ALC: Fundamentos de Álgebra Linear

Profs. Alexandre Salles da Cunha e Ana Paula Couto

Universidade Federal de Minas Gerais
Departamento de Ciência da Computação
Belo Horizonte, Brasil

acunha@dcc.ufmg.br
ana.coutosilva@dcc.ufmg.br

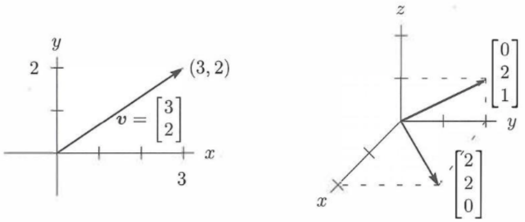
UFMG
UNIVERSIDADE FEDERAL
DE MINAS GERAIS

UFMG - ICEx
DEPARTAMENTO DE CIÊNCIA DA
COMPUTAÇÃO

Definição

Lista ordenada de n números reais.

- Cada número da lista é chamado de **coordenada** do vetor.
- Todo vetor será representado como um vetor coluna $\in \mathbb{R}^{n\times 1}$.
- Exemplos: $\begin{bmatrix}1\\2\end{bmatrix} \in \mathbb{R}^2$; $\begin{bmatrix}-\pi\\2\\\sqrt{3}\end{bmatrix} \in \mathbb{R}^3$; $\begin{bmatrix}1\\2\\3\\4\end{bmatrix} \neq \begin{bmatrix}2\\1\\3\\4\end{bmatrix} \in \mathbb{R}^4$



Introduction to Linear Algebra - Gilbert Strang

Soma

Associa a todo par $\{u, v\}$ de vetores um novo vetor $z \in \mathbb{R}^n$:

$$z = u + v = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$$

Soma

Associa a todo par $\{u, v\}$ de vetores um novo vetor $z \in \mathbb{R}^n$:

$$z = u + v = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$$

Elemento neutro ou origem ou vetor nulo

$$0_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \text{ que satisfaz } u + 0_n = u; \forall\, u \in \mathbb{R}^n$$

Multiplicação por escalar

Dado um escalar $d \in \mathbb{R}$ e um vetor $u \in \mathbb{R}^n$:

$$du = d \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} du_1 \\ du_2 \\ \vdots \\ du_n \end{bmatrix}$$

Definição

O produto interno dos vetores u e v no $\mathbb{R}^n$ é conhecido como produto linha-coluna de dois vetores ou produto escalar:

$$\langle u, v\rangle = u^T v = \sum_{i=1}^{n} u_i v_i$$

Definição

O produto interno dos vetores u e v no $\mathbb{R}^n$ é conhecido como produto linha-coluna de dois vetores ou produto escalar:

$$\langle u, v \rangle = u^T v = \sum_{i=1}^{n} u_i v_i$$

- Vetores u e v tais que $\langle u, v \rangle = 0$ são perpendiculares entre si.
- A norma Euclidiana de um vetor é dada por $||u||_2 = \sqrt{(u^T u)}$.
- O vetor u é um vetor unitário quando $u^T u = 1$.

Definição

x é a combinação linear de uma coleção de vetores $\{x^1, x^2, \cdots, x^m\}$, se pode ser escrito como:

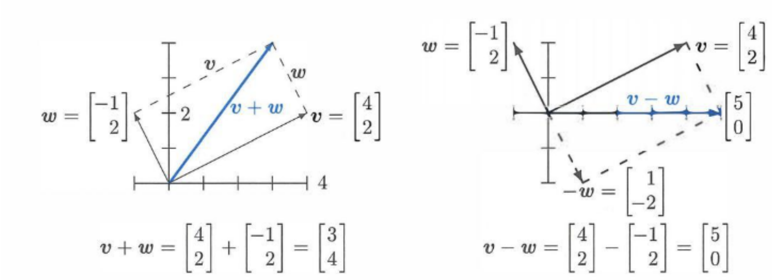
$$x = \sum_{i=1}^{m} \alpha_i x^i, \text{ para } \alpha_i \in \mathbb{R}.$$

Quatro combinações lineares "especiais":

- $1v + 1w$: soma de vetores.
- $1v - 1w$: diferença entre vetores.
- $0v + 0w$: vetor nulo.
- $cv + 0w$: vetor cv na direção de v.

Quatro combinações lineares "especiais":

- $1v + 1w$: soma de vetores.
- $1v - 1w$: diferença entre vetores.
- $0v + 0w$: vetor nulo.
- $cv + 0w$: vetor αv na direção de v.



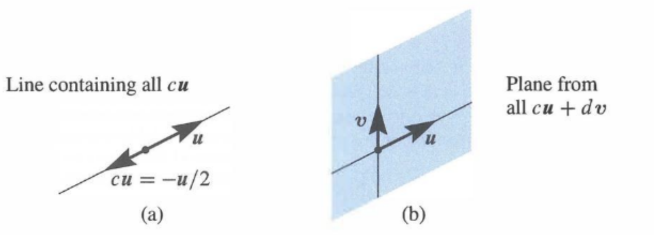
Introduction to Linear Algebra - Gilbert Strang

Dados os vetores $u, v, w \in \mathbb{R}^3$ e os escalares $c, d, f \in \mathbb{R}$:

- Qual a representação geométrica de todas as combinações cu?
- Qual a representação geométrica de todas as combinações $cu + dv$?
- Qual a representação geométrica de todas as combinações $cu + dv + fw$?

Dados os vetores $u, v, w \in \mathbb{R}^3$ e os escalares $c, d, f \in \mathbb{R}$:

- Qual a representação geométrica de todas as combinações cu?
- Qual a representação geométrica de todas as combinações $cu + dv$?
- Qual a representação geométrica de todas as combinações $cu + dv + fw$?



Introduction to Linear Algebra - Gilbert Strang

Independência linear

Uma coleção de vetores $\{x^1, \ldots, x^m\}$ de um espaço vetorial $\mathcal{X}$ é linearmente independente se a igualdade

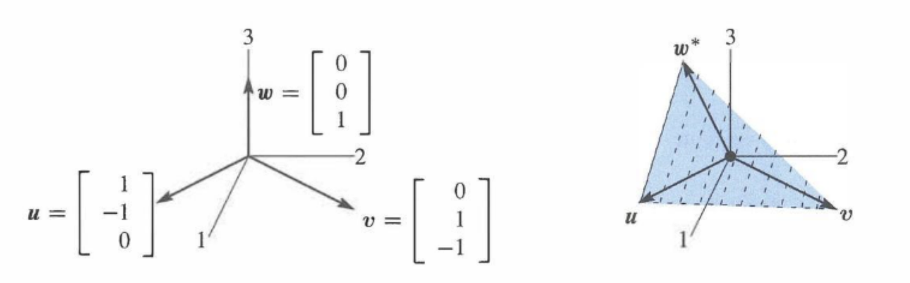
$$\sum_{i=1}^{m} \alpha_i x^i = 0$$

somente é verdadeira para $\alpha_i = 0; i = 1, \ldots, m$

Se existem $\alpha_i : i = 1, \ldots, m$, nem todos nulos tais que $\sum_{i=1}^{m} \alpha_i x^i = 0$, estes vetores são denominados linearmente dependentes.

Dimensão

A dimensão do espaço vetorial associado a um conjunto $\{x^1, \ldots, x^m\}$ é a cardinalidade do maior subconjunto linearmente independente de $\{x^1, \ldots, x^m\}$.



Introduction to Linear Algebra - Gilbert Strang

- Espaço vetorial: Conjunto não vazio $\mathcal{X}$, cujos elementos são denominados vetores, fechado nas operações de:
    - soma, que associa a todo par u, v de vetores de $\mathcal{X}$ um novo vetor $u + v \in \mathcal{X}$
    - multiplicação por escalar, para todo par $\alpha \in \mathbb{R}$ e $v \in \mathcal{X}$: $\alpha v \in \mathcal{X}$.
- Propriedades no espaço vetorial, para todo $u, v, w \in \mathcal{X}$:
    - 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: $\alpha(u + v) = \alpha u + \alpha v$, $(\alpha + \beta)u = \alpha u + \beta u$, $(\alpha\beta)u = \alpha(\beta u), 1u = u$.

Definição

Um subconjunto $\mathcal{V}$ não vazio de um espaço vetorial é um subespaço vetorial se e somente se $\mathcal{V}$ é fechado na soma de seus elementos e na multiplicação por escalar. Isto é, dados $u, v \in \mathcal{V}$ e $\alpha, \beta \in \mathbb{R}$, $\alpha u + \beta v \in \mathcal{V}$.

- Observação: se $\mathcal{V}$ é um subespaço vetorial, o vetor nulo $0 \in \mathcal{V}$.

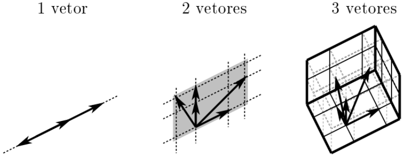
Span

O span ou subespaço gerado pelo conjunto $S = \{x^1, \ldots, x^m\}$, span(S), corresponde a todas as possíveis combinações lineares $\sum_{i=1}^{m} \alpha_i x^i$ (os pesos são $\alpha_i : i = 1, \ldots, m$).

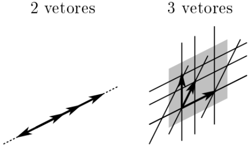
Base de um subespaço (espaço) vetorial

Uma base para um subespaço $\mathcal{V}$ (espaço $\mathcal{X}$) é um conjunto de vetores $S' = \{x^1, \ldots, x^d\}$ linearmente independentes tais que $\mathcal{V} = \text{span}(S')$. Neste caso, a dimensão do subespaço é a dimensão deste conjunto, d.

- Espaços gerados por vetores linearmente independentes.



- Espaços gerados por vetores linearmente dependentes.



Curso de Álgebra Linear - Fundamentos e Aplicações - Marco Cabral e Paulo Goldfeld

- $v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ e $w = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$: geram todo o espaço $\mathbb{R}^2$.
- $v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $w = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ e $u = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$: também geram todo o espaço $\mathbb{R}^2$.
- $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ e $w = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$: geram somente uma linha em $\mathbb{R}^2$.
- $v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ e $w = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ e $u = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$: geram todo o espaço $\mathbb{R}^3$.

Norma

Uma norma em um espaço vetorial é uma função real que associa a todo elemento $x \in \mathcal{X}$ um valor $\|x\|$ satisfazendo:

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

Exemplos de normas para $\mathcal{X} = \mathbb{R}^n$:

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

Casos particulares: $p = 2$ (norma Euclideana), $p = 1$ (norma soma de valores absolutos) e $p = \infty$ (máximo módulo).

$$u = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \; v = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \; w = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

- Combinação linear: $x_1 u + x_2 v + x_3 w$:

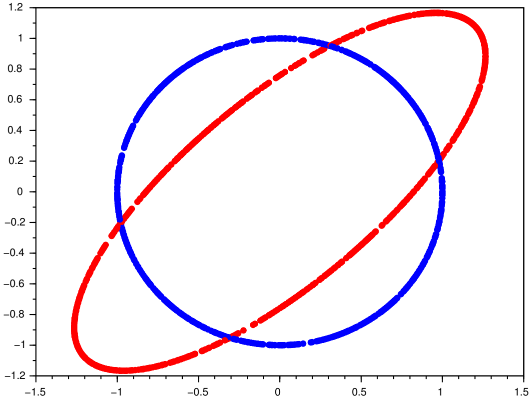
$$x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix}$$

- Representação Matricial:

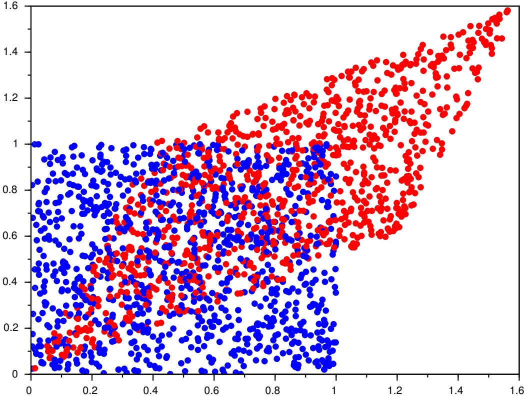
$$Ax = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = b$$

- b é a combinação linear das colunas de A.

Para $A = \begin{bmatrix} 1.2 & 0.4 \\ 0.6 & 1 \end{bmatrix}$, tomando alguns pontos aleatoriamente escolhidos $x \in \mathbb{R}^2 : \|x\|_2 = 1$ temos a seguinte transformação do disco:

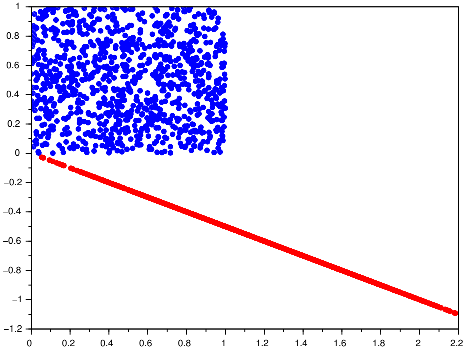


Tranformação do quadrado unitário



Área do quadrado $= 1$, área do losango $|a_{11}a_{22} - a_{21}a_{12}| = |\det(A)|$

$$A = \begin{pmatrix} 1.5 & 0.8 \\ -0.75 & -0.4 \end{pmatrix}$$



- Visão elemento a elemento: $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$

- Visão por linhas: $A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix}, a_i = \begin{bmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{bmatrix}$
  - a_i: vetor coluna associado a i-ésima linha de A.
  - a_i^T: i-ésima linha de A.

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

Produto Escalar:

$$b = Ax = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_m^T x \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n a_{1j} x_j \\ \sum_{j=1}^n a_{2j} x_j \\ \vdots \\ \sum_{j=1}^n a_{mj} x_j \end{bmatrix}$$

Combinação linear das colunas de A:

$$b = Ax = \begin{bmatrix} A_1 & A_2 & \dots & A_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} =$$

$$\begin{bmatrix} A_1x_1 + A_2x_2 + \cdots + A_nx_n \end{bmatrix}$$

Combinação linear das colunas de A:

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

Produto interno:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \;\; i = 1, \cdots, m; \; j = 1, \cdots, p$$

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

Colunas de C como uma combinação linear das colunas de A:

$$C = AB = A\begin{bmatrix} B_1 & B_2 & \dots & B_p \end{bmatrix} = \begin{bmatrix} AB_1 & AB_2 & \dots & AB_p \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{k=1}^{n} A_k b_{k1} & \sum_{k=1}^{n} A_k b_{k2} & \dots & \sum_{k=1}^{n} A_k b_{kp} \end{bmatrix}$$

Colunas de C como uma combinação linear das colunas de A:

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

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

Linhas de C como uma combinação linear das linhas de B:

$$C = AB = \begin{bmatrix} a_1^T \\ a_2^T \\ \dots \\ a_m^T \end{bmatrix} B = \begin{bmatrix} a_1^T B \\ a_2^T B \\ \dots \\ a_m^T B \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^n b_k^T a_{1k} \\ \sum_{k=1}^n b_k^T a_{2k} \\ \vdots \\ \sum_{k=1}^n b_k^T a_{mk} \end{bmatrix}$$

Linhas de C como uma combinação linear das linhas de B:

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

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

$$A = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & \cdots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \cdots & u_2 v_n \\ \vdots & \vdots & \vdots & \vdots \\ u_m v_1 & u_m v_2 & \cdots & u_m v_n \end{bmatrix} =$$

$$\begin{bmatrix} u_1 \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \\ u_2 \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \\ \vdots \\ u_m \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} v_1 & \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} v_2 \cdots & \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} v_n \end{bmatrix}$$

$$A = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & \cdots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \cdots & u_2 v_n \\ \vdots & \vdots & \vdots & \vdots \\ u_m v_1 & u_m v_2 & \cdots & u_m v_n \end{bmatrix} =$$

$$\begin{bmatrix} u_1 \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \\ u_2 \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \\ \vdots \\ u_m \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} v_1 & \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} v_2 \cdots & \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} v_n \end{bmatrix}$$

Exemplo:

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

Matrizes podem ser divididas em blocos

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

- Matriz 4×6: Blocos 2×2 resultam em uma matriz em blocos 2×3.
- Se B é uma matriz 4×6, com mesma divisão em blocos, $A + B$ é calculado somando cada bloco separadamente.
- Se os blocos de uma matriz A podem multiplicar blocos de uma matriz B, temos que:

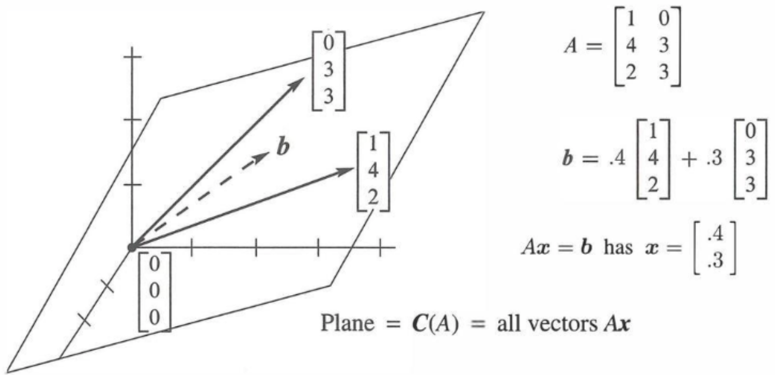
$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix}$$

Definição

1. $\mathcal{C}(A) = \{Ax | x \in \mathbb{R}^n\}$.
2. Espaço vetorial gerado pelas colunas de A.

Base de $\mathcal{C}(A)$

1. Conjunto de vetores LI que gera $\mathcal{C}(A)$.
2. A dimensão de $\mathcal{C}(A)$ é o total de vetores coluna LI de A.



Introduction to Linear Algebra - Gilbert Strang

- Pré-multiplicando A por y^T, onde $y \in \mathbb{R}^m$, geramos uma linha (n dimensional) que é uma *agregação* das m linhas de A.
- $y^T A = \begin{bmatrix} y_1 & y_2 & \dots & y_m \end{bmatrix} \begin{bmatrix} a_1^T \\ a_2^T \\ \dots \\ a_m^T \end{bmatrix} = \sum_{i=1}^m y_i a_i^T$.
- Ou, transpondo A, temos o conjunto dos vetores $A^T y$ para algum $y \in \mathbb{R}^m$.
- Ou seja, o espaço linha de A é o espaço coluna de A^T.

Definições

1. Conjunto formado pelos vetores $A^T y$, para algum y.
2. Espaço vetorial gerado pelas linhas linearmente independentes de A, i.e., $\mathcal{C}(\mathcal{A}^T) = \text{span}\left(\{a_1^T, a_2^T, \cdots, a_m^T\}\right)$.
3. Conjunto de vetores v_i resultantes das combinações lineares das linhas de A.

Base de $\mathcal{C}(A^T)$

1. Conjunto das linhas da matriz A que são linearmente independentes.
2. A dimensão de $\mathcal{C}(A^T)$ ($\dim \mathcal{C}(A^T)$) será o total de linhas linearmente independentes.

Definição

$r(A) =$ dimensão de $\mathcal{C}(A) =$ dimensão de $\mathcal{C}(A^T)$.

Propriedades

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

- Uma matriz A possui rank completo quando $r(A) = min(m, n)$.
- Deficiência de posto ou de rank: quando $min(m, n) - r(A) > 0$.

A matriz C é formada pelas colunas linearmente independentes de A

1. Incluir coluna 1 de A em C, caso não seja o vetor nulo.
2. Incluir coluna 2 de A em C, caso não seja múltipla da coluna 1.
3. Incluir coluna 3 de A em C, caso não seja combinação linear das colunas 1 e 2. *Continue*.

- C é formada por r colunas de A, $(r \leq n)$.
- As colunas de C formam UMA base para $\mathcal{C}(A)$.
- Colunas não incluídas são combinações lineares da base escolhida para $\mathcal{C}(A)$, i.e. das colunas LI de A.

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

$A = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 0 & 4 \end{bmatrix}, C = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, r = \mathbf{2}$ (fatoração CR alternativa à apresentada pelo algoritmo)

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}, C = A,\ r = \mathbf{3},\ \mathcal{C}(\mathcal{A}) = \mathbb{R}^3$$

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 1 & 2 & 5 \\ 1 & 2 & 5 \end{bmatrix}, C = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, r = \mathbf{1}$$

Suponha que as r primeiras colunas de A sejam as r colunas LI:

$$A = \begin{bmatrix} A_1 & A_2 & \cdots & A_r \end{bmatrix} R = \begin{bmatrix} A_1 & A_2 & \cdots & A_r \end{bmatrix} \begin{bmatrix} I_r & ? \end{bmatrix}$$

A matriz R é a forma escada reduzida da matriz A

1. Para as colunas em C, incluir nas colunas de R as entradas da matriz identidade.
2. Para as demais colunas em R, incluir as combinações para obter as colunas em A que não estão em C.

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

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

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

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

Se $r(D) = n$ então D é igual à soma de n matrizes de rank 1 (Produto externo de 2 vetores):

$$D = AB = \begin{bmatrix} A_1 & A_2 & \dots & A_n \end{bmatrix} \begin{bmatrix} b_1^T \\ b_2^T \\ \vdots \\ b_n^T \end{bmatrix} = \sum_{i=1}^{n} A_i b_i^T$$

Exemplo:

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

Definição

A matrix $B \in \mathbb{R}^{n\times n}$ que satisfaz as condições

$$AB = I \text{ e } BA = I$$

é chamada matriz inversa de A, denotada por $B = A^{-1}$.

- Nem toda matriz $n \times n$ possui inversa.
- Matrizes inversíveis são chamadas de matrizes não singulares.
- Matrizes não inversíveis são chamadas de matrizes singulares.
- Quando uma matriz possui inversa, esta é única.

A inversa de A existe se:

1. A é não singular.
2. $r(A) = n$. A é uma matriz de rank completo.
3. $det(A) \neq 0$.
4. $Ax = 0_n \implies x = 0_n$.

As afirmativas acima são equivalentes.

Toda matriz $A \in \mathbb{R}^{m\times n}$ gera 4 subespaços fundamentais

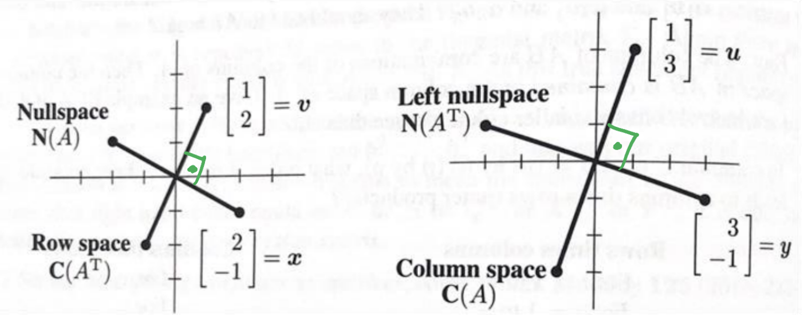
1. Espaço coluna $\mathcal{C}(A)$: todas as combinações das colunas de A. Subespaço de $\mathbb{R}^m$.
2. Espaço linha $\mathcal{C}(A^T)$: todas as combinações das colunas de A^T (linhas de A). Subespaço de $\mathbb{R}^n$.
3. Espaço nulo $\mathcal{N}(A)$: todas as soluções x para $Ax = 0_m$. Subespaço de $\mathbb{R}^n$.
4. Espaço nulo à esquerda $\mathcal{N}(A^T)$: todas as soluções y para $A^T y = 0_n$. Subespaço de $\mathbb{R}^m$.

Exemplo: $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$, $m = 2$, $n = 2$.

Subespaços gerados pertencem ao $\mathbb{R}^2$ (linhas infinitas):

1. Espaço coluna $\mathcal{C}(A)$ é a linha em $u = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$
   $\mathcal{C}(A) = \{\alpha u : \alpha \in \mathbb{R}\}$.
2. Espaço $\mathcal{C}(A^T)$ é a linha em $v = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathcal{C}(A^T) = \{\alpha v : \alpha \in \mathbb{R}\}$.
3. Espaço nulo $\mathcal{N}(A)$ é a linha em $x = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$
   $\mathcal{N}(A) = \{\alpha x : \alpha \in \mathbb{R}\}$.
4. Espaço nulo à esquerda $\mathcal{N}(A^T)$ é a linha em $y = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$;
   $\mathcal{N}(A^T) = \{\alpha y : \alpha \in \mathbb{R}\}$.

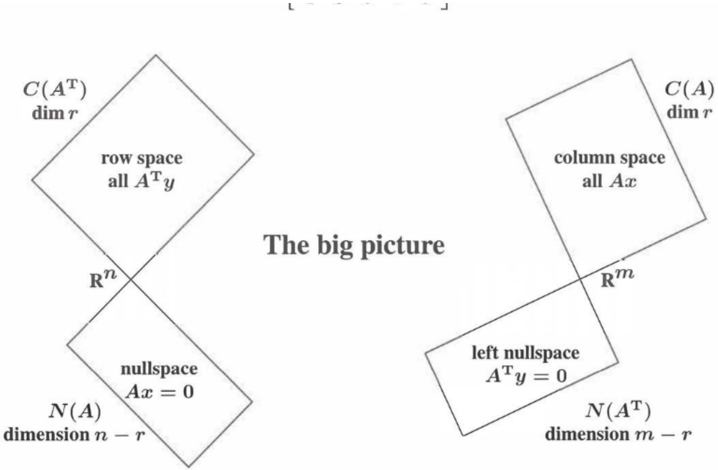
Exemplo: $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$, $m = 2$, $n = 2$.



Linear Algebra and Learning from Data - Gilbert Strang

Dimensões

1. Dimensão de $\mathcal{C}(A)$: $r = rank(A)$.
2. Dimensão de $\mathcal{C}(A^T)$: $r = rank(A)$.
3. Dimensão de $\mathcal{N}(A)$: $n - r$.
4. Dimensão de $\mathcal{N}(A^T)$: $m - r$.



Introduction to Linear Algebra - Gilbert Strang

Ortogonalidade significa perpendicular.

1. Vetores ortogonais. $x^T y = 0$.
2. Bases ortogonais de um subespaço. $v_i^T v_j = 0, \forall \{v_i, v_j\}$.
   - Bases ortonormais de um subespaço. $v_i^T v_i = 1, \forall v_i$
3. Subespaços ortogonais $\mathcal{R}$ e $\mathcal{N}$. Todo vetor em $\mathcal{R}$ é ortogonal a todo vetor em $\mathcal{N}$.
4. Matrizes *esbeltas* Q $(m >> n)$ com colunas ortonormais.
   - $Q^T Q = I_n$.
   - $\|Qx\|^2 = (Qx)^T (Qx) = x^T Q^T Qx = x^T x = \|x\|^2 \rightarrow \|Qx\| = \|x\|$.
5. Matrizes ortogonais. Matrizes ortonormais $n \times n$.
   - $Q^T = Q^{-1}$.
   - $Q^T Q = I$, $QQ^T = I$.
   - $(Qx)^T (Qy) = x^T y, \forall x, y$.
   - Norma Euclideana e ângulos não são alterados pela matriz Q.
   - As colunas de uma matriz ortogonal são uma base ortonormal para $\mathbb{R}^n$.
   - As linhas de uma matriz ortogonal são uma base ortonormal (provavelmente diferente) para $\mathbb{R}^n$
   - Computar com Q não acarreta *overflow*.

- Vetores Ortogonais e Teorema de Pitágoras.
  Se $x^T y = 0 \rightarrow ||x+y||^2 = ||x||^2 + ||y||^2$.

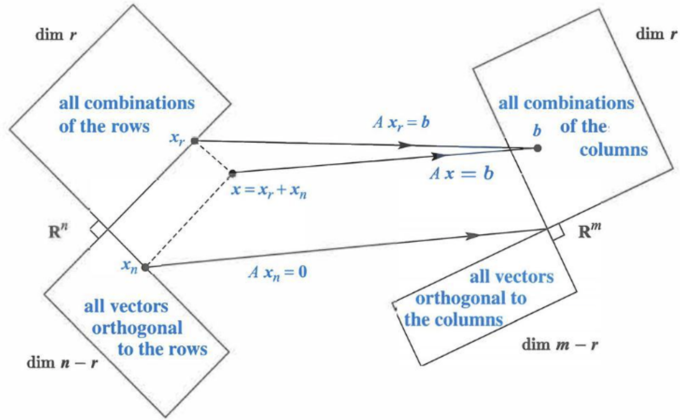
- Bases ortogonais.

  Base Canônica do $\mathbb{R}^3$. $i = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, j = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, k = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

- Subespaços Ortogonais.

  $Ax = 0_m$: Todas as linhas de A e suas combinações são ortogonais aos vetores $x \in \mathcal{N}(A)$.

  $A^T y = 0_n$: Todas as colunas de A e suas combinações são ortogonais aos vetores $x \in \mathcal{N}(A^T)$.



Introduction to Linear Algebra - Gilbert Strang

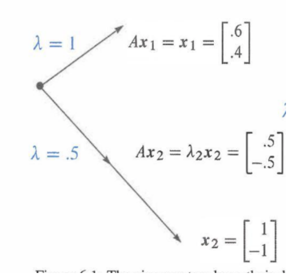
- Matrizes *esbeltas* Q com colunas ortonormais: $Q^TQ = I$.

$$Q_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \; Q_2 = \frac{1}{3}\begin{bmatrix} 2 & 2 \\ 2 & -1 \\ -1 & 2 \end{bmatrix}$$

Definição

Todo vetor x linearmente dependente com Ax é chamado de autovetor de $A \in \mathbb{R}^{n \times n}$. Ou seja: $Ax = \lambda x$.

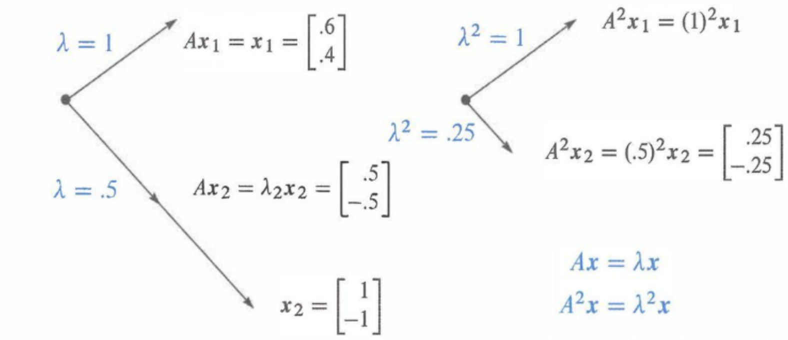
- x é um autovetor de A.
- $\lambda \in \mathbb{C}$ é um autovalor de A (o autovalor pode ser complexo).
- (λ, x) é autopar de A.



Introduction to Linear Algebra - Gilbert Strang

Os autovetores da matriz A também são autovetores da matriz A^k.

$$k = 2 : A(Ax) = A(\lambda x) = \lambda(Ax) = \lambda^2 x.$$
$$k = 3, \cdots : A^k x = \lambda^k x.$$



Introduction to Linear Algebra - Gilbert Strang

- Os autovetores da matriz A também são autovetores da matriz A^{-1}:

$$(A^{-1}A)x = A^{-1}(\lambda x)$$
$$x = A^{-1}(\lambda x)$$
$$A^{-1}x = \frac{1}{\lambda}x$$

- Para matrizes com n autovetores LI, $x^1, \cdots, x^n$ com n autovalores $\lambda_1 > \lambda_2 > \cdots > \lambda_n$:

$$\forall v \in \mathbb{R}^n, v = c_1x^1 + c_2x^2 + \cdots + c_nx^n$$
$$Av = A(c_1x^1 + c_2x^2 + \cdots + c_nx^n)$$
$$Av = c_1Ax^1 + c_2Ax^2 + \cdots + c_nAx^n$$
$$Av = c_1\lambda_1x^1 + c_2\lambda_2x^2 + \cdots + c_n\lambda_nx^n$$
$$A^kv = c_1\lambda_1^kx^1 + c_2\lambda_2^kx^2 + \cdots + c_n\lambda_n^kx^n$$

Se $|\lambda_1| > 1$, o componente $c_1\lambda_1^kx^1$ crescerá com o aumento de k. Se $|\lambda_2| < 1$ então o componente $c_2\lambda_2^kx^2 \to 0$ quando $k \to \infty$ e A^kv é dominado por $c_1\lambda^kx^1$.

- Se A é deslocada de $A + sI$, 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).

- 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.

Exemplo:

$S = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ tem como autovetores $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, com autovalor $\lambda_1 = 3$ e $v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, com autovalor $\lambda_1 = 1$.

Propriedades

- Traço: A soma $\lambda_1 + \lambda_2 = 4$ é igual a soma dos elementos na diagonal de S (2+2=4).
- Determinante: O produto $\lambda_1 \lambda_2 = 3$ é igual ao determinante de S (4-1=3).
- Matrizes simétricas $S = S^T$ possuem sempre autovalores reais.
- Matrizes com $\lambda_1 \neq \lambda_2 \neq \cdots \neq \lambda_n$ possuem autovetores li.

- Os autovalores de $A + B$ normalmente diferem da soma dos autovalores de A, B. Isto é, $\lambda(A + B)$ difere de $\lambda(A) + \lambda(B)$.
- Os autovalores $\lambda(AB)$ de AB normalmente diferem de $\lambda(A)\lambda(B)$.
- Sobre autovalores com multiplicidade algébrica maior que 1 (repetidos): Autovalores $\lambda_1 = \lambda_2$ podem ou não ter autovetores independentes.
- Os autovetores de A são ortogonais se e somente se $A^T A = AA^T$.

Temos:

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

- Os autovetores de A geram o espaço $\mathcal{N}(A - \lambda I)$.
- Se $(A - \lambda I)x = 0_n$ tem solução com $x \neq 0_n$, $(A - \lambda I)$ não é inversível (singular) e $det(A - \lambda I) = 0$.
- Resolva $det(A - \lambda I) = 0$, i.e., encontre as n raízes do polinômio característico $det(A - \lambda I)$ que são os n autovalores de A.
- Para cada autovalor λ resolva $(A - \lambda I)x = 0_n$ ou $Ax = \lambda x$ para encontrar o autovetor correspondente.

Exemplo: Encontre os autovalores e autovetores da matriz

$$A = \begin{bmatrix} 8 & 3 \\ 2 & 7 \end{bmatrix}.$$

$$det(A-\lambda I) = \begin{bmatrix} 8-\lambda & 3 \\ 2 & 7-\lambda \end{bmatrix} = \lambda^2 - 15\lambda + 50 = (\lambda-10)(\lambda-5) = 0$$

$$\lambda_1 = 10, x_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \text{ e } \lambda_2 = 5, x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Matrizes simétricas (i.e., $S = S^T$) possuem as seguintes propriedades:

1. Todos os n autovalores λ são números reais.
2. Os n autovetores q podem ser escolhidos de tal forma que sejam ortonormais.
3. Toda matriz simétrica S pode ser fatorada na forma de $S = Q\Lambda Q^T$ (encontrar os fatores Q, Λ consiste na diagonalização de S), onde Q é ortogonal e Λ é diagonal.

   $\Rightarrow$ Isto é, toda matriz simétrica é diagonalizável, ou seja, é similar a uma matriz diagonal.

   $\Rightarrow$ Portanto o espectro de S é o espectro de Λ.

Matrizes SPD possuem as seguintes propriedades (se uma propriedade é satisfeita, todas as demais também são):

1. Teste 1: Todos os autovalores são positivos.
2. Teste 2: A forma quadrática (também chamada energia) $x^T S x$ é positiva para todos os vetores $x \neq 0_n$.
3. Teste 3: S admite uma fatoração $S = A^T A$ (forma de Gram), para A com colunas independentes. (Obs: Em particular, S admite uma fatoração de Cholesky $S = LL^T$, onde L é triangular inferior com diagonal positiva).
4. Teste 4: Determinantes das submatrizes principais são positivos.
5. Teste 5: Possuem pivôs positivos (processo de Eliminação).

Teste 2: Energia

$$x^T S x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 8x_1x_2 + 9x_2^2.$$

- Este valor é positivo para todo x_1 e x_2, exceto para $(x_1, x_2) = (0, 0)$? Sim, pois é a soma de quadrados.

$$x^T S x = 2x_1^2 + 8x_1x_2 + 9x_2^2 = 2(x_1 + 2x_2)^2 + x_2^2 > 0.$$

- Se $\lambda > 0$ então $x^T S x > 0$:

$$Sx = \lambda x \rightarrow x^T S x = \lambda x^T x. \text{ Então } \lambda > 0 \text{ resulta em } x^T S x > 0.$$

Uso do teste de Energia para definir se uma matriz é SPD:

Se S_1 e S_2 são matrizes SPD, então $S_1 + S_2$ é uma matriz SPD.

Prova:

$$x^T(S_1 + S_2)x = x^T S_1 x + x^T S_2 x > 0 + 0.$$

Os autovalores e autovetores de $S_1 + S_2$ são computacionalmente caros para serem calculados.

Teste 3: Se S é SPD então $S = A^T A$, para A com colunas independentes.

$$x^T S x = x^T A^T A x = (Ax)^T (Ax) = ||Ax||^2.$$

- Energia é o quadrado da norma Euclideana de Ax.
- A energia é positiva se $Ax \neq 0$.
- $Ax \neq 0$ para algum $x \neq 0$ somente quando as colunas de A são linearmente independentes.

Por exemplo:

$$S = A^T A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 5 & 7 \\ 4 & 7 & 10 \end{bmatrix}.$$

não é positiva definida:

- Coluna 2 de A é combinação linear das colunas 1 e 3 (coluna 1 $+$ coluna 3 $= 2\times$ coluna 2).
- $A^T A$ possui um autovalor com $\lambda = 0$. Neste caso $S = A^T A$ é semipositiva definida.
- Teste 3 garante que $S = A^T A$ seja, pelo menos, uma matriz simétrica semipositiva definida, pois $x^T S x = ||Ax||^2$ nunca será negativa.

Matrizes Simétricas Positivas Definidas (SPD)

Teste 4 e Teste 5: Determinantes e Pivôs.

- O teste do determinante é pouco custoso para matrizes pequenas.

$$S = \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{bmatrix} \quad \text{has} \quad \begin{array}{ll} \text{1st determinant} & D_1 = \mathbf{2} \\ \text{2nd determinant} & D_2 = \mathbf{3} \\ \text{3rd determinant} & D_3 = \mathbf{4} \\ \text{4th determinant} & D_4 = \mathbf{5} \end{array}$$

Linear Algebra and Learning from data - Gilbert Strang

- Pivôs (valores na diagonal da matriz após o processo de eliminação) são positivos, pois o k-ésimo pivô é dado por:

$$\frac{D_k}{D_{k-1}}$$