na forma acima, de <u>uma única forma</u> (os u_j's são únicos). Neste caso, dizemos que o vetor é a soma direta destes m subespaços.
Definição: Suponha que $U_1, \ldots, U_m$ sejam subespaços de V. A soma $U_1 + U_2 + \cdots + U_m$ é chamada de soma direta, se cada elemento u de $U_1 + U_2 + \cdots + U_m$ puder ser escrito de uma única forma $u_1 + u_2 + \cdots + u_m$, onde cada $u_j \in U_j$. Se $U_1 + U_2 + \cdots + U_m$ é uma soma direta, representamos como $U_1 \oplus U_2 \oplus \cdots \oplus U_m$.

Alguns resultados adicionais:

1. $U + U^{\perp}$ formam uma soma direta de V, se U é subespaço de V.

2. Se U, W são subespaços de V, então $U + W$ é uma soma direta se e somente se $U \cap W = \{0\}$.

3. Se $U_1, U_2, \ldots, U_m$ são subespaços de V então $U_1 + U_2 + \cdots + U_m$ é uma soma direta se e somente se a única forma de escrevermos o vetor 0 (zero) como uma soma de $u_1 + u_2 + \cdots + u_m$ é tomando cada um dos u_j's como o próprio vetor 0.

Questão 20: Responda se a soma dos U's abaixo formam somas diretas.

1. $U_1 = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}$, $U_2 = \{(0, 0, z) \in \mathbb{R}^3 : z \in \mathbb{R}\}$.

   Solução: Sim. $U_1 \cap U_2 = \{0\}$

2. $U_1 = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}$, $U_2 = \{(0, 0, z) \in \mathbb{R}^3 : z \in \mathbb{R}\}$, $U_3 = \{(0, y, y) \in \mathbb{R}^3 : y \in \mathbb{R}\}$.

Solução: Não. Veja que podemos escrever qualquer vetor $v = (\alpha, \gamma, \beta)^T \in U_1 + U_2 + U_3$ de pelo menos duas formas distintas:

- $v = (\alpha, \gamma, 0)^T + (0, 0, \beta)^T + (0, 0, 0)^T$
- $v = (\alpha, \gamma - \beta, 0)^T + (0, 0, 0)^T + (0, \beta, \beta)^T$

Questão 21 Para $k \geq 2$ calcule A^k para:

1. $A = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}$ Esta matriz possui autovalores distintos $\lambda_1 = 1$ e $\lambda_2 = 4$ e seus autovetores, respectivamente, $x^1 = (-1, 1)^T$ e $x^2 = (1, 2)^T$, geram o $\mathbb{R}^2$. Então, vamos seguir os seguintes passos:

   - Passo 1: Vamos escrever as colunas de A, A_1, A_2, em função de x^1, x^2. $A_1 = -\frac{2}{3}x^1 + \frac{4}{3}x^2$ e $A_2 = \frac{1}{3}x^1 + \frac{4}{3}x^2$.
   - Passo 2: Vamos escrever $A^k = A^{k-1}[A_1, A_2]$, em função de λ_1, λ_2 e x^1, x^2. $A^{k-1}A_1 = -\frac{2}{3}\lambda_1^{k-1}x^1 + \frac{4}{3}\lambda_2^{k-1}x^2 = \frac{1}{3}\begin{bmatrix} 2 + 2^{2k} \\ -2 + 2^{2k+1} \end{bmatrix}$
     $A^{k-1}A_2 = \frac{1}{3}\lambda_1^{k-1}x^1 + \frac{4}{3}\lambda_2^{k-1}x^2 = \frac{1}{3}\begin{bmatrix} -1 + 2^{2k} \\ 1 + 2^{2k+1} \end{bmatrix}$.

     Logo, temos $A^k = \frac{1}{3}\begin{bmatrix} 2 + 2^{2k} & -1 + 2^{2k} \\ -2 + 2^{2k+1} & 1 + 2^{2k+1} \end{bmatrix}$

2. $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

   Solução:: $A = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}$, facilmente demonstrável por indução.

Questão 22: Responda verdadeiro ou falso e justifique.

1. $A^k = 0$ para todo inteiro positivo $k \geq 2$, então $A = 0$.

   Solução: Falso. Contra-exemplo: $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$

2. $A^k = 0$ para algum inteiro positivo k, então $\sum_i a_{ii} = 0$.

   Solução: Verdadeiro. Uma matriz A tal que $A^k = 0$ para algum inteiro k é chamada *nilpotente*. Veja que qualquer potência de A possui os mesmos autovetores de A e, como autovalores, $\lambda(A)_i^k$, onde $\lambda(A)_i : i = 1, \ldots, n$ são os n autovalores de A. Então temos que para algum x (autovetor não nulo de A):

$0 = A^k x = \lambda_i(A)^k x$. Portanto, todos os autovalores de A^k são todos nulos, o polinômio característico de A^k e de A é λ^k. Em resumo, os autovalores de A são todos nulos. Logo, $tr(A) = \sum_{i=1}^{n} \lambda_i(A) = 0$.

3. Se $\sum_i a_{ii} = 0$, então $|A| = 0$ (determinante de A é zero).

   Solução: Falso. Contra exemplo: considere a matriz $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ cujo determinante é -1 e autovalores são $1, -1$.

4. Se A, B são similares, $det(A) = det(B)$.

   Solução: Verdadeiro, pois os seus autovalores são iguais e temos que a multiplicação dos autovalores éigual ao determinante da matriz.

5. Se A, B são similares, então as duas matrizes possuem os mesmos autovalores.

   Solução: Verdadeiro, veja prova no slide 63 (fundamentos de álgebra linear).

6. Se A, B possuem os mesmos autovalores, então são similares.

   Solução: Falso. Duas matrizes podem ter o mesmo conjunto de autovalores, mas com diferentes multiplicidades e portanto são diferentes. Exemplo: I_2 e I_3. Recorde-se de que para que duas matrizes sejam similares, deve existir uma matriz C, que possua inversa, tal que $A = CBC^{-1}$. Não há uma transformação similar para I_2, I_3 pois não existe uma C quadrada que faça a transformação similar.

7. Se A, B possuem o mesmo polinômio característico, então possuem os mesmos autovalores.

   Solução: Verdadeiro. As raízes do polinômio característico são os autovalores destas matrizes.

8. Se A, B possuem os mesmos autovalores, então possuem o mesmo polinômio característico.

   Solução: Falso. A multiplicidade algébrica de um mesmo autovalor pode ser diferente, resultando em polinômios diferentes.

9. $diag\{1, 2, \ldots, n\}$ é similar a $diag\{n, n-1, \ldots, 1\}$ (se verdadeira, encontre a matriz B e sua inversa que garantem a similaridade).

   Verdadeiro. A matriz de permutação P associada a $\pi = (n, n-1, n-2, \ldots, 1)$. A transformação é $P(diag\{1, 2, \ldots, n\})P^T = diag\{n, n-1, \ldots, 1\}$. Recorde-se de que a inversa de uma matriz de permutação é sua transposta.

10. Se A possui autovalores repetidos, A é não diagonalizável.

    Solução: Falso, pois depende da multiplicidade geométrica dos seus autovetores.

11. Se A é unitariamente diagonalizável, então A é normal.

    Solução: Verdadeira. A matriz A éuma matriz normal se e somente se $AA^T = A^TA$. Portanto, seus autovalores são ortogonais e a matriz éunitariamente diagonalizável. Uma demonstração elegante será apresentada quando estudarmos fatoração SVD.

12. Se A possui r autovalores não nulos, então $rank(A) \geq r$.

    Verdadeiro. O número de autovalores não nulos, sem contar a multiplicidade algébrica de cada um no polinômio característico, fornece um limite inferior válido para a dimensão do espaço coluna de A. Se A for não defectiva, por exemplo, o número de autovalores não nulos, contando suas multiplicidades, será a dimensão de $C(A)$ e, portanto o seu $rank$. Nesse caso, a dimensão do auto-espaço associado a cada autovalor é exatamente a multiplicidade algébrica de cada autovalore e $rank(A)$ = número de autovalores não nulos, contando suas multiplicidades.

Questão 23: Por que a matriz identidade I é a única matriz simétrica positiva definida com $\lambda_{min} = \lambda_{max} = 1$? Quais matrizes A são perfeitamente condicionadas, ou seja, $\kappa(A) = 1$? Importante: A matriz identidade é a matriz que possui o menor valor de $\kappa(A)$ possível.

Solução: Se $\lambda_{min} = \lambda_{max} = 1$, temos que a matriz diagonal Λ com autovalores de A é a matriz identidade. Como estamos assumindo que a matriz A é simétrica positiva, ela pode ser fatorada como a multiplicação dos seus autovalores (ortonormais) e seus autovalores (matriz diagonal). Assim:

$$A = QIQ^T = QQ^T = I \ .$$

Toda matriz A, tal que $A^TA = I$, terá $\kappa(A) = 1$. Se A é uma matriz ortogonal, para o cálculo de κ, temos:

$$\kappa(A) = \tfrac{\sigma_{max}(Q^TQ)}{\sigma_{min}(Q^TQ)} = \tfrac{\sqrt{\lambda_{max}}}{\sqrt{\lambda_{min}}} = 1.$$

Questão 24: Mostre que A e A^{-1} possuem o mesmo número de condição.

Solução:

$$\kappa(A) = ||A^{-1}||||(A^{-1})^{-1}| = ||A^{-1}||||A|| = \kappa(A) \ .$$

Questão 25: Matrizes ortogonais possuem norma $||Q||_2 = 1$. Se a matriz A pode ser fatorada como $A = QR$, mostre que $||A|| \leq ||R||$ e $||R|| \leq ||A||$. O que podemos concluir?

Solução: Pela propriedade $||AB|| \leq ||A||||B||$*, temos:*

1. $||A|| \leq ||Q||||R|| \leq ||R||$.
2. $||R|| \leq ||Q^T||||A|| \leq ||A||$

Podemos concluir que $||A|| = ||R||$.

A.2 Capítulo 3

Questão 01: Qual a matriz M que transforma A em uma matriz triangular superior U ($MA = U$)? Multiplique por $M^{-1} = L$ para fatorar $A = LU$.

Solução: Para esta matriz A, será necessário somente a definição do multiplicador m_{31} e uma única operação entre as linhas 1 e 3 da matriz A para encontrar a matriz U. Assim, temos a definição da matriz $M_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$.

Questão 02: Quais são as duas matrizes de multiplicação M_1 e M_2 que transformam a matriz A que uma matriz triangular superior U ($M_2M_1A = U$)? Multiplique a matriz U pelas inversas de $M_1 e M_2$ para fatorar A em $A = LU$.

Solução: Para obter a matriz U, a partir da matriz A, teremos que fazer a eliminação considerando as colunas 1 e 2 de A. Desta forma, para obter a matriz L serão necessárias duas matrizes M_1 e M_2 e a definição dos multiplicadores m_{21}, m_{31}, m_{32}:
$M_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$, $M_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$.

Questão 03: Defina as matrizes L e U para a matriz simétrica A. Quais são as condições em a, b, c, d que definem os pivôs na diagonal da matriz U para que A seja fatorada em LU?

Solução: $U = \begin{bmatrix} a & a & a & a \\ 0 & b-a & b-a & b-a \\ 0 & 0 & c-b & c-b \\ 0 & 0 & 0 & d-c \end{bmatrix}$. $L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$. Para que a matriz seja decomposta em LU com quatro pivôs definidos, $a \neq 0, b \neq a, c \neq b, d \neq c$.

Questão 04: Considere as matrizes L, U e o vetor b. Resolva $Lc = b$. Então encontre a solução de $Ux = c$. Encontre a matriz A, do sistema original $Ax = b$.

Solução: $c = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$, $x = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, $A = LU = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}$.

Questão 05: Uma das aplicações da solução de sistemas lineares é no cálculo da inversa da matriz A. Considere a matriz $A = \begin{bmatrix} 1 & -3 & 2 \\ -2 & 8 & -1 \\ 4 & -6 & 5 \end{bmatrix}$ e a sua fatoração em $PA = LU$. Encontre a primeira coluna da matriz V que é a inversa de A através da solução de um sistema linear, usando explicitamente os fatores P, L, U. Lembre-se que $AA^{-1} = I$.

Solução: A questão 09 da lista mostra o algoritmo para obtenção de cada uma das colunas da inversa $V = X$, com $B = I$. Para a matriz dada, temos: $V_1 = \begin{bmatrix} -1.41 \\ -0.25 \\ 0.83 \end{bmatrix}$.

Questão 06: Utilize a Decomposição de Cholesky (baixa abstração) para determinar se a matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 10 \\ 3 & 10 & 16 \end{bmatrix}$ é positiva definida.

Solução: A matriz A não é positiva definida. Elemento $l_{33} = \sqrt{-9}$.

Questão 07: Utilize a Decomposição outer Cholesky para fatorar a matriz: $A = \begin{bmatrix} 4 & -2 & 4 & 2 \\ -2 & 10 & -2 & -7 \\ 4 & -2 & 8 & 4 \\ 2 & -7 & 4 & 7 \end{bmatrix}$.

Solução: $L = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -1 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & -2 & 1 & 1 \end{bmatrix}$, $L^T = \begin{bmatrix} 2 & -1 & 2 & 1 \\ 0 & 3 & 0 & -2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

Questão 08: Resolva o sistema $Ax = b$ com: $A = \begin{bmatrix} 0 & 4 & 1 \\ 1 & 1 & 3 \\ 2 & -2 & 1 \end{bmatrix}$ e $b = \begin{bmatrix} 9 \\ 6 \\ -1 \end{bmatrix}$, utilizando a fatoração $PA = LU$.

Solução: $L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & 1/2 & 1 \end{bmatrix}$, $U = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 4 & 1 \\ 0 & 0 & 2 \end{bmatrix}$, $x = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$.

Questão 09: Suponha que $A \in \mathbb{R}^{n \times n}$ seja não singular e $B \in \mathbb{R}^{n \times p}$. Considere o problema de encontrar a matriz $X \in \mathbb{R}^{n \times p}$ tal que $AX = B$. Construa um algoritmo que encontre X em não mais que $O(\max\{pn^2, n^3\})$ operações aritméticas de ponto flutuante.

Solução: Assuma que X_k denote a $k-$ésima coluna de X. Idem para B_k em relação a B.

1. *Fatore $PA = LU$ - custo $O(n^3)$.*

2. *Para $k = 1, \ldots, p$ faça:*

   (a) *Resolva $Ly = PB_k$ e obtenha y - custo $O(n^2)$.*

   (b) *Resolva $AX_k = y$ - Custo $O(n^2)$.*

Questão 10: Deseja-se resolver o sistema linear $A^k x = b$ sem computar a matriz A^k (k é um inteiro qualquer). Sabe-se que a matriz A é não singular. Construa um algoritmo que resolva este sistema linear sem explicitamente avaliar A^k.

Solução:

$$\begin{aligned} A(A \ldots AA)x &= b \\ (A \ldots AA)x &= z \\ \text{Resolva } Az &= b \\ \text{Faça } b &\leftarrow z \\ \Rightarrow \text{repita a ideia k vezes} \end{aligned}$$

Então implemente o seguinte algoritmo:

1. Fatore $PA = LU$.

2. Para $i = 1, \ldots, k$

   (a) Resolva $Ax = b$ isto é:

      i. Resolva $Ly = Pb$

      ii. Resolva $Ux = y$

   (b) Atualize $b \leftarrow x$

Ao final do algoritmo você dispõe da solução x do sistema $A^k x = b$.

Implementação recursiva - Chame o procedimento abaixo `x = Resolve(P,L,U,k,b)`

1. `Resolve(P,L,U,k,b)`

   Se $k = 0$, retorne b.

Caso contrário:

Resolva $Ly = Pb$

Resolva $Ux = y$

```
Resolve(P,L,U,k-1,x)
```

Questão 11: Suponha que dispomos de $A \in \mathbb{R}^{n \times n}$, $d \in \mathbb{R}^n$, $c \in \mathbb{R}^n$ e que desejemos encontrar $s = c^T A^{-1} d$. Uma abordagem seria computar A^{-1} conforme o exercício 1 acima sugere e depois calcular $s = cXd$. Entretanto, há uma forma mais econômica de se proceder. Identifique esta forma mais econômica.

Solução: Não é necessário avaliar a inversa explicitamente. Chame $A^{-1}d$ de z. Então $Az = d$. Fatore a matriz A obtendo $PA = LU$. Resolva o sistema linear $Az = d$ via usando os fatores de $PA = LU$, obtendo z. Calcule $c^T z$.

A.3 Capítulo 4

Questão 01: Encontre a matriz de projeção P_C no espaço coluna da matriz e a matriz de projeção P_R no espaço linha de A. O que podemos dizer da matriz $B = P_C A P_R$?

$$A = \begin{bmatrix} 3 & 6 & 6 \\ 4 & 8 & 8 \end{bmatrix}.$$

Solução: Como $r(A) = 1$, a projeção no espaço coluna da matriz é a projeção no vetor coluna A_1. As colunas A_2, A_3 são combinações lineares de A_1. Temos:

$$P_C = \tfrac{A_1 A_1^T}{A_1^T A_1} = \tfrac{1}{25} \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}.$$

O espaço linha da matriz (A) também possui dimensão igual a 1. A projeção deverá ser feita em uma linha. Para definir a base do espaço linha, podemos fatorar A como:

$$A = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \end{bmatrix}.$$

A projeção será feita no vetor $[1\,2\,2]$, que é uma das bases do espaço linha de A:

$$P_R = \tfrac{1}{9} \begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix}.$$

A matriz $B = P_C A P_R = A$, ou seja, colunas de A são projetadas nas próprias colunas de A. $AP_R = P_R^T A^T = A^T$, ou seja, linhas de A são projetadas nas próprias linhas de A.

Questão 02: Considere vetor b e o vetor p que é a combinação de $A_1, \cdots, A_n$ pertencentes à $\mathbb{R}^m$. Como podemos verificar se p é uma projeção de b no subespaço gerado pelos vetores de A_i?

Solução: Verificar se o vetor $e = b - p \in N(A^T)$ é perpendicular a todos os vetores $A_1, \cdots, A_n \in C(A)$. A matriz P é o projetor ortogonal de b em $C(A)$. A matriz $I - P$ projeta b no espaço perpendicular à $C(A)$, chamada de complemento ortogonal de P.

Questão 03: Considere o vetor b. Suponha que P_1 seja a matriz de projeção no subespaço $\mathbb{R}^1$ gerado pela primeira coluna da matriz $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$. Suponha que P_2 seja a matriz de projeção no espaço coluna de A. Qual é o resultado do produto P_2P_1 ?

Solução: Primeiramente, vamos projetar o vetor b em A_1, utilizando a matriz de projeção P_1 definida por:

$$P_1 = \frac{A_1A_1^T}{A_1^TA_1} = \frac{1}{5}\begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

O vetor $p = P_1b \in C(A)$. A projeção dada pela matriz P_2 irá projetar um vetor que já pertence ao espaço coluna de A. Assim, $P_2P_1 = P_1$ (propriedade de matriz idempotente $P^2 = P$. Uma matriz para ser uma matriz de projeção deve ter esta propriedade).

Questão 04: Se A é uma matriz quadrada e inversível, qual é matriz de projeção P no espaço gerado pelas colunas de A?

Solução:

$$P = A(A^TA)^{-1}A^T = AA^{-1}(A^T)^{-1}A^T = I$$

Se A possui matriz inversa, o espaço coluna de A gera todo o $\mathbb{R}^n$. Ou seja, a projeção de qualquer vetor b em $C(A)$ é o próprio vetor b ($Pb = Ib = b$). O mesmo ocorre para matrizes A que são ortogonais.

Questão 05: Seja E uma matriz $m \times m$, com $Ex = \frac{x+Fx}{2}$ onde F é uma matriz $m \times m$ que transforma $[x_1, \cdots, x_m]$ em $[x_m, \cdots, x_1]$. A matriz E é um projetor ortogonal, um projetor oblíquo ou não é um projetor?

Solução: Para ser uma matriz de projeção, $E^2 = E$ e para ser um projetor ortogonal, $E = E^T$. Iremos verificar se E segue ou não estas propriedades.

$$Ex = \frac{x+Fx}{2} = \left(\frac{I+F}{2}\right)x \rightarrow E = \left(\frac{I+F}{2}\right)$$

Temos que E^2 é igual a:

$$\left(\frac{I+F}{2}\right)^2 = \frac{1}{4}(I^2 + 2F + F^2)$$

Para E ser um matriz projetora, $E^2 = E$. Se mostrarmos que $(I^2 + 2F + F^2)$ $= 2I{+}2F$, teremos $E^2 = E$. Para tal, devemos verificar se $F^2 = FF = I$.

Como $F \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} x_m \\ x_{m-1} \\ \vdots \\ x_1 \end{bmatrix}$, $F = \begin{bmatrix} e_m^T \\ e_{m-1}^T \\ \vdots \\ e_1^T \end{bmatrix}$, *e* $FF = I$.

Portanto, E é um projetor.

Por fim, vamos verificar se o projetor é ortogonal. Para tal, $E = E^T$.

$$E = \left(\frac{I+F}{2}\right) = \frac{1}{2}\begin{bmatrix} 1 & 0 & \cdots & 1 \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & 0 & \cdots & 1 \end{bmatrix} = E^T.$$

Questão 06: Se P é um projetor ortogonal, então $I - 2P$ é uma matriz unitária.

Solução: Considerando o conjunto dos números reais, um matriz U é uma matriz unitária se:

$$U^T U = UU^T = I.$$

Esta propriedade se verifica para matrizes U onde $U^T = U^{-1}$. Por exemplo, matrizes ortogonais são matrizes unitárias. Vamos verificar se $(I - 2P)$ é unitária:

$$(I - 2P)^T (I - 2P) = (I - 2P)^2 = I^2 - 4P + 4P^2 = I,$$

dado que a matriz P é um projetor e a propriedade $P^2 = P$ se aplica.

- Matrizes unitárias preservam norma Euclideana: então $\|(I - 2P)b\|_2 = \|b\|_2$.
- $(I - 2P)b = b - Pb - Pb$ significa um primeiro deslocamento de b até $b - Pb$ e um segundo até $I - 2Pb$. O erro de projetar em $C(A)$ é $b - Pb$ ou seja, este segmento é ortogonal ao $C(A)$. Daquele ponto nos deslocamos $-Pb$ acionais e temos uma reflexão.

Questão 07: Suponha que as colunas de A não sejam independentes. Como podemos definir uma matriz B tal que $P = B(B^T B)^{-1} B^T$ seja a matriz de projeção no espaço coluna de A?

Solução: A matriz B deve ser definida a partir das colunas independentes de A. Assim, $B^T B$ será inversível.

Questão 08: Considere um conjunto de valores t_i, deslocado da média $\hat{t} = (t_1 + \cdots + t_m)/m$ para obter $T_i = t_i - \hat{t}$, sabendo que $\sum T_i = 0$. A partir desta transformação, qual a estrutura da nova matriz A? Qual a relação entre as novas colunas de A, que representa a equação de ajuste da equação $C + DT$? Quais os valores dos parâmetros C e D?

Solução: A transformação torna a base do espaço coluna de A em uma base ortogonal, $A^T A$ é uma matriz diagonal, com entradas $T_1^2 + \cdots + T_m^2$. $A^T b$, possui entradas $b_1 + \cdots + b_m$ e $T_1 b_1 + \cdots + T_m b_m$. $C = \frac{b_1 + \cdots + b_m}{m}$ e $D = \frac{b_1 T_1 + \cdots + b_m T_m}{T_1^2 + \cdots + T_m^2}$.

A.4 Capítulo 5

Questão 01: Aplicando o algoritmo de Gram-Schmidt, encontre a fatoração completa q_1, q_2, q_3 da matriz $A = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{bmatrix}$ tal que q_1, q_2 sejam uma base para $C(A)$
.

Resposta: $q_1 = 1/3[1, 2, -2]^T, q_2 = 1/3[2, 1, 2]^T, q_3 = 1/3[2, -2, -1]^T$

Questão 02: Aplicando o algoritmo de Gram-Schmidt, encontre uma base ortonormal para o espaço coluna da matriz $A = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}$. Calcule a projeção do vetor $b = \begin{bmatrix} -4 \\ -3 \\ 3 \\ 0 \end{bmatrix}$ neste subespaço.

Resposta: $q_1 = 1/2[1, 1, 1, 1]^T, q_2 = 1/\sqrt{52}[-5, -1, 1, 5]^T, p = 1/2[-7, -3, -1, 3]^T$

Questão 03: Aplicando o algoritmo de Gram-Schmidt, encontre uma base ortonormal para o espaço coluna da matriz $A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix}$. Escreva $A = QR$.

Resposta: $q_1 = [1, 0, 0]^T, q_2 = [0, 0, 1]^T, q_3 = [0, 1, 0]^T$

Questão 04: Se Q tem colunas ortonormais, qual é a solução $\hat{x}$ para o ajuste linear $Qx = b$? *Resposta:* $Q^TQ\hat{x} = Q^Tb \rightarrow \hat{x} = Q^Tb$

Questão 05: Calcule a matriz de projeção $P = QQ^T$ quando $q_1 = \begin{bmatrix} 0.8 \\ 0.6 \\ 0 \end{bmatrix}$ e $q_2 = \begin{bmatrix} -0.6 \\ 0.8 \\ 0 \end{bmatrix}$.

$$\textit{Resposta: } Q = \begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \\ 0 & 0 \end{bmatrix}, P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Questão 06: Se A é uma matriz $m \times n$, com $r(A) = n$ e após a sua fatoração em QR é produzida uma matriz $Q = [Q_1\, Q_2]$ quadrada de ordem m e uma matriz $R = [R\, 0]^T$ $m \times n$, com 0 uma matriz nula, responda:

1. As n colunas de Q_1 formam uma base ortonormal para qual subespaço fundamental?

2. As $m - n$ colunas de Q_2 formam uma base ortonormal para qual subespaço fundamental?

3. Como as colunas de Q_2 devem ser obtidas.

Resolução: As colunas de Q_1 formam uma base ortonormal obtida, por exemplo, pela versão clássica do algoritmo de Gram-Schmidt, para o espaço coluna de A. As $m-n$ colunas de Q_2 formam uma base ortonormal para o espaço nulo à esquerda de A (ou espaço nulo de A^T). As colunas de Q formam uma base ortornormal para $\mathbb{R}^m$. As colunas de Q_2 devem ser obtidas por duas etapas. Primeiro, obtém-se uma base B para $N(A^T)$ (já vimos como fazer isso diversas vezes). Na sequência, aplicamos algum algoritmo para fazer a fatoração $B = Q_2R_2$, por exemplo, Gram-Schmidt revisado.

Questão 07: A matrix $P = QQ^T$ é a matriz de projeção no espaço coluna de $Q^{m\times n}$. Agora adicione uma nova coluna a, fazendo $A = [Q\, a]$. A coluna a é substituída por qual nova coluna q, após a aplicação do algoritmo de Gram-Schmidt?

Resolução: O algoritmo de Gram-Schmidt irá calcular a nova coluna a partir da matriz QQ^T ou seja, não será realizada a projeção independente em cada coluna pertencente à Q. Calcule $p = QQ^Ta$, calcule o erro da projeção $e = a - QQ^Ta$ e divida pela sua norma euclidiana. Ou seja, $q_{n+1} = \frac{e}{||e||}$.

Questão 08:

1. Encontre os vetores ortonormais q_1, q_2, q_3 tais que q_1 e q_2 gerem o espaço coluna de $A = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{bmatrix}$.

2. Qual dos 4 espaços fundamentais contém q_3?

3. "Resolva"$Ax = [1\,2\,7]^T$ usando mínimos quadrados (ou ajuste).

Resolução:

1. $q_1 = \frac{1}{3}[1\,2\ -2]^T, q_2 = \frac{1}{3}[2\,1\,2]^T, q_3 = \frac{1}{3}[2\ -2\ -1]^T$. *Para encontrar* q_3*, resolva* $A^T y = [0\,0\,0]^T$*, com a variável livre* $y_3 = -1$.

2. $q_3 \in N(A^T)$.

3. $\hat{x} = (A^T A)^{-1} A^T [1\,2\,7]^T$. $\hat{x} = [1\,2]^T$.

Questão 09: Qual o múltiplo α de $a = [4\,5\,2\,2]^T$ tal que αa é o vetor mais próximo de $b = [1\,2\,0\,0]$? Encontre os vetores ortornormais q_1 e q_2 no plano gerado por a e b.

Resolução:Encontre a projeção de b em a. $p = \frac{2}{7}a$. *Logo* $\alpha = \frac{2}{7}$. $q_1 = \frac{1}{7}[4\,5\,2\,2]$ *e* $q_2 = \frac{1}{7}[-1\,4\ -4\ -4]$.

Questão 10: Considere que $A \in \mathbb{R}^{m\times n}$ possui posto completo igual a n. Considere a fatoração $A = QR$ (reduzida) onde $Q \in \mathbb{R}^{m\times n}$ é ortonormal e $R \in \mathbb{R}^{n\times n}$ é triangular superior, com a diagonal positiva. Mostrar que

1. A fatoração é única

2. A matriz R é o fator triangular superior da fatoração de Cholesky de $A^T A$.

Resolução: $A^T A = (R^T Q^T)(QR) = R^T R$. *Então, de fato R é o fator de Cholesky de* $A^T A$. *Como este fator é único na fatoração de Cholesky, se impusermos* $r_{ii} > 0$*, a fatoração QR de A também é uma vez que* $Q = AR^{-1}$ *e a inversa de R existe e é única.*

Questão 11: A matriz $A \in \mathbb{R}^{m\times n}$ de posto completo n foi fatorada $A = QR$ (reduzida). Deseja-se "resolver"o sistema linear $Ax = b$, isto é, encontrar o ponto $p \in C(A)$ que minimiza $\|p-b\|_2$. Conhecendo-se o vetor $\hat{x}$ que combina as colunas de A e obtém o ponto p, seria possível determinar algum vetor $\hat{y}$ que combina as colunas de Q e leva ao mesmo ponto p ? Em caso positivo, justifique sua resposta e apresente o vetor y. Em caso negativo, indique a razão pela qual não se pode obter tal vetor y.

Resolução: Sim, é possível já que $C(A) = C(Q)$. *Então* $p = A\hat{x} = QR\hat{x} = Q(R\hat{x}) = Q\hat{y}$.

A.5 Capítulo 6

Questão 01: Para as matrizes a seguir, encontre os valores singulares e os vetores singulares à esquerda e à direita: $A_1 \begin{bmatrix} 0 & 4 \\ 0 & 0 \end{bmatrix}, A_2 \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix}$. Verifique: $A_1 = U\Sigma V^T, A_2 = U\Sigma V^T$.

Resposta: $A_1 : \sigma_1 = 4, V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.
$A_2 : \sigma_1 = 4, \sigma_2 = 1, V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Questão 02: O espaço linha de $A = \begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix}$ tem dimensão igual a 1. Encontre v_1 no espaço linha de A e u_1 no espaço coluna. Qual o valor de σ_1? Escreva a matriz A como $A = U\Sigma V^T$.
Resposta: $\sigma_1 = \sqrt{20}, v_1 = 1/\sqrt{2}[1,1]^T, u_1 = 1/\sqrt{10}[1,3]^T$

Questão 03: Encontre bases ortonormais para os quatro subespaços fundamentais da matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$.
Resposta: $C(A^T) : 1/\sqrt{5}[1,2]^T, N(A) : 1/\sqrt{5}[2,-1]^T, C(A) : 1/\sqrt{10}[1,3]^T, N(A^T) : 1/\sqrt{10}[3,-1]^T$

Questão 04: Seja a seguinte matriz construída a partir de dados coletados $X_0 = \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix}$. Encontre a média de cada uma das variáveis e encontre a matriz centralizada $\tilde{X}$. Calcule a matriz de covariaância amostral S e encontre os autovalores λ_1, λ_2. Qual a linha que passa pela origem e que é mais próxima das 5 amostras da matriz X_0?

Resposta: $S = \frac{1}{5}\begin{bmatrix} 10 & 0 \\ 0 & 4 \end{bmatrix}, \lambda_1 = 10/5, \lambda_2 = 4/5$.

Questão 05: Considere a matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$.

- Calcule a matriz A^TA e seus autovalores e autovetores. Encontre os valores singulares de A.
- Calcule a matriz AA^T e seus autovalores e autovetores.
- Verifique que $Av_1 = \sigma_1 u_1$. Fatore a matriz A usando a fatoração SVD reduzida e completa.

- Calcule a pseudoinversa de A.

Resposta: $A^+ = 1/50 \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$

Questão 06: Suponha que a matriz A tenha colunas ortogonais $w_1, w_2, \cdots, w_n$ com normas $\sigma_1, \sigma_2, \cdots, \sigma_n$, respectivamente. Descreva as matrizes U, Σ, V da fatoração $A = U\Sigma V^T$.

Solução: Se a matriz A tem colunas ortogonais $w_1, w_2, \cdots, w_n$ com normas $\sigma_1, \sigma_2, \cdots, \sigma_n$, A^TA será uma matriz diagonal com elementos $\sigma_1^2, \cdots, \sigma_n^2$ e os valores σ_i, para $i = 1, \cdots, n$ serão os valores singulares de A. Os autovetores da matriz A^TA serão as colunas da matriz I, então $V = I$. Cada $u_i = \frac{Av_i}{\sigma_i}$, ou seja, u_i é o vetor unitário $\frac{w_i}{\sigma_i}$. Assim, $A = U\Sigma V^T = (A\Sigma^{-1})(\Sigma)(I)$.

Questão 07: Mostre que se v é um autovetor de A^TA, então Av é um autovetor de AA^T.

Solução: $Como\ (AA^T)A = A(A^TA), temos:$

$$(AA^T)Av = A(A^TA)v = A\lambda v = \lambda Av.$$

Questão 08: Aplicando a fatoração SVD, mostre que as matrizes A^TA e AA^T possuem os mesmos autovalores diferentes de zero.

Solução: Se $A = U\Sigma V^T$, então $A^T = V\Sigma^TU^T$ e $A^TA = V\Sigma^T\Sigma V^T = V\Lambda V^T$, que á a diagonalização da matriz A^TA, com $\Lambda = \Sigma^T\Sigma$ ($\sigma_i^2 = \lambda_i$). Similarmente, $AA^T = U\Sigma\Sigma^TU^T$ é a diagonalização AA^T. Podemos verificar que os autovalores em $\Sigma\Sigma^T$ são os mesmos ($\sigma_i^2 = \lambda_i$).

Questão 09: Suponha que $u_1, \cdots, u_n$ e $v_1, \cdots, v_n$ formam bases ortonormais para $\mathbb{R}^n$. Defina a matriz $A = U\Sigma V^T$ que transforme cada v_j em u_j tal que $Av_1 = u_1, \cdots, Av_n = u_n$.

Solução: $A = UV^T$, já que $\sigma_j = 1, \forall j$, ou seja, $\Sigma = I$.

Questão 10: Suponha que A seja uma matriz simétrica 2×2 com autovetores unitários u_1 e u_2 e autovalores $\lambda_1 = 3$ e $\lambda_2 = -2$. Quais são as matrizes U, Σ, V^T da fatoração $A = U\Sigma V^T$?

Solução: Como $A = A^T$, temos que $\sigma_1^2 = \lambda_1^2$ e $\sigma_2^2 = \lambda_2^2$ [1] e $\sigma_1 = 3$ e $\sigma_2 = 2$.

[1] Veja uma prova em https://rampure.org/resources/data100/notes/eigen-singular.html

Adicionalmente, temos $u_1 = v_1$ *(*$A^T A = AA^T$*) e* $u_2 = -v_2$*, dado que* $\sigma_2 = -\lambda_2$.

A.6 Capítulo 7

Questão 01: Falso ou Verdadeiro ou Não é possível dizer? Justifique. Se os autovalores de uma matriz A são iguais a 2,2,5 então a matriz é: (a) inversível; (b) diagonalizável; (c) não diagonalizável.

Solução: (a) Verdadeiro. Matriz não possui autovalor igual a zero. A é singular se det(A) = 0 = $\lambda_1, \lambda_2, ..., \lambda_n$*, ou seja, A tem pelo menos um autovalor* $\lambda_i = 0$*. (b) Não é possível dizer. Matrizes com todos autovetores diferentes, são diagonalizáveis. No caso de autovalores iguais, temos que verificar se os autovetores são LI. Neste caso, o autovalor* $\lambda = 2$ *pode ter somente uma linha de autovetores. (c) Não é possível dizer. Autovalores repetidos podem ter autovetores independentes.*
IMPORTANTE: A existência da inversa da matriz é caracterizada pelos seus autovalores. A diagonalização da matriz é caracterizada pelos seus autovetores.

Questão 02: Falso ou Verdadeiro ou Não é possível dizer? Justifique: Se os autovetores de uma matriz A são múltiplos do vetor $[1 \ \ 4]^T$, então A: (a) não tem inversa; (b) tem um autovalor repetido; (c) não é diagonalizável.
Solução: (a) Não é possível dizer. Não sabemos se A tem autovalor igual a zero. (b) Verdadeiro. Temos uma linha de autovetores, que ocorre somente quando temos autovalores repetidos. (c) Verdadeiro. Autovetores são LD.

Questão 03: Seja $A^k = X\Lambda^k X^{-1}$. Em quais casos $A^k = X\Lambda^k X^{-1}$ se aproxima de uma matriz nula, ou seja, $A^k \to 0$?
Solução: A^k *se aproxima de uma matriz nula quando seus autovalores possuem valores absolutos menores que 1.*

Questão 04: Qual o valor de b na matriz $A = \begin{bmatrix} 2 & b \\ 1 & 0 \end{bmatrix}$ tal que: (a) $A = Q\Lambda Q^T$ exista? (b) A é não diagonalizável? (c) A é singular?

Solução: (a) Se $b = 1$*, A é simétrica, por tanto* $A = Q\Lambda Q^T$ *sempre existe. (b) Se* $b = -1$*,* $\lambda_1 = \lambda_2 = 1$ *e* X^{-1} *não existe. (c) Se* $b = 0$*, det(A) = 0.*

Questão 05: Se a matriz A é uma matriz ortogonal, quais são as matrizes Q e R da fatoração QR? Neste caso, o algoritmo QR para o cálculo dos autovalores de A irá convergir?

Solução: Se A é ortogonal, $Q = A$ *e* $R = I$*. Assim* $A_1, A_2, A_k = RQ = A$ *e o algoritmo não converge.*

Questão 06: Para a matriz $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ aplique o método da potência com $x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Para qual autovetor os vetores x_k estão se aproximando? E para qual autovalor? Divida x_k por $||x_k||$.

Solução: O método está convergindo para o autovetor $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, *com* $\lambda = 3$. *Divida* x_k *por* $||x_k||$.

Questão 07: Aplicando o algoritmo QR, calcule os autovalores da matriz:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 9 & 2 \\ 0 & -1 & 2 \end{bmatrix}$$

Questão 08: Quais são os valores singulares, diferentes de zero e em ordem decrescente, de $A - A_k$?

Solução: Considerando $r(A) = r$ *e* $r(A_k) = k$, *as matrizes* A *e* A_k *podem ser escritas como soma de matrizes de posto 1:*

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$$
$$A_k = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_k u_k v_k^T$$

A matriz diferença é definida como:

$$A - A_k = \sigma_{k+1} u_{k+1} v_{k+1}^T + + \cdots + \sigma_r u_r v_r^T$$

Assim, temos que os valores singulares da matriz $A - A_k$ *são:* $\sigma_{k+1}, \cdots, \sigma_r$.

Questão 09: Encontre a melhor aproximação de rank-1 da matriz:
$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. *Solução: Como* A *é uma matriz diagonal, seus valores singulares são os elementos da diagonal principal,* $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1$ *e os vetores*

singulares são colunas da matriz I. Assim:

$$A_1 = 3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

Questão 10: Quais matrizes de rank igual a 3 possuem $||A - A_1||_2 = ||A - A_2||_2$?

Solução: Pelo resultado da Questão 08, temos:

$$||A - A_1||_2 = \sigma_2 \text{ e } ||A - A_2||_2 = \sigma_3$$

Assim, toda matriz de rank igual a 3 com $\sigma_2 = \sigma_3$ satisfaz a condição $||A - A_1||_2 = ||A - A_2||_2$.

Questão 11: Por que as matrizes A e A^+ possuem o mesmo rank? Se A é uma matriz quadrada, A e A^+ possuem os mesmos autovalores? Caso não possuam, quais são os autovalores de A^+?

Solução: Considerando a fatoração SVD de $A = U\Sigma V^T$, temos:

$$A^+ = V\Sigma^+U^T = \textstyle\sum_{i=1}^{r} \frac{v_i u_i^T}{\sigma_i}.$$

Cada termo $v_i u_i^T$ do somatório (matriz de rank-1) é multiplicado por um valore singular σ_i da matriz A. Assim, A^+ é soma de r matrizes de rank-1, sendo $r(A) = r(A^+) = r$.

Para os autovalores temos:

$$\begin{aligned} Av_i &= \lambda v_i \\ A^+Av_i &= A^+\lambda v_i \\ A^+Av_i = v_i &= A^+\lambda_i v_i \\ A^+Av_i = v_i &= A^+\lambda_i v_i \\ A^+v_i &= \tfrac{1}{\lambda_i} v_i \end{aligned}$$

A e A^+ possuem os mesmos autovetores e os autovalores de A^+ são o inverso dos autovalores de A.

Questão 12: Suponha que a matriz A tenha colunas independentes ($r(A) = r = n$ e o espaço nula de A possui somente o vetor nulo). Descreva a matriz $\Sigma^{m\times n}$ de $A = U\Sigma V^T$.

Solução: Como a matriz A possui rank completo $r(A) = n$, a matriz Σ possui dois blocos distintos: as primeiras n linhas é uma matriz diagonal, com elementos diferentes de zero na sua diagonal. As últimas $m - n$ linhas possuem somente elementos iguais a zero.

Questão 13: Mostre que A^T tem os mesmos valores singulares que a matriz A (diferentes de zero).

Solução:

$$A = U\Sigma V^T$$
$$A^T = V\Sigma^T U^T$$

Para cada par de vetores singulares v_i, u_i, os valores singulares associados são os valores singulares da matriz A.

Questão 14: Quais são os valores singulares de AA^TA?

Solução:

$$AA^TA = (U\Sigma V^T)(V\Sigma^T U^T)(U\Sigma V^T) = U\Sigma\Sigma^T\Sigma V^T.$$

Os valores singulares da matriz AA^TA são iguais a $\sigma_1^3, \cdots, \sigma_r^3$, para $r(A) = r$.

Capítulo B

Avaliações de Semestres Anteriores

B.1 Prova 1 - 2022.2

Questão 01: [40%] Considere o sistema linear $Ax = b$ representado na forma $[A|b]$, como indicado, e responda: $\left[\begin{array}{cccc|c} 1 & -1 & 0 & 2 & 4 \\ 0 & 2 & 1 & 3 & 1 \\ 4 & -6 & -1 & 5 & 15 \end{array}\right]$

1. (15%) Qual é o posto de A ? Fazendo a eliminação em $[A|b]$ temos: $\left[\begin{array}{cccc|c} 1 & -1 & 0 & 2 & 4 \\ 0 & 2 & 1 & 3 & 1 \\ 0 & -2 & -1 & -3 & -1 \end{array}\right]$

   e, na sequência, $\left[\begin{array}{cccc|c} 1 & -1 & 0 & 2 & 4 \\ 0 & 2 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$.

   Portanto, a terceira linha é combinação linear das duas primeiras. O sistema linear admite solução, pois a eliminação mostrou que $b \in \mathcal{C}(A)$. **Resposta:** $r(A) = 2$.

2. (15%) Quais as dimensões dos subespaços $\mathcal{C}(A), \mathcal{C}(A^T), \mathcal{N}(A), \mathcal{N}(A^T)$? Em função da resposta acima, $r(A) = 2$, temos que as dimensões de $\mathcal{C}(A), \mathcal{C}(A^T), \mathcal{N}(A), \mathcal{N}(A^T)$ são, respectivamente, $r(A) = 2, r(A) = 2, n - r(A) = 4 - 2 = 2$ e $m - r(A) = 3 - 2 = 1$.

3. (35%) Caracterize bases $\{v_i : i = 1, \ldots, k\}$ e $\{u_i : i = 1, \ldots, k\}$ respectivamente para os subespaços $\mathcal{C}(A)$ e $\mathcal{C}(A^T)$, tais que $A = \sum_{i=1}^{k} v_i u_i^T$, indicando o valor correto de k.

Uma fatoração $A = CR$ é: $A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 4 & -6 \end{bmatrix} \begin{bmatrix} 1 & 0 & \frac{1}{2} & \frac{7}{2} \\ 0 & 1 & \frac{1}{2} & \frac{3}{2} \end{bmatrix}$. Assim sendo, $v_1 = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}, v_2 = \begin{bmatrix} -1 \\ 2 \\ -6 \end{bmatrix}, u_1^T = \begin{bmatrix} 1 & 0 & \frac{1}{2} & \frac{7}{2} \end{bmatrix}$,

$u_2^T = \begin{bmatrix} 0 & 1 & \frac{1}{2} & \frac{3}{2} \end{bmatrix}$, onde $k = r(A) = 2$.

4. (35%) Caso $\mathcal{N}(A) \neq 0$ (vetor de zeros), forneça uma base para $\mathcal{N}(A)$, obtida fixando as variáveis livres em -1.

Como a matriz A possui deficência de posto, $\mathcal{N}(A) \neq 0$. A dimensão deste espaço é 2, portanto precisamos de uma base com dois elementos.

Partimos da forma escalonada de A após a eliminação, sem naturalmente considerar a terceria linha que é combinação das duas primeiras.

O ponto central a ser observado é que um vetor $y \in \mathcal{N}(A)$ é ortogonal a $\mathcal{C}(A^T)$.

Logo, resolvemos o sistema linear: $\left[\begin{array}{cccc|c} 1 & -1 & 0 & 2 & 0 \\ 0 & 2 & 1 & 3 & 0 \end{array}\right]$, fixando $y_3 = -1, y_4 = 0$ para um dos vetores da base e $y_3 = 0, y_4 = -1$ para o outro.

Sistema Linear I (em y_1, y_2) $\left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 2 & 1 \end{array}\right]$, cuja solução é $y_2 = \frac{1}{2}, y_1 = \frac{1}{2}$. Logo, um vetor da base é $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -1 & 0 \end{bmatrix}^T$.

Para a outra variável livre, $y_4 = -1$ e $y_3 = 0$, temos o sistema linear II $\left[\begin{array}{cc|c} 1 & -1 & 2 \\ 0 & 2 & 3 \end{array}\right]$, cuja solução é é $y_1 = \frac{3}{2}, y_2 = \frac{7}{2}$. Portanto, o segundo vetor na base é $\begin{bmatrix} \frac{7}{2} & \frac{3}{2} & 0 & -1 \end{bmatrix}^T$.

Questão 02: [30%] Considere a função $b(x) = \frac{1}{\alpha e^{\beta x}}$ e um conjunto de dados a serem ajustados $\{(x_i, b_i) : i = 1, \ldots, n\}$.
(No enunciado original, a função $b(x)$ foi escrita como $b(x) = \frac{1}{\alpha + e^{\beta x}}$). Esta resolução deverá ser entregue no Moodle até dia 28/10/2022, 20:00 horas, valendo 3 pontos extras. Responda:

1. No ajuste linear dos dados acima pela função $b(x)$ escolhida, é calculado um vetor de parâmetros $\hat{x}$ que minimiza a norma Euclideana do erro $r(\hat{x}) = z - A\hat{x}$, para A e z correspondentes ao ajuste. Identifique A e z em função dos dados.
‘

2. O processo de identificar $\hat{x}$ que minimiza $r(\hat{x})$ pode ser entendido como um processo de projeção. Considerando os dados $\{(x_i, b_i) : i = 1, \ldots, n\}$ disponíveis, o quê é projetado e onde é projetado ? Seja preciso em função dos dados.

3. Identifique o sistema linear que permite encontrar $\hat{x}$, isto é, defina claramente a matriz de coeficientes e o termo independente do sistema linear, em função dos dados.

Questão 03: [30%] Considere o conjunto $\mathcal{C}_1 = span\{A_1, A_2\}$ e $\mathcal{C}_2 = span\{B_1, B_2\}$ tais que $A_1^T A_2 = 0$ e $B_1^T B_2 \neq 0$ e responda:

1. Com no máximo 5 linhas de argumentação, explique em que difere projetar b em $\mathcal{C}_1$ de b em $\mathcal{C}_2$. Respostas com mais de 5 linhas não serão consideradas.
O sistema de equações normais relativo à projeção em $\mathcal{C}_1$, $(A^T A)\hat{x} = A^T b$ é um sistema linear diagonal, enquanto que o o sistema linear $(B^T B)\hat{z} = B^T y$ não é, dado que $B_1^T B_2 \neq 0$. Em outras palavras, no caso de $\mathcal{C}_1$, podemos projetar independentemente, primeiro em A_1 e depois em A_2. O mesmo não pode ser dito em relação às colunas B_1, B_2.

2. Considerando os vetores abaixo, calcule as projeções deixando evidente a diferença que identificou na resposta acima. $A_1 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T$, $A_2 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$, $B_1 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T$, $B_2 = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}^T$, $b = \begin{bmatrix} 1 & 2 & -3 \end{bmatrix}^T$.

<u>Projetando em $\mathcal{C}_1$:</u>
$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$, cuja solução é $\hat{x}_1 = 2, \hat{x}_2 = 0$. Então $p = 2A_1 = [2 \quad 0 \quad -2]^T$.

<u>Projetando em $\mathcal{C}_2$:</u>
$\begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} \hat{z}_1 \\ \hat{z}_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$, cuja solução é $\hat{z}_1 = 0, \hat{z}_2 = 2$. Então $p = 2B_2 = [2 \quad 2 \quad -2]^T$.

B.2 Prova 2 - 2022.2

Questão 01: Abaixo são apresentadas duas implementações da fatoração QR via Gram-Schmidt: `QR_GS_1(A) e QR_GS_2(A)`. Recordando: As instruções `A(:,k)`,`A(k,:)` respectivamente denotam a $k-$ ésima coluna e linha da matriz A e o comando `size(A)` retorna o número de linhas e colunas de A, nesta ordem. Por sua vez, a transposta de `A` é representada como

```
A'
```

1. (35%) Qual implementação é a revisada ? Justifique distiguindo-a da clássica. (max 3 linhas).
   `QR_GS_1(A)`. Na implementação clássica (`QR_GS_2(A)`), a coluna A_j permanece inalterada até que todas as colunas $q_1, \ldots, q_{j-1}$ de Q sejam calculadas. Na revisada, assim que uma coluna q_j de Q é calculada, as colunas $A_{j+1}, \ldots, A_n$ são modificadas, descontando destas colunas sua projeção em $span\{q_j\}$. Isso pode ser feito pois as colunas de Q são ortogonais.

2. (15%) Elas são matematicamente equivalentes (Sim/Não) ?
   Sim, ambas garantem que $span\{q_1, \ldots, q_i\} = span\{A_1, \ldots, A_i\}$ para todo $i = 1, \ldots, n$.

3. (15%) Elas são numericamente equivalentes (Sim/Não) ?
   Não, produzem resultados numéricos distintos, uma vez que utilizamos aritmética de precisão finita.

4. (35%) Existe alguma vantagem de uma sobre a outra (Sim/Não) ? Justifique (max 5 linhas).
   Sim, a revisada produz melhores resultados numéricos. Sendo $\hat{Q}$ e Q as matrizes produzidas pela revisada e pela clássica, normalmente temos $\|I - Q^T Q\| >> \|I - \hat{Q}^T \hat{Q}\|$. Como há perda de ortogonalidade quando as colunas de Q são calculadas, projetar as colunas de A assim que uma coluna q_k é disponível ajuda a reduzir os erros.

```
function [Q,R] = QR_GS_1(A)
   [m,n] = size(A)
   R = zeros(n,n)
   Q = zeros(m,n)
   V = A
   for i = 1:n
       R(i,i) = norm(V(:,i),2)
       Q(:,i) = V(:,i)/R(i,i)
       for j = (i+1):n
           R(i,j) = Q(:,i)'*V(:,j)
           V(:,j) = V(:,j)-R(i,j)*Q(:,i)
       end
   end
endfunction
```

```
function [Q,R] = QR_GS_2(A)
   [m,n] = size(A)
   R = zeros(n,n)
   Q = zeros(m,n)
   for j = 1:n
       V = A(:,j)
       for i = 1:j-1
           R(i,j) = Q(:,i)'*A(:,j)
           V = V - R(i,j)*Q(:,i)
       end
       R(j,j) = norm(V,2)
       Q(:,j) = 1.0/R(j,j) * V
   end
endfunction
```

Questão 02: Na Fase I dos algoritmos que fatoram $A = QTQ^*$ (A é quadrada, Q unitária), são feitas operações similares em A, de forma a transformá-la em uma forma conveniente para aplicação da fase subsequente, a Fase II, que é o algoritmo QR. Considerando a matriz A identificada abaixo, responda:

1. Qual é a forma da matriz similar a A obtida ao final da Fase I ? Seja o mais específico que puder e justifique (máx. 3 linhas).
   Para uma matriz A qualquer, o resultado é uma Hessenberg superior, isto é, uma matriz que possui elementos não nulos na parte triangular superior e na subdiagonal abaixo da diagonal principal. Para a matriz em questão, a Hessenberg é uma tridiagonal, dado que $A = A^T$.

2. Qual é a forma da matriz T obtida ao final da Fase II ? Seja o mais específico que puder e justifique (máx. 3 linhas).
   Quando A é uma matriz qualquer, a matriz T é triangular superior. No caso em questão, para A simétrica, T é diagonal.

3. Caracterize a primeira transformação similar necessária desta Fase I, calculando as 2 matrizes que devem ser empregadas e como devem ser empregadas.
   Vamos construir uma tranformação similar $Q_1 A Q_1^*$ por meio de uma matriz Q_1 unitária, definida como $Q_1 = \begin{bmatrix} 1 & 0_3^T \\ 0_3 & F_1 \end{bmatrix}$, onde $F_1 \in \mathbb{R}^{3\times 3}$ é um refletor de Householder que reflete $x = A(2:4,1)$ em $r = \begin{bmatrix} \sqrt{6} & 0 & 0 \end{bmatrix}^T$ ou em seu simétrico.

```
-->Q1= eye(4,4);
```

```
-->x = A(2:4,1);
-->v = sign(x(1))*norm(x,2)*eye(3,1)+x;
-->F = eye(3,3) - 2 * v * v'/(v' * v);
-->Q1(2:4,2:4) = F;
-->Q1
 Q1  =

   1.   0.          0.          0.
   0.  -0.8164966   0.4082483  -0.4082483
   0.   0.4082483   0.9082483   0.0917517
   0.  -0.4082483   0.0917517   0.9082483
-->A2 = Q1*A*Q1'
 A2  =
   4.         -2.4494897   0.          0.
  -2.4494897   2.8333333   0.4457058   0.7790391
   0.          0.4457058   8.2575679   1.9166667
   0.          0.7790391   1.9166667   0.9090987
```

4. Descreva a Fase II do algoritmo para se obter $A = QTQ^*$ (max 5 linhas).
   Assuma que H seja a matriz produzida na Fase I. A fase II consiste em fazer a fatoração QR de H, isto é $H = QR$. Na sequência, atualizamos a matriz H como $H = RQ$ e repetimos o processo, até que a parcela de H, abaixo da subdiagonal, seja suficientemente próximo de zero. Assintoticamente, o produto RQ será uma matriz triangular superior (ou diagonal, no caso da matriz A dada).

5. Se A é uma matriz de grande porte, qual é a justificativa para aplicação da Fase I antes do algoritmo QR ? (max 5 linhas).
   São duas justificativas, ambas visando redução do custo computacional. A primeira é reduzir o número de iterações necessárias para que a Fase II produza uma matriz suficientemente triangular. A segunda é permitir que a fatoração QR seja acelerada, explorando a estrutura (tridiagonal, por exemplo) da matriz de entrada da Fase II.

$$A = \begin{bmatrix} 4 & 2 & -1 & 1 \\ 2 & 3 & 1 & 1 \\ -1 & 1 & 8 & 2 \\ 1 & 1 & 2 & 1 \end{bmatrix}$$

Questão 3: Deseja-se ajustar a função $b(x) \approx \alpha + \beta x$ para os dados da tabela abaixo. Sabe-se os coeficientes ótimos α, β do ajuste podem ser calculados resolvendo-se o sistema de Equações Normais, $A^TA\hat{x} = A^Tb$, onde $\hat{x}^T = [\alpha \quad \beta]$ e A é obtida a partir dos dados e modelo a serem ajustados.

1. Considerando os dados apresentados na tabela e a fatoração QR de A (isto é, Q, R e Q^Tb), encontre os valores ótimos α, β.

$$\begin{aligned} A^TA\hat{x} &= A^Tb \\ (QR)^TQR\hat{x} &= (QR)^Tb \\ R^TQ^TQR\hat{x} &= R^TQ^Tb \\ R^TR\hat{x} &= R^TQ^Tb \\ (R^{-T})R^TR\hat{x} &= (R^{-T})R^TQ^Tb \\ R\hat{x} &= Q^Tb \end{aligned}$$

   No desenvolvimento acima, R^{-T} existe, assumindo-se A com posto coluna completo (o caso não completo é tratado na sequência). Portanto, basta resolvermos o sistema triangular superior $R\hat{x} = Q^Tb$.

```
R  =
  -2.236068  -3.354102
   0.          0.7905694
QTb  =
  -4.2127521
   0.4016093
-->inv(R)*QTb
 ans  =
   1.122
   0.508
```

2. Justifique o uso do método empregado na questão acima (max 5 linhas).
   A matriz A^TA do sistema de equações normais é usualmente malcondicionada (seu número de condição é muito pior que o de A) e seu cálculo explícito para resolução do problema de Mínimos Quadrados deve ser evitado.

3. Explique como você encontraria α, β ótimos se a matriz A for singular (max. 3 linhas).
   Um caminho é fazer a fatoração SVD de $A = U\Sigma V^T$ e calcular uma solução $\hat{x}$ para o problema de Mínimos Quadrados dada por $\hat{x} = A^+b = V\Sigma^+U^Tb$.

i	x_i	b_i
1	1.00	1.63
2	1.25	1.76
3	1.50	1.88
4	1.75	2.01
5	2.00	2.14

```
->Q =
  -0.4472136  -0.6324555
  -0.4472136  -0.3162278
  -0.4472136   0.0000000
  -0.4472136   0.3162278
  -0.4472136   0.6324555
->R =
  -2.236068  -3.354102
   0.         0.7905694
->Q'*b =
  -4.2127521
   0.4016093
```

Questão 04: Utilizando a *abordagem inocente* apresentada no curso, realize a fatoração SVD da seguinte matriz, identificando os fatores pertinentes:

$$A = \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix}.$$

```
 A  =


   0.   4.
   1.   0.

-->ATA = A'*A
 ATA  =


   1.   0.
   0.   16.

-->[Q,S] = spec(ATA)
 Q  =


   1.   0.
   0.   1.
 S  =


   1.   0.
   0.   16.
```

Invertendo a ordem dos autovalores e autovetores S já que $\sigma_1 \geq \sigma_2$ na fatoração SVD:

```
->S(1,1) = 16;S(2,2)=1
 S  =


   16.   0.
   0.    1.
-->Sigma = sqrt(S)
 Sigma  =


   4.   0.
   0.   1.
-->V(:,1) = Q(:,2);V(:,2) = Q(:,1)
 V  =
   0.   1.
   1.   0.
-->U = A*V
 U  =


   4.   0.
   0.   1.
-->U(:,1) = U(:,1)/Sigma(1,1);
-->U(:,2) = U(:,2)/Sigma(2,2);
 U  =
   1.   0.
   0.   1.
```

B.3 Prova 1 - 2023.1

Questão 01: [25%] Sobre ajuste de curvas, responda:

1. (33,3%) Descreva como você encontraria a parábola $C + Dt + Et^2$ que resulta no menor erro de projeção do vetor $b = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \end{bmatrix}^T$ nos pontos $t = -2, -1, 0, 1, 2$. Apresente, com base nos dados fornecidos, o sistema utilizado para obter a solução proposta.

   Solução: Para a solução deste problema, o sistema de equações normais $A^T A\hat{x} = A^T b$ deverá ser resolvido.

   Com base nos dados apresentados, temos: $A = \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}$, $A^T A = \begin{bmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{bmatrix}$,

   $A^T b = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

2. (33,3%) O sistema de equações normais pode ser resolvido pela Fatoração de Cholesky, que é mais barata que, por exemplo, a fatoração $A = QR$. Porém, não raro, resolver o sistema de equações normais via QR é mais adequado numericamente. Esta frase é verdadeira, parcialmente verdadeira, ou falsa ? Justifique (assuma que A tem posto completo).

   Solução: A frase é verdadeira. A matriz $A^T A$ é simétrica definida positiva. Portanto, pode ser fatorada utilizando a Fatoração de Cholesky, que é mais barata computacionalmente. No entanto, caso a matriz A possua um número de condição elevado, a matriz $A^T A$ possuirá um número de condição ainda mais elevado, incorrendo em problemas numéricos. Nestes casos, realizar a fatoração Fatoração QR da matriz A, evitando o cálculo explícito do termo $A^T A$.

3. (33,3%) A matriz P_r é um projetor ortogonal em $C(A^T)$ e P_c é um projetor ortogonal em $C(A)$. Então $P_c A P_r = I$. Verdadeiro ou falso ? Justifique.
   Solução: Esta afirmativa é claramente falsa pois não é sequer dimensionalmente correta se $A \in \mathbb{R}^{m\times n}$ não for quadrada.
   Veja o resultado da aplicação dos projetores P_C, P_r à esquerda e à direita de

A:

$$
\begin{aligned}
P_c[A_1,\dots,A_n]P_r &= \\
[P_cA_1,\dots,P_cA_n]P_r &= \\
AP_r &= \\
\begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix} P_r &= \\
\begin{bmatrix} a_1^TP_r \\ a_2^TP_r \\ \vdots \\ a_m^TP_r \end{bmatrix} &= A
\end{aligned}
$$

Questão 02: [25%] Sabendo que a fatoração QR completa de $A(m \times n)$, com posto completo igual a n é $A = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix}$, onde 0 representa uma matriz $(m-n) \times n$ de zeros, responda, justificando sua resposta com no máximo 3 linhas:

1. (25%) As n colunas de Q_1 formam uma base ortonormal para qual subespaço fundamental?
   Solução: As colunas de Q_1 formam uma base ortonormal obtida, por exemplo, pela versão clássica do algoritmo de Gram-Schmidt, para o espaço coluna de A.

2. (25%) As $m-n$ colunas de Q_2 formam uma base ortonormal para qual subespaço fundamental?
   Solução: As $m-n$ colunas de Q_2 formam uma base ortonormal para o espaço nulo à esquerda de A (ou espaço nulo de A^T). As colunas de Q formam uma base ortornormal para $\mathbb{R}^m$

3. (50%) Podemos afirmar que $I = Q_2Q_2^T + Q_1Q_1^T$? Justifique.
   A expressão é verdadeira. $Q_1Q_1^T$ projeta em $C(A)$ e $Q_2Q_2^T$ projeta em $N(A^T)$ que são espaços ortogonais. Portanto $I - Q_1Q_1^T = Q_2Q_2^T$.

Questão 03:[22%]Considerando as matrizes A_1, A_2 dadas, forneça:

1. (50%) Uma base ortonormal para $C(A_1^T) \cap C(A_2)$.

   Solução: $C(A_1^T) \cap C(A_2) = span \left\{ \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \right\}$.

   Observações:

   - *A interseção de subespaços é um subespaço.*
   - *As matrizes A_1 e A_2 já estão escritas por meio de fatorações que revelam seu posto. Portanto, já são conhecidas bases para seus espaços coluna e linha.*
   - *Basta colocar as linhas de A_1 e as colunas em A_2 lado a lado e verificar que qualquer vetor que pertença a $C(A_1^T) \cap C(A_2)$ deve ser múltiplo do vetor indicado acima.*

2. (50%) O projetor ortogonal que projeta em $(C(A_1^T) \cap C(A_2))^{\perp}$.

   Solução: $P = I - \frac{1}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \end{pmatrix}$

   - $\frac{1}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \end{pmatrix}$ *projeta em* $C(A_1^T) \cap C(A_2)$.
   - *Portanto, para se projetar no complemento ortogonal, basta construir o projetor P indicado.*
   - *Um caminho alternativo (envolvendo mais operações aritméticas do que o necessário) para a solução é encontrar os dois vetores que geram o subespaço ortogonal ao subespaço dado pela interseção, resolvendo, por exemplo, o sistema* $\frac{1}{\sqrt{6}} \begin{bmatrix} 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = 0$ *e calcular a matriz de projeção no subespaço gerado pela base encontrada.*

Justifique os passos necessários para obter suas respostas.

$$A_1 = \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 2 \end{bmatrix}, \; A_2 = \begin{bmatrix} -3 & -1 \\ 3 & 0 \\ -6 & -7 \end{bmatrix} \begin{bmatrix} -3 & 1 & -3 & 0 \\ 0 & 1 & 2 & -1 \end{bmatrix}$$

Questão 04: [28%] Considere os algoritmos clássico e revisado de Gram-Schmidt para produzir a fatoração QR (reduzida) de uma matriz A. Seja A_i, Q_i respectivamente a i-ésima coluna de A e de Q. Denote por R_i e r_i^T respectivamente a $i-$ésima coluna e linha de R.
Utilizando no máximo 5 linhas para cada questão, responda:

1. (25%) Qual é a diferença na ordem em que as entradas de Q, R (isto é Q_i, R_i, r_i^T) são computadas entre os algoritmos ?
   Solução: O algoritmo Clássico calcula $Q_1, R_1, Q_2, R_2, \ldots$. *O algoritmo revisado calcula* $Q_1, r_1^T, Q_2, r_2^T, \ldots$.

2. Considere a matriz A abaixo e sua fatoração obtida por um dos métodos citados. Sabe-se que $A_3 = A_1 + 2A_2$ e $A_4 = A_3 - 3A_1$. Pede-se:

   (a) (25%) Qual algoritmo foi empregado ? Justifique.
   Solução: Clássico. Pelo enunciado $A_3, A_4 \in span\{A_1, A_2\} = span\{Q_1, Q_2\}$. *A diferença entre a coluna* A_3 *e sua projeção em* $span\{A_1, A_2\}$ *é praticamente um vetor de zeros (veja* r_{33}*). Porém, ao se normalizar a diferença, o vetor* Q_3 *deixou de ser um vetor de zeros para ser um vetor lineamente independente de* Q_1, Q_2. *O valor de* r_{44} *deveria ser mas não é próximo de zero pois a coluna* A_4 *foi tardiamente projetada em* $span\{Q_1, Q_2, Q_3\}$, *sofrendo o efeito de* Q_3 *muito distinto de zero.*

   (b) (25%) Há algum problema com os resultados numéricos obtidos ? Em caso positivo, identique-os e justifique sua resposta.
   Solução: Há muitos problemas, essencialmente causados pela perda de ortogonalidade de Q_3 *em relação a* Q_2, Q_1, *como explicado acima.*

   - *A fatoração sugere um posto de 3, quando o posto de A é 2.*
   - *As colunas* Q_3, Q_4 *não tem nenhum significado neste caso, pois foi usada a fatoração reduzida e não há detecção de posto numérico.*

   (c) (25%) Apresente uma base para $C(A)$ a partir da fatoração e indique uma medida numérica para a qualidade desta base (não é necessário calcular a medida de qualidade, apenas apresente sua expressão matemática).
   Solução: As colunas Q_1, Q_2 *fornecem uma base aproximada para* $C(A)$. *Para avaliar sua qualidade, basta calcular* $\|I - [Q_1, Q_2]^T [Q_1, Q_2]\|_2$. *Se esta quantidade for da ordem de* $\|A\|_\infty \epsilon$ *(*ϵ *é a precisão da máquina), a base possui boa qualidade. Caso contrário, procedemos a uma reortogonalização (veja slides 43-47 do curso de Fatoração QR).*

```
A =
```

```
   1.    3.    7.     4.
   4.    3.    10.   -2.
   0.    4.    8.     8.
   2.    3.    8.     2.
Q  =
   0.2182179    0.4264014    0.4229444  -0.4229444
   0.8728716  -0.2132007    0.0704907  -0.0704907
   0.           0.8528029    0.8458889  -0.8458889
   0.4364358    0.2132007    0.3172083  -0.3172083
 R  =
   4.5825757    4.5825757    13.747727    0.
   0.           4.6904158    9.3808315    9.3808315
   0.           0.           6.300D-15    8.9523237
   0.           0.           0.           8.9523237
```

B.4 Prova 2 - 2023.1

Questão 01: Sobre a fatoração SVD e considerando que A_k denota a melhor aproximação de A de rank igual ou menor que k, responda:

1. (33,3%) Quais são os valores singulares, diferentes de zero e em ordem decrescente, de $A - A_k$, sendo $r(A) = r > k$? **Resolução:** A matriz A possui r valores singulares diferentes, de forma que $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_r$. Como $A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$ e $A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T$, temos que $A - A_k = \sum_{i=k+1}^{r} \sigma_i u_i v_i^T$ e $A - A_k = \sigma_i u_i v_i^T$. Assim os valores singulares de $A - A_k$ são $\sigma_{k+1} \geq \sigma_{k+2} \cdots \geq \sigma_r$.

2. (33,3%) Caracterize as matrizes de rank igual a 4 satisfazendo $||A - A_2||_2 = ||A - A_3||_2$?
Pelo exposto acima, são aquelas que satisfazem $\sigma_3 = \sigma_4$.

3. (33,3%) Suponha que a matriz $A^{m \times n}$ tenha n colunas independentes ($r(A) = n$). Descreva as matrizes da fatoração completa de $A = U\Sigma V^T$, considerando as bases oferecidas pela fatoração, bem como a forma da matriz Σ.
As primeiras r colunas de U fornecem uma base ortonormal para $C(A)$ e as últimas $m - r$ colunas de U para $N(A^T)$. As primeiras r linhas de V^T fornecem uma base ortonormal para $C(A^T)$, enquanto as últimas $n - r$ linhas de V^T fornecem uma base ortonormal para $N(A)$. A matriz Σ possui m linhas e n colunas, sendo uma matriz de zeros, exceto pelas suas entradas $\sigma_{i,i} = \sigma_i$ que recebem os valores singulres de A ordenados em magnitude não crescente.

Questão 02: Sobre a decomposição espectral, responda:

1. (50%) Seja $A^k = X\Lambda^k X^{-1}$. Em quais casos $A^k = X\Lambda^k X^{-1}$ se aproxima de uma matriz nula, ou seja, $A^k \to 0$? Quando o raio espectral de A é menor que a unidade, isto é, quando todos seus autovalores possuem módulo inferior a 1.

2. (50%) Qual o valor de b na matriz $A = \begin{bmatrix} 2 & b \\ 1 & 0 \end{bmatrix}$ tal que: (a) $A = Q\Lambda Q^T$ exista? Quando $A^T A = AA^T$ (isto é, A é normal), que implica: $A^T A = \begin{bmatrix} 4+b^2 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 4+1 & 2b \\ 2b & b^2 \end{bmatrix} = AA^T$. O sistema linear acima admite solução quando $b^2 = 1$ e $2b = 2$, ou seja, apenas quando $b = 1$. (b) A seja singular? Justifique a escolha dos valores de b. Quando a segunda coluna é linearmente dependente da primeira. Neste caso, $b = 0$, pois para qualquer outro valor, são LI.

Questão 03: Sobre a fatoração de matrizes quadradas e do cálculo dos seus autovalores e valores singulares, responda:

1. (20%) Qual a principal vantagem em se utilizar o método de duas Fases para o cálculo dos autovalores de uma matriz ? **Resolução:** A principal vantagem é que a transformação em uma Hessenberg (resultado da primeira etapa) permite diminuir o número de iterações da fase seguinte, QR. Também permite reduzir o custo computacional de cada iteração da fase 2, se a estrutura da matriz de entrada for usada adequadamente.

2. (20%) Ao final da Fase I, quais são os possíveis formatos da matriz resultante ? Justifique. **Resolução:** Se a matriz de entrada for hermitiana, o resultado é uma tridiagonal hermitiana. Caso contrário, é uma Hessenberg (trinagular superior mais a primeira subdiagonal abaixo da principal com elementos possivelmente distintos de zero).

3. (20%) É possível que a fatoração de Schur também seja a fatoração SVD de uma matriz ? Se sim, como se relacionam os fatores nas fatorações ? **Resolução:** Sim, se a matriz A fatorada for normal isto é $A^*A = AA^*$ (um caso particular de matriz normal é a matriz simétrica/hermitiana). Neste caso, os vetores singulares à direita e à esquerda são iguais e os valores singulares são os módulos dos autovalores distintos de zero, fornecidos na fatoração espectral.

4. (40%) Para a matriz A fatorada como a seguir, responda:
$A = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} -\frac{2}{3} & -\frac{1}{4} \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$

   (a) É possível obter autovalores autovetores de A por meio da fatoração acima ? Em caso positivo, indique qual é (são) o (os) autovetores com seus correspondentes autovalores e justifique sua resposta. **Resolução:** Os autovalores podem ser lidos na diagonal da matriz tringular superior indicada, similar à matriz A. Porém, a fatoração de Schur acima não é uma fatoração espectral pois a matriz similar à matriz A é triangular superior e não diagonal. Portanto, apenas um dos autovetores, associados ao elemento $T_{1,1} = -2/3$ pode ser lido diretamente, sem cálculos adicionais. Em resumo, os autovalores de A são $-2/3, 1/2$ e, associado ao primeiro destes valores, temos o autovetor $[\sqrt{2}/2 \quad \sqrt{2}/2]^T$.

   (b) À partir da fatoração acima, como você calcularia a norma espectral de A ? Justifique sua resposta.

   Resolução: Como a matriz T similar à A não é hermitiana (não é diagonal, e sim triangular), precisamos calcular os autovalores de T^TT, tomando como norma espectral de A seu maior valor singular, isto é, a raiz quadrada do maior autovalor em módulo de $T^T T$.

(c) $C(A) = C(A^T)$? Responda sim ou não e justifique.
Resolução: Sim, é verdadeira. A fatoração de Schur revela que A é não singular, pois a diagonal de T é não nula. Logo seus espaços coluna e linha geram o $\mathbb{R}^n$.

Questão 04: Em uma de suas etapas, o algoritmo que determina a fatoração SVD de uma matriz A implementa a bidiagonalização de A. Em um determinado momento da aplicação do algoritmo, a matriz A foi transformanda na matriz G abaixo indicada.

1. (60%) Indique qual é a próxima transformação que deve ser implementada na matriz, de forma a concluir aquela etapa do algoritmo. Não é necessário implementar a transformação, indique os fatores envolvidos, como foram calculados, justificando os seus cálculos. **Resolução:** A primeira etapa do algoritmo transforma a matriz de entrada A em uma matriz bidiagonal, ortogonalmente equivalente à matriz A. A primeira coluna da matriz G indica que já foi feita a transformação à esquerda de A, por meio de uma matriz ortogonal E_1. A primeira linha de G indica que a transformação à direita precisa ser feita, usando um refletor de Householder adequado.

```
->G = [3 0 1 1;0 5 2 3;0 -1 4 -1;0 3 1 2]
 G  =
   3.   0.   1.   1.
   0.   5.   2.   3.
   0.  -1.   4.  -1.
   0.   3.   1.   2.
-->x = G(1,2:4)'
 x  =
   0.
   1.
   1.
-->v = norm(x,2)*eye(3,1) - x
 v  =
   1.4142136
  -1.
  -1.
-->D1 = eye(4,4)
```

```
 D1  =
 
    1.   0.   0.   0.
    0.   1.   0.   0.
    0.   0.   1.   0.
    0.   0.   0.   1.
-->D1(2:4,2:4) = eye(3,3) - 2*v*v'/(v'*v)
 D1  =
 
    1.   0.          0.          0.
    0.  -2.220D-16   0.7071068   0.7071068
    0.   0.7071068   0.5        -0.5
    0.   0.7071068  -0.5         0.5
-->G*D1
 ans  =
 
    3.   1.4142136   0.          0.
    0.   3.5355339   3.0355339   4.0355339
    0.   2.1213203   1.7928932  -3.2071068
    0.   2.1213203   1.6213203   2.6213203
```

2. (40%) Em que consiste a segunda etapa do procedimento que calcula os fatores U, V, Σ de A, que se inicia após a bidiagonalização ?

   Resolução: Assumindo que a matriz de entrada usada na primeira fase é quadrada como no exemplo (A é quadrada ou seu fator R em $A = QR$ foi empregado como entrada para a bidiagonalização), foi obtida a matriz bidiagonal B ao final da primeira fase: $B = EAD$, onde E e D são ortogonais de dimensões conformáveis a A. Construímos a matriz $H = \begin{bmatrix} 0 & B^* \\ B & B \end{bmatrix}$ de dimensão $2n \times 2n$ (assumindo que A tem ordem $n \times n$) e fazemos sua fatoração espectral $HQ = Q\Sigma$. Os autovalores de H aparecem aos pares, $\sigma_i, -\sigma_i$, sendo $\sigma_i > 0$ valores singulares de B e de A. Os autovetores de H possuem a forma $\begin{bmatrix} v \\ u \end{bmatrix}$ e $\begin{bmatrix} v \\ -u \end{bmatrix}$, sendo que $v \in \mathbb{R}^n$ e $u \in \mathbb{R}^n$ fornecem os vetores singulares à direita e à esquera de B. Organizamos estes vetores em matrizes V e U respectivamente, após normalização dos v's e u's. Para recuperar a fatoração SVD de A usamos $B = EAD \rightarrow A = E^T BD^T = E^T(U\Sigma V^T)D^T = (E^T U)\Sigma(V^T D^T)$. de forma que $U_A = (E^T U)$ e $V_A^T = (V^T D^T)$. Caso a matriz de entrada fosse retangular, teríamos feito na primeira fase a bidiagonalização de R e então na fatoração SVD de A teríamos o vator Q pré-multiplicando os demais, ist é, $U_A = QE^T U$.

$$G = \begin{bmatrix} 3 & 0 & 1 & 1 \\ 0 & 5 & 2 & 3 \\ 0 & -1 & 4 & -1 \\ 0 & 3 & 1 & 2 \end{bmatrix}.$$

B.5 Prova 1 - 2023.2

Questão 01: Considere a matriz A indicada e responda:

$$A = \begin{bmatrix} 2 & 0 & 2 & -2 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 1 \end{bmatrix}$$

1. (20%) Qual o posto de A ? Justifique.

2. (30%) Forneça bases para $C(A), C(A^T), N(A)$.

3. (30%) Os sistemas lineares $Ax = b$ e $Ax = c$ para $b = (3, 2, 1)^T$ e $c = (2, 2, 1)^T$ admitem solução única, infinitas soluções ou não admitem solução ? Justifique sua resposta, conectando-a com as dimensões dos espaços fundamentais de A.

4. (20%) Pré-multiplicar A por uma matriz de permutação altera $C(A)$ e/ou $C(A^T)$? Sim ou não e justifique sua resposta. Caso altere, apresente algum elemento na diferença entre os subespaços antes e depois da permutação.

Resolução da questão 1:

1. A fatoração indicada $A = CR$ revela o rank da matriz, 2, pois é o número de linhas linearmente independentes de R e de colunas li de C.

2. As colunas de C fornecem uma base para A e as linhas de R uma base para $C(A^T)$. Uma base para $N(A)$ pode ser obtida resolvendo-se o sistema linear homogêneo $Rx = 0$. Há duas variáveis livres, x_3 e x_4 pois a dimensão de $N(A)$ é $4 - 2 = 2$. Fixando $x_3 = -1$ e $x_4 = 0$ obtemos um elemento da base cmo $(1, -1, -1, 0)^T$. Fixando $x_3 = 0, x_4 = -1$, obtemos o outro elemento da base $(-1, 1, 0, -1)^T$.

3. $b \notin C(A)$ pois $b \in \mathbb{R}^3 \setminus C(A)$. Veja que as duas últimas entradas de b são a soma das duas colunas de C na fatoração e o mesmo não pode ser dito para a primeira entrada de b, que é 3. Já o sistema $Ax = c$ admite infinitas soluções, por exemplo $x = (1, 1, 0, 0)^T + v$ para qualquer $v \in N(A)$. Como $N(A) \neq \{0\}$, temos infinitas soluções.

4. A permutação das linhas de A preserva $C(A^T)$ mas pode alterar $C(A)$. Por exemplo se $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ o vetor $v = (2, 1, 0)^T \in C(A)$ e $v \notin C(PA)$.

Questão 02: Sabe-se que a matriz real simétrica S é similar à matriz Λ de forma que $S = X\Lambda X^{-1}$ onde $\Lambda = \begin{bmatrix} a & 0 & 0 \\ 0 & 5-a & 0 \\ 0 & 0 & a-3 \end{bmatrix}$. Pede-se:

1. (40%) Para quais valores de a a matriz S admite uma Fatoração de Cholesky ?

2. (40%) Para quais valores de a a matriz S é negativa definida ?

3. (20%) $X^{-1} = X^T$? Justifique.

<u>**Resolução da questão 2:**</u>

1. A matriz simétrica S admite fatoração se seus autovalores forem positivos. Como S e Λ são similares, os autovalores de Λ (seus elementos na diagonal principal) devem ser positivos. Então temos que satisfazer as seguintes condições para a: $a > 0, 5 - a > 0 \rightarrow a < 5, a - 3 > 0 \rightarrow a > 3$. Logo $a \in (3, 5) \iff S$ é positiva definida e admite fatoração de Cholesky.

2. Para que seja negativa definida, seus autovalroes devem ser todos negativos. Então devemos satisfazer $a < 0$ e $a > 5$. Portanto, não há valores de a que tornem S negativa definida.

3. Sim, a simetria de S garante que $X^{-1} = X^T$. As matrizes reais simétricas são unitariamente diagonalizáveis, seus autovetores formam uma base para $\mathbb{R}^n$ e podem ser ortogonalizados.

Questão 03 Sobre as fatorações básicas:

1. Considere a matriz simétrica $S = \begin{bmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{bmatrix}$ e responda.

   (a) (45%) Usando a fatoração $S = LU$, escreva S como uma soma de matrizes de rank-1.

   (b) (10%) Indique as condições necessárias (em a, b, c, d) para que não haja pivot nulo no processo de fatoração e, desta forma, possamos fatorar $S = LU$.

2. (45%) Quais são as matrizes de multiplicadores M e de permutação P tais que $A = P^T M^{-1}\hat{A}$ para as matrizes $A, \hat{A}$ indicadas abaixo ? $A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & 2 \\ 4 & 8 & -1 \end{bmatrix}, \hat{A} = \begin{bmatrix} 0 & -2 & 9/4 \\ 4 & 8 & -1 \\ 0 & -3 & -1/2 \end{bmatrix}$

Resolução da questão 3:

1. Fatorando a matriz S, *outer LU* (você poderia ter usado a *inner LU*, visão linhas).

- $S = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} a & a & a & a \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & b-a & b-a & b-a \\ 0 & b-a & c-a & c-a \\ 0 & b-a & c-a & d-a \end{bmatrix}$
- $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & b-a & b-a & b-a \\ 0 & b-a & c-a & c-a \\ 0 & b-a & c-a & d-a \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & b-a & b-a & b-a \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & c-b & c-b \\ 0 & 0 & c-b & d-b \end{bmatrix}$
- $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & c-b & c-b \\ 0 & 0 & c-b & d-b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & c-b & c-b \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d-c \end{bmatrix}$
- $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d-c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & d-c \end{bmatrix}$

1a) Logo $S = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} a & a & a & a \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & b-a & b-a & b-a \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & c-b & c-b \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & d-c \end{bmatrix}$

1b) $a \neq 0, b \neq a, c \neq b, d \neq c$.

2. $A = P^T M^{-1} \hat{A} \to MPA = \hat{A}$. A matriz de permutação P faz alguma troca de linhas e a matriz de multiplicadores M cria zeros na primeira coluna. Veja que a terceira linha de $\hat{A}$ é a linha 2 de A. Para zerar os elementos nas demais linhas de A usando como pivot sua linha 3 (linha 2 de $\hat{A}$), usamos os multiplicadores $-1/2$, $-1/4$.

$$\begin{bmatrix} 1 & -1/4 & 0 \\ 0 & 1 & 0 \\ 0 & -1/2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & 2 \\ 4 & 8 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -2 & 9/4 \\ 4 & 8 & -1 \\ 0 & -3 & -1/2 \end{bmatrix}$$

Questão 04: Responda verdadeiro ou falso e justifique.

1. Para uma matriz A simétrica, $C(A)$ é perpendicular a $N(A)$.

2. $\mathcal{X} = \{(x, y) \in \mathbb{R}^2 : y = |x|, x \in \mathbb{R}\}$ é um subespaço.

3. A matriz $A = \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}$ possui 2 autovetores linearmente independentes.

4. Se A e B são similares, seus determinantes são iguais.

5. Para duas matrizes A, B temos que $AB = 0$. Então, as colunas de B pertendem a $C(A)$ e as linhas de A pertencem a $C(B^T)$.

6. Considere as matrizes A, B, D tais que $D = AB$. Então, temos que $C(A) \subseteq C(D)$.

Resolução da questão 4:

1. Verdadeiro. Temos que para qualquer A, $C(A) \perp N(A^T)$. Para A simétrica temos $C(A) = C(A^T)$ e $N(A) = N(A^T)$. Então $C(A) \perp N(A)$.

2. Falso. Não o conjunto não é fechado para a soma. Tome $(1, 1)^T, (-1, 1)^T$ e veja que a soma dos dois, $(0, 2)^T \notin \mathcal{X}$.

3. Falso. A matriz A possui um autovalor $\lambda = 5$, com multiplicidade 2. O único autovetor associado é $(1, 0)^T$.

4. Verdadeiro. Matrizes similares tem os mesmos autovalores e, devido à expressão de similaridade ($A = XBX^{-1}$ para alguma matriz X inversível), possuem a mesma ordem e multiplicidade de autovalores. Como o determinante de qualquer matriz é o produto de seus autovalores, o resultado segue.

5. Falso. As linhas de A estão em $N(B^T)$ e as colunas de B em $N(A)$.

6. Verdadeiro se $posto(B)$ é completo, uma vez que $C(A) = C(D)$. Porém, se $posto(B)$ for deficiente, é falso. Por exemplo, se $B = 0$ (matriz identicamente nula), A é identicamente nula, $N(A) = \{0\}$ pode ser diferente de $C(A)$.

B.6 Prova 2 - 2023.2

Questão 01 Responda à seguintes questões:

1. Considere a seguinte matriz $A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & -1 & 4 \\ -2 & 4 & -4 \end{bmatrix}$ e sua fatoração completa $A = QR$.

   (a) (15%) As colunas de q_1, q_2, q_3 de Q são base(s) para qual(is) subespaço(s) fundamental(is) ?

   (b) (15%) Qual o menor valor da norma Euclidiana entre os vetores linha de R?

2. (70%) Considere duas matrizes $A \in \mathbb{R}^{m \times n_A}$ e $B \in \mathbb{R}^{m \times n_B}$ e sejam $\{A_k : k = 1, \ldots, n_A\}$ e $\{B_k : k = 1, \ldots, n_B\}$ as colunas de A e B respectivamente. As matrizes A e B podem ter colunas linearmente dependentes. Construa um teste ou algoritmo que permita responder se $C(A) \subseteq C(B)$ ou se $C(A) \not\subseteq C(B)$. Não é necessário escrever o pseudo-código do algoritmo/teste.

Resolução da questão 1.1:

1. $\{q_1, q_2\}$ é uma base para $C(A)$. q_3 é uma base para $N(A^T)$. $||r_3||_2 = 0$.

Questão 02 Sobre projeções responda:

1. (40%) Sejam P e $(I - P)$ matrizes de projeção ortogonais e as suas respectivas transformações lineares Py e $(I - P)z$, para vetores z, y quaisquer. Qual o ângulo formado entre os vetores Py e $(I - P)z$? Justifique algebricamente a sua resposta.

2. Considere as matrizes $P = A(A^TA)^{-1}A^T$ e $I - P$ onde $A \in \mathbb{R}^{m \times n}$ $(m > n)$ possui posto completo. Então responda:

   (a) (20%) P é projetor ? Em caso positivo, P projeta onde ? Em caso positivo, é projetor ortogonal ? Justifique suas respostas.

   (b) (20%) $posto(P) = m - n$? Justifique.

   (c) (20%) Se $\|(I - P)b\| = 0$, b é linearmente independente das colunas de A. Responda verdadeiro ou falso e justifique.

Resolução da questão 2:

1. Parte 1

$$\begin{aligned}(Py)^T(I-P)z &= \\ y^TP^T(I-P)z &= \\ y^T(P^T-P^2)z &= \\ y^T(P-P)z &= 0 \qquad\qquad \text{usando } P=P^T, P=P^2\end{aligned}$$

Portanto, os dois vetores são ortogonais, formando um ângulo de $\frac{\pi}{2}$.

2. Parte 2

   (a) Sim, P é projetor ortogonal, pois satisfaz a idempotência:

$$\begin{aligned}P^2 &= \\ A(A^TA)^{-1}A^TA(A^TA)^{-1}A^T &= \\ A(A^TA)^{-1}A^T &\end{aligned}$$

   e simetria

$$\begin{aligned}P^T &= \\ (A(A^TA)^{-1}A^T)^T &= \\ A(A^TA)^{-T}A^T &= \\ A(A^TA)^{-1}A^T & \qquad \text{já que } A^TA, (A^TA)^{-1} = (A^TA)^{-T} \text{ são simétricas}\end{aligned}$$

   P projeta em $C(A)$. Veja tome um $z \in \mathbb{R}^m$ qualquer:

$$\begin{aligned}A(A^TA)^{-1}A^Tz &= \\ A((A^TA)^{-1}A^Tz) &= \\ Au &\in C(A)\end{aligned}$$

   (b) Falso, pois $posto(P) = n$, a dimensão do espaço em que projeta.

   (c) Falso. $\|(I-P)b\| = 0 \iff (I-P)b = 0$ para qualquer norma vetorial e, portanto, $b \in C(A)$. Portanto, b é linearmente dependente das colunas de A.

Questão 03 Considere uma matriz $A \in \mathbb{R}^{5\times 5}$ satisfazendo $A_3, A_5 \in span\{A_1, A_2, A_4\} = C(A)$ (A_k representa a $k-$ésima coluna de A). Após alguma permutação P das colunas de A foi obtida a seguinte fatoração $AP = QR$, onde os fatores são abaixo discriminados.

```
ALGO
Q  =
   0.3779645   0.5070926   0.4743416
   0.          0.7888106  -0.0527046
  -0.3779645   0.2817181  -0.5270463
   0.7559289  -0.1690309  -0.1581139
   0.3779645   0.1126872  -0.6851602
R  =
   2.6457513  -0.7559289  -0.7559289   1.8898224   0.
   0.          2.5354628   1.3522468   1.3522468   1.183216
   0.          0.          1.2649111   1.2649111  -1.2649111
pivot  = 1. 5. 3. 2. 4.
```

Além disso, foram empregados dois algoritmos para produzir a fatoração $A = QR$ de A: GS Clássico, GS Revisado. As saídas destes algoritmos (não necessariamente nesta ordem) é apresentada abaixo. Responda, apresentando suas justificativas:

1. (10%) Com base apenas no enunciado da questão, sem levar em conta a saída de qualquer uma das três fatorações apresentadas, apresente um limite superior para $posto(A)$.

2. (50%) Identifique qual saída corresponde a qual algoritmo (GS Clássico e Revisado).

3. (30%) Para cada algoritmo que você caracterizar, identifique se na saída do algoritmo há alguma informação nas linhas ou colunas das correspondentes Q, R que devem ser desconsideradas em decorrência de erros numéricos.

4. (10%) Qual é o posto de A ? Forneça uma base para $C(A)$ à partir das fatorações.

```
ALG1
Q1  =
   0.3779645   0.6943651  -0.2637522  -0.0100482  -0.1653954
   0.          0.5400617   0.1318761   0.5827932   0.
  -0.3779645  -0.1543033   0.3296902   0.5903293   0.
   0.7559289  -0.231455   -0.7253185  -0.0276324   0.4410543
   0.3779645  -0.3857584  -0.5275044   0.5576728  -0.8821086
 R1  =
   2.6457513   1.8898224  -0.7559289   0.         -0.7559289
   0.          1.8516402   1.8516402  -5.551D-17   1.8516402
   0.          0.          8.419D-16  -0.065938   -0.065938
   0.          0.          0.          1.7307952   1.7307952
   0.          0.          0.          0.          1.259D-16
```

```
ALG2
Q2  =
   0.3779645   0.6943651   0.          2.230D-17  -2.181D-17
   0.          0.5400617   0.4282302   0.5519368  -0.3932963
  -0.3779645  -0.1543033   0.2141151   0.5651966  -0.1787711
   0.7559289  -0.231455   -0.695874    0.0430942   0.6972071
   0.3779645  -0.3857584  -0.5352877   0.6116057   0.5720674
 R2  =
   2.6457513   1.8898224  -0.7559289   0.         -0.7559289
   0.          1.8516402   1.8516402  -5.551D-17   1.8516402
   0.          0.          1.037D-15   0.1070575   1.7664495
   0.          0.          0.          1.728739    1.625976
   0.          0.          0.          0.          1.6625709
```

Resolução da questão 3:

1. O enunciado diz que $C(A) = span\{A_1, A_2, A_4\}$, logo A_3, A_5 são desnecessárias para descrever o espaço e $posto(A) \leq 3$. Veja que $C(A) = span\{A_1, A_2, A_4\}$ não garante que $posto(A) = 3$.

2. ALG1 é GS revisado e ALG2 é GS clássico. A entrada r_{55} da matriz $R2$ erroneamente sugere que A_5 seja linearmente independente de A_1, A_2, A_3, A_4. Além disso, a fatoração de ALG2 sugere que o posto seja 4, quando isso não pode ser verdade, face à resposta dada para o item anterior. Já a fatoração produzida pelo ALG1 corretamente identifica o posto da matriz e a dependência linear entre A_3, A_5 das demais colunas de A.

3. Os dois algoritmos, ALG1 e ALG2, forneceram fatorações onde Q possui o mesmo número de colunas de A. Porém, como A tem posto incompleto, todas as colunas de Q, nos dois algoritmos, ALG1 e ALG2, associadas a A_3, A_5 devem ser desconsideradas, pois deveriam ser identicamente nulas. Em resumo, na deficiência de posto, estas colunas não deveriam ser retornadas pelo algoritmo. Além disso, as entradas ao longo das linhas de $R1, R2$ referentes a estes índices, $3, 5$ também não têm significado, pois as colunas q_3, q_5 não devem ser empregadas para se representar as demais. Assim, também não deveriam ser retornadas. Veja que mesmo as entradas $r_{3,4}, r_{3,5}$ de R1, produzidas pelo algoritmo revisado, não dizem nada, deveriam ser nulas, pois A_3 é linearmente dependente de A_1, A_2, segundo as entrada $r_{3,3}$ da mesma matriz. Por sua vez, A_5 é linearmente dependente das demais e $r_{3,5}$ também deveria ser zero.

4. O posto é 3, pois a fatoração com permutação de colunas produziu uma fatoração QR reduzida, capaz de caracterizar o posto de A.

Questão 04 Responda verdadeiro ou falso e justifique.

1. A solução do sistema de equações normais quando $A = Q$, isto é, quando A tem colunas ortonormais, é dada por $\hat{x} = Q^T b$.

2. Uma matriz é perfeitamente condicionada quando seu número de condição, para alguma norma matricial induzida qualquer, é inferior à unidade.

3. Todo projetor possui pelo menos um autovalor nulo e um autovalor um.

4. Os algoritmos Gram-Schmidt clássico e Gram-Schmidt revisado são matematicamente equivalentes, mas não são numericamente equivalentes.

5. O sistema de equações normais pode ser resolvido pela Fatoração de Cholesky, que é mais barata que, por exemplo, a fatoração $A = QR$. Porém, não raro, resolver o sistema de equações normais via QR é mais adequado numericamente. Assuma que A tem posto completo.

6. Sejam ALG1 e ALG2, algoritmos propostos para fatorar a matriz A em QR. Para ALG1, temos $||I - Q_1^T Q_1|| = 10^{-08}$ e para ALG2, temos $||I - Q_2^T Q_2|| = 10^{-02}$. ALG2 é o algoritmo mais estável numericamente.

7. Se o projetor P satisfaz $Px = x$ para todo $x \in \mathbb{R}^n$ então $P = I$.

Resolução da questão 4:

1. Verdadeiro. A solução do sistema de equações normais se reduz a:

$$\begin{aligned} A^T A\hat{x} &= A^T b \\ Q^T Q\hat{x} &= Q^T b \\ \hat{x} &= Q^T b \end{aligned}$$

2. Falso, pois $k(A) \geq 1$ para qualquer matriz, em qualquer norma induzida por norma vetorial. Não há matriz cujo número de condição é menor que a unidade. Matrizes perfeitamente condicionadas são aquelas que possuem $\kappa(A) = 1$.

3. Falso. Contra-exemplo: $P = I$, para o qual todos os autovalores são iguais a 1. Para qualquer outro projetor distinto da matriz nula (outro contra-exemplo, que só tem zero como autovalor), a afirmativa é verdadeira.

4. Verdadeiro. Dado que o GS revisado desconta as projeções assim que as colunas de Q são calculadas, e GS clássico só projeta (ou ortogonaliza) A_k quando todas as colunas $q_1, q_2, \ldots, q_{k-1}$ são disponíveis, os dois algoritmos produzem resultados numéricos distintos, diante de aritmética de precisão finita.

5. Verdadeiro. A matriz A^TA é simétrica definida positiva. Portanto, pode ser fatorada utilizando a Fatoração de Cholesky, que é mais barata computacionalmente. No entanto, caso a matriz A possua um número de condição elevado, a matriz A^TA possuirá um número de condição ainda mais elevado, incorrendo em problemas numéricos. Nestes casos, é recomendável realizar a fatoração Fatoração QR da matriz A, evitando o cálculo explícito do termo A^TA.

6. Falso. Para a fatoração $A = QR$, o algoritmo mais estável numericamente é aquele que garante, o máximo possível, a ortogonalidade entre as colunas geradas. Assim, a norma da matriz resultante da diferença entre a matriz I e matriz da Q^TQ tem que ser a menor possível, ou seja, o algoritmo mais estável numericamente é o ALG1.

7. Verdadeiro, pois nesse caso $C(P) = \mathbb{R}^n$.

B.7 Prova 3 - 2023.2

Questão 01. Responda:

1. (25%) Defina o refletor de Householder que reflita o vetor $a = (1, 0, 1, 2)^T$ no sentido positivo da linha $e_4 = (0, 0, 0, 1)^T$. Não é necessário calcular o refletor, apenas o vetor de Householder normalizado e como o refletor se relaciona com o vetor de Householder.

2. (25%) Sendo u o vetor de Householder normalizado, F o refletor e b o ponto a ser refletido, represente graficamente Fb, $uu^T b$ e $(I - uu^T)b$, em relação aos subespaços $span\{u\}$ e $span\{u\}^\perp$.

3. (50%) Considere uma matriz A de ordem 4×6. Seria possível, utilizando matrizes de permutação e refletores de Householder, construir uma sequência de transformações ortogonais que, após aplicadas em A, tenham o efeito de zerar as entradas de A armazenadas em $A_{2,1}, A_{2,6}$ e de substituir o conteúdo anteriormente existente em $A_{2,3}$ pela quantidade $\sqrt{A_{2,1}^2 + A_{2,3}^2 + A_{2,6}^2}$? Em caso negativo, justifique. Em caso positivo, detalhe os passos destas transformações, as dimensões das matrizes envolvidas e como estas matrizes deveriam ser definidas para se obter o efeito desejado.

Resolução da questão 01:

1. Para mais detalhes, ver notas de aula fatoração QR, exemplo 7.

```
-->a = [1;0;1;2]
-->r = [0;0;0;1]*norm(a,2)
 r  =
   0.
   0.
   0.
   2.4494897
-->v = a - r
 v  =
   1.
   0.
   1.
  -0.4494897
-->u = v/norm(v,2)
 u  =
```

```
   0.6738873
   0.
   0.6738873
  -0.3029054
-->F = eye(4,4)-2*u*u'
 F  =
   0.0917517   0.  -0.9082483   0.4082483
   0.          1.   0.          0.
  -0.9082483   0.   0.0917517   0.4082483
   0.4082483   0.   0.4082483   0.8164966
-->F*a
 ans  =
  -7.772D-16
   0.
  -7.772D-16
   2.4494897
```

2. Ver o desenho da Figura 8 das notas de aula de fatoração SVD. Fb é a reflexão de b, obtida por meio de $(I - 2uu^T)b$. O ponto uu^Tb é a projeção de b em $span\{v\}^\perp$ e este ponto fica *no meio do caminho*, entre b e sua reflexão Fb. O vetor $b - uu^Tb$ pertence ao $span\{u\}$.

3. Sim, é possível e há mais de uma maneira de se proceder, dependendo de como a permutação de colunas é realizada. Vamos fazer uma sequência de 3 transformações, do tipo $\hat{A} = APFP^T$, onde $\hat{A}$ é a matriz com a propriedade desejada do enunciado. A matriz P troca as colunas de A de forma que as colunas $1, 3, 6$ fiquem contíguas e possamos assim empregar um refletor de Householder. Este é o aspecto imprescindível, as colunas $1, 3, 6$ precisam ficar contíguas na matriz antes da aplicação do refletor. A matriz F é o refletor de Householder. Por fim, aplicamos a matriz P^T para restaurar as colunas de A à sequência original. Na nossa resolução, vamos colocar as colunas que devemos alterar nas últimas 3 posições. Então P será definida por $pivot = (2, 4, 5, 1, 3, 6)$. Veja o exemplo numérico abaixo. Todas as matrizes P, P^T, F são quadradas de ordem 6. Embora no enunciado tenhamos soliciado que ao final das transformações o conteúdo de $\hat{A}_{2,3}$ seja $\sqrt{A_{2,1}^2 + A_{2,3}^2 + A_{2,6}^2}$, na execução abaixo, colocamos o simétrico, -$\sqrt{A_{2,1}^2 + A_{2,3}^2 + A_{2,6}^2}$, por ser a opção mais estável neste caso.

```
A  =
```

```
    4.   3.   2.   7.   6.   1.
    7.   5.   6.   2.   5.   3.
    3.   5.   8.   4.   2.   9.
    5.   1.   0.   8.   8.   8.
-->pivot = [2;4;5;1;3;6];
-->P = zeros(6,6);
-->for i = 1:6
-->P(pivot(i),i) =1
-->end
 P  =
    0.   0.   0.   1.   0.   0.
    1.   0.   0.   0.   0.   0.
    0.   0.   0.   0.   1.   0.
    0.   1.   0.   0.   0.   0.
    0.   0.   1.   0.   0.   0.
    0.   0.   0.   0.   0.   1.
-->F = eye(6,6);
-->Ahat = A*P
 Ahat  =
    3.   7.   6.   4.   2.   1.
    5.   2.   5.   7.   6.   3.
    5.   4.   2.   3.   8.   9.
    1.   8.   8.   5.   0.   8.
-->x = Ahat(2,4:6)'
 x  =
    7.
    6.
    3.
-->normax = norm(x,2)
 normax  =
    9.6953597
-->e2 = [0;1;0]
 e2  =
    0.
    1.
    0.
-->a = normax*e2
 a  =
```

```
    0.
    9.6953597
    0.
-->v = x + sign(x(1))*normax*e2
 v  =
    7.
    15.69536
    3.
-->u = v / norm(v,2)
 u  =
    0.4012504
    0.8996813
    0.1719644
-->F2 = eye(3,3) - 2*u*u'
 F2  =
    0.6779963  -0.7219949  -0.1380016
   -0.7219949  -0.6188527  -0.3094264
   -0.1380016  -0.3094264   0.9408565
-->F(4:6,4:6) = F2
 F  =
    1.   0.   0.   0.           0.           0.
    0.   1.   0.   0.           0.           0.
    0.   0.   1.   0.           0.           0.
    0.   0.   0.   0.6779963  -0.7219949  -0.1380016
    0.   0.   0.  -0.7219949  -0.6188527  -0.3094264
    0.   0.   0.  -0.1380016  -0.3094264   0.9408565
-->A*P*F*P'
 ans  =
    1.1299938   3.  -4.4351114   7.   6.  -0.2300027
   -2.220D-16   5.  -9.6953597   2.   5.   0.
   -4.9839844   5.  -9.901644    4.   2.   5.5782924
    2.2859687   1.  -6.0853854   8.   8.   6.8368437
```

Questão 02. Considere um problema de projeção em $C(A)$ para uma matriz $A \in \mathbb{R}^{m\times n}$ e assuma que A possua posto $r < n = \min\{m, n\}$. Responda justificando:

1. (30%) Considere o sistema $A^T A\hat{x} = A^T b$. $(A^T A)^{-1}$ existe ?

2. (30%) O que representa o vetor $z = A\hat{x}$, caso $\hat{x}$ exista ?

3. (40%) Como você poderia usar a fatoração SVD de A para obter uma solução do sistema $A^TA\hat{x} = A^Tb$, caso exista ?

Resolução da questão 2:

1. Não existe, pois A possui posto incompleto r e A^TA possui ordem n e posto r, também incompleto.

2. Existe a solução $\hat{x}$ e não é única. z representa o ponto de $C(A)$ mais próximo de b, na norma Euclideana.

3. Com a fatoração SVD $A = U\Sigma V^T$, podemos obter a pseudo-inversa $A^+ = V\Sigma^+U^T$ de A e com ela uma solução $\hat{x}$ por meio de $\hat{x} = A^+b = V\Sigma^+U^Tb$.

Questão 03. Sobre fatorações as fatorações matriciais vistas, responda:

1. (40%) Durante o curso, discutimos duas formas de se fazer a fatoração SVD de A não quadrada. Uma delas é pouco estável e a outra é estável. Descreva a não estável e justifique por quais motivos é pouco estável (máximo 8 linhas).

2. (30%) Aplicou-se o algoritmo de duas fases para fornecer a fatoração $A = Q\Lambda Q^T$ das matrizes A, B indicadas abaixo. Qual a forma das matrizes obtidas após a aplicação da primeira e da segunda fase ? Justifique considerando as transformações ortogonais empregadas (máximo 8 linhas).

3. (30%) Considerando o algoritmo empregado no item acima, é possível assegurar que sempre conseguiremos recuperar os n autovetores da matriz fatorada, com a fatoração obtida ? Sim ou não ? Justifique com no máximo 8 linhas.

```
A  =
    17.    14.    16.    12.    8.
    14.    22.    29.    18.    20.
    16.    29.    41.    21.    31.
    12.    18.    21.    18.    12.
    8.     20.    31.    12.    26.
B  =
    2.    3.    3.    5.    9.
    2.    9.    6.    4.    0.
    2.    2.    5.    3.    5.
    9.    3.    3.    6.    3.
    7.    4.    6.    4.    4.
```

Resolução da questão 3:

1. Assumimos que o posto de A é r. <u>Não estável:</u> calcula-se A^TA e fatoramos $A^TA = Q\Lambda Q^T$. Os vetores singulares à direita de A são as colunas $q_i : i = 1, \ldots, r$ associados aos r autovalores $\lambda_i > 0$ de A. Os valores singulares de A são $\sigma_i = \sqrt{\lambda_i} : i = 1, \ldots, r$. Os vetores singulares à esquerda de A são $u_i = \frac{1}{\sigma_i} A v_i : i = 1, \ldots, r$. A fatoração é não estável pois A^TA é pior condicionada, e os valores singulares muito pequenos de A serão difíceis de serem computados, pois serão avaliados através de λ_i^2, grandezas ainda menores.

2. Matriz A, real simétrica. Primeira fase produz uma Hessenberg superior. Como A é simétrica, e as operações ortogonais à direita e a esquerda de A são simétrias, o resultado é tridiagonal. Na segunda fase, aplicamos o algoritmo QR iterativamente, que produz uma triangular, com os autovalores na diagonal. Matriz B não simétrica. Primeira fase produz uma Hessenberg superior e segunda fase uma triangular superior, fatoração de Schur.

3. Se a matriz de entrada for real simétrica sim, pois a forma da matriz similar obtida com a fatoração é diagonal e não há falta de autovetores (é não defectiva). Entretanto, para matrizes não simétricas, a fatoração obtida é uma Schur e apenas um autovetor estará disponível através da fatoração. Além disso, pode ser o caso de que haja falta de autovetores (não somam a dimensão do espaço) caso a matriz seja defectiva.

Questão 04. Responda Verdadeiro ou Falso justificando sua resposta. Atribuições verdadeiras ou falsas adequadamente dadas, mas com justificativas erradas não serão consideradas. Todos os itens são igualmente valorados.

1. Toda matriz de posto igual a 3 com $\sigma_2 = \sigma_3$ satisfaz a condição $||A - A_1||_2 = ||A - A_2||_2$.

2. Os valores singulares de AA^TA são iguais a $\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2$, quando o posto de A é igual a r.

3. Toda matriz real simétrica $A = Q\Lambda Q^T$ tem sua fatoração SVD escrita da seguinte forma $\Sigma = \Lambda$ e $U = Q$, e $V = Q$.

4. Toda transformação ortogonalmente equivalente é uma transformação similar, pois preserva os valores singulares.

Resolução da questão 4:

1. Verdadeiro. Sabendo que $A - A_k = \sigma_{k+1} u_{k+1} v_{k+1}^T + \cdots + \sigma_r u_r v_r^T$, temos que $||A - A_k||_2 = \sigma_{k+1}$. Deste resultado: $||A - A_1||_2 = \sigma_2$ e $||A - A_2||_2 = \sigma_3$. Assim, se $||A - A_1||_2 = ||A - A_2||_2$ então $\sigma_2 = \sigma_3$.

2. Falso. $AA^T A = (U\Sigma V^T)(V\Sigma^T U^T)(U\Sigma V^T) = U\Sigma\Sigma^T\Sigma V^T$. Assim, os valores singulares da matriz $AA^T A$ são iguais a $\sigma_1^3, \cdots, \sigma_r^3$, para uma matriz A de posto igual a r.

3. Falso. A matriz simétrica, a não ser que seja positiva definida ou semi-positiva definida, o que não foi especificado no enunciado, pode ter autovalores negativos. Porém, os valores singulares de A em sua fatoração SVD são sempre positivos. Desta forma, para um autovalor $\lambda_i < 0$ temos que $v_i = q_i$ onde q_i é o autovetor associado a λ_i e $u_i = -v_i$.

4. Falso. Uma transformação similar é escrita como $A = XBX^{-1}$ e a transformação ortogonalmente equivalente é $A = EBD$ para E, D ortogonais. Se X for ortogonal, a transformação similar é simplificada para $A = XBX^T$. Assim sendo, se a matriz X for ortogonal, a transformação similar é também uma transformação ortogonalmente equivalente. Porém, para o sentido inverso, se matriz A não for quadrada ou se $E \neq D^{-1}$, a transformação ortogonalmente equivalente $A = EBD$ não é similar.

B.8 Prova 1 - 2024.1

Questão 01: Considere a matriz A singular indicada e responda justificando: $A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

1. (25%) Qual é o posto de A ?

2. (25%) $\lambda_i = 0$ é autovalor de A ?

3. (25%) Apresente bases para os quatro espaços fundamentais de A.

4. (25%) A matriz A admite fatoração de Cholesky ? Em caso positivo, apresente a fatoração.

Resolução: Questão 01

1. Posto de $A = 2$. A matriz A foi apresentada já fatorada $A = MM^T$, por meio de uma fatoração que revela seu posto, correspondendo ao número de colunas (linhas) linearmente independentes de M (de M^T).

2. Verdadeiro. A matriz A possui posto incompleto, é singular (conforme antecipado pelo enunciado), $det(A) = \prod_i \lambda_i = 0 \Rightarrow \exists \lambda_i = 0$.

3. A matriz A é simétrica e, pela fatoração apresentada $C(A) = C(A^T) = span\{m_1, m_2\}$ onde $m_1 = [2 \quad 1 \quad 0]^T, m_2 = [0 \quad 1 \quad 1]^T$ são as linhas de M^T. Pelos mesmos motivos, $N(A) = N(A^T)$. Para determinar um destes espaços, digamos $N(A)$, usamos o fato de que $N(A) \perp C(A^T)$. Então devemos resolver o sistema linear $M^T w = 0$: $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$. Fixando a variável livre $w_3 = -1$, temos a solução $w_2 = 1, w_1 = -\frac{1}{2}$. Portanto, $N(A) = N(A^T) = span\{[-\frac{1}{2} \quad 1 \quad -1]^T\}$.

4. Não. A matriz A é singular, não é positiva definida, portanto não admite fatoração de Cholesky.

Questão 02 Sobre as fatorações básicas:

1. Considere a matriz simétrica $S = \begin{bmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{bmatrix}$.

   (a) (20%) Assuma que S seja positiva definida e considere sua fatoração de Cholesky $S = LL^T = \sum_{i=1}^{4} L_i L_i^T$. Apresente L_1.

   (b) (30%) Quais as relações que devem ser satisfeitas por a, b, c, d para que a matriz S tenha posto igual a 1 ?

2. (10%) A fatoração $PA = LU$ de A é fornecida abaixo e emprega a representação do vetor *pivot* para P. Quais são os mutiplicadores empregados para a linha $i = 3$ da matriz A ?
   $pivot = \begin{bmatrix} 2 & 1 & 4 & 3 & 5 \end{bmatrix}^T$.
   $$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -0,111 & 1 & 0 & 0 & 0 \\ -0,125 & -0.125 & 1 & 0 & 0 \\ -0,333 & -0,125 & -0,200 & 1 & 0 \\ -0,200 & 1.000 & -0.500 & -0.200 & 1 \end{bmatrix}$$
   $$U = \begin{bmatrix} 3.000 & -8.000 & 12.000 & -9.000 & 6.000 \\ 0 & 7.000 & 0.000 & -8.000 & -12.000 \\ 0 & 0 & 5.000 & 8.000 & 12.000 \\ 0 & 0 & 0 & -7.000 & 5.000 \\ 0 & 0 & 0 & 0 & -1.000 \end{bmatrix}$$

3. (40%) Quais são as matrizes de multiplicadores M e de permutação P tais que $\hat{A}M^{-1}P^T = A$ para as matrizes $A, \hat{A}$ indicadas abaixo ? $A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & 2 \\ 8 & 8 & -1 \end{bmatrix}$, $\hat{A} = \begin{bmatrix} 0 & 0 & 2 \\ 5/2 & -1/2 & 1 \\ 3 & 4 & 8 \end{bmatrix}$

Resolução: Questão 02

1. $r_{11} = \sqrt{a}, s = \frac{1}{\sqrt{a}}[a \quad a \quad a]^T$, Logo $L_1 = \begin{bmatrix} \sqrt{a} \\ \frac{a}{\sqrt{a}} \\ \frac{a}{\sqrt{a}} \\ \frac{a}{\sqrt{a}} \end{bmatrix}$, $L_1 L_1^T = \begin{bmatrix} a & a & a & a \\ a & a & a & a \\ a & a & a & a \\ a & a & a & a \end{bmatrix}$,

2. $\sum_{i=2}^{n} L_i L_i^T = A - L_1 L_1^T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & b-a & b-a & b-a \\ 0 & b-a & c-a & c-a \\ 0 & b-a & c-a & d-a \end{bmatrix} = 0_{4\times 4}$. Logo, $b = a = c = d$.

3. $m_{31} = 0,333, m_{32} = 0,125, m_{33} = 0,200$.

4. Dados os fatores, temos que $\hat{A} = APM$, o que indica operações de permutação e combinação de colunas de A. A coluna 1 de A foi preservada na terceira posição de $\hat{A}$, então a terceira coluna de P é a primeira coluna da identidade. O resultado na primeira linha de $\hat{A}$ é, exceto pelo elemento na posição $1,3$, uma linha de zeros. Para se obter estes zeros foram feitas combinações lineares, usando a coluna pivot. Observe que $(-0.5)A_1 + A_2 = \hat{A}_2$ e, portanto, a segunda coluna de P é a segunda coluna da identidade. Observe também que $(0.5)A_1 + A_3 = \hat{A}_1$ e, portanto, a primeira coluna de P é a terceira da identificade. Então temos as seguintes matrizes: $P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & -1/2 & 1 \end{bmatrix}$

Questão 03: Responda verdadeiro ou falso e justifique.

1. Dados dois subsepaços $\mathcal{V}_1 = span\{[1 \quad 0 \quad 1]^T\}$, $\mathcal{V}_2 = span\{[2 \quad 1 \quad 0]^T\}$, o conjunto $\mathcal{V}_1 \cup \mathcal{V}_2$ é um subespaço.

2. Os autovetores de $A = \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}$ formam uma base para $\mathbb{R}^2$.

3. Se A e B são similares, um autovetor de A é autovetor de B.

4. Para duas matrizes A, B temos que $A^T B = 0$. Então, as colunas de B pertendem a $N(A)$.

5. Ao adicionarmos uma coluna b em A criando uma matriz $[A|b]$, a dimensão do espaço coluna da nova matriz aumenta quando b é linearmente dependente das demais colunas de A.

Resolução da Questão 3:

1. $z = b(x)$; $A = \begin{bmatrix} e^{x_1} & e^{-x_1} \\ e^{x_2} & e^{-x_2} \\ \vdots & \vdots \\ e^{x_m} & e^{-x_m} \end{bmatrix}$.

2. O vetor $b = z$ é projetado em $p \in C(A)$.

3. O sistema linear a ser resolvido é $A^T A\hat{x} = A^T b$, com:

   $A^T A = \begin{bmatrix} \sum_{i=1}^m e^{2x_i} & m \\ m & \sum_{i=1}^m e^{-2x_i} \end{bmatrix}$;

   $\hat{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$;

   $A^T b = \begin{bmatrix} \sum_{i=1}^m e^{x_i} b_i \\ \sum_{i=1}^m e^{-x_i} b_i \end{bmatrix}$.

Questão 04:

1. (50%) Explorando os quatro espaços fundamentais, discuta a existência e unicidade de solução para o sistema linear $Ax = b$ onde $A \in \mathbb{R}^{6\times 4}$ quando:

   (a) posto de A é completo.

   (b) posto de A é incompleto.

2. (50%) Defina a matriz A solicitada que atenda ao estabelecido em cada questão. Justifique quando não for possível.

   (a) Matriz A onde $Ax = [1\ 1\ 1]^T$ tem solução e $A^T[1\ 0\ 0]^T = [0\ 0\ 0]^T$.

   (b) Matriz A onde $[1\ 1\ 0]^T, [0\ 0\ 1]^T \in C(A)$ e $[1\ 2]^T, [2\ 5]^T \in C(A^T)$.

Resolução: Questão 04

1. (a) Quando posto A é completo, as dimensões de $C(A), C(A^T)$ são 4 e de $N(A)$ e $N(A^T)$ são, respectivamente 0 e 2. Ou seja, $N(A) = \{0_4\}$. Então, temos dois casos a considerar:

   - $b \notin C(A)$. Essa é uma condição possível, pois $C(A) \neq \mathcal{R}^6$.
   - O outro caso possível é $b \in C(A)$, ou seja, o sistema admite alguma solução. Suponha então que x_p seja um certificado de pertinência deste fato. Temos $Ax_p = b$ e como qualquer outra solução alternativa do sistema linear pode ser escrita como $x_p + x_n$ onde $x_n \in N(A)$, a única alternativa é $x_n = 0_4$. Logo, temos que quando há solução, a solução é única.

(b) Seja $r < 4$ o posto da matriz A com deficiência de posto. Então as dimensões de $C(A), C(A^T)$ são r e de $N(A), N(A^T)$ são $4 - r > 0$ e $6 - r > 0$, respectivamente. Desta forma, podemos ter o caso em que $b \notin C(A)$ (como explicado acima) e como $N(A) \neq \{0_4\}$, quando há solução para o sistema linear, temos infinitas soluções.

2. Impossível pois o primeiro o vetor (vetor de 1's) pertence a $C(A)$ mas não é ortogonal a $N(A^T)$ (não é ortogonal ao vetor $[1 \ \ 0 \ \ 0]^T$).

3. Existem infinitas matrizes A que atendem ao solicitado, dentre elas segue uma alternativa $A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$.

B.9 Prova 2 - 2024.1

Questão 01: Considere a matriz que possui suas duas primeiras colunas A_1 e A_2 linearmente independentes. Responda justificando.

1. Sabendo que P_1, P_2 são os projetores que projetam em $span\{A_1\}$ e $span\{A_2\}$, defina P_1P_2.

   Resposta:

$$\begin{aligned} P_1 &= \frac{A_1A_1^T}{A_1^TA_1} \\ P_2 &= \frac{A_2A_2^T}{A_2^TA_2} \\ P_1P_2 &= \frac{A_1A_1^T}{A_1^TA_1}\frac{A_2A_2^T}{A_2^TA_2} \\ &= \frac{(A_1^TA_2)}{(A_1^TA_1)(A_2^TA_2)}A_1A_2^T \end{aligned} \tag{B.1}$$

2. Qual é o posto de P_1P_2 ?

   Resposta: A expressão (B.1) mostra que P_1P_2 é o produto de um escalar $\frac{(A_1^TA_2)}{(A_1^TA_1)(A_2^TA_2)}$ por uma matriz $A_1A_2^T$ cujo posto é 1, tem espaço coluna é dado por $span\{A_1\}$ e espaço linha é dado por $span\{A_2\}$. Logo, o posto de P_1P_2 é 1.

3. P_1P_2 é um projetor ? Em caso positivo, é projetor ortogonal ou oblíquo ?

   Resposta: Falso. Para que seja projetor P_1P_2 deve ser idempotente, isto é: $P_1P_2P_1P_2 = P_1P_2$. Então $P_1(P_2P_1)P_2 = P_1P_2$ e portanto $P_2P_1 = I$. Isso é impossível, pois por analogia com a expressão (B.1), $P_2P_1 = \frac{(A_1^TA_2)}{(A_1^TA_1)(A_2^TA_2)}A_2A_1^T$ e também é uma matriz de posto 1. Como I possui posto completo isso não pode ocorrer.

4. Considere agora que P_1, P_2 são os projetores ortogonais que projetam em $span\{A_1\}$ e $span\{A_1, A_2\}$, respectivamente. Indique claramente o resultado de P_2P_1.

   Resposta: $P_2P_1 = P_1$ pois toda coluna de P_1 pertence a $span\{A_1, A_2\}$, espaço em que o projetor P_2 projeta.

Questão 02: A fatoração QR de uma matriz A é apresentada abaixo. Para a fatoração, foi empregado o algoritmo de Gram-Schimid revisado com pivoteamento. Responda, justificando sua resposta com base nos resultados apresentados pela fatoração:

```
Q  =
   0.3552925  -0.4133617
   0.6661734   0.2534585
   0.3108809  -0.7851962
   0.2220578   0.1651537
   0.5329387   0.3479684
R  =
   22.51666  -1.6432277   12.079944   10.436716
   0.         8.2643695  -4.1321847   4.1321847
pivot  =    2.    4.    3.    1.
```

1. Quais as dimensões de A e o posto de A ?

   Resposta: O número de linhas de A é o número de linhas de Q e o número de colunas de A é o número de colunas de R. Portanto, A possui 5 linhas e 4 colunas. A matriz A possui posto 2, que é o número de linhas de R e colunas de Q.

2. Por que $r_{11} \geq r_{22}$?

   Resposta: Foi empregado pivotemento de colunas. Isso significa que a primeira coluna de A que é ortogonalizada é aquela de maior norma Euclideana. A partir daí, na iteração $j > 1$, a coluna de A que é ortogonalizada é, dentre as não ortogonalizadas até então, aquela coluna A_k cuja projeção em $span\{q_1, q_2, \ldots, q_{j-1}\}^{\perp}$ possui maior norma. Isto é, na iteração j escolhe-se a coluna que maximiza $\|A_k - \sum_{i=1}^{j-1}(q_i^T A_k)q_i\|_2$. Por esta razão, sempre ortogonalizamos um vetor cuja norma é não superior à norma dos vetores anteriormente ortogonalizados e estas grandezas (as normas dos erros) são armazenadas na diagonal de R.

3. Escreva as colunas de $A_1, A_2, \ldots$ de A em função das colunas de Q, $q_1, q_2, \ldots$.

   Resposta: $AP = QR$, onde P é uma matriz $n \times n$ $(n = 4)$ de permutação. Para escrever esta expressão de forma conveniente, definimos $A_{pivot(k)}$ como a $pivot(k)-$ésima coluna de A para todo $k = 1, 2, 3, 4$. Considerando que o posto da matriz A é 2, temos

$$A_{pivot(k)} = \sum_{i=1}^{\min\{posto(A), k\}} r_{ik} q_i,$$

que resulta em:

$$\begin{aligned}
A_2 &= 22.51666(q_1) && k = 1\\
A_4 &= -1.6432277(q_1) + 8.2643695(q_2) && k = 2\\
A_3 &= 12.079944(q_1) - 4.1321847(q_2) && k = 3\\
A_1 &= 10.436716(q_1) + 4.1321847(q_2) && k = 4
\end{aligned}$$

4. Como calcularia a fatoração QR completa de A à partir da fatoração apresentada ? Quais informações adicionais esta fatoração produziria e como estas informações seriam organizadas nas novas matrizes de fatores (Q, R, etc...) ?
Resposta:

passo 1 Adicionaria 3 linhas em R pois esta é a dimensão de $N(A^T)$.

passo 2 Encontraria uma base para $N(A^T)$. Vamos definir esta base pelas colunas $\tilde{q}_i : i = 3, 4, 5$, que devem ser encontradas resolvendo-se o sistema linear homogêneo com duas restrições e 5 variáveis (isto é, 3 variáveis livres) definido por $q_1^T \tilde{q}_i = 0, q_2^T \tilde{q}_i = 0$ para $i = 3, 4, 5$.

passo 3 Orgonalizamos as colunas $\tilde{q}_3, \tilde{q}_4, \tilde{q}_5$, usando fatoração $\tilde{Q} = QR$, obtendo q_3, q_4, q_5, 3 colunas ortonormais, base ortonormal para $N(A^T)$.

passo 4 Justapomos as colunas q_1, q_2, q_3, q_4, q_5 formando a nova matriz $Q : 5 \times 5$ desejada.

A fatoração completa informa uma base ortonormal para $C(A)$ (dada pelas primeiras posto(A) colunas de Q) e uma base ortonormal para $N(A^T)$ dada pelas últimas $n - posto(A)$ colunas de Q.

Questão 03: Considere a função $b(x) = \alpha e^x + \beta e^{-x}$ e um conjunto de dados a serem ajustados $\{(x_i, b_i) : i = 1, \ldots, m\}$, satisfazendo $x_i \neq x_j$, $i \neq j$. Responda:

1. No ajuste linear dos dados acima pela função $b(x)$ escolhida, é calculado um vetor de parâmetros $\hat{x}$ que minimiza a norma Euclideana do erro $r(\hat{x}) = z - A\hat{x}$, para A e z correspondentes ao ajuste. Identifique A e z em função dos dados.
Resposta: $A = \begin{bmatrix} e^{x_1} & e^{-x_1} \\ e^{x_2} & e^{-x_2} \\ \vdots & \vdots \\ e^{x_m} & e^{-x_m} \end{bmatrix}$, $z = b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$ já que o sistema é linear nos parâmetros α, β.

2. O processo de identificar $\hat{x}$ que minimiza $r(\hat{x})$ pode ser entendido como um processo de projeção. Considerando os dados $\{(x_i, b_i) : i = 1, \ldots, m\}$ disponíveis, o quê é projetado e onde é projetado ? Seja preciso em função dos dados.
   Resposta: b é projetado em $span\{A_1, A_2\} = C(A)$.

3. Identifique o sistema linear que permite encontrar $\hat{x}$, isto é, defina claramente a matriz de coeficientes e o termo independente do sistema linear, em função dos dados.
   Resposta: O sistema é o sistema de equações normais $A^T A\hat{x} = A^T b$,
   onde $A^T A = \begin{bmatrix} \sum_{i=1}^m e^{2x_i} & m \\ m & \sum_{i=1}^m e^{-2x_i} \end{bmatrix}$, $A^T b = \begin{bmatrix} \sum_{i=1}^m b_i e^{x_i} \\ \sum_{i=1}^m \frac{b_i}{e^{x}i} \end{bmatrix}$

Questão 04: Responda verdadeiro ou falso e justifique.

1. Se P é um projetor e $Pb = b$ então P é a matriz identidade.
   Resposta: Falso. $Pb = b$ é verdadeira quando $b \in C(P)$, mesmo quando $P \neq I$ e portanto P possui posto deficiente.

2. O produto de dois projetores ortogonais é um projetor ortogonal.
   Resposta: Falso, veja o contra-exemplo $P_1 P_2$ dado que questão 1 desta avaliação.

3. Sejam ALG1 e ALG2, algoritmos propostos para fatorar a matriz A em QR. Para ALG1, obtemos a matriz Q_1 e temos $||I - Q_1^T Q_1|| = 10^{-08}$ e para ALG2, obtemos Q_2 e $||I - Q_2^T Q_2|| = 10^{-02}$. ALG2 apresentou melhores resultados numéricos.
   Resposta: Falso. ALG1 produziu uma matriz Q_1 mais próxima de ter colunas de fato ortonormais que ALG2 pois, $\|I - Q_1^T Q_1\|_2 <<< \|I - Q_2^T Q_2\|_2$.

4. Se a fatoração $QR = A$ de uma matriz $A \in \mathbb{R}^{m\times n}, m \geq n$ de posto completo é conhecida, o sistema de equações normais que resolve $\min\|Ax - b\|_2$ pode ser formulado como $Rx = Q^T b$ e pode ser resolvido ao custo de $O(n^2)$ operações elementares.
   Resposta: Verdadeiro.

$$\begin{aligned} A^T Ax &= A^T b \\ (R^T Q^T QR)x &= R^T Q^T b \\ Rx &= Q^T b \qquad\qquad R^{-T} \text{ existe e } Q^T Q = I \end{aligned}$$

   O sistema linear $Rx = Q^T b$ possui ordem n e é triangular superior, podendo ser resolvido em $O(n^2)$ operações aritméticas.

5. Considere que $\hat{x}$ seja uma solução para o Problema de Mínimos Quadrados (PMQ) $\min\|Ax - b\|_2$. Se o posto de A é incompleto, $\hat{x}$ é único e também é único o ponto $p = A\hat{x}$ em $C(A)$ que minimiza a distância de b a $C(A)$.
   Resposta: Falso. O sistema linear $A^TA\hat{x} = A^Tb$ possui infinitas soluções pois $N(A^TA) \neq \{0\}$, a dimensão de $N(A^TA) \geq 1$. Portanto se $\hat{x}$ é uma solução para o sistema de equações normais, $\hat{x} + x_N$ para qualquer $x_N \in N(A^TA)$ também é. O ponto de projeção $A\hat{x}$ é de fato é único, conforme o enunciado da questão.

B.10 Prova 3 - 2024.3

Questão 01: Dois algoritmos ALG1 e ALG2 foram usados para transformar a matriz A abaixo indicada em uma matriz similar B. Após $k = 3$ iterações do algoritmo QR, foram coletadas as aproximações de B abaixo indicadas. Um dos dois algoritmos não empregou uma fase de pre-processamento, anterior ao algoritmo QR. Responda às questões propostas, apresentando suas justificativas.

```
A = -2.    0.   4.  -3.    0.
  -15.   4.   2.   8.    1.
  -8.    2.   7.   3.   -3.
  -15.  -1.   2.   10.   1.
  -11.   2.   1.   4.    2.
ALG1
11.535344  -15.814475  -0.0098925   17.542151  -12.039947
   0.2484158   6.8120458   0.7182519   2.3324697   1.2548657
   0.0071437   1.0802719   4.9441015   3.6996861  -1.720255
   0.078128   -0.4117166   1.0261165  -1.6331624   3.8166157
   0.0003491   0.0022462  -0.0009539  -0.0158491  -0.6583285
ALG2
   11.535344  -9.8357265   17.38408   -12.453403   12.196314
   0.2605102   6.6893101  -0.6741826  -0.6960549  -2.2805135
   5.672D-17  -3.6360701  -1.2419462   3.4176443  -3.7334895
  -2.362D-16   1.152D-14   1.0832425   4.6946447   0.3735902
   2.337D-18  -1.143D-16  -3.126D-16  -0.0459834  -0.6773523
```

1. Em que consiste a fase de pré-processamento ?

   Resposta: A fase de pré-processamento consiste em transformar a matriz A em uma Hessenberg superior (isto é uma triangular superior + subdiagonal abaixo da principal), para acelerar a Fase II, de natureza iterativa, o algoritmo QR.

2. Algum algoritmo empregou a fase de pré-processamento ? Em caso positivo, qual ?

   Resposta: Sim, os resultados numéricos indicam que ALG2 empregou a transformação de A em uma Hessenberg H, pois as entradas abaixo da subdiagonal inferior possuem magnitude bastante inferior às demais entradas da matriz coletada após 3 iterações de QR. Com isso, com as mesmas $k = 3$ iterações, ALG2 produziu uma matriz mais próxima de uma triangular superior que ALG1. Isso é o que se espera do efeito da fase de pré-processamento.

3. Para k suficientemente grande, qual a forma da matriz B retornada pelos algoritmos ?

   Resposta: A matriz A não é simétrica, portanto o algoritmo QR produzirá uma matriz assintoticamente triangular superior. Caso A fosse simétrica, B seria assintoticamente diagonal.

4. Sabendo que os autovalores da matriz A são -2.135676, -0.7025164, 10.424842, 8.3359461 e 5.0774047, quantos autovetores linearmente independentes A possui ?

   Resposta: Como os autovalores informados são distintos, A é garantidamente não-defectiva (a multiplicidade geométrica e algébrica de cada autovalor é a mesma), portanto a matriz possui $n = 5$ autovetores linearmente independentes. Em resumo, a matriz não defectiva possui autovetores que geram o $\mathbb{R}^n$.

5. Os algoritmos ALG1 e ALG2 também podem apresentar os autovetores de A ?

   Resposta: Como a matriz A não é simétrica, a fatoração computada por ALG1 e ALG2 é a faturação de Schur e apenas o autovetor associado a λ_1 será produzido. Ou seja, assintoticamente teremos $QAQ^* = T$, onde T e uma triangular superior. Para o cálculo dos demais autovetores, é necessário esforço computacional adicional, empregando-se outros métodos, à partir dos autovalores já calculados.

Questão 02: Sobre as matrizes reais, responda verdadeiro ou falso, e justifique.

1. Toda matriz quadrada é unitariamente similar a uma matriz diagonal.

   Resposta: Falso. As matrizes unitárias A (aquelas em que se observa $AA^* = A^*A$) são as matrizes unitariamente similares a diagonais. Nem toda matriz A quadrada é ortogonalmente similar a uma diagonal. Um exemplo já mencionado aqui é das matrizes defectivas, uma vez que seus autovetores não fornecem uma base para $\mathbb{R}^n$.

2. A matriz $A \in \mathbb{R}^{2\times 2}$, com os autovalores $\lambda_1 = 5 - i$ e $\lambda_2 = 5 + i$ é simétrica.

   Resposta: Falso. É propriedade das matrizes reais simétricas admitirem autovalores (e autovetores) reais.

3. Seja A uma matriz $n \times n$, defectiva. Os autovetores de A geram uma base para $\mathbb{R}^n$.

   Resposta: Falso. Matrizes defectivas não são similares a uma diagonal. Portanto, os autovetores linearmente independentes não são suficientes para escrever o espaço.

4. Dada A^+ a pseudoinversa de A, os valores singulares de AA^+A são iguais a $\sigma_1^3, \sigma_2^3, \ldots, \sigma_r^3$, quando o posto de A é igual a r.

   Resposta: Falso. Os valores singulares de AA^+A são os próprios valores singulares de A: $AA^+A = U\Sigma V^T V\Sigma^+ U^T U\Sigma V^T = U\Sigma\Sigma^+\Sigma V^T = U\Sigma V^T$.

5. Seja $A \in \mathbb{R}^{m\times n}$ de posto n. Não é possível subtrair de A uma matriz de posto 1 de forma que a nova matriz, B, satisfaça $posto(B) = posto(A) - 1$.
   Resposta: Falso. Considere a fatoração SVD de $A = \sum_{i=1}^{n} \sigma_i u_i v_i^T$. Escolha qualquer $k : 1 \leq k \leq n$ e faça $B = A - \sigma_k u_k v_k^T$, de forma que $B = \sum_{i=1, i\neq k}^{r} \sigma_i u_i v_i^T$ e veja que B é a soma de $n-1$ matrizes de posto 1, portanto possui posto $n-1$.

Questão 03: A matriz R dada é ortogonalmente equivalente à matriz A, retangular. Responda às questões colocadas justificando. Não é necessário calcular explicitamente as entradas das matrizes pedidas. Basta indicar como são construídas.

$$R = \begin{bmatrix} 3 & -1 & 0 & 2 \\ 0 & 5 & 2 & 3 \\ 0 & 0 & 4 & -1 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

1. Na Fase I do algoritmo que computa a fatoração SVD de $A = U\Sigma V^T$, é necessária a aplicação de uma ou mais operações ortogonais na matriz R dada, de forma que seja ortogonalmente equivalente a uma bidiagonal. Quantas transformações ortogonais adicionais são indispensáveis para a conclusão dessa fase e, assim, obter uma matriz Z apropriada, para acelerar a fase subsequente ?

   Resposta: Como a matriz R é triangular superior, as transformações ortogonais que relacionam Z e R são $Z = ERD$, onde $E = I$ não é necessária. Portanto, apenas $n-2 = 2$ matrizes D_1, D_2 são necessárias, de forma que $D = D_1 D_2$, $D_i^T D_i = I$ para $i = 1, 2$.

2. Qual é a matriz ortogonal empregada na primeira transformação ?

Resposta: A matriz $D_1 \in \mathbb{R}^{4\times 4}$ possui a forma $D_1 = \begin{bmatrix} 1 & 0_{1\times 3} \\ 0_{3\times 1} & F_1 \end{bmatrix}$ e $F_1 \in \mathbb{R}^{3\times 3}$ é o refletor de Householder $F_1 = I_3 - 2\frac{vv^T}{v^Tv}$ para $v = [-1, 0, 2]^T - \sqrt{5}[1, 0, 0]^T$.

3. Qual a forma da matriz Z obtida ?

   Resposta: Como mencionado no enunciado, $Z = RD_1D_2$ é bidiagonal, após a aplicação do pré-processamento.

4. A fatoração SVD de R e Z são idênticas ?

   Resposta: Não, Z e R possuem os mesmos valores singulares, e as fatorações SVD se relacionam da seguinte forma: se $Z = U_Z\Sigma_Z V_Z^T$, $U_R = U_Z$, $\Sigma_R = \Sigma_Z$ e $V_R^T = V_Z^T D^T$.

5. Considere a matriz $H(X) = \begin{bmatrix} 0 & X^* \\ X & 0 \end{bmatrix}$. As matrizes $H(Z)$ e $H(R)$ são similares ? Essas matrizes possuem os mesmos autovetores ? Justifique.

   Resposta: Sim, são similares pois $H(R) = XH(Z)X^{-1}$ onde $X = \begin{bmatrix} D & 0 \\ 0 & I \end{bmatrix}$ admite inversa $X^{-1} = X^T = \begin{bmatrix} D^T & 0 \\ 0 & I \end{bmatrix}$. As duas matrizes $H(Z), H(R)$ possuem os mesmos autovalores $\sigma_i, -\sigma_i$, mas suas fatorações espectrais são distintas, ou seja não possuem os mesmos autovetores. Sejam $\begin{bmatrix} v_i \\ u_i \end{bmatrix}$ e $\begin{bmatrix} v_i \\ -u_i \end{bmatrix}$ os autovetores de $H(Z)$ associados a $\lambda_i = \sigma_i, \lambda_{i+n} = -\sigma_i$. Os autovetores de $H(R)$ associados a $\sigma_i, -\sigma_i$ são $\begin{bmatrix} Dv_i \\ u_i \end{bmatrix}$ e $\begin{bmatrix} Dv_i \\ -u_i \end{bmatrix}$.

Questão 04: Considere $v \in \mathbb{R}^2$. **Resposta:** Considere a figura:

f1f2a.jpg

Figura B.1: Representação geométrica de F_1F_2a.

1. (35%) Quais os refletores de Householder F_1, F_2 cujos hiperplanos de reflexão são $span\{v\}^\perp$ e $span\{v\}$, respectivamente ?

   Resposta: Sem perda de generalidade, podemos assumir que $\|v\|_2 = 1$. Sendo v o construtor do refletor, a reflexão definida por $F_1 = I - 2vv^T$ se dá sobre $span\{v\}^\perp$. Tomando $w \in span\{v\}^\perp$, ou seja $w^Tv = 0$ como o construtor para o segundo refletor, também com norma unitária, temos que $F_2 = I - 2ww^T$ é o refletor sobre $span\{v\}$.

2. (35%) Quais são os projetores P_1, P_2 que projetam a em $span\{v\}^\perp$ e $span\{v\}$, respectivamente ?
   Resposta: $P_2 = vv^T$. $P_1 = I - P_2 = I - vv^T$ (Observação: $I - vv^T = ww^T$).

3. (30%) Dado $a \in \mathbb{R}^2$, qual é o resultado de F_1F_2a ? Justifique.
   Resposta: $F_1F_2a = -a$. Mostramos este resultado por dois caminhos: algébrico e geométrico.
   $\Rightarrow$ Demonstração algébrica: Veja:

$$\begin{aligned}F_1F_2 &= (I - 2vv^T)(I - 2ww^T)\\ &= I - 2vv^T - 2ww^T + 4vv^Tww^T\\ &= I - 2vv^T - 2ww^T\end{aligned}$$

Vamos mostrar que $vv^T + ww^T = I$ para esse caso do $\mathbb{R}^2$. Como qualquer vetor $z \in \mathbb{R}^2$ pode ser escrito como $z = \alpha v + \beta w$, temos que

$$\begin{aligned}(vv^T + ww^T)z &= (vv^T + ww^T)(\alpha v + \beta w)\\ &= vv^T(\alpha v) + (ww^T\beta w)\\ &= \alpha v + \beta w\\ &= Iz\end{aligned}$$

Como z é qualquer, $vv^T + ww^T = I$ e, portanto, $F_1F_2a = (I - 2vv^T - 2ww^T)a = -a$.

Observe que esse resultado não é válido para o $\mathbb{R}^n$, pois uma base para $span\{v\}^\perp$ não é composta por apenas um vetor w ortogonal a v.

$\Rightarrow$Para a "demonstração geométrica", veja a Figura B.1.

B.11 Prova 1 - 2024.2

Questão 01: A matriz A foi fatorada na forma $A = ZDZ^{-1}$ onde D é uma matriz diagonal 4×4 com as seguintes entradas na diagonal, $a, b, 0, 0$, nesta ordem, onde $a > b > 0$. Assuma que $C(A), N(A), C(A^T), N(A^T)$ são os espaços associados à A, que z_i^T, z_i^{-T} sejam as linhas de Z, Z^{-1} respectivamente e que Z_i, Z_i^{-1} sejam as colunas de Z, Z^{-1}, respectivamente. Para todas as questões, justifique sua resposta.

1. (20%) Qual é o posto de A ?
   A fatoração revela o posto. O posto é o número de entradas não nulas (autovalores não nulos) na diagonal de D, 2.

2. (20%) Quais as dimensões dos quatro espaços fundamentais ?
   Como a matriz é 4×4 e o posto é 2, todos os espaços possuem dimensão $2 = 4 - 2$.

3. (20%) Caracterize $C(A), N(A)$ (apresente bases para).
   $C(A) = span\{Z_1, Z_2\}$ e $N(A) = span\{Z_3, Z_4\}$. A matriz é diagonalizável e os autovetores associados aos autovalores não nulos geram $C(A)$ enquanto os autovetores associados ao autovalor $\lambda = 0$ (caso exista) geram $N(A)$.

4. (20%) Caracterize o subespaço $N(A - aI)$.
   $N(A-aI)$ é o autoespaço associado ao autovetor cujo autovalor é a. Portanto, $N(A - aI) = span\{Z_1\}$.

5. (20%) E possível afirmar que $N(A) = N(D)$?
   Não é possível, pois $N(A) = span\{Z_3, Z_4\}$ enquanto que $span\{e_3, e_4\} = N(D)$.

Questão 02: Considere a matriz $A = \begin{bmatrix} a^2 & 0 & 0 \\ 0 & c^2 & cd \\ 0 & dc & d^2 \end{bmatrix}$, onde $a, c > 0$. Responda justificando.

1. Apresente uma fatoração de A que revele seu posto r, na forma $A = \sum_{j=1}^{r} L_j L_j^T$ onde as colunas L_j satisfazem $L_{jj} > 0$ para toda coluna $j = 1, \ldots, r$.
   Por exemplo, via Cholesky, temos: $A = \begin{bmatrix} a & 0 \\ 0 & c \\ 0 & d \end{bmatrix} \begin{bmatrix} a & 0 & 0 \\ 0 & c & d \end{bmatrix}$, que revela o posto $r = 2$ de A. Logo $A = \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} a & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c \\ d \end{bmatrix} \begin{bmatrix} 0 & c & d \end{bmatrix}$.

2. Qual é o valor do posto r ?
   $r = 2$, veja a fatoração acima.

3. Defina bases para $C(A), C(A^T)$ e $N(A), N(A^T)$ à partir de a, c, d. A matriz A é simétrica, logo $C(A) = C(A^T) = span\left\{\begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ c \\ d \end{bmatrix}\right\}$. Da mesma forma, $N(A) = N(A^T) = span\left\{\begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ c \\ d \end{bmatrix}\right\}^{\perp}$. Usando este fato, encontramos uma base para $N(A)$ resolvendo o sistema linear homogêneo $\begin{bmatrix} a & 0 & 0 \\ 0 & c & d \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = 0$. Por exemplo, $y = \begin{bmatrix} 0 \\ -\frac{d}{c} \\ 1 \end{bmatrix}$.

Questão 3 Responda verdadeiro ou falso e justifique.

1. Os autovetores de uma matriz $A \in \mathbb{R}^{n \times n}, A = ZDZ^{-1}$, similar a uma matriz diagonal D, fornecem uma base para $\mathbb{R}^n$.
   Verdadeiro, $A = ZDZ^{-1} \iff AZ = ZD$ e Z possui colunas linearmente independentes, todas elas autovetores de A, com autovalores correspondentes na diagonal de D.

2. Considere o conjunto $C = \{(x, t) : \|x\|_2 \leq t\} \subseteq \mathbb{R}^{n+1}$ e a Figura B.2 que ilustra o caso em que $x \in \mathbb{R}^2$. O conjunto C define um subsepaço vetorial ? Em caso positivo, apresente uma base para C.
   Falso, não é fechado à multiplicação por escalar. Tome $(\hat{x}, \hat{t}) \in C$ e $\alpha < 0$. Então $\alpha(\hat{x}, \hat{t}) = (\alpha\hat{x}, \alpha\hat{t})$. $\|\alpha\hat{x}\| = |\alpha|\|\hat{x}\| > \alpha\hat{t}$ pois $\alpha\hat{t} < 0$. Logo, $(\alpha\hat{x}, \alpha\hat{t}) \notin C$.

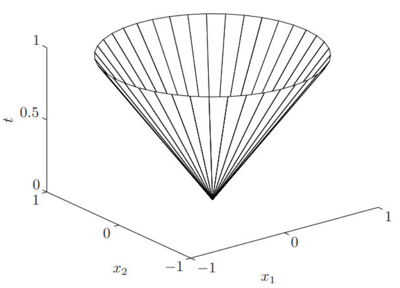

Figura B.2: Ilustração do conjunto C quando $x \in \mathbb{R}^2$.

3. Se $A = Q\Lambda Q^T$ onde Q é ortogonal e Λ diagonal, então $det(\Lambda) \neq 0$.
   Falso, se A singular, $det(A) = det(\Lambda) = 0$, pois Λ e A possuem pelo menos um autovalor nulo.

4. Se a fatoração $A = Q\Lambda Q^T$ da matriz quadrada de posto completo A é disponível (onde Q é ortogonal e Λ é uma matriz diagonal), a inversa de A^k é $A^{-k} = Q\Lambda^{-k}Q^T$, para algum $k \geq 1$ inteiro.
   Verdadeiro. $A^k A^{-k} = (Q\Lambda^k Q^T)(Q\Lambda^{-k}Q^T) = I$. Observe que Λ^{-k} existe pois A possui posto completo.

5. O vetor $v^T = (1, 0, -1)$ pode ser uma linha de uma matriz A e também pertencer à $N(A)$.
   Falso, pois $C(A^T) \perp N(A)$ e $v \neq 0 = C(A^T) \cap N(A)$.

Questão 04 Resolva as questões abaixo, justificando e dizendo se são verdadeiras ou falsas.

1. Se a terceira coluna de uma matriz B é um vetor de zeros, a terceira coluna de EB será um vetor de zeros, para qualquer E.
   Verdadeiro, pois a terceira coluna de EB será a soma de zero $\times$ cada coluna de E.

2. Se a terceira linha de B é toda de zeros, a terceira linha de EB pode não ser um vetor de zeros.
   Verdadeiro. A única observação que pode ser feita é que a terceria coluna de E não terá efeito no produto EB. Em particular, a terceira linha de EB pode ter elementos distintos de zero, a depender dos outros pesos em B e das outras colunas de E. Um exemplo em que isso pode ocorrer é: E é uma matriz de 1's e B uma matriz de 1's também, exceto pela sua terceira linha.

3. A matriz A foi fatorada em $A = UPV$, onde U é uma matriz que tem como colunas U_1, U_2, U_3, U_4 e V tem como linhas $v_1^T, v_2^T, v_3^T, v_4^T$ e $P = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$.
   Seria possível apresentar a matriz A como uma soma de matrizes de posto 1 ? Em caso negativo, justifique. Em caso positivo, escreva a soma.
   Sim, é possível. Fazendo $A = (UP)V$ ou $A = U(PV)$ (permutando as colunas de U ou as linhas de V) temos $A = u_4v_1^T + u_3v_2^T + u_2v_3^T + u_1v_4^T$.

B.12 Prova 2 - 2024.2

Questão 01: A matriz R na fatoração $AP = QR$ completa da matriz $A \in \mathbb{R}^{5\times 5}$ de posto $r = 2$ é apresentada abaixo. Sabe-se que $Q = [\overline{Q}, \hat{Q}]$, onde $\overline{Q} = [q_1, q_2] \in \mathbb{R}^{5\times 2}$, $\hat{Q} = [q_3, q_4, q_5] \in \mathbb{R}^{5\times 3}$, $Q^T Q = I_5$, $pivot = (3, 2, 1, 5, 4)^T$.

```
R  =

   19.33908  -16.598515  -9.1214269  -2.6268055  -2.7638337
   0.          5.5217121   1.1043424   0.5521712   0.2760856
   0.          0.          7.238D-16   2.469D-16  -2.959D-16
   0.          0.          0.         -2.555D-17   2.555D-17
   0.          0.          0.          0.          1.253D-16
```

Responda às questões abaixo, apresentando sua justificativa.

Observações gerais sobre o enunciado e os dados da questão: A fatoração $AP = QR$ completa fornece nas primeiras $r = posto(A) = 2$ colunas de Q uma base para $C(A) = C(AP)$ e em suas últimas $m - posto(A)$ colunas uma base para $C(A)^\perp = N(A^T)$. Portanto $\overline{QQ}^T$ é um projetor que projeta em $C(A)$ e seu projetor complemento ortogonal $(I - \overline{QQ}^T)$ projeta $N(A^T) = C(A)^\perp$.
Resumindo: $\hat{Q}\hat{Q}^T$ projeta em $span\{q_3, q_4, q_5\} = N(A^T)$.
$\overline{QQ}^T$ projeta em $C(AP) = C(A) = span\{q_1, q_2\}$.

1. Se $b = 2q_1 + q_4$, $\|(I - \overline{QQ}^T)b\|_2^2$=?
   Pelo explicado acima, $q_1 \perp span\{q_3, q_4, q_5\}$. Logo:
$$\begin{aligned} \hat{Q}\hat{Q}^T b &= \\ &= \hat{Q}\hat{Q}^T(2q_1 + q_4) \\ &= 2\hat{Q}(\hat{Q}^T q_1) + \hat{Q}\hat{Q}^T q_4 \\ &= 0_5 + q_4 \\ \|q_4\| &= 1 \end{aligned}$$

2. $\hat{Q}\hat{Q}^T q_3$ =?
   Como mencionamos acima $\hat{Q}\hat{Q}^T$ projeta no $C(A)^\perp = N(A^T)$. E como q_3 é um dos elementos da base de $N(A^T)$, $\hat{Q}\hat{Q}^T q_3 = q_3$.

3. $\hat{Q}\hat{Q}^T q_1$ =?
   q_1 é ortogonal ao espaço onde $\hat{Q}\hat{Q}^T$ projeta, logo $\hat{Q}\hat{Q}^T q_1 = 0_5$.

4. Se $v = A_3 + A_2$, $\|v\|_2^2$=?

$$\begin{aligned} v &= A_3 + A_2 \\ &= (r_{11}q_1) + (r_{12}q_1 + r_{22}q_2) \\ &= (r_{11} + r_{12})q_1 + r_{22}q_2 \end{aligned}$$

Como $q_1 \perp q_2$ e ambos têm norma Euclideana unitária, temos $\|v\|_2^2 = (r_{11} + r_{12})^2 + r_{22}^2$ (não é necessário ir além disso e realizar os cálculos). Mas, fazendo-os temos: $\|v\|_2^2 = (19.33908 - 16.598515) + 5.5217121 \approx 38.0$.

5. $q_2q_2^TA_5$ =?

$$\begin{aligned} A_5 &= r_{14}q_1 + r_{24}q_2 \\ q_1 &\perp q_2 \\ q_2q_2^TA_5 &= q_2q_2^T(r_{14}q_1 + r_{24}q_2) \\ &= r_{24}q_2(q_2^Tq_2) \\ &= 0.5521712q_2 \end{aligned}$$

6. Qual o posto de $I - \hat{Q}\hat{Q}^T$?
$I - \hat{Q}\hat{Q}^T = \overline{QQ}$ projetor que projeta em $C(A)$ que tem dimensão = 2. Logo posto de $I - \hat{Q}\hat{Q}^T$ é 2.

7. $(I - \overline{QQ}^T)(\hat{Q}\hat{Q}^T)$ é um projetor ? Em caso positivo, é ortogonal, qual seu posto e onde projeta ?

$$\begin{aligned} (I - \overline{QQ}^T)(\hat{Q}\hat{Q}^T) &= \hat{Q}(\hat{Q}^T\hat{Q})\hat{Q}^T \\ &= \hat{Q}\hat{Q}^T \end{aligned}$$

Observe que demonstramos acima a idempotência de $\hat{Q}\hat{Q}^T$. Portanto, é projetor, que projeta em $N(A^T)$, é ortogonal pois $\hat{Q}\hat{Q}^T$ é simétrica e seu posto é 3, a dimensão de $N(A^T)$.

8. $span\{q_3, q_4, q_5\} = C(A)$?

Falso $span\{q_3, q_4, q_5\} = C(A)^{\perp} = N(A^T)$.

Questão 02 Considere o algoritmo abaixo e assuma que A é uma matriz $m \times n, m \geq n$. Em qualquer iteração do algoritmo, assuma que A_k é a $k-$ésima coluna de A atualizada.

- Para $j = 1, \ldots, n$, faça a atualização da matriz A segundo:

$$A \leftarrow A \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \frac{1}{\|A_j\|_2} & \frac{-A_j^T A_{j+1}}{\|A_j\|_2^2} & \ldots & \frac{-A_j^T A_n}{\|A_j\|_2^2} \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

1. O que o algoritmo faz, assumindo que sua execução seja bem sucedida ?
   O algoritmo ortogonaliza a matriz A fazendo n transformações lineares triangulares superiores, no espírito do algoritmo Gram-Schmidt (GS) revisado. Ou seja, a cada iteração j, normaliza o vetor armazenado em A_j e desconta das demais colunas de A (atualizada, não a matriz A original que é substituída por sua ortogonalização) a projeção destas colunas no $span\{A_j\}$. Veja que na diagonal da matriz triangular superior que multiplica a matriz A atualizada a cada iteração j temos a entrada $\frac{1}{\|A_j\|}$ que corresponde ao inverso da entrada r_{jj} de R na fatoração QR via GS. Já a entrada na coluna k da linha j da matriz triangular superior corresponde a $\frac{-A_j^T A_k}{\|A_j\|_2^2}$. Na iteração j, a coluna $A_k : k \geq j+1$ armazena a coluna k inicial em A (antes da ortogonalização) descontada de todas as projeções em $q_1, q_2, \ldots, q_{j-1}$. Então, o termo $\frac{-A_j^T A_k}{\|A_j\|_2^2}$ só faz descontar da coluna A_k a sua projeção em $q_j = \frac{A_j}{\|A_j\|}$ na iteração j do algoritmo dado. Veja que $\frac{-A_j^T A_k}{\|A_j\|_2^2} A_j$ equivale a $\frac{-A_j^T A_k}{r_{jj}} \frac{A_j}{r_{jj}} = -(q_j^T A_k) q_j$ que é a parcela de A_k original da matriz A relativa a $span\{q_j\}$.

2. O algoritmo acima possui alguma condição de falha ? Em caso positivo, o que esta condição de falha caracteriza ? O algoritmo possui alguma restrição de uso ?
   Sim, não é capaz de lidar com deficiência de posto, pois quando r_{jj} for muito pequeno, comparado às demais entradas em R, temos a indicação de dependência linear. Este problema é resolvido incorporando pivoteamento de colunas.

3. Este algoritmo se assemelha a algum algoritmo visto durante o curso ? Em caso positivo, qual algoritmo ? Caracterize estas semelhanças.
   A menos do armazenamento das entradas do fator R que não é explicitado, o algoritmo acima é Gram-Schmidt revisado. Semelhanças:
   - Normalização de A_j a cada iteração.
   - Desconta das demais colunas de A, de índice $j+1,\ldots,n$ (atualizada, não a matriz A original que é substituída por sua ortogonalização) a projeção destas colunas no $span\{A_j\}$ que acabou de produzir a coluna q_j na fatoração.

Questão 03: Considere o problema de mínimos quadrados (PMQ) $\min_x \|QRx-b\|_2^2$, sabendo que $A = QR$, $Q \in \mathbb{R}^{m\times n}, Q^TQ = I_n$, $R \in \mathbb{R}^{n\times n}$ é uma triangular superior, satisfazendo $r_{ii} \neq 0$ para todo $i = 1,\ldots,n$.

1. Há solução para o PMQ ? Em caso positivo, apresente a solução e discuta sua unicidade.

   A matriz A possui posto n, completo, pois sua fatoração $A = QR$ é tal que todos os elementos na diagonal de R são não nulos. Sabemos que nesse caso, o sistema equações normais admite solução única. Vamos mostrar isso. Buscar um x que minimize $\min_x \|QRx-b\|_2^2$ equivale a buscar um y que minimize $\min_y \|Qy-b\|_2^2$, definindo-se $y = Rx$. Recorde-se que $C(Q) = C(A)$. Então, o sistema de equações normais em y pode ser escrito como $Q^TQy = Q^Tb$ cuja solução é $y = Q^Tb$. Portanto, $Rx = y = Q^Tb$ e $x = R^{-1}Q^Tb$ é a solução única de PMQ, pois R admite inversa.

2. Assuma agora que $Q \in \mathbb{R}^{m\times m}$ onde $N(Q^T) = \{0_m\}$ e $R \in \mathbb{R}^{m\times m}$, satisfazendo $r_{11} \geq r_{22} \geq \cdots r_{mm} > 0$. O PMQ admite solução ? Em caso positivo, qual é a solução ? Obtenha uma expressão para $\|QRx-b\|_2^2$ neste caso.
   Nesse caso, $C(Q) = \mathbb{R}^m$, a matriz $A : m \times m$ possui posto completo m. Portanto, qualquer $b \in \mathbb{R}^m$ é combinação linear das colunas de A ou de Q, respectivamente com pesos x ou y adequados. Logo, $\|QRx-b\|_2^2 = 0$ e a solução do PMQ é, na verdade, a solução única do sistema linear $QRx = b$, $x = R^{-1}Q^Tb$.

Questão 04: Responda verdadeiro ou falso e justifique.

1. Uma matriz é perfeitamente condicionada quando seu número de condição, para alguma norma matricial induzida qualquer, é inferior à unidade.
   Falso. O número de condição de uma matriz nunca é inferior à unidade, que é o número de condição de uma matriz identidade ou de uma indentidade por um escalar.

2. Para uma matriz P ser uma matriz de projeção ortogonal, a única condição é $P = P^T$.
   Falso, um projetor precisa ser idempotente, isto é, satisfazer $P = P^2$. Para ser ortogonal, é necessário ser uma matriz simétrica.

3. A matriz $P = Q \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix} Q^T$ para $Q^TQ = I$ é um projetor ortogonal.
   Falso, depende dos valores de a, b que, como sabemos são os autovalores de P (por similaridade). Sabemos que um projetor possui autovalores 1 ou 0. Veja o desenvolvimento a seguir, que complementa a resposta dada até agora. Sendo q_1, q_2 as duas colunas ortonormais de Q, temos que $P = aq_1q_1^T + cq_2q_2^T$. Como $q_1 \perp q_2$, $P^2 = a^2q_1q_1^T + c^2q_2q_2^T$. Portanto, para que seja projetor ($P^2 = P$) $a, c \in \{0, 1\}$.

4. Se $Q_AR_A = A^TP_A$ e $Q_BR_B = B^TP_B$ são fatorações QR reduzidas, P_A, P_B são matrizes de permutação e $C(A^T) \subseteq C(B^T)$ então $Q_AQ_A^TQ_BQ_B^T = Q_AQ_A^T$.
   Resolução: Afirmativa verdadeira. Vamos assumir que o posto de A^T é r e de B^T é $r+1$ para que $C(A^T) \subset C(B^T)$, $C(B^T) \neq C(A^T)$, ou seja, para que haja o pertencimento estrito. Podemos assumir que as r primeiras colunas de B^T gerem o mesmo espaço das primeiras r colunas de A^T e que A^T possua exatamente r colunas. Assim sendo, vamos assumir que $Q_A = [q_1, \ldots, q_r]$ e

$Q_B[q_1, \ldots, q_r, q_{r+1}]$, com $q_{r+1} \perp span\{q_1, \ldots, q_r\}$. Então temos:

$$\begin{aligned}
Q_A Q_A^T Q_B Q_B^T &= (\sum_{i=1}^{r} q_i q_i^T)(\sum_{j=1}^{r+1} q_j q_j^T) \\
&= (\sum_{i=1}^{r} q_i q_i^T)(\sum_{i=1}^{r} q_j q_j^T + q_{r+1} q_{r+1}^T) \\
&= (\sum_{i=1}^{r} q_j q_j^T)(\sum_{i=1}^{r} q_j q_j^T) + (\sum_{i=1}^{r} q_j q_j^T)(q_{r+1} q_{r+1}^T) \\
&= Q_A Q_A^T + \sum_{i=1}^{r} (q_i q_i^T q_{r+1} q_{r+1}^T) \\
&= Q_A Q_A^T + \sum_{i=1}^{r} q_i (q_i^T q_{r+1}) q_{r+1}^T \\
&= Q_A Q_A^T
\end{aligned}$$

B.13 Prova 3 - 2024.2

Questão 01: Considere $w \in \mathbb{R}^n$, $\|w\|_2 = 1$. Responda justificando sua resposta.

1. Apresente uma fatoração QR reduzida de $\alpha w w^T$, onde $\alpha \neq 0$.

   A fatoração QR reduzida apresenta uma base ortonormal para o column space da matriz, e o vetor w define esta base. Então $Q = w$, $R = \alpha w^T$ é uma fatoração. Se $\alpha < 0$, podemos ter $R = |\alpha|(-w^T)$. Chegamos à mesma conclusão empregando o algoritmo de Gram-Schmidt.

2. Quais são os autovalores de $\alpha w w^T$? Caso haja algum autovalor distinto de zero, indique qual é o autovetor associado.

   ww^T é um projetor ortogonal (matriz simétrica) e como tal admite autovalores 1 (associados a uma base para o espaço no qual projeta) e 0 associados ao seu espaço nulo (lembre-se que o projetor é simétrico). Então, supondo que (λ, y) seja um autopar de $\alpha w w^T$, com $\lambda \neq 0$, podemos assumir que $\|y\|_2 = 1$ e temos:

   $$\begin{aligned}(\alpha w w^T) y &= \lambda y \\ w(\alpha w^T y) &= \lambda y\end{aligned}$$

   Portanto, autovetores associados a autovalores distintos de zero λ são os próprios w. Observando que $w^T y = 1$ temos:

   $$w(\alpha) = (\lambda) w$$

   e portanto, o único autovalor distinto de zero é $\lambda = \alpha$, associado ao autovetor w.

3. Quais são os autovalores do refletor $F = I - 2ww^T$? Qual a multiplicidade algébrica dos autovalores ?

   Lembrando que os autovalores $\lambda(A)$ de $A = B + Is$ são $\lambda(B) + s$, temos que os autovalores de F são $\lambda(F) = 1 + \lambda(-2ww^T)$.
   Pelo resultado acima, temos que $\lambda(F)$ são $1 + \{0, -2\}$ e portanto F admite

$\{1, -1\}$ como autovalores. A multiplicidade do autovalor 1 de F é a multiplicidade do autovalor 0 para $-2ww^T$, que é $n-1$. A multiplicidade do autovalor -1 de F é 1, pois é a multiplicidade do autovalor -2 de $-2ww^T$.

4. É possível dizer que o refletor F seja similar a uma matriz diagonal com determinante distinto de zero ? Justifique.
   Sim, F é uma matriz simétrica, e portanto ortogonalmente similar a uma diagonal. Isto é, admite uma fatoração espectral.

5. Mostre que para qualquer $u \in span\{w\}^\perp$, F funciona como a matriz identidade e que para $u \in span\{w\}$, como o simétrico da identidade.

   Tomando $u = \beta w$, temos $(I - 2ww^T)(\beta w) = \beta w - 2\beta w(w^T w) = -\beta w = (-I)(\beta w)$. Tomando $z \in span\{w\}^\perp$, temos $(I-2ww^T)z = Iz-2w(w^T z) = Iz$.

Questão 02: Considere $w \in \mathbb{R}^n, \|w\|_2 = 1$, $F = I - 2ww^T$ e $P = ww^T$. Assuma que a matriz $(FP)^k$ para $k \geq 1$ inteiro corresponda ao produto de k fatores FP.

Observação para todas as questões:

$$\begin{aligned} FP &= (I - 2ww^T)(ww^T) \\ &= ww^T - 2w(w^T w)w^T \\ &= ww^T - 2ww^T \\ &= -ww^T \\ &= -P \end{aligned}$$

1. Caracterize $C(FP)$ (apresente uma base para), indicando também a dimensão do espaço.
   $C(FP) = span\{w\}$, pois $FP = -ww^T$ é uma fatoração que revela o posto 1 para $FP = -ww^T$.

2. Qual é o resultado de $(FP)^k$ para todos os possíveis valores de k ?

   Observe que $(FP)^2 = (-ww^T)(-ww^T) = ww^T$. Portanto, para qualquer $k = 2p$ para p inteiro (isto é para k par) temos que $(FP)^{2p}$ corresponde ao produto de p fatores $(FP)^2$ e assim $(FP)^{2p} = ww^T$.

   Por outro lado, para qualquer $k = 2p + 1$ (isto é para k ímpar), temos $(FP)^{2p+1} = (FP)^{2p}(FP) = ww^T(-ww^T) = -ww^T$.

3. FP é uma matriz normal ?

   Sim. Uma matriz A é normal se e somente se $AA^T = A^TA$. Como $FP = -ww^T$ é uma matriz simétrica $(FP)(FP)^T = FP^2 = (FP)^T(FP)$. Portanto, FP é uma matriz normal.

4. FP é projetor ?
   Não. $FP = -ww^T$ enquanto $(FP)^2 = ww^T$. Como não se verifica a idempotência, não é projetor.

5. FP é um refletor de Householder ?

   Uma matriz X é um refletor de Householder se existe $u \in \mathbb{R}^n$, $\|u\|_2 = 1$ tal que $X = I - 2uu^T$. Uma matriz deste tipo satisfaz $X^TX = XX = I$, portando admite inversa e tem determinante distinto de zero.

   Então, o refletor X é simétrico e não singular pois $X^TX = X^2 = I$. Porém, $det(FP) = det(F)det(P)$, como P é um projetor de rank 1, $det(FP) = det(F)det(P) = 0$, pois $det(P) = 0$ (possui rank 1) e, assim sendo, FP não pode ser refletor.

Questão 03: A matriz $A \in \mathbb{R}^{5\times5}$ foi fatorada $A = U\Sigma V^T$, onde $\Sigma \in \mathbb{R}^{3\times3}$ é uma matriz de zeros, a menos da diagonal que contém os elementos $7, 6, 5$ (nesta ordem), $U^TU = I_3$ e $V^TV = I_3$. Sabe-se que $UU^TB = QR$, $Q^TQ = I_5$ e R triangular superior indicada abaixo. As 3 primeiras colunas de Q são $q_1 = U_1, q_2 = U_2, q_3 = U_3$ ($U_i : i = 1, 2, 3$ são as colunas de U). Assuma que as colunas de B sejam $B_1, \dots, B_5$ e que as de Q sejam $q_1, \dots, q_5$.

```
R  =
    0.2113249    0.3303271    0.8497452    0.    0.
    0.           0.6653811    0.685731     0.    0.
    0.           0.           0.8782165    0.    0.
    0.           0.           0.           2.    1.
    0.           0.           0.           0.    1.
```

1. Quais são os autovalores de A^TA e de AA^T ?

   $A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^2V^T$. Portanto, seus autovalores são os quadrados dos valores singulares de A: 49, 36, 25. $AA^T = U\Sigma^2U^T$ e a mesma observação se aplica aqui.

2. Defina bases para $C(AA^T)$ e $N(AA^T)$.

   $AA^T = U\Sigma U^T$ tem o mesmo espaço coluna de A que é dado pelas colunas de U. Como informação adicional, UU^TB projeta as colunas de B em $C(U) = C(A)$. Este espaço coluna é dado pelas colunas q_1, q_2, q_3 da fatoração QR de UU^TB. Portanto, $C(A) = span\{U_1, U_2, U_3\} = span\{q_1, q_2, q_3\}$.

   A fatoração QR de UU^TB revela que q_4, q_5 são ortogonais a q_1, q_2, q_3 e estas produzem uma base para $C(AA^T)$. Com AA^T é simétrica, $N(AA^T) = N((AA^T)^T) = span\{q_1, q_2, q_3\}^\perp$. Portanto $N(AA^T) = span\{q_4, q_5\}$.

3. Qual é a solução $X \in \mathbb{R}^{5\times5}$ do Problema de Mínimos Quadrados $z = \min_{X\in\mathbb{R}^{5\times5}} \|AX - B\|_F^2$? Qual o valor ótimo z ? Lembre-se que $\|Z\|_F^2 = \sum_{i,j} |Z_{ij}|^2$.

   Assuma que X_i, B_i sejam as colunas $i = 1, \ldots, 5$ de X e de B, respectivamente. Então, aplicando A em cada coluna de X e comparando com a coluna pertinente de B, podemos escrever:

$$
\begin{aligned}
z &= \min_{X\in\mathbb{R}^{5\times5}} \|AX - B\|_F^2 \\
&= \sum_{i=1}^{5} \min_{X_i\in\mathbb{R}^5} \|AX_i - B_i\|_2^2
\end{aligned}
$$

   Vamos designar por X^* a matriz ótima que resolve o Problema de Mínimos Quadrados.

   Agora observe que as 3 primeiras colunas de B foram projetadas em $C(U) = C(A)$ e o erro de projeção foi zero. Portanto, $X_i^* = q_i : i = 1, \ldots, 3$ resolve as primeiras 3 colunas da solução ótima X^* do Problema de Mínimos Quadrados.

   Por outro lado, observe agora que a projeção das colunas B_4, B_5 de B em $C(U) = span\{q_1, q_2, q_3\}$ é o vetor zero, pois as entradas de R nas colunas 4 e 5 e linhas 1 a 3 são nulas. Portanto, B_4, B_5 são ortogonais a $C(U)$.

   Desta forma $X_4^* = r_{44}q_4, X_5^* = r_{45}q_4 + 55q_5$. Com isto, temos a definição completa da matriz X^* ótima.

Em relação ao valor ótimo z: Como as três primeiras colunas de B não contribuem com erro de projeção, o valor de z depende exclusivamente da projeção B_4, B_5, cujas normas são 2^2 e $(1^2) + (1^2)$, respectivamente.

Portanto, $z = 4 + 2 = 6$.

Questão 04: Responda verdadeiro ou falso e justifique.

1. Na Fase I do algoritmo que calcula a fatoração espectral de A, obtém-se uma matriz Hessenberg superior ortogonalmente equivalente a A.

   Verdadeiro. A Fase I emprega transformações do tipo QAQ^T e portanto entrega uma Hessenberg Superior similar à A. Similaridade implica em equivalência ortogonal (transformações ortogonalmente equivalentes de matrizes quadradas não são necessariamente similares). No caso em que A é normal (por exemplo, simétrica), a forma particular da Hessenberg Superior é tridiagonal.

2. A matriz A defectiva foi submetida ao algorimto iterativo QR para cálculo de autovalores, obtendo uma fatoração $A = \hat{Q}G\hat{Q}^T$. Os autovetores de A são as colunas de $\hat{Q}$.

   Falso. A matriz defectiva não é ortogonalmente similar a uma diagonal, ou seja, não admite n autovetores linearmente independentes. Portanto, as colunas de $\hat{Q}$, exceto pela primeira coluna, não fornecem autovetores para A.

3. A forma da matriz G da questão acima é diagonal, por isso, os autovalores de A são as entradas na diagonal de G.
   Falso. A matriz defectiva admite fatoração de Schur, portanto a matriz G é triangular superior e não diagonal.

4. Considere a matriz $H(X) = \begin{bmatrix} 0 & X^T \\ X & 0 \end{bmatrix}$, $X \in \mathbb{R}^{n \times n}$ de posto completo. Se A e Z são ortogonalmente equivalentes então a fatoração SVD de A pode ser obtida a partir da fatoração espectral de $H(Z)$. Em caso positivo, indique como a fatoração SVD de A pode ser recuperada.

   Verdadeiro. A fatoração espectral de $H(X)$ é:

$$\begin{bmatrix} 0 & X^* \\ X & 0 \end{bmatrix} \begin{bmatrix} V_X & V_X \\ U_X & -U_X \end{bmatrix} = \begin{bmatrix} V_X & V_X \\ U_X & -U_X \end{bmatrix} \begin{bmatrix} \Sigma_X & 0 \\ 0 & -\Sigma_X \end{bmatrix}.$$

Após normalização das n colunas de V_X e de U_X, temos a fatoração SVD de X: $X = U_X \Sigma_X V_X^T$. Como A e X são OE, temos $A = EXD$ para $E^T E = I$, $D^T D = I$. Então temos $A = (EU_X)\Sigma_X(V_X^T D) = U\Sigma V^T$ para $U = EU_X$, $V^T = V_X^T D$.

5. Considere a matriz $H(X)$ acima. Se A é ortogonalmente equivalente à $X = I$ (matriz identidade) a fatoração espectral de $H(I)$ vai revelar autovalores todos iguais a 1.

   Falso. A matriz $H(X)$ (para qualquer X não singular) é não singular e admite $2n$ autovalores não nulos aos pares: $\sigma_i, -\sigma_i$. Portanto, seus autovalores não podem ter o mesmo sinal. Complemetando, a matriz $H(X)$ possui n autovalores $1, -1$, com multiplicidade algébrica e geométrica iguais a 1 em ambos os casos.

B.14 Prova 1 - 2025.1

Questão 01: Considere as matrizes A e $\hat{A}$: $\hat{A} = \begin{bmatrix} a & a & a \\ a & b & b \\ a & b & c \end{bmatrix}$, $A = P\hat{A}P^T$, onde P é uma matriz de permutação dada por $[e_3, e_2, e_1]$ e $e_i : i = 1, \ldots, 3$ é um vetor de zeros, exceto pela posição i que contém uma entrada 1. Assuma que qualquer par de valores em $\{a, b, c\}$ sejam distintos e que $0 \notin \{a, b, c\}$. Responda, sempre justificando.

1. Empregando uma fatoração LU ou de Cholesky para A ou para $\hat{A}$ (a que você entender mais apropriada), apresente uma fatoração para A que revele seu posto, e que um dos fatores tenha pelo menos duas colunas distintas das colunas de uma matriz identidade de ordem 3. Indique claramente qual é o posto.

2. É possível garantir que a matriz A admite uma fatoração do tipo $A = MM^T$, para alguma matriz M ?

3. A matriz A admite uma fatoração do tipo $A = Q\Lambda Q^T$ onde Q é ortogonal e Λ é uma matriz diagonal ?

4. $\hat{A}$ e A possuem os mesmos autovalores e autovetores ?

5. É possível estabelecer condições necessárias e suficientes sobre a, b, c de forma que A admita fatoração $A = LL^T$ onde $l_{ii} > 0, i = 1, 2, 3$? Em caso positivo, quais são elas ? Em caso negativo, justifique a impossibilidade.

Resolução da Questão 1:

Veja que $A = \begin{bmatrix} c & b & a \\ b & b & a \\ a & a & a \end{bmatrix}$ e que a fatoração de LU de $\hat{A}$ deve envolver multiplicadores mais simples (apenas 1s) que a de A.

1. Faremos uma fatoração $\hat{A} = LU$ que revele o posto de $\hat{A}$ e então escrevemos que $A = (PL)(UP^T)$. Fazendo a fatoração de $\hat{A}$ temos: $\hat{A} = LU = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a & a & a \\ 0 & b-a & b-a \\ 0 & 0 & c-b \end{bmatrix}$. Logo $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a & a & a \\ b-a & b-a & 0 \\ c-b & 0 & 0 \end{bmatrix}$.
Esta fatoração revela que o posto de A é 3, pois a, b, c são distintos entre si.

2. Não. Para que valha a fatoração $A = MM^T$, temos que garantir que A seja simétrica semipositiva definida. Porém, não é possível garantir a positividade

ou semi-positividade de A ou de $\hat{A}$, considerando os dados fornecidos. Por exemplo, se $a < 0$ ou se $c < 0$, ambas possuem autovalores negativos. Portanto, pode não haver M tal que $A = MM^T$.

3. Sim, tanto A quanto $\hat{A}$ são simétricas e, portanto, diagonalizáveis. Matrizes simétricas são casos particulares de matrizes normais ($AA^T = A^TA$ que é exatamente a classe de matrizes ortogonalmente similares a matrizes diagonais).

4. A e $\hat{A}$ são similares, portanto tem os mesmos autovalores. Porém, os autovetores de matrizes similares não são os mesmos. No caso em questão, $A = P\hat{A}P^T$. Se $\hat{A}x = \lambda x$, isto é, λ, x é autopar de $\hat{A}$, $(P^TAP)x = \lambda x \to APx = P\lambda x \to A(Px) = \lambda(Px)$. Portanto, Px é autovetor de A, associado ao autovalor λ.

5. A admite fatoração de Cholesky se e somente se $\hat{A}$ admitir. A matriz $\hat{A}$ deve ser positiva definida para admitir uma fatoração de Cholesky. Assumindo que este seja o caso, obtemos a fatoração de Cholesky, onde $L = \begin{bmatrix} \sqrt{a} & 0 & 0 \\ \sqrt{a} & \sqrt{b-a} & 0 \\ \sqrt{a} & \sqrt{b-a} & \sqrt{c-b} \end{bmatrix}$, o fator de Cholesky de $\hat{A}$, que só ocorre quanto $a > 0, b - a > 0, c - b > 0$.

Questão 02: A matriz A de posto incompleto foi fatorada na forma $A = QQ^T$, onde $Q \in \mathbb{R}^{m \times n}$, $1 < n < m$ e $Q^TQ = I_n$. Responda verdadeiro ou falso, justificando.

1. A^k para quaisquer valores de $k \geq 2$ (k inteiro) e A possuem os mesmos autovalores e autovetores.

2. A^k para quaisquer valores inteiros de $k \geq 1$ possui $\lambda = 0$ como autovalor.

3. Se (λ, x) é um autopar de A para $\lambda \neq 0$, então $x \in C(Q)$.

4. O sistema linear $Ax = b$ admite solução sempre que b for uma combinação linear dos autovetores de A.

5. $N(A - \lambda I) \subseteq C(Q)$ para qualquer autovalor λ de A.

Resolução da Questão 2:
Observação geral:

$$\begin{aligned} A^2 &= QQ^TQQ^T \\ &= A \\ A^k &= A \end{aligned}$$

1. Verdadeiro, pois $A^k = A$, uma vez que $Q^TQ = I_n$.

2. Como A tem posto incompleto (pelo enunciado), $det(A) = 0$ e pelo menos um autovalor de A ou de sua potência inteira k qualquer, é $\lambda = 0$.

3. Verdadeiro. $Ax = \lambda x \neq 0$ se $\lambda \neq 0$. Assim, x certifica que $\lambda x \in C(A)$. Logo $x \in C(A)$.

4. Falso. Por exemplo, se A é simétrica e $b(b \neq 0)$ é autovetor associado ao autovalor $\lambda = 0$, $b \in N(A)$. Portanto, $b = Ax \rightarrow b \in C(A) \perp N(A)$ e temos uma contradição. Veja o contra exemplo em que $b = (0,0,1,0)^T \in N(A)$ é autovetor de $A = vv^T + uu^T$ onde $v^T = \frac{1}{\sqrt{2}}(1,1,0,0)$, $u^T = \frac{1}{\sqrt{2}}(1,-1,0,0)$.

5. Falso, para $\lambda = 0$ temos $N(A - \lambda I) = N(A) \perp C(Q) = C(A)$ (posto de $Q =$ posto de $A = n$).

Questão 3 Uma fatoração para a matriz A de posto incompleto é $A = \begin{bmatrix} 1 & 0 & x \\ 2 & 1 & y \\ 3 & 0 & z \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ a & b & c \end{bmatrix}$.

Responda:

1. Apresente valores admissíveis para x, y, z, a, b, c.

2. Apresente uma matriz A que atenda ao enunciado da questão.

3. Dada a escolha acima, escreva A^T como uma soma de $r = posto(A)$ matrizes de posto 1.

4. Caracterize $C(A), C(A^T), N(A), N(A^T)$, definindo claramente os elementos da base e suas dimensões.

Resolução da Questão 3:

1. Basta fazer $x = y = z = a = b = c = 0$. Esta não é a única alternativa. Qualquer resposta correta em que as grandezas x, y, z, a, b, c não são todas nulas precisa escrever $(x, y, z)^T = \alpha(1,2,3)^T + \beta(0,1,0)$ e $(a, b, c) = \gamma(1,1,2) + \mu(0,1,3)$, pois A tem claramente posto igual a superior a dois e pelo enunciado tem posto incompleto, o que a impede de ter posto 3.

2. Considerando a alternativa $x = y = z = a = b = c = 0$, temos $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 7 \\ 3 & 3 & 6 \end{bmatrix}$, que revela que A possui posto incompleto igual a 2.

3. $A^T = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}.$

4. A fatoração revela que o posto é 2 e portanto $dim(C(A)) = dim(C(A^T)) = 2$. Consequentemente, $dim(N(A)) = dim(N(A^T)) = 3 - 2 = 1$.

$$C(A) = span \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

$$C(A^T) = span \left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} \right\}$$

$$N(A^T) = C(A)^\perp \rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow N(A^T) = span \left\{ \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix} \right\}.$$

Analogamente, $N(A) = C(A^T)^\perp \rightarrow \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow N(A) =$

$$span \left\{ \begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix} \right\}.$$

Questão 4 Responda Verdadeiro ou Falso e justifique sua resposta.

1. Considere $A \in \mathbb{R}^{n \times n}$. Se o sistema linear $Ax = e_i$ admite solução para todo $i = 1, \ldots, n$ (e_i é o vetor de zeros a menos da entrada i que é 1), então o sistema linear $A^T A y = b$ admite solução única, mas não pode ser resolvido via Fatoração de Cholesky.

2. Dada uma matriz A qualquer, os vetores b que não pertencem a $C(A)$ formam um subespaço.

3. Se $AB = 0$, então as colunas da matriz $B \in C(A)$ e as linhas de $A \in C(B^T)$.

4. O espaço coluna da matriz $C = AB$ contém o espaço coluna de A.

5. Ao adicionarmos uma coluna b em A criando uma matriz $[A|b]$, a dimensão do espaço coluna da nova matriz aumenta quando b é linearmente dependente das demais colunas de A.

Resolução da Questão 4:

1. Falso. Se $Ax = e_i$ admite solução para todo $i = 1, \ldots, n$, $C(A) = \mathbb{R}^n$. Logo a matriz $A^T A$ é simétrica e positiva definida. Portanto, o sistema linear pode ser resolvido via fatoração de Cholesky.

2. Falso, pois o vetor zero precisa pertencer a qualquer subespaço.

3. Falso. As colunas da matriz B pertencem ao $N(A)$ e as linhas de A pertencem ao $N(B^T)$.

4. Falso. O que é verdadeiro é que o espaço coluna de C está contido no espaço coluna de A. Como contra exemplo, considere $C = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$.

5. Falso, $dim(C([A|b])) > dim(C(A))$ apenas se $b \notin C(A)$. Caso contário, isto é, $b \in C(A)$, os subespaços são os mesmos e as dimensões são iguais.

B.15 Prova 2 - 2025.1

Questão 1: Responda verdadeiro ou falso e justifique.

1. M_1, M_2 são projetores ortogonais de mesma ordem. $M = M_1 + M_2$ é projetor ortogonal ? Em caso positivo, estabeleça condições necessárias e suficientes. Em caso negativo, justifique.

   Resposta: Falso. $M = M_1 + M_2$ é projetor se e somente se $M^2 = M$. Isso não é observado no caso geral. Veja:

   $$\begin{aligned}(M_1 + M_2)(M_1 + M_2) &= M_1^2 + M_1M_2 + M_2M_1 + M_2^2 \\ &= M_1 + M_2 + M_1M_2 + M_2M_1\end{aligned}$$

   Portanto, para que seja projetor, é necessário que $M_1M_2 + M_2M_1 = 0$ (uma matriz de zeros). Como M_2, M_1 são simétricas, a condição $M_1M_2 + M_2M_1 = 0$ implica que $C(M_1) \perp C(M_2)$ e esta condição não é sempre satisfeita entre dois projetores ortogonais. Portanto, a resposta é falsa.

2. Uma matriz $A \in \mathbb{R}^{m\times n}$ de posto $r < min\{m, n\}$ foi fatorada $AP = QR$ onde Q possui r colunas ortonormais e P é uma matriz de permutação. Sabe-se que o sistema linear $Ax = b$ admite solução. Então $\|QQ^Tb - b\|_2 \neq 0$.

   Resposta: Falsa. Como Q possui r colunas ortonormais, a fatoração QR dada é reduzida, foi obtida via permutação de colunas e fornece $C(A) = C(Q)$. Assim sendo, QQ^T é o projetor que projeta em $C(AP) = C(A) = C(Q)$. Portanto, se $Ax = b$, $b \in C(A)$, $QQ^Tb = b$, $\|QQ^Tb - b\|_2 = 0$.

3. É possível haver duas matrizes simétricas distintas $A, P \in \mathbb{R}^{n\times n}$ com os quatro espaços fundamentais idênticos e satisfazendo $A^2 \neq A, P^2 = P$. Em caso positivo, ilustre com um exemplo.

   Resposta: Verdadeiro.
   É possível. Faça $P = I$ e $A = \alpha I$, para $\alpha \neq 0, \alpha \neq 1$. Veja que $A^2 = \alpha^2 I \neq A$.

4. Seja E uma matriz $m \times m$, com $Ex = \frac{x+Fx}{2}$ onde F é uma matriz $m \times m$ que transforma $[x_1, \cdots, x_m]$ em $[x_m, \cdots, x_1]$. Então a matriz E é um projetor ortogonal.

Resposta: Verdadeiro.
$Ex = \frac{x+Fx}{2} \to E = \frac{1}{2}(I+F)$. Pelo enunciado temos que $F = [e_m, e_{m-1}, \ldots, e_1]$, ou seja, F inverte as coordenadas de x. Portanto, $F^2 = I$.

E é projetor se e somente se $E^2 = E$.

$$\begin{aligned} E^2 &= \frac{1}{4}(I+F)^2 \\ &= \frac{1}{4}(I^2 + 2F + F^2) \\ &= \frac{1}{4}(2I + 2F) \\ &= \frac{1}{2}(I+F) \\ &= E \end{aligned}$$

Portanto, é projetor e é projetor ortogonal pois E é simétrica.

Questão 2: A fatoração $A = QR$ de A com posto completo foi realizada com o algoritmo Gram-Schmidt Clássico e Revisado, cujas implementações são dadas ou pelo `Algoritmo_X` ou pelo `Algortimo_Y` abaixo. Um algoritmo produziu a fatoração $A = Q_1R_1$ e o outro $A = Q_2R_2$. 2

```
function [M3,M4] = Algoritmo_X(C)
   [m,n] = size(C)
   M3 = zeros(n,n)
   M4 = zeros(m,n)
   u = C
   for i = 1:n
       M3(i,i) = norm(u(:,i),2)
       M4(:,i) = u(:,i)/M3(i,i)
       for j = (i+1):n
           M3(i,j) = M4(:,i)'*u(:,j)
           u(:,j) = u(:,j) - M3(i,j)*M4(:,i)
       end
   end
endfunction
```

```
function [M1,M2] = Algoritmo_Y(B)
   [m,n] = size(B)
   M1 = zeros(n,n)
   M2 = zeros(m,n)
   for j = 1:n
       u = B(:,j)
       for i = 1:j-1
           M1(i,j) = M2(:,i)'*B(:,j)
           u = u - M1(i,j)*M2(:,i)
       end
       M1(j,j) = norm(u,2)
       M2(:,j) = 1.0/M1(j,j) * u
   end
endfunction
```

```
->I - Q1'*Q1 =
   2.220D-16  -1.684D-15   8.465D-15  -1.776D-14   2.456D-14  -5.704D-14   3.501D-13  -3.353D-12
  -1.684D-15   3.331D-16   4.229D-14  -1.927D-13   4.883D-13  -1.205D-12   3.579D-12  -1.321D-11
   8.465D-15   4.229D-14  -2.220D-16  -1.365D-12   5.832D-12  -1.907D-11   7.025D-11  -2.946D-10
  -1.776D-14  -1.927D-13  -1.365D-12  -2.220D-16   3.156D-11  -1.758D-10   9.279D-10  -5.236D-09
   2.456D-14   4.883D-13   5.832D-12   3.156D-11  -2.220D-16  -8.876D-10   8.682D-09  -7.301D-08
  -5.704D-14  -1.205D-12  -1.907D-11  -1.758D-10  -8.876D-10   1.110D-16   6.525D-08  -0.000001
   3.501D-13   3.579D-12   7.025D-11   9.279D-10   8.682D-09   6.525D-08   0.          -0.0000158
  -3.353D-12  -1.321D-11  -2.946D-10  -5.236D-09  -7.301D-08  -0.000001   -0.0000158   2.220D-16
->I - Q2'*Q2 =
   2.220D-16  -1.684D-15   8.573D-15  -1.825D-14   2.578D-14  -5.877D-14   3.377D-13  -3.104D-12
  -1.684D-15   3.331D-16  -1.056D-15  -3.059D-15   1.772D-14  -4.495D-14   4.812D-14   5.608D-13
   8.573D-15  -1.056D-15   0.          1.161D-15  -1.218D-14   6.834D-14  -3.853D-13   2.262D-12
  -1.825D-14  -3.059D-15   1.161D-15   2.220D-16  -2.949D-16   9.714D-16  -1.263D-15  -6.466D-14
   2.578D-14   1.772D-14  -1.218D-14  -2.949D-16   2.220D-16  -3.886D-16   4.330D-15  -1.427D-14
  -5.877D-14  -4.495D-14   6.834D-14   9.714D-16  -3.886D-16   0.          -8.327D-16   1.746D-14
   3.377D-13   4.812D-14  -3.853D-13  -1.263D-15   4.330D-15  -8.327D-16  -2.220D-16   3.803D-15
  -3.104D-12   5.608D-13   2.262D-12  -6.466D-14  -1.427D-14   1.746D-14   3.803D-15  -2.220D-16
```

Considerando os algoritmos e resultados numéricos obtidos acima, responda:

1. Dentre as matrizes $\{M_1, M_2, M_3, M_4\}$ há alguma que corresponda a Q_1, Q_2 ?

   Resposta: Sim. $Q_2 = M_4$ e $Q_1 = M_2$.

   O algoritmo X implementa Gram-Schmidt revisado pois em cada iteração i (ortogonalização de uma coluna) faz uma projeção de posto 1 nas colunas de u de índice $i+1$ até n. Ao final restorna a matriz M_4 que armazena a Q na fatoração. Já o algoritmo Y implementa Gram-Schmidt clássico pois a cada iteração j faz uma projeção de posto $j-1$ na coluna j de A armazenada no vetor u. Ao final retorma a matriz M_2 que retorna a sua Q na fatoração. Já os resultados numéricos de $I - Q_1^T Q_1$ e $I - Q_2^T Q_2$ mostram que a matriz Q_1 possui colunas menos ortogonais que as colunas de Q_2, uma vez que suas entradas possuem magnitudes maiores, várias ordens de grandeza superiores à precisão da máquina, 10^{-16}.

   Portanto, a matriz Q_1 corresponde à matriz retornada pelo algoritmo Clássico, $Q_1 = M_2$ e Q_2 corresponde à matriz Q retornada pelo algoritmo Revisado, $Q_2 = M_4$.

2. Existe alguma diferença notável entre o perfil de perda de ortogonalidade entre as colunas de Q produzidas pelos dois algoritmos, X ou Y ? Justifique.

   Resposta: Sim, existe.

   O resultado numérico apresentado para $I - Q_1^T Q_1$ e $I - Q_2^T Q_2$ indica exatamente a perda de ortogonalidade. Enquanto no algoritmo revisado as entradas de $I -$

$Q_2^T Q_2$ são mais homogêneas e menores, para o algoritmo clássico, as entradas da matriz $I - Q_1^T Q_1$ apresentam valores maiores e erros maiores associados às colunas de maiores índices, isto é, no canto inferior direito da matriz $I - Q_1^T Q_1$.

3. A menos de erros numéricos, seria possível construir um projetor ortogonal a partir daquilo que cada um dos dois algoritmos retorna (M_1, M_2, M_3, M_4)? Em caso positivo, diga como e seja P_1, P_2 os projetores obtidos por meio da saída dos algoritmos 1 e 2. Explicite os projetores, indicando seu posto e seu espaço coluna.

   Resposta: Sim, seria possível. $P_1 = M_2 M_2^T = Q_1 Q_1^T$ e $P_2 = M_4 M_4^T = Q_2 Q_2^T$ são projetores ortogonais que projetam em $C(A)$. Possuem posto igual ao posto de A.

4. Diferencie os algoritmos X e Y em relação ao processo de ortogonalização, diferenciando o posto do projetor empregado em cada momento em que ocorre alguma etapa de projeção em cada um deles. Identifique claramente a partir da indexação dos algoritmos.

   Resposta: O algoritmo X (revisado) faz $n - (i + 1)$ projeções de posto 1 logo após computar a coluna q_i de Q. O conteudo em u, que armazena $A_j - \sum_{k=1}^{i-1} q_k q_k^T A_j$ para $j \geq i + 1$, é submetido a estas projeções. Já o algoritmo Y (Clássico) faz uma projeção de posto $i - 1$, na coluna A_i, tão logo as colunas $q_1, \ldots, q_{i-1}$ tenham sido computadas. Ou seja, preserva a coluna A_i intacta até o momento da projeção de posto $i - 1$. Esta segunda opção, a clássica, produz resultados numéricos piores.

5. Qual algoritmo X ou Y é mais apropriado para a introdução de pivoteamento de colunas ?

   Resposta: O algoritmo X que implementa Gram-Schmidt revisado é mais apropriado, pois a cada iteração j, temos a indicação do erro de projeção, que indica o quão linearmente independentes as colunas de A que não foram ortogonalizadas são das colunas de Q já computadas. No algoritmo Y que implementa o Gram-Schmidt clássico esta informação não está disponível, pois a projeção é feita em um único passo.

Questão 3: O Problema de Mínimos Quadrados $\min \|Ax - b\|_2$ deve ser resolvido para $A \in \mathbb{R}^{m \times n}$ para se ajustar a função $g(z) = a + bz + c \log_{10}(z)$ aos dados

$\{(z_i, y_i) : i = 1, \ldots, m\}$. Sabe-se que as abscissas z_i são: $1, 10, 0.1, 0.01, 10000$, nesta ordem. Responda:

1. Qual é a matriz A ?

   Resposta: $A = \begin{bmatrix} 1 & 1 & \log_{10}(1) \\ 1 & 10 & \log_{10}(10) \\ 1 & 0.1 & \log_{10}(0.1) \\ 1 & 0.01 & \log_{10}(0.01) \\ 1 & 10000 & \log_{10}(10000) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 10 & 1 \\ 1 & 0.1 & -1 \\ 1 & 0.01 & -2 \\ 1 & 10000 & 4 \end{bmatrix}.$

2. Sabendo que os valores singulares de A são $\{l_i : i = 1, 2, 3\}$ que satisfazem $l_1 > l_2 > l_3 > 0$ qual é o valor de $\kappa_2(A^T A)$?

   Resposta: A matriz A é não simétrica, portanto $\sigma_i(A) = \sqrt{\lambda_i(A^T A)}$ e $\lambda_i(A^T A) = \sigma_i(A)^2$. Portanto $\kappa_2(A^T A) = \frac{l_1^2}{l_3^2}$.

3. Discuta a existência e unicidade de soluções do Sistema de Equações Normais.

   Resposta: O sistema de equações normais $A^T Ax = A^T y$ é definido por uma matriz de coeficientes $A^T A$ de posto completo, pois A possui posto completo, posto(A) $= 3$. Portanto, a solução x do sistema existe e é única: $A^T y \in C(A^T A)$, $N(A^T A) = \{0\}$.

4. Como a fatoração QR de A pode ser empregada para resolvê-lo ? Há alguma vantagem em assim procedermos ?

   Resposta: Pode e deve ser usada pois é uma fatoração mais estável que a alternativa mais barata, fatoração de Cholesky de $A^T A$. Basta resolver o sistema via $Rx = Q^T b$.

   No desenvolvimento abaixo, observe que usamos $Q^{-1} = Q^T$ e que podemos multiplicar por R^{-T}, pois R^T é quadrada e não singular, uma vez que A possui posto completo.

$$\begin{aligned} A^T Ax &= A^T b \\ R^T Q^T QRx &= R^T Q^T b \\ R^T Rx &= R^T Q^T b \\ Rx &= Q^T b \end{aligned}$$

5. Considere que z_5 foi substituído por $z_3 = 0.1$ (e y_5 por y_3). Quais as dificuldades em usar QR clássico resolver o sistema de equações normais ?

   Resposta: Mesmo com a substituição, neste caso A continua com posto completo. Assim, as dificuldades de se usar o algoritmo Clássico para a fatoração são as mesmas caso a substituição não tivesse sido feita. As dificuldades são que a matriz Q retornada pelo algoritmo clássico não é tão precisa quanto a produzida por outras implementações da fatoração QR.

Questão 4: Considere o conjunto $Y = \{x \in \mathbb{R}^3 : Ax = b\}$, os pontos $z = (1, 2, 4)^T$ e $x^0 = (1, 1, 1)^T \in Y$, $A = \begin{bmatrix} 3 & -1 & 0 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2/3 \\ 0 & 1 & 2 \end{bmatrix}$ e $b = (2, 3)^T$.

Responda:

1. (10%) Y é um subespaço vetorial ? Justifique.

   Resposta: Não é subespaço vetorial, pois $0_3 \notin Y$.

2. (25%) O problema de projetar um ponto qualquer em Y equivale a um problema de projeção em subespaço vetorial ? Em caso positivo, indique claramente em qual espaço vetorial e a equivalência da problema. Em caso negativo, apresente uma justificativa.

   Resposta: O conjunto Y pode ser reescrito como

   $$Y = x^0 + span\{v_1, \ldots, v_d\}$$

   onde $x^0 \in Y$ (foi dado) e $span\{v_1, \ldots, v_d\} = N(A)$. No caso em questão, considerando a matriz A dada e sua fatoração, temos que $d = dim(N(A)) = 1$ e $v = (2/3, 2, -1)^T$ fornece uma base para $N(A)$.

   Esta é uma propriedade de qualquer conjunto afim Y, que é uma translação de um subespaço vetorial, uma translação de $N(A)$ onde A é a matriz que define Y. Então, projetar um ponto z qualquer em Y corresponde a projetar $z - x^0$ em $N(A)$. Isso será elaborado em detalhes ainda maiores na próxima questão.

3. (40%) Qual é o ponto u de Y de mínima distância Euclideana de z ?

   Resposta: Primeiro observe que $z \notin Y$. Assim sua projeção u em Y é diferente de z.

O ponto u resolve

$$\min_{u\in Y}\|z-u\|_2 = \min_{\alpha\in\mathbb{R}}\|z-(x^0+\alpha v)\|_2 = \min_{\alpha\in\mathbb{R}}\|(z-x^0)-\alpha v)\|_2.$$

Para este último, resolvedo o sistema de Equações Normais temos: $v^T v\alpha = v^T(z-x^0)$. $v^T v = 49/9, z-x^0 = (0,1,3)^T, v^T(z-x^0) = -1 \rightarrow \alpha = -9/49$.
Logo $u = x^0+\alpha v = \frac{1}{49}\begin{bmatrix} 43 \\ 31 \\ 58 \end{bmatrix}$.

4. (25%) Seria possível escrever o ponto u como $u = w + Pz$ onde P é um projetor ortogonal e w é um vetor convenientemente escolhido ? Em caso negativo justifique. Em caso positivo, apresente as propriedades que devem satisfeitas por w e P. Para este último indique seu posto e seu espaço coluna.

Resposta: Sim, seria. Bastaria que P projete em $N(A)$ (logo $C(P) = N(A)$) e $w \in N(A)^\perp \cap Y$. A título de diferenciação entre o pedido na questão logo acima e o que pedido aqui, quando na questão acima escrevemos $u = x^0 + \alpha v$, o termo αv não é (necessariamente) a projeção de z em $N(A)$ tanto quanto não temos necessariamente que $x^0 \in N(A)^\perp$. Mas como x^0 pode ser qualquer ponto em Y, podemos escolher $x^0 = w \in N(A)^\perp \cap Y$, $u = w + Pz$ para um projetor P que projeta em $N(A)$. Esta forma de escrever u resulta naturalmente das duas escolhas. Podemos fazer isso pois $\mathbb{R}^n$ é a soma direta de $N(A)$ e de seu complemento ortogonal. Veja a Figura B.3 que ilustra a projeção em $N(A)$ e em Y.

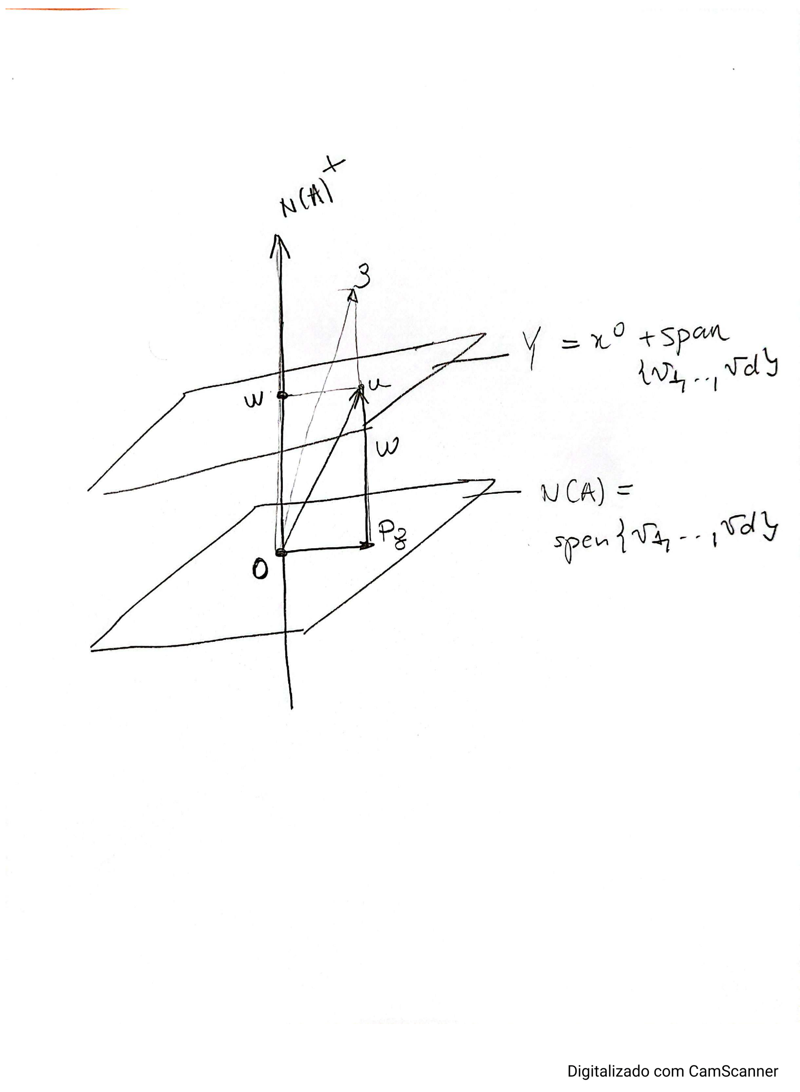

Figura B.3: $u = w + Pz$

B.16 Prova 3 - 2025.1

Questão 1: Usando seus conhecimentos sobre a fatoração a $A = U\Sigma V^T$, responda verdadeiro ou falso e justifique. Para as questões desta avaliação, considere que a função traço $tr(B) : B \in \mathbb{R}^{n\times n} \to \mathbb{R}$ corresponde à soma dos elementos na diagonal principal de B.

1. $tr(A^TA) = tr(Z^TA^TAZ)$ quando Z é unitária.
   Resposta: Verdadeiro. Podemos argumentar de diversas formas. A^TA e Z^TA^TAZ são similares ($Z^{-1} = Z^T$), portanto possuem os mesmos autovalores. Logo, $\sum_i \lambda_i(Z^TA^TAZ) = tr(Z^TA^TAZ) = tr(A^TA) = \sum_i \lambda_i(A^TA) = \sum_i \sigma_i^2(A)$.
   Por outro caminho de demonstração, (sendo o posto de A igual a r) desenvolvemos:

$$\begin{aligned} tr(Z^TA^TAZ) &= tr(Z^T(V\Sigma^TU^TU\Sigma V^T)Z) \\ &= tr((Z^TV)\Sigma^2(V^TZ)) \\ &= tr(V\Sigma^2V^T) \\ &= tr(\Sigma^2) \\ &= \sigma_1^2 + \dots \sigma_r^2 \\ &= tr(V\Sigma^2V^T) \\ &= tr(V\Sigma^TU^TU\Sigma V^T) \\ &= tr(A^TA) \end{aligned}$$

2. Se A é quadrada e $C = A^+BA$, então C e B possuem os mesmos autovalores.
   Resposta: Falso. Se A admite posto incompleto, $A^+ \neq A^{-1}$ e C, B não são similares. Veja o contra exemplo abaixo, onde B é uma diagonal com entradas 3 e 2 na diagonal e $A = uu^T$ onde $u = \frac{\sqrt{2}}{2}(1,1)^T$. Veja que C tem um autovalor zero e o autovalor positivo não é um dos autovalores de B.

```
B  =
   2.   0.
   0.   3.
 A  =
   1.   1.
   1.   1.
u = [sqrt(2]/2;sqrt(2)/2]
u  =
```

```
    0.7071068
    0.7071068
 A  =
    1.   1.
    1.   1.
-->Amais = u*(1/2)*u'
 Amais  =
    0.25   0.25
    0.25   0.25
-->C = Amais*B*A
 C  =
    1.25   1.25
    1.25   1.25
-->spec(C)
 ans  =
    0.
    2.5
```

3. A^+b resolve o sistema de equações normais associado a projetar b em $C(A)$.
Resposta: Verdadeiro: $A^+ = V\Sigma^+U^T$. Substituindo $x^+ = A^+x$ em A^TAx temos:

$$\begin{aligned}
A^TAx^+ &= (V\Sigma^TU^TU\Sigma V^T)(V\Sigma^+U^T)b \\
&= V\Sigma^2(V^TV)\Sigma^+U^Tb \\
&= V\Sigma^2\Sigma^+U^T)b \\
&= (V\Sigma U^T)b \\
&= A^Tb
\end{aligned}$$

4. $AA^+b \in C(A)$.
Resposta: Verdadeiro. $A(A^+b) = Ax^+ = p$, logo x^+ certifica que p, projeção de b em $C(A)$, está em $C(A)$.

Questão 2: É dada $A = \begin{pmatrix} d & 1 \\ 0 & \frac{1}{d} \\ d & -1 \end{pmatrix}$, com colunas ortogonais e $d \in \mathbb{R}, d \neq 0$.

Observação geral para a resolução da questão: Como as colunas de A são ortogonais, são linearmente independentes, e fornecem uma base para $C(A)$. Para obter $A = U\Sigma V^T$ onde $U = [U_1, U_2]$ possui duas colunas ortonormais, base para $C(A)$, basta dividir cada coluna de A por sua norma Euclideana. O valor das normas Euclideanas destas colunas são os valores singulars. Para completar a fatoração SVD de A, basta verficiar que $V^T = I_2$.

1. (25%) Quais são os valores singulares de A ? **Resposta:** A matriz A possui dois valores singulares, pois seu posto é 2. Normalmente atribuímos a σ_1 o maior valor singular. Como este valor depende do valor assumido por d, vamos simplesmente indicar os valores, sem que sejam ordenados.
   Valor singular 1: $\sigma = \|(d, 0, d)^T\|_2 = |d|\sqrt{2}$.
   Valor singular 2: $\sigma = \|(1, 1/d, -1)^T\|_2 = \sqrt{2 + 1/d^2}$.
   $\sigma_1 = \max\{|d|\sqrt{2}, \sqrt{2 + 1/d^2}\}$
   $\sigma_2 = \min\{|d|\sqrt{2}, \sqrt{2 + 1/d^2}\}$.

2. (25%) $A^+ = ?$
   Pelo desenvolvimento acima, sem ordenar os valores singulares, temos $A = U\Sigma V^T$ onde $V^T = I_2$, $\Sigma = \begin{pmatrix} |d|\sqrt{2} & 0 \\ 0 & \sqrt{2 + 1/d^2} \end{pmatrix}$, $U = \begin{pmatrix} \frac{d}{|d|\sqrt{2}} & \frac{1}{\sqrt{2+1/d^2}} \\ 0 & \frac{1}{d\sqrt{2+1/d^2}} \\ \frac{d}{|d|\sqrt{2}} & -\frac{1}{\sqrt{2+1/d^2}} \end{pmatrix}$.
   Então temos: $A^+ = I_2\Sigma^+U^T$ onde: $\Sigma^+ = \begin{pmatrix} \frac{1}{|d|\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2+1/d^2}} \end{pmatrix}$ e
   $U^T = \begin{pmatrix} \frac{d}{|d|\sqrt{2}} & 0 & \frac{d}{|d|\sqrt{2}} \\ \frac{1}{\sqrt{2+1/d^2}} & \frac{1}{d\sqrt{2+1/d^2}} & -\frac{1}{\sqrt{2+1/d^2}} \end{pmatrix}$

3. (50%) Seja A_1 a matriz de rank-1 que melhor aproxima A, na norma $\|\cdot\|_F$. É possível $C(A_1) = span\{(1, 0, 1)^T\}$ (responda SIM ou NÃO) ? Em caso positivo, quais condições devem ser observadas para que isso ocorra ? Em caso negativo, apresente uma justificativa.
   Resposta: Sim, é possível, basta que $\sigma_1 = \max\{|d|\sqrt{2}, \sqrt{2 + 1/d^2}\} = |d|\sqrt{2}$ e que o primeiro valor singular seja estritamente maior que o segundo (a norma Euclideana da primeira coluna de A seja maior que a da segunda). Então temos:

$$
\begin{aligned}
|d|\sqrt{2} &> \sqrt{2 + 1/d^2} \\
2d^2 &> 2 + 1/d^2 \\
2d^4 - 2d^2 - 1 &> 0 \\
2z^2 - 2z - 1 &> 0
\end{aligned}
$$

Então temos $z \in (-\infty, \frac{1-\sqrt{3}}{2}) \cup (\frac{1+\sqrt{3}}{2}, +\infty)$. O primeiro intervalo não nos interessa pois $z = d^2$ e $d \in \mathbb{R}$.

Portanto, $d > \frac{\sqrt{1+\sqrt{3}}}{\sqrt{2}}$ garante que a norma da primeira coluna seja estritamente maior que norma da segunda coluna de A.

Questão 3: Seja $H \in \mathbb{R}^{m\times m}$ um refletor de Householder cujo hiperplano de reflexão é $span\{u\}^{\perp}, u \in \mathbb{R}^m, \|u\|_2 = 1$ e P uma matriz de permutação. Responda justificando.

1. Quais são os autovalores (com as suas multiplicidades algébricas e geométricas) e autovetores de H ?
   Resposta: $F = I - 2uu^T$, F é simétrica e não possui autovalor defectivo. Então:

   $$(I - 2uu^T)x = \lambda x$$

   Temos dois casos a considerar:

   - Para $x \in span\{u\}$, temos $(I - 2uu^T)u = u - 2u = -u$. Portanto, u é autovetor com $\lambda = -1$ sendo seu autovalor, com multiplicidade algébrica e geométrica 1.
   - Os demais autovetores de $F = I - 2uu^T$ pertencem a $span\{u\}^{\perp}$. Basta tomar uma base ortonormal para este espaço vetorial $m-1$ dimensional que termos $m-1$ outros autovetores associados ao autovalor 1. Veja: para $x \in span\{u\}^{\perp}$, temos $(I - 2uu^T)x = x - 2u(u^Tx) = x$. A multiplicidade algébrica e geométrica do autovalor 1 é $m-1$.

2. Qual é o determinante de H ?
   Resposta: $det(H) = \prod_{i=1}^{m} \lambda_i = (-1)^1(1)^{m-1} = -1$.

3. Para qualquer $a \in \mathbb{R}^m$, $(Ha + uu^Ta) \in span\{u\}^{\perp}$. Falso ou verdadeiro ?
   Resposta: Verdadeiro.

   $$\begin{aligned}(Ha + uu^Ta) &= (H + uu^T)a \\ &= (I - 2uu^T + uu^T)a \\ &= (I - uu^T)a\end{aligned}$$

   $(I - uu^T)$ é um projetor ortogonal de posto $m-1$ que projeta em $span\{u\}^{\perp}$.

4. PHP^T é um refletor de Householder ? Falso ou Verdadeiro ? Em caso positivo, indique o hiperplano de reflexão.
   Resposta: Verdadeiro.
   PHP^T é um refletor de Householder se existe algum $w \in \mathbb{R}^m, \|w\|_2 = 1$ tal

que

$$PHP^T = I - 2ww^T$$

$$\begin{aligned} PHP^T &= P(I - 2uu^T)P^T \\ &= PIP^T - 2(Pu)(u^TP^T) \\ &= I - 2ww^T \end{aligned}$$

onde $w = Pu$. O refletor de Householder que tem Pu como construtor tem como hiperplano de reflexão o subespaço $span\{Pu\}^{\perp}$.

5. Seja Q um projetor ortogonal. Então $I - HQH$ é um projetor ? Falso ou verdadeiro ? Em caso positivo, diga em qual espaço projeta e qual a sua dimensão.

   Resposta: Verdadeiro.

   $I - HQH$ é projetor se e somente se HQH for projetor:

$$\begin{aligned} (HQH)^2 &= HQ(HH)QH \\ &= HQQH \\ &= HQH \end{aligned}$$

   No desenvolvimento acima, usamos o fato de que $Q^2 = Q, H^2 = I$. Então $I - HQH$ projeta no complemento ortogonal de $C(HQH)$. Como H possui posto completo, $posto(HQH) = posto(Q)$. Portanto, $dim(C(HQH)) = dim(C(Q))$ e $dim(C(Q)^{\perp}) = m - dim(C(Q))$.

Questão 4: Assuma que $H(X) = \begin{bmatrix} 0 & X^T \\ X & 0 \end{bmatrix}$, para qualquer $X \in \mathbb{R}^{n\times n}$.

1. É possível empregar a fatoração espectral de $H(X)$ para alguma X adequadamente escolhida, para encontrar a fatoração SVD de $A \in \mathbb{R}^{m\times n}$ com $m \neq n$ (responda SIM ou NÃO) ? Em caso positivo diga como e diga quais propriedades X deve possuir.

   Resposta: Sim, se X for ortogonalmente equivalente a A, por exemplo $X = R$ onde $A = QR$ é a fatoração QR de A. Através dos autovetores de $H(R)$ e da matriz ortogonal Q são recuperdos os vetores singulares de A, à esquerda e à direita. Os valores singulares de A são os autovalores positivos de $H(R)$.

2. Seja σ_i e σ_j dois valores singulares distintos de A. Então $N(H(A) - \sigma_i I) \perp N(H(A) - \sigma_j I)$?

   Resposta: Verdadeiro. $N(H(A)-\sigma_i I)$ é o auto-espaço associado ao autovalor $\lambda_i = \sigma_i$ de $H(A)$. Auto espaços de autovalores distintos são sempre ortogonais.

3. Se $A \in \mathbb{R}^{2\times 2}$ é simétrica e possui autovalores $\lambda_1 = 5$, $\lambda_2 = 3$, com os auto espaços associados a estes autovalores sendo, respectivamente, $span\{y^1\}, span\{y^2\}$, onde $y^1 = (\cos(\theta), -\sin(\theta))^T$ e $y^2 = (\sin(\theta), \cos(\theta))^T$ (para algum $\theta \in (0, \frac{\pi}{2})$), qual é a fatoração espectral de $H(A)$?
   Resposta: A fatoração espectral de $H(A)$ fornece os valores singulares de A. Além disso, das $2n$ entradas correspondentes aos autovetores de $H(A)$ extraímos, após devida normalização, os vetores singulares u, v de A. Isto é, os autovalores de $H(A)$ são $\{\sigma_i, -\sigma_i : i = 1, \ldots, 2\}$, onde σ_i é o $i - simo$ valor singular de A. Como A é simétrica positiva definida (os dois autovalores informados são positivos), $\sigma_1 = \lambda_1 = 5, \sigma_2 = \lambda_2 = 3$.

   Para fazermos o processo inverso, isto é, para compor os autovetores de $H(A)$ a partir de y^1, y^2, observamos que a simetria de A garante que a mesma tem vetores singulares a esquerda e a direita iguais ($U = V$ na fatoração SVD).

   Então temos $H(A)Q = Q\Lambda$, onde $\Lambda = Diagonal(5, 3, -5, -3)$ e $Q \in \mathbb{R}^{4\times 4}$ é a matriz ortogonal, contendo em suas colunas os autovetores de $H(A)$, definida como:
   $$Q = \left(\begin{array}{cc|cc} \frac{1}{\sqrt{2}}y^1 & \frac{1}{\sqrt{2}}y^2 & \frac{1}{\sqrt{2}}y^1 & \frac{1}{\sqrt{2}}y^2 \\ \hline \frac{1}{\sqrt{2}}y^1 & \frac{1}{\sqrt{2}}y^2 & -\frac{1}{\sqrt{2}}y^1 & -\frac{1}{\sqrt{2}}y^2 \end{array}\right)$$
   Verifique, por meio da expressão de Q apresentada acima, que $Q^T Q = I_4$.

Referências Bibliográficas

[1] Sheldon Jay Axler. *Linear Algebra Done Right.* Undergraduate Texts in Mathematics. Springer, New York, 1997.

[2] G. Strang. *Introduction to Linear Algebra.* Wellesley, 2016.

[3] Gilbert Strang. *Linear Algebra and Learning from Data.* Wellesley-Cambridge Press, Wellesley, MA, 2019.

[4] Lloyd N. Trefethen and David Bau. *Numerical Linear Algebra.* SIAM, 1997.

[5] David S. Watkins. *Fundamentals of Matrix Computations.* Wiley, 1991.