

PNL: Revisão de Fundamentos

Prof. Alexandre Salles da Cunha

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

acunha@dcc.ufmg.br

2022/1



UNIVERSIDADE FEDERAL
DE MINAS GERAIS



$$\begin{aligned} & \min f_0(x) \\ & f(x) \leq 0 \\ & h(x) = 0 \end{aligned}$$

Hipóteses:

- $f_0(x) : \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $h(x) : \mathbb{R}^n \rightarrow \mathbb{R}^q$ são funções contínuas, com derivadas de primeira ordem contínuas ($f_i : i = 0, \dots, m$, $h_i \in \mathbb{C}^1$).
- Quando conveniente, também assumiremos que suas derivadas de segunda ordem existem ($f_i, h_i \in \mathbb{C}^2$)

- **Espaço vetorial:** Dado um **corpo F** cujos elementos são denominados escalares, um espaço vetorial sobre F é um **conjunto não vazio \mathcal{X}** , cujos elementos são denominados vetores, em conjunto com **duas operações**:
 - **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 $\{r, v\} : F \times \mathcal{X}$ associa o vetor $rv \in \mathcal{X}$.
- **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: $r(u + v) = ru + rv$, $(a + b)u = au + bu$, $(ab)u = a(bu)$, $1u = u$.

Espaço vetorial de interesse: \mathbb{R}^n .

Definição

Um subconjunto \mathcal{S} não vazio de um espaço vetorial é um subespaço vetorial se e somente se \mathcal{S} é **fechado na soma de seus elementos e na multiplicação por escalar**. Isto é, dados $u, v \in \mathcal{S}$ e $a, b \in F$, $au + bv \in \mathcal{S}$.

- Observação: se \mathcal{S} é um subespaço vetorial, $0 \in \mathcal{S}$.

Independência linear

Uma coleção de vetores $\{x^1, \dots, x^m\}$ de um espaço vetorial \mathcal{X} é linearmente independente se a única solução para o sistema linear

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

é a solução trivial $\alpha_i = 0 : i = 1, \dots, m$

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

Dimensão

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

Span

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

Base de um subespaço (espaço) vetorial

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

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 $\mathcal{X} = \mathbb{R}^n$:

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

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

Bola unitária na norma p

O conjunto de vetores cuja norma p não excede a unidade é chamado de bola unitária na norma p :

$$\mathcal{B}_p = \{x \in \mathbb{R}^n : \|x\|_p \leq 1\}$$

Exemplos: $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_\infty$.

Bola unitária com centro em x_c e raio r

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

Norm cone

$$\{(x, t) \in \mathbb{R}^{n+1} : \|x\|_p \leq t\}$$

Operação fundamental sobre elementos de um espaço vetorial. Dados $x, y, z \in \mathcal{X}$ e um escalar $\alpha \in F$, o produto interno $\langle \cdot, \cdot \rangle$ é uma função que satisfaz:

- $\langle x, x \rangle \geq 0$
- $\langle x, x \rangle = 0 \iff x = 0$
- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- $\langle \alpha x, z \rangle = \alpha \langle x, z \rangle$
- $\langle z, x \rangle = \langle x, z \rangle$

Se um produto interno é definido para \mathcal{X} , dizemos que é um espaço vetorial equipado com produto interno (epi).

O produto interno padrão no \mathbb{R}^n é conhecido como o produto linha-coluna de dois vetores: $\langle x, y \rangle = x^T y = \sum_{i=1}^n x_i y_i$

Observações:

- Em um espaço vetorial epi, a função $\sqrt{\langle x, x \rangle}$ define uma norma, designada simplesmente por $\|x\| = \|x\|_2$.
- Outros produtos internos podem ser definidos no \mathbb{R}^n .
- A definição de produto interno pode ser estendida a outros espaços vetoriais, distintos do \mathbb{R}^n , por exemplo, espaço de matrizes.

Ortogonalidade

Dado um espaço vetorial \mathcal{X} e $x, y \in \mathcal{X}$ são ortogonais se $\langle x, y \rangle = 0$. A ortogonalidade entre x, y é expressa via $x \perp y$.

Um conjunto de vetores $\{x^1, \dots, x^d\}$ é mutuamente ortogonal se $\langle x^i, x^j \rangle = 0$ para qualquer $i \neq j$.

Proposição

Vetores mutuamente ortogonais são linearmente independentes.

Um conjunto de vetores $\{x^1, \dots, x^d\}$ é ortonormal se são ortogonais e possuem norma unitária:

$$\langle x^i, x^j \rangle = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i = j \end{cases}$$

Base ortonormal de \mathcal{V}

Uma base $\{x^1, \dots, x^d\}$ formada por vetores ortonormais.

Definição

Seja \mathcal{V} um subespaço vetorial de um espaço vetorial \mathcal{X} equipado com produto interno $\langle \cdot, \cdot \rangle$. Então, o complemento ortogonal \mathcal{V}^\perp de \mathcal{V} é o conjunto:

$$\mathcal{V}^\perp = \{x \in \mathcal{X} : \langle x, y \rangle = 0 \text{ para qualquer } y \in \mathcal{V}\}$$

Além disto, vale $(\mathcal{V}^\perp)^\perp = \mathcal{V}$.

- Se \mathcal{V} é um subespaço vetorial de um espaço vetorial \mathcal{X} epi, então todo elemento $x \in \mathcal{X}$ pode ser decomposto como a soma de um elemento $v \in \mathcal{V}$ e outro no complemento ortogonal $s \in \mathcal{V}^\perp$.

Denotamos $\mathcal{X} = \mathcal{V} \oplus \mathcal{V}^\perp$ (soma direta).

Dados $x, y \in \mathcal{X}$ epi, $\|x\| = \sqrt{\langle x, x \rangle}$:

- $|\langle x, y \rangle| \leq \|x\| \|y\|$ (Cauchy-Schwartz).
- $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$ (paralelogramo)
- Se $x \perp y$, $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ (Pitágoras)
- $\mathcal{X} = \mathcal{V} \oplus \mathcal{V}^\perp$
- $\dim(\mathcal{X}) = \dim(\mathcal{V}) + \dim(\mathcal{V}^\perp)$

O produto interno $x^T y$ do \mathbb{R}^n relaciona-se com o ângulo θ entre x e y . Definindo $z = x - y$ temos

$$\begin{aligned}\|z\|_2^2 &= (\|y\|_2 \sin(\theta))^2 + (\|x\|_2 - \|y\|_2 \cos(\theta))^2 \\ &= \|x\|_2^2 + \|y\|_2^2 - 2\|x\|_2\|y\|_2 \cos(\theta)\end{aligned}$$

$$\|z\|_2^2 = \|x - y\|_2^2 = (x - y)^T (x - y) = x^T x + y^T y - 2x^T y$$

Logo $\cos(\theta) = \frac{x^T y}{\|x\|_2\|y\|_2}$

- Se $x \perp y \rightarrow \theta = \frac{\pi}{2}, x^T y = 0$.
- Se x e y são linearmente dependentes, $\cos(\theta) = 1, \theta = 0$.

Dado um conjunto fechado \mathcal{S} e um $x \in \mathcal{X}$ epi (por exemplo $\mathcal{X} = \mathbb{R}^n$ com produto interno $x^T y$), a projeção de x em \mathcal{S} consiste em um ponto que minimiza a distância

$$\text{Proj}(x)_{\mathcal{S}} = \arg \min \|y - x\| : y \in \mathcal{S}$$

onde a norma considerada aqui é induzida pelo produto interno
 $\|x - y\| = \sqrt{\langle x - y, x - y \rangle}$

Teorema da Projeção

Seja \mathcal{X} um espaço vetorial epi (evepi), $x \in \mathcal{X}$ e \mathcal{S} um subespaço de \mathcal{X} . Então existe um único vetor x^* que resolve

$$\min_{y \in \mathcal{S}} \|y - x\|$$

Além disto, condições necessárias e suficientes para que x^* seja a solução do problema de otimização acima são:

- $x^* \in \mathcal{S}$
- $(x - x^*) \perp \mathcal{S}$.

- Uma vez que $\mathcal{X} = \mathcal{S} \oplus \mathcal{S}^\perp$, um ponto $x \in \mathcal{X}$ pode ser escrito como

$$x = u + z, \quad u \in \mathcal{S}, \quad z \in \mathcal{S}^\perp.$$

- Para qualquer $y \in \mathcal{S}$:

$$\begin{aligned}\|y - x\|^2 &= \|(y - u) - z\|^2 \\ &= \|y - u\|^2 + \|z\|^2 - 2\langle y - u, z \rangle \\ &= \|y - u\|^2 + \|z\|^2\end{aligned}\tag{1}$$

já que $y - u \in \mathcal{S}, z \in \mathcal{S}^\perp$.

- A solução de $\min_{y \in \mathcal{S}} \{\|y - u\|^2 + \|z\|^2\}$ é ($x^* =)y = u$
- Com a escolha acima, $z = x - y \in \mathcal{S}^\perp$.

- **Linha** entre dois pontos x^1, x^2 : $x(\theta) = \theta x^1 + (1 - \theta)x^2$ para $\theta \in \mathbb{R}$.
- **Conjunto afim**: Contém a linha entre quaisquer dois pontos no conjunto. Por exemplo, um conjunto de restrições lineares:
$$\mathcal{A} = \{x \in \mathbb{R}^n : Ax = b\}.$$
- Alternativamente, todo conjunto afim pode ser representado como uma solução de um sistema de equações lineares.
- Uma linha é um conjunto afim unidimensional.

- Um conjunto afim pode ser definido como uma **translação de um subespaço vetorial**:

$$\mathcal{A} = \{x \in \mathbb{R}^n : x = x^0 + v, v \in \mathcal{S}\}$$

onde \mathcal{S} é um subespaço vetorial do \mathbb{R}^n e $x^0 \in \mathbb{R}^n$.

- Dado $\mathcal{A} = \{x \in \mathbb{R}^n : Ax = b\} \neq \emptyset$ (isto é, $b \in \mathcal{R}(A)$), podemos obter a representação acima escolhendo x^0 como uma solução particular de $Ax = b$ ($A \in \mathbb{R}^{m \times n}$, $m \leq n$) e \mathcal{S} sendo o espaço nulo $\mathcal{N}(A)$ de A . Esta representação pode ser obtida via fatoração $A = QR$.
- A **dimensão do conjunto afim** é a dimensão do seu subespaço gerador \mathcal{S} , no exemplo acima, $\dim(\mathcal{N}(A))$.

Corolário

Seja \mathcal{X} um espaço, $x \in \mathcal{X}$ e \mathcal{A} um conjunto afim, descrito por $\mathcal{A} = x^0 + \mathcal{S}$. Isto é, o conjunto afim foi obtido por uma translação de \mathcal{S} pelo vetor x^0 . Então, existe um único vetor $x^* \in \mathcal{A}$ que resolve:

$$\min_{y \in \mathcal{A}} \|y - x\|$$

Além disto, condições necessárias e suficientes para que x^* seja a solução deste problema de otimização são:

- $x^* \in \mathcal{A}$
- $x - x^* \perp \mathcal{S}$.

- Assumimos que o conjunto afim é dado por $\mathcal{A} = \{y \in \mathbb{R}^n | y = x^0 + z, z \in \mathcal{S}\}$. Logo

$$\min_{y \in \mathcal{A}} \|y - x\| = \min_{z \in \mathcal{S}} \|z - (x - x^0)\|.$$

- O problema então consiste em projetar o ponto $(x - x^0) \in \mathcal{X}$ em \mathcal{S} (projeção de ponto em sub-espacô linear).
- Pelo Teorema da Projeção, as condições de otimalidade para este problema são:
 - $z^* \in \mathcal{S} \rightarrow x^* = x^0 + z^*, x^* \in \mathcal{A}$.
 - $z^* - (x - x^0) \perp \mathcal{S} \rightarrow (x^* - x) \perp \mathcal{S}$

- Vamos supor que $\mathcal{S} \subseteq \mathcal{X}$ seja um sub-espaco vetorial e que $\{x^1, x^2, \dots, x^d\}$ uma base para \mathcal{S} .
- O Teorema da Projeção estabelece que a projeção de x em \mathcal{S} é $x^* \in \mathcal{S}$ tal que $(x^* - x) \perp \mathcal{S}$.
- Escrevendo $x^* = \sum_{i=1}^d \alpha_i x^i$, temos que $x^* - x \perp \mathcal{S}$ implica em $\langle x^* - x, x^k \rangle = 0$, $k = 1, \dots, d$.
- Encontrar a projeção x^* corresponde a encontrar os $\alpha_i \in \mathbb{R} : i = 1, \dots, d$ resolvendo o sistema linear:

$$\sum_{i=1}^d \alpha_i \langle x^i, x^k \rangle = \langle x, x^k \rangle, \quad k = 1, \dots, d$$

- Se $\{x^1, x^2, \dots, x^d\}$ for uma base ortonormal, temos que $\alpha_i = \langle x^i, x \rangle$ e $x^* = \sum_{i=1}^d \langle x^i, x \rangle x^i$.

$$\begin{pmatrix} (x^1)^T x^1 & \dots & (x^1)^T x^d \\ \vdots & \ddots & \vdots \\ (x^d)^T x^1 & \dots & (x^d)^T x^d \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} = \begin{pmatrix} (x^1)^T x \\ \vdots \\ (x^d)^T x \end{pmatrix}$$
$$G = X^T X$$

Dada uma base $\{x^1, x^2, \dots, x^d\}$ para $\mathcal{S} = \text{span}\{x^1, x^2, \dots, x^d\} \subseteq \mathcal{X}$, desejamos encontrar uma base ortonormal para o mesmo espaço.

- Procedimento muito simples
- Importante para a construção de resultados
- Não figura entre os procedimentos mais estáveis.
A fatoração QR (ortogonal-triangular), por exemplo, é mais estável e é muito relacionada ao procedimento de Gram-Schmidt.

- Tome um vetor qualquer na base original: $\zeta^1 = x^1$ e faça $z^1 = \frac{\zeta^1}{\|\zeta^1\|}$ (observe que $\|\zeta^1\| \neq 0$, dado que $\{x^1, x^2, \dots, x^d\}$ são l.i.)
- Tomamos o próximo vetor x^2 . Sua projeção sobre $\text{span}(\{z^1\})$ é $\tilde{x}^2 = \langle x^2, z^1 \rangle z^1$, uma vez que z^1 é uma base ortonormal para este subespaço. Logo, $(x^2 - \langle x^2, z^1 \rangle z^1) \perp \text{span}(\{z^1\})$. Fazemos:

$$\zeta^2 = x^2 - \langle x^2, z^1 \rangle z^1, \quad z^2 = \frac{\zeta^2}{\|\zeta^2\|}$$

Observe que $\|\zeta^2\| \neq 0$ pois $\{x^2, z^1\}$ são l.i. (já que $\{x^2, x^1\}$ são l.i.)

Dispomos de um conjunto ortonormal $\{z^1, \dots, z^{k-1}\}$ ($k-1$ dimensional)

- Projete x^k em $\text{span}(\{z^1, \dots, z^{k-1}\})$:

- A projeção é $\tilde{x}^k = \sum_{i=1}^{k-1} \langle x^k, z^i \rangle z^i$.

- A diferença $\zeta^k = (x^k - \sum_{i=1}^{k-1} \langle x^k, z^i \rangle z^i)$ satisfaz

$$\zeta^k \perp \text{span}(\{z^1, \dots, z^{k-1}\})$$

- Normalize: $z^k = \frac{\zeta^k}{\|\zeta^k\|}$ (veja que $\|\zeta^k\| \neq 0$)

Ao final, dispomos de conjunto ortonormal $\{z^1, \dots, z^k\}$ k dimensional.

O termo **mapeamento** é usualmente empregado para funções cuja **saída** é um vetor e não um escalar. Exemplo: $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$

Dada uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$:

- $\text{dom } f = \{x \in \mathbb{R}^n : -\infty < f(x) < \infty\}$ é conjunto de pontos onde a função é finita. Ex: $f(x) = \log(x)$, $\text{dom } f = \mathbb{R}_{++}$ (**domínio**).

Veja que $g(x) = \begin{cases} 1/x & \text{se } x \neq 0 \\ +\infty & \text{cc} \end{cases}$ e

$h(x) = \begin{cases} 1/x & \text{se } x > 0 \\ +\infty & \text{cc} \end{cases}$ são distintas, embora sejam definidas pela mesma expressão em seus respectivos domínios.

- $\text{graph } f = \{(x, f(x)) \in \mathbb{R}^{n+1} : x \in \mathbb{R}^n\}$ (**gráfico**)
- $\text{epi } f = \{(x, t) \in \mathbb{R}^{n+1} : x \in \mathbb{R}^n, t \geq f(x)\}$ (**epigraph**)
- $C_f(t) = \{x \in \mathbb{R}^n : f(x) = t\}$ (**conjunto de nível**).
- $L_f(t) = \{x \in \mathbb{R}^n : f(x) \leq t\}$ (**conjunto de sub-níveis**).

Linear

Preserva escala e adição do argumento de entrada.

$f : \mathbb{R}^n \rightarrow \mathbb{R}$ é linear sse:

- $f(\alpha x) = \alpha f(x)$ para quaisquer $\alpha \in \mathbb{R}$, $x \in \mathbb{R}^n$.
- $f(x_1 + x_2) = f(x_1) + f(x_2)$ para quaisquer $x_1, x_2 \in \mathbb{R}^n$

Afim

$f : \mathbb{R}^n \rightarrow \mathbb{R}$ é afim sse $f(x) - f(0)$ é linear (afim = linear + constante).

Ambas podem ser convenientemente escritas por meio de produto interno: A função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é afim sse puder ser expressa como $f = a^T x + b$ para um único par (a, b) , $a \in \mathbb{R}^n$, $b \in \mathbb{R}$. f é linear sse $b = 0$.

É um conjunto de nível de uma função definida como um produto escalar. H é um hipérplano sse existem $a \in \mathbb{R}^n$, $a \neq 0$, $b \in \mathbb{R}$ tais que:

$$H = \{x \in \mathbb{R}^n : a^T x = b\}.$$

- Se $b = 0$, H é o conjunto de pontos ortogonais ao vetor a . Neste caso, H é um subespaço linear $(n - 1)$ dimensional. Tome $x \in H$, isto é, $x : a^T x = 0$ e veja que $x \in \text{span}(\{a\})^\perp$ que é um espaço $n - 1$ dimensional.
- Se $b \neq 0$, H é uma translação do subespaço \mathcal{S} , um conjunto afim $(n - 1)$ dimensional.

Tomando $x_0 \in H$, para qualquer outro $x \in H$, $x - x_0 \in \mathcal{S}$ e $a^T(x - x_0) = 0$. Logo:

$$H = \{x \in \mathbb{R}^n : a^T(x - x_0) = 0\}$$

- Conjuntos afins $n - 1$ dimensionais que generalizam o plano no \mathbb{R}^3 .
- Caracterizações:

$$\begin{aligned} H &= \{x \in \mathbb{R}^n : a^T x = b\} \\ &= \{x \in \mathbb{R}^n : a^T(x - x_0) = 0\} \\ &= x_0 + \mathcal{S} \\ &= x_0 + \text{span}(\{u_1, \dots, u_{n-1}\}) \end{aligned}$$

onde $\{u_1, \dots, u_{n-1}\}$ é uma base para $\mathcal{S} = \text{span}(\{a\})^\perp$ e $x_0 : a^T x_0 = b$.

Um hiperplano H separa o espaço em duas regiões:

- $H_- = \{x : a^T x \leq b\}$ e
- $H_{++} = \{x : a^T x > b\}$.

Estas duas regiões são chamadas semi-espacos (*half-space*). H_- é um conjunto fechado enquanto H_{++} é um conjunto aberto.

- Visão por colunas

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}$$

- Visão por linhas

$$A = \begin{pmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_m^T \end{pmatrix}$$

- Transformação das colunas de B .

$$AB = A \begin{pmatrix} b_1 & b_2 & \dots & b_p \end{pmatrix} = \begin{pmatrix} Ab_1 & Ab_2 & \dots & Ab_p \end{pmatrix}$$

- Transformação das linhas de A

$$AB = \begin{pmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_m^T \end{pmatrix} B = \begin{pmatrix} \alpha_1^T B \\ \alpha_2^T B \\ \vdots \\ \alpha_m^T B \end{pmatrix}$$

- Produto de matrizes de posto-1

$$AB = \sum_{i=1}^n a_i \beta_i^T$$

- Conjunto das matrizes $A \in \mathbb{R}^{m \times n}$ define um espaço vetorial.
- Qualquer norma vetorial do \mathbb{R}^{mn} pode ser usada para *medir* A . Podemos equipar este espaço com o produto interno padrão do \mathbb{R}^{mn} , empilhando as colunas de A em um vetor mn dimensional.
- Para $A, B \in \mathbb{R}^{m \times n}$

$$\langle A, B \rangle = \text{traço}(A^T B) = \sum_{i=1}^n \sum_{j=1}^m a_{ji} b_{ji}$$

$$(\text{traço}(X) = \sum_{i=1}^n x_{ii})$$

- Este produto interno induz a **norma de Frobenius** (análoga à norma vetorial $\|\cdot\|_2$)

$$\sqrt{\langle A, A \rangle} = \sqrt{\text{traço}(AA^T)} = \|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^m a_{ji}^2}$$

Norma matricial

Uma norma matricial é uma função real que associa a todo elemento $X \in \mathbb{R}^{m \times n}$ um valor $\|X\|$ satisfazendo:

- $\|X\| \geq 0$ para todo $X \in \mathbb{R}^{m \times n}$ e $\|X\| = 0 \iff X = 0_{m,n}$.
- $\|X + Y\| \leq \|X\| + \|Y\|$ para todo $X, Y \in \mathbb{R}^{m \times n}$.
- $\|\alpha X\| = |\alpha| \|X\|$ para todo $\alpha \in \mathbb{R}, X \in \mathbb{R}^{m \times n}$.

A norma de Frobenius é uma norma matricial.

Caso particular de normas: normas matriciais induzidas por normas vetoriais.

$\|A\|_p$

A norma matricial p induzida por norma vetorial p é o menor escalar C tal que $\|Ax\|_p \leq C\|x\|_p$ para qualquer $x \neq 0$

$$C = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p} = \sup_{\|x\|_p=1} \|Ax\|_p$$

Normas matriciais induzidas por normas vetoriais são normas.

Norma de Frobenius não é induzida por norma vetorial, porém é consistente com a norma vetorial Euclideana

$$\|A\|_p = \sup_{x \in \mathbb{R}^n, \|x\|_p=1} \|Ax\|_p \text{ para } p = 1, 2, \dots, \infty$$

Alguns casos particulares de interesse:

- $\|A\|_1 = \max_{j=1, \dots, n} \|a_j\|_1$ (máxima norma 1 coluna)
- $\|A\|_\infty = \max_{j=1, \dots, m} \|\alpha_j^T\|_1$ (máxima norma 1 linha)
- $\|A\|_2 = \begin{cases} |\lambda_n(A)| & A = A^T \\ \sigma_n(A^T A) & A \neq A^T \end{cases}$
(norma espectral, não é a norma de Frobenius).

$$\sigma_n(A^T A) = \sqrt{|\lambda_n(A^T A)|}$$

$|\lambda_n(A)|$: módulo do maior autovalor em módulo de A .

Vamos tomar $A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$ e considerar a imagem Ax da bola unitária, medida por cada uma das normas vetoriais abaixo.

Tomando $x^1 = (1 \ 0)^T$, $x^2 = (0 \ 1)^T$ temos

$$Ax^1 = (1 \ 0)^T, \quad Ax^2 = (2 \ 2)^T$$

- $\|Ax\|_1$: fator máximo de majoração $\|A\|_1 = 4$, $x^* = (0 \ 1)^T$
- $\|Ax\|_2$: fator máximo de majoração $\|A\|_2 = 2,9208$.
 $x^* \approx (0.26 \ 0.97)^T$ via SVD.
- $\|Ax\|_\infty$: fator máximo de majoração $\|A\|_\infty = 3$, $x^* = (1 \ 1)^T$

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

- $\mathcal{R}(A) = \{Ax : x \in \mathbb{R}^n\}$ (range de A ou imagem de A) é um subespaço vetorial.
- $\dim(\mathcal{R}(A)) = \text{posto}(A)$ é o número de colunas linearmente independentes de A .
- $\text{posto}(A) = \text{posto}(A^T)$, ou seja, o número de linhas li iguala o número de colunas li.
- $\mathcal{N}(A) = \{x \in \mathbb{R}^n : Ax = 0\}$ é o espaço nulo de A . (é um subespaço vetorial)

- Tomando um vetor $x \in \mathcal{R}(A^T)$: $x = A^T y$ para algum $y \in \mathbb{R}^m$.
- Tomando um vetor $z \in \mathcal{N}(A)$: $Az = 0$. Então $x^T z = y^T A^T z = 0$.

$$\mathcal{R}(A^T) \perp \mathcal{N}(A) \text{ ou } \mathcal{R}(A^T) = \mathcal{N}(A)^\perp$$

$$\mathcal{R}(A^T) \perp \mathcal{N}(A) \implies \mathcal{R}(A^T) = \mathcal{N}(A)^\perp$$

$$\mathcal{R}(A^T) = \{x \in \mathbb{R}^n : x = A^T y \text{ para algum } y \in \mathbb{R}^m\}$$

$\mathcal{N}(A)^\perp = \{x \in \mathbb{R}^n : z^T x = 0 \text{ para todo } z : Az = 0\}$ Tomando todos os $y \in \mathbb{R}^m$ e sua imagem $x = A^T y$ temos $z^T A^T y = 0$ para todo $z \in \mathcal{N}(A)$ e então $\mathcal{R}(A^T)$ e $\mathcal{N}(A)^\perp$ são o mesmo espaço.

$$z \in \mathcal{N}(A) \Leftrightarrow Az = 0 \Leftrightarrow \langle y, Az \rangle = 0 \forall y \in \mathbb{R}^m \Leftrightarrow \langle z, A^T y \rangle = 0 \forall y \in \mathbb{R}^m \Leftrightarrow z \in \mathcal{R}(A^T)^\perp$$

Decomposição ortogonal

$$\mathbb{R}^n = \mathcal{R}(A^T) \oplus \mathcal{N}(A)$$



Por argumento similar:

- Tomando um vetor $x \in \mathcal{R}(A) : x = Ay$ para algum $y \in \mathbb{R}^n$.
- Tomando um vetor $z \in \mathcal{N}(A^T) : A^T z = 0$. Então
 $z^T x = z^T A y = 0$.

$$\mathcal{R}(A) \perp \mathcal{N}(A^T) \text{ ou } \mathcal{R}(A) = \mathcal{N}(A^T)^\perp$$

$$\mathbb{R}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^T)$$

Para qualquer $A \in \mathbb{R}^{m \times n}$, vale que $\mathcal{R}(A^T) \perp \mathcal{N}(A)$ e $\mathcal{R}(A) \perp \mathcal{N}(A^T)$ e portanto:

$$\mathbb{R}^n = \mathcal{R}(A^T) \oplus \mathcal{N}(A)$$

$$\mathbb{R}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^T)$$

$$n = \text{posto}(A) + \dim(\mathcal{N}(A))$$

$$m = \text{posto}(A) + \dim(\mathcal{N}(A^T))$$

Assim sendo, qualquer vetor $x \in \mathbb{R}^n$ pode ser decomposto como uma soma direta de dois vetores ortogonais, um em $\mathcal{R}(A^T)$ e outro em $\mathcal{N}(A)$:

$$x = A^T \zeta + z \quad \text{para } z \in \mathcal{N}(A)$$

Similarmente, $w \in \mathbb{R}^m$ pode ser escrito como:

$$w = A\phi + \zeta \quad \text{para } \zeta \in \mathcal{N}(A^T)$$

Uma matriz quadrada é chamada projetor (ou matriz de projeção, ou matriz idempotente) se $P = P^2 = PP$.

- Para um projetor dito ortogonal, deve valer ainda $P = P^T$ (atenção: P projetor ortogonal não é uma matriz ortogonal !)
- Um projetor (ou projetor oblíquo, isto é, não ortogonal) não precisa satisfazer $P = P^T$.

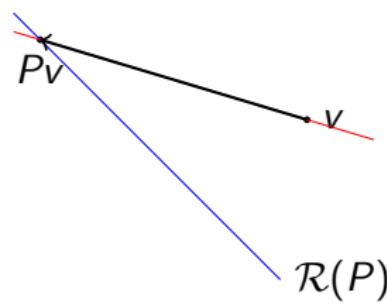
Dado uma matriz P , projetor:

- Para um dado $v \in \mathcal{R}(P)$ temos

$$Px = v \rightarrow PPx = Pv \rightarrow Px = Pv \rightarrow Pv = v$$

Ou seja, a aplicação de P em v não altera v .

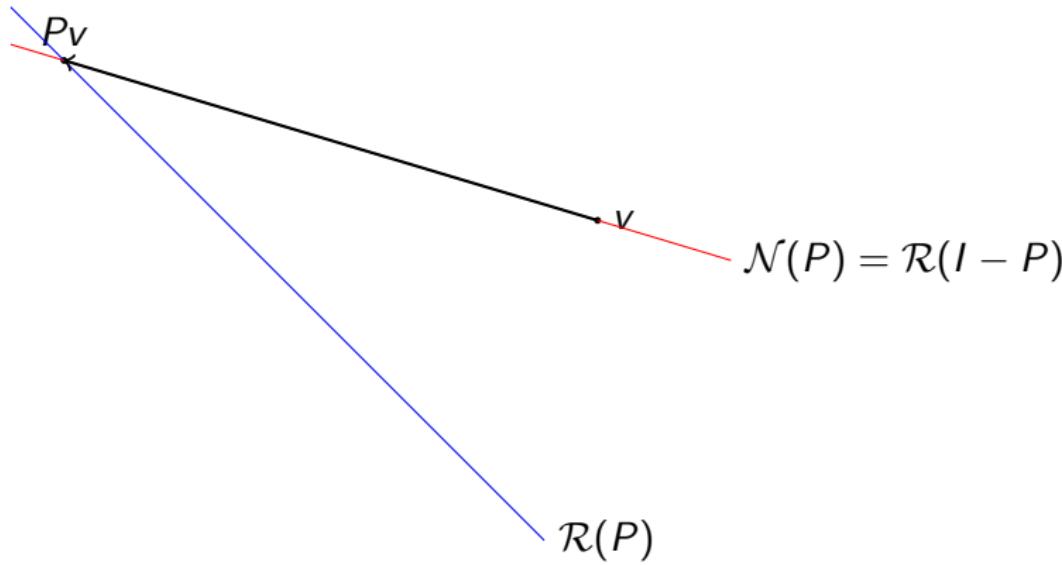
- A transformação linear Pv pode ser vista como a operação de iluminar $\mathcal{R}(P)$, a partir da direção $Pv - v$.
- Isto é, Pv consiste na sombra de v , ao se iluminar $\mathcal{R}(P)$, a partir da direção $Pv - v$.



Se P é um projetor, $I - P$ é um projetor, dito projetor complementar de P .

- De fato, $(I - P)(I - P) = I - 2P + PP = I - P$, e $I - P$ é projetor.
- Por um lado P projeta v em $\mathcal{R}(P)$. Em qual subespaço linear $I - P$ mapeia? Em $\mathcal{N}(P)$.

- $Pv \in \mathcal{R}(P)$ (óbvio)
- $P(Pv - v) = P^2v - Pv = 0 \Rightarrow (Pv - v) \in \mathcal{N}(P)$



$$\mathcal{N}(P) = \mathcal{R}(I - P)$$

- $\mathcal{R}(I - P) \subseteq \mathcal{N}(P)$:

Para qualquer v : $(I - P)v = v - Pv \in \mathcal{N}(P)$ (já que $Pv - PPv = 0$)

- $\mathcal{N}(P) \subseteq \mathcal{R}(I - P)$:

$v \in \mathcal{N}(P) \iff Pv = 0$. Então $(I - P)v = v$, logo $v \in \mathcal{R}(I - P)$.

Tomando o complementar P de $I - P$: $P = I - (I - P)$, temos o resultado complementar:

$$\mathcal{R}(P) = \mathcal{N}(I - P).$$

Por um lado, se P é projetor, P separa \mathbb{R}^n em S_1, S_2 tais que:

- $S_1 = \mathcal{R}(P), S_2 = \mathcal{N}(P) = \mathcal{R}(I - P)$:
- $S_1 + S_2 = \mathbb{R}^n$ (dado $v \in \mathbb{R}^n$, existem $v_1 \in S_1, v_2 \in S_2$ (únicos) tais que $v = v_1 + v_2$)
- $S_1 \cap S_2 = \{0\}$.

Veja:

- $\mathcal{N}(P) \cap \mathcal{N}(I - P) = \{0\}$ (qualquer vetor v em ambos satisfaz: $v = v - Pv = (I - P)v = 0$)
- $\mathcal{R}(P) \cap \mathcal{N}(P) = \{0\}$

Se S_1, S_2 são subespaços vetoriais tais que:

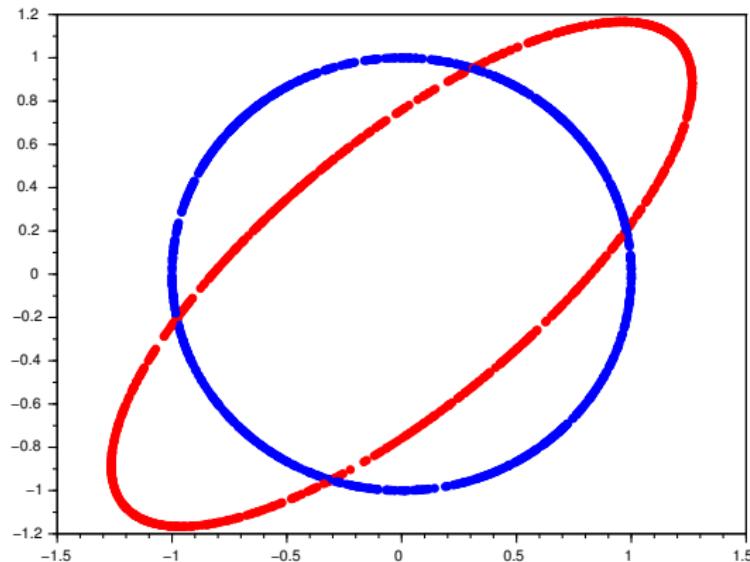
- $S_1 + S_2 = \mathbb{R}^n$
- $S_1 \cap S_2 = \{0\}$.

Então, existe um projetor P tal que $S_1 = \mathcal{R}(P), S_2 = \mathcal{N}(P)$

- Veja que, dado um $v \in \mathbb{R}^n$, fazendo $v_1 = Pv$ e $v_2 = (I - P)v$ temos $Pv + (I - P)v = v$.
- Estes vetores v_1, v_2 são únicos. Caso contrário, toda solução desta decomposição de ser do tipo $Pv + v_3 \in \mathcal{R}(P)$ e $((I - P)v - v_3) \in \mathcal{N}(P)$ para algum $v_3 \in \mathcal{R}(P) \cap \mathcal{N}(P)$. Claramente, $v_3 = 0$ e os vetores são únicos.

Transformação linear $y = Ax$

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



elipse1.sce

$A = \begin{pmatrix} 1.2 & 0.4 \\ 0.6 & 1 \end{pmatrix}$ possui dois vetores invariantes à transformação linear:

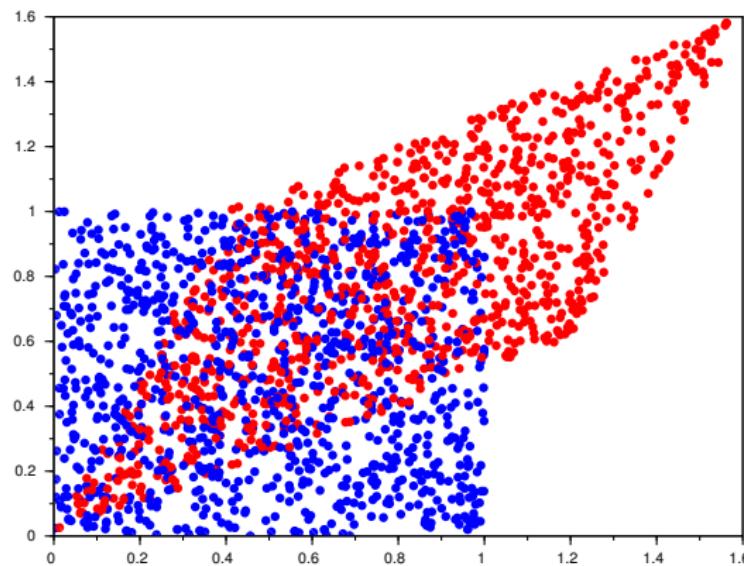
- $x^1 = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, Ax^1 = 1.6x^1$

- $x^2 = \frac{2}{\sqrt{13}} \begin{pmatrix} 1 \\ -\frac{3}{2} \end{pmatrix}, Ax^2 = 0.6x^2$

$$Ax = \lambda x$$

(λ, x) autopares de A , $\lambda \in \mathbb{C}^n$ é autovalor de A e $x \in \mathbb{C}^n$ seu correspondente autovetor.

Transformação do quadrado unitário



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

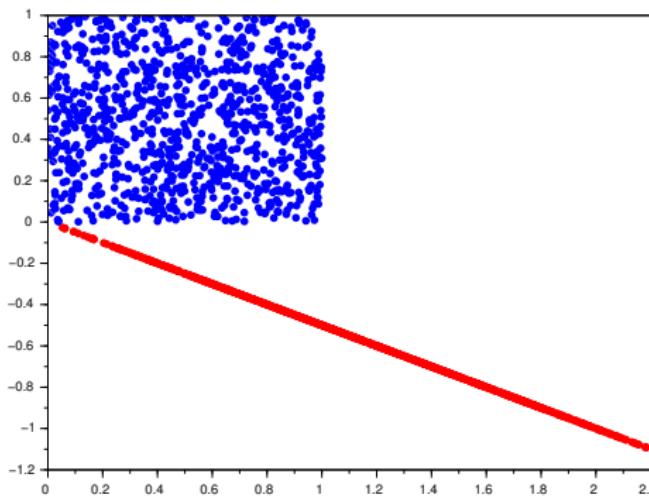
Definição - Expansão de Laplace

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

Pode ser provado que o módulo do determinante fornece o volume do sólido obtido pela transformação dos vértices do hipercubo unitário do \mathbb{R}^n .

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

$$\det(A) = 0 \iff \mathcal{N}(A) \neq \{0\}, \mathcal{R}(A) \neq \mathbb{R}^n$$



- $A \in \mathbb{R}^{n \times n}$, $\det(A) \neq 0$, existe A^{-1} dita inversa de A , tal que

$$AA^{-1} = A^{-1}A = I_n.$$

- $\det(A) = \frac{1}{\det(A^{-1})}$, caso A^{-1} exista.
- $(A^{-1})^T = (A^T)^{-1} = A^{-T}$

Inversas à esquerda e à direita para matrizes $A \in \mathbb{R}^{m \times n}$ retangulares:

- Se $m \geq n$, $A^{li} : A^{li}A = I_n$ é a inversa à esquerda de A .
- Se $m \leq n$, $A^{ri} : AA^{ri} = I_m$ é a inversa à direita de A .

Pseudo-inversa

De um modo geral, A^{pi} é uma pseudo-inversa de A se $AA^{pi}A = A$.

Obs: Uma pseudoinversa de $A \in \mathbb{R}^{m \times n}$ pode ser obtida via fatoração SVD de A (pseudo-inversa de Moore-Penrose)

$A \in \mathbb{C}^{n \times n}$ (resp. $A \in \mathbb{R}^{n \times n}$) é normal se $A^*A = AA^*$ (resp. $A^T A = AA^T$), isto é, a matriz e a sua transposta conjugada (resp. transposta) comutam.

Uma matriz $X = [x_1, \dots, x_n] \in \mathbb{R}^{n \times n}$ é ortogonal se suas colunas formam uma base ortonormal para \mathbb{R}^n . Então:

- $x_i^T x_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$
- $X^T X = I_n = X X^{-1}$
- $\|Xy\|_2^2 = (Xy)^T (Xy) = y^T X^T X y = y^T y = \|y\|_2^2$
(a transformação linear preserva o comprimento, $\|\cdot\|_2$)
- $\|UAV\|_F = \|A\|_F$ (norma de Frobenius também é preservada, U e V ortogonais)

Se $X = [x_1, \dots, x_n] \in \mathbb{C}^{n \times n}$ complexa satisfaz as propriedades acima é chamada unitária.

Matrizes unitárias ou ortogonais são casos particulares de matrizes normais.

Duas matrizes $A, B \in \mathbb{R}^{n \times n}$ são similares se existe uma matriz $P \in \mathbb{R}^{n \times n}$ tal que $B = P^{-1}AP$.

- $y = Ax$ mapeia \mathbb{R}^n em \mathbb{R}^n .
- $\mathcal{R}(P) = \mathbb{R}^n$ e então existem \tilde{x}, \tilde{y} tais que $x = P\tilde{x}, y = P\tilde{y}$.
- $y = Ax \rightarrow P\tilde{y} = AP\tilde{x} \rightarrow \tilde{y} = P^{-1}AP\tilde{x} = B\tilde{x}$

$B = P^{-1}AP$ representa o mapeamento linear $y = Ax$ em outra escolha de base para \mathbb{R}^n , definida pelas colunas de P .

- Matrizes similares possuem os mesmos autovalores.

$$\begin{aligned}\det(\lambda I - B) &= \det(\lambda I - P^{-1}AP) \\ &= \det(P^{-1}(\lambda I - A)P) \\ &= \det(P^{-1}) \det(\lambda I - A) \det(P) \\ &= \det(\lambda I - A)\end{aligned}$$

- $Ax = \lambda x$, autovetor x é invariante em ângulo à transformação linear.
- $Ax = \lambda x \rightarrow (A - \lambda I_n)x = 0$.

Para que haja solução não trivial $x \neq 0$, precisamos $\mathcal{N}(A - \lambda I_n) \neq \{0\}$. Ou seja, λ precisa ser raiz do polinômio característico de A :

$$\det(A - \lambda I_n) = 0$$

- Reinterpretando: λ torna $A - \lambda I_n$ singular e $x \in \mathcal{N}(A - \lambda I_n)$.

Fatorando o polinômio característico

$$P_n(\lambda) = c_n \prod_{i=1}^k (\lambda - \lambda_i)^{\mu_i}$$

k denota o número de autovalores distintos de A e $\mu_i : i = 1, \dots, k$ as multiplicidades algébricas dos autovalores.

A multiplicidade algébrica de λ_i é número de vezes μ_i que λ_i é raiz do polinômio característico de A .

Teorema Fundamental da Álgebra

Toda matriz $A \in \mathbb{R}^{n \times n}$ possui n autovalores, contando suas multiplicidades. Logo, $\sum_{i=1}^k \mu_i = n$.

- A cada autovalor $\lambda_i : i \in \{1, \dots, k\}$ corresponde um **subespaço linear**

$$\phi_i = \mathcal{N}(\lambda_i I_n - A),$$

denominado **auto-espaço** associado ao autovalor λ_i .

A multiplicidade geométrica de λ_i é a dimensão de $\mathcal{N}(\lambda_i I_n - A)$, que representa o número de autovetores linearmente independentes associados à λ_i .

$$A = \begin{pmatrix} 2 & & \\ & 2 & \\ & & 2 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{pmatrix}$$

- Polinômio característico de A, B : $P_3(\lambda) = (\lambda - 2)^3$
- A possui 3 autovetores li: e_1, e_2, e_3 . A multiplicidade geométrica também é 3.
- B possui apenas um autovetor e_1 , sua multiplicidade geométrica é 1.
- Matriz defectiva: possui autovalor cuja multiplicidade algébrica excede a multiplicidade geométrica. (autovalor defectivo)

Matrizes diagonalizáveis ou não defectivas

As matrizes não defectivas são diagonalizáveis, isto é, $A = U \Lambda U^{-1}$, colunas de U são autovetores de A e Λ é a matriz diagonal com os correspondentes autovalores. São portanto similares a uma matriz diagonal.

Teorema

Seja $\{\lambda_i : i = 1, \dots, k\}$ o conjunto de autovalores distintos de $A \in \mathbb{R}^{n \times n}$ e $\phi_i = \mathcal{N}(A - \lambda_i I_n)$ o autoespaço associado a λ_i . Seja x^i um vetor não nulo tal que $x^i \in \phi_i$.

Então:

- ① Quaisquer vetores x^i, x^j não nulos satisfazendo $x^j \in \phi_j, x^i \in \phi_i$ são linearmente independentes.
- ② Além disto, qualquer conjunto de não nulos $\{x^i \in \phi_i : i = 1, \dots, k\}$ é um conjunto de vetores linearmente independentes.

Obsevação inicial

Note que $x^i \notin \phi_j, i \neq j$ ($\phi_i \cap \phi_j = \{0\}$ para $i \neq j$), caso contrário $Ax^i = \lambda_j x^i = \lambda_i x^i \rightarrow \lambda_i = \lambda_j$ (contradição, $\lambda_i \neq \lambda_j$).

- Por absurdo, se forem l.d, $x^1 = \sum_{i=2}^k \alpha_i x^i$ e então:

$$\sum_{i=2}^k \alpha_i \lambda_i x^i = \sum_{i=2}^k A\alpha_i x^i = Ax^1 = \lambda_1 x^1 = \sum_{i=2}^k \alpha_i \lambda_1 x^i$$

$$\sum_{i=2}^k \alpha_i (\lambda_i - \lambda_1) x^i = 0$$

dado que $\lambda_1 \neq \lambda_i, x^i : i = 2, \dots, k$ seriam l.d.

- Repetindo o raciocínio, por indução concluiríamos que x^3, \dots, x^k são l.d até que x^{k-1}, x^k seriam l.d e portanto $x^{k-1} \in \phi_k$ (contradição, dada a observação inicial)

Matriz triangular em blocos

$$\begin{pmatrix} A_{11} & A_{12} & \dots & A_{1p} \\ 0 & A_{22} & \dots & A_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & A_{pp} \end{pmatrix}$$

Cada matriz $A_{ii} : i = 1, \dots, p$ é quadrada.

Corolário

Qualquer matriz $A \in \mathbb{R}^{n \times n}$ é similar a uma matriz triangular em blocos que possui o bloco $\lambda_i I_{v_i}$ na diagonal, onde λ_i é um autovalor distinto de A (possivelmente complexo) e v_i a dimensão de seu subespaço ϕ_i . Ou seja, existem P, P^{-1} e B triangular em blocos tais que $A = P^{-1}BP$

- $v_i = \dim(\phi_i)$, $X^i = [x_1^i, \dots, x_{v_i}^i]$ fornece uma base para ϕ_i .
 - Podemos assumir que $x_1^i, \dots, x_{v_i}^i$ são ortonormais (cc, aplicamos Gram-Schmidt): $(X^i)^T X^i = I_{v_i}$.
 - Seja $Q_i \in \mathbb{R}^{n \times n-v_i}$ uma matriz com colunas ortonormais que geram $\mathcal{N}(A - \lambda_i I_n)^\perp = \phi_i^\perp$.
 - A matriz $P^i = [X^i \ Q^i]$ é uma matriz ortogonal:
 - $P^i (P^i)^T = I_n$, $(P^i)^T P^i = I_n$, P^i admite inversa
 - suas colunas geram uma base ortonormal para \mathbb{C}^n
 - Como $AX^i = \lambda_i X^i$ temos ($\lambda_i \in \mathbb{C}$ é um escalar)
 - $(X^i)^T A X^i = \lambda_i (X^i)^T X^i = \lambda_i I_{v_i}$:
 - $(Q^i)^T A X^i = \lambda_i (Q^i)^T X^i = 0$.
 - e então $(P^i)^{-1} A P^i = (P^i)^T A P^i = [X^i \ Q^i]^T A [X^i \ Q^i] = \begin{pmatrix} \lambda_i I_{v_i} & (X^i)^T A Q^i \\ 0 & (Q^i)^T A Q^i \end{pmatrix}$ e o resultado segue
- (obs: pode-se aplicar o mesmo raciocínio à matriz quadrada $(Q^i)^T A Q^i$)

Teorema

Sejam $\lambda_i : i = 1, \dots, k$ os autovalores distintos de $A \in \mathbb{R}^{n \times n}$, $\mu_i : i = 1, \dots, k$ suas multiplicidades algébricas e $\phi_i : i = 1, \dots, k$ seus autoespaços. Seja $X^i = [x_1^i, \dots, x_{v_i}^i]$ uma matriz cujas colunas geram uma base para ϕ_i ($v_i = \dim(\phi_i)$). Então:

- ① $v_i \leq \mu_i : i = 1, \dots, k$
- ② Se $v_i = \mu_i : i = 1, \dots, k$, $X = [X^1, \dots, X^k]$ admite inversa, A pode ser fatorada na forma $A = X\Lambda X^{-1}$ onde

$$\Lambda = \begin{pmatrix} \lambda_1 I_{v_1} & 0 & \dots & 0 \\ 0 & \lambda_2 I_{v_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_k I_{v_k} \end{pmatrix}$$

Parte 1: $v_i \leq \mu_i$

Ao final da prova anterior, obtivemos

$$(P^i)^{-1}AP^i = \begin{pmatrix} \lambda_i I_{v_i} & (X^i)^T AQ^i \\ 0 & (Q^i)^T AQ^i \end{pmatrix}$$

Uma vez que matrizes similares possuem os mesmos autovalores, se $v_i > \mu_i$, teríamos mais de μ_i ocorrências de λ_i na diagonal da matriz triangular superior acima. Logo a multiplicidade algébrica não seria μ_i , mas um valor superior a μ_i (uma contradição).

Parte 2: diagonalização

- Consideramos então que $\mu_i = v_i : i = 1, \dots, k$ (A é não defectiva)
- Os vetores $x_1^i, \dots, x_{v_i}^i$ são li, para qualquer $i = 1, \dots, k$
- Qualquer par de vetores $x_j^i \in \phi_i, x_u^z \in \phi_z$ para $i \neq z$, $i, z \in \{1, \dots, k\}$ são li.
- $X = [X^1, \dots, X^k]$ possui posto completo n e admite inversa ($\sum_{i=1}^k v_i = \sum_{i=1}^k \mu_i = n$).
- Para todo $i = 1, \dots, k$, $Ax_j^i = \lambda_i x_j^i : j = 1, \dots, \mu_i$
- Sistematizando:

$$AX^i = \lambda_i X^i \quad i = 1, \dots, k$$

$$AX = X\Lambda$$

$$A = X\Lambda X^{-1}$$

$A \in \mathbb{R}^{m \times n}$ é um *dyad* se $A = uv^T$ para algum $u \in \mathbb{R}^m, v \in \mathbb{R}^n$.

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

- ① A transformação linear $Ax = (uv^T)x = (v^T x)u$ sempre tem a imagem em $\text{span}(u)$, independentemente de x .
- ② $\text{posto}(A) = 1$
- ③ Linhas (colunas) de um *dyad* são múltiplos umas das outras.
- ④ Um *dyad* quadrado $A \in \mathbb{R}^{n \times n}$ possui um único autovalor distinto de zero, $\lambda = v^T u$, com correspondente autovetor u :
 $Au = (uv^T)u = (v^T u)u = \lambda u$.

- Se A é quadrada, bloco diagonal, i.e., $A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}$:
 - o espectro de A é $\lambda(A) = \lambda(A_{11}) \cup \lambda(A_{22})$.
 - A admite inversa se e somente se A_{11}, A_{22} admitem inversa,

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22}^{-1} \end{pmatrix}$$
- Para A quadrada, bloco triangular inferior ou superior:
 - $\lambda(A) = \lambda(A_{11}) \cup \lambda(A_{22})$.
 - $A = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}, A^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{21}A_{11}^{-1} & A_{22}^{-1} \end{pmatrix}$
 - $A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, A^{-1} = \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{pmatrix}$

Para $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, definimos os complementos de Schur:

- $S_1 := A_{11} - A_{12}A_{22}^{-1}A_{21}$
- $S_2 := A_{22} - A_{21}A_{11}^{-1}A_{12}$

$$\begin{aligned} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} &= \begin{pmatrix} S_1^{-1} & -A_{11}^{-1}A_{12}S_2^{-1} \\ -A_{22}^{-1}A_{21}S_1^{-1} & S_2^{-1} \end{pmatrix} \\ &= \begin{pmatrix} S_1^{-1} & -S_1^{-1}A_{12}A_{22}^{-1} \\ -S_2^{-1}A_{21}A_{11}^{-1} & S_2^{-1} \end{pmatrix} \end{aligned}$$

O cálculo da inversa de uma matriz pode ser simplificado se os blocos tiverem uma estrutura conveniente.

- Fatorar uma matriz $A \in \mathbb{R}^{m \times n}$ consiste em escrevê-la como o produto de outras matrizes, desejavelmente com alguma estrutura conveniente.
- Diversos atributos de A podem avaliados mais eficientemente após sua fatoração (autovalores, bases para subespaços associados a A).

Exemplos:

- **Fatoração LU:** $A \in \mathbb{R}^{n \times n}$, $PA = LU$, onde P é uma matriz de permutação, L triangular inferior com diagonal unitária, U triangular superior: $|\det(A)| = |\prod_{i=1}^n u_{ii}|$.
- **Cholesky:** $A \in \mathcal{S}_{++}^n$ (simétrica positiva definida), $A = LL^T$ onde L é uma triangular inferior com diagonal positiva.
- QR, SVD,...

- $A \in \mathbb{R}^{m \times n}$ ($m \geq n$) pode ser fatorada como $A = QR$ onde $Q \in \mathbb{R}^{m \times n}$ é uma matriz ortogonal ($Q^T Q = I_n$) e $R \in \mathbb{R}^{n \times n}$ é uma triangular superior.
- A fatoração é única, assumindo-se que as entradas na diagonal de R são positivas e que posto de A é n , completo.
- O método de Gram-Schmidt fornece uma fatoração $A = QR$, porém métodos numericamente mais estáveis são o Método de Transformação de Householder, Método de Givens.
- Importante para resolver sistemas lineares malcondicionados:
 $Ax = b$, $QRx = b$, $Rx = Q^T b$.

O posto de A é o número de entradas não nulas na diagonal de R

$$(A_1 \mid A_2 \mid \dots \mid A_n) = (Q_1 \mid Q_2 \mid \dots \mid Q_n) \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & r_{nn} \end{pmatrix}$$

Reinterpretando a fatoração

$$\text{span}(A_1, \dots, A_k) = \text{span}(Q_1, \dots, Q_k), \text{ para } k = 1, \dots, n$$

- $Q_1 = \frac{A_1}{\|A_1\|_2}$
- $k \geq 2$: Q_k é obtido projetando-se A_k em $\text{span}(Q_1, \dots, Q_{k-1})$. Os pesos $r_{1k} = \langle A_k, Q_1 \rangle, \dots, r_{k-1,k} = \langle A_k, Q_{k-1} \rangle$ fornecem a combinação linear que determina a projeção $\sum_{i=1}^{k-1} \langle A_k, Q_i \rangle Q_i$ e r_{kk} normaliza a diferença $\left(A_k - \sum_{i=1}^{k-1} \langle A_k, Q_i \rangle Q_i \right) \in \text{span}(Q_1, \dots, Q_{k-1})^\perp$

Motivação

A imagem da esfera unitária (na norma 2) do \mathbb{C}^n diante de uma transformação linear de qualquer matriz de ordem $m \times n$ é uma hiperelipse em \mathbb{C}^m .

É aplicável tanto para matrizes em $\mathbb{R}^{m \times n}$ quanto em $\mathbb{C}^{m \times n}$.

SVD (reduzido), $A \in \mathbb{R}^{m \times n}$, $m \geq n$, $\text{posto}(A) = n$.

Tomando n vetores li em \mathbb{R}^n formando uma base $\{v_1, \dots, v_n\}$ ortonormal.

Imagem da transformação linear Av_j . Para $j = 1, \dots, n$: $Av_j = \sigma_j u_j$ onde $\sigma_j > 0$ é o comprimento (valor singular) na direção do eixo principal u_j da resultante em \mathbb{R}^m . Então

$A[v_1 \ v_2 \ \dots \ v_n] = AV = [u_1 \ u_2 \ \dots \ u_n]\hat{\Sigma} = \hat{U}\hat{\Sigma}$ e

$$A = \hat{U}\hat{\Sigma}V^T$$

- Se $\text{posto}(A) = r < n$, a forma da fatoração ainda é válida. Basta completarmos \hat{U} com $m - r$ colunas li adicionais e $\hat{\Sigma}$ com $m - r$ linhas de zero.
- Neste caso, $n - r$ colunas de V pertencem a $\mathcal{N}(A)$.

SVD (completo)

As n colunas de $\hat{U} \in \mathbb{R}^{m \times n}$ não formam uma base para \mathbb{R}^m , mas podemos expandir \hat{U} com $m - n$ colunas adicionais, de forma que o resultado U seja ortonormal. Por consistência, adicionamos $m - n$ linhas de zeros à $\hat{\Sigma}$, obtendo Σ . Então $AV = [u_1 \ u_2 \ \dots \ u_m]\Sigma = U\Sigma$ ou $A = U\Sigma V^T$.

A e Σ são ortogonalmente equivalentes pois existem V , U ortogonais tais que $A = U\Sigma V^T$.

$$A = \begin{pmatrix} 1.2 & 0.4 \\ 0.6 & 1 \end{pmatrix}$$

[U,S,V] = svd(A);

V =

-0.8012766 -0.5982941

-0.5982941 0.8012766

S =

1.6144381 0.

0. 0.5946342

U =

-0.7438189 -0.6683812

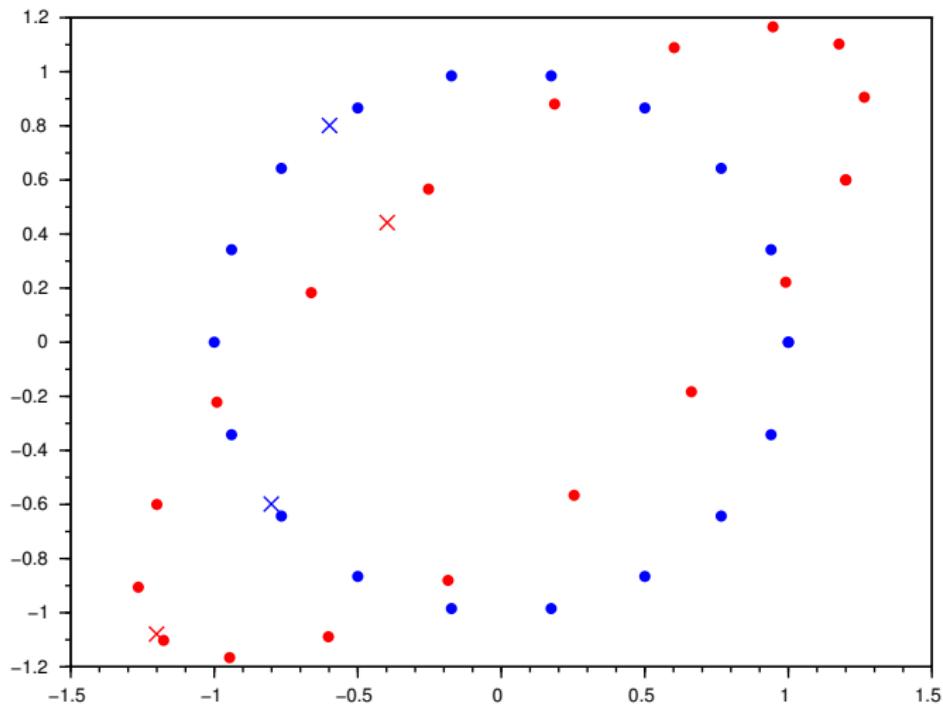
-0.6683812 0.7438189

A*V-U*S

ans =

0. 0.

-2.220D-16 0.



- $v_i : i = 1, \dots, r = \text{posto}(A)$ são os vetores singulares à direita de A .
- $u_i : i = 1, \dots, r$ são os vetores singulares à esquerda de A .
- $\sigma_i > 0 : i = 1, \dots, r$. Uma vez que $\|u_i\|_2 = 1$, σ_i fornece o comprimento do vetor singular.

A transformação linear Av_i fornece como imagem o vetor u_i com norma unitária, corrigido em comprimento por σ_i .

Qualquer matriz $A \in \mathbb{R}^{m \times n}$ ($A \in \mathbb{C}^{m \times n}$) pode ser fatorada na forma $A = U\Sigma V^T$ onde $V \in \mathbb{R}^{n \times n}$, $U \in \mathbb{R}^{m \times m}$ são ortogonais (unitárias) e $\Sigma \in \mathbb{R}^{m \times n}$ é uma matriz diagonal que possui as primeiras $r = \text{posto}(A)$ entradas na diagonal iguais aos valores $\sigma_1, \dots, \sigma_r$ positivos e não crescentes em magnitude, todas as demais sendo nulas.

$$A = U\Sigma V^T \rightarrow AV = U\Sigma \rightarrow Av_i = \begin{cases} \sigma_i u_i & i = 1, \dots, r \\ 0 & i = r+1, \dots, n \end{cases}$$

$$A^T = V\Sigma^T U^T \rightarrow A^T U = V\Sigma^T \rightarrow A^T u_i = \begin{cases} \sigma_i v_i & i = 1, \dots, r \\ 0 & i = r+1, \dots, m \end{cases}$$

- $\mathcal{R}(A) = \text{span}\{u_1, \dots, u_r\}$
- $\mathcal{N}(A) = \text{span}\{v_{r+1}, \dots, v_n\}$
- $\mathcal{R}(A^T) = \text{span}\{v_1, \dots, v_r\}$
- $\mathcal{N}(A^T) = \text{span}\{u_{r+1}, \dots, u_m\}$

- ① A fatoração revela muita informação sobre o mapeamento linear: sua imagem, o espaço nulo, posto de A , norma espectral, número de condição.
- ② Importante em compressão de dados, PCA (principal component analysis) para a resolução de sistemas lineares (malcondicionados) definidos por A ,

$b = Ax \rightarrow \hat{b} = \Sigma \hat{x}$: uma transformação linear induzida por uma matriz diagonal, na base certa.

SVD completo: $A = U\Sigma V^T$

- Escrevendo b como uma combinação linear das colunas de U : $b = U\hat{b}$, temos $U\hat{b} = Ax$.
- Escrevendo x como uma combinação linear das colunas de V : $x = V\hat{x}$, $U\hat{b} = AV\hat{x}$.
- Logo $\hat{b} = U^T AV\hat{x} = \Sigma \hat{x}$

Interpretação da transformação linear: $Ax = U\Sigma V^T x$

- $y = V^T x$: rotação/reflexão (mudança de ângulo) de x
- $z = \Sigma y$: mudança de escala, sendo que dimensões podem ser removidas ou adicionadas.
- $w = Uz$: outra transformação ortogonal no espaço de saída.

$$AV = \Sigma U(\text{completa})$$

- Número de valores singulares $\sigma_i > 0 : i = 1, \dots, r$ é o posto de A .
- $\{v_{r+1}, \dots, v_n\}$ base ortonormal para $\mathcal{N}(A)$
- $\{u_1, \dots, u_r\}$ base ortonormal para $\mathcal{R}(A)$
- $\|A\|_F = \sqrt{\sum_{i=1}^r \sigma_i^2}$
- $\|A\|_2 = \sigma_1$
- $\sigma_i = \sqrt{\lambda_i(A^T A)} = \sqrt{\lambda_i(A A^T)} : i = 1, \dots, r.$
(Veja que

$$A^T A = (U \Sigma V^T)^T (U \Sigma V^T) = V \Sigma^T U^T U \Sigma V^T = V (\Sigma^T \Sigma) V^T$$

(SVD) $A = U\Sigma V^T \iff A = \sum_{i=1}^r \sigma_i u_i v_i^T$ onde r é o posto de A .

Problema de aproximar uma matriz

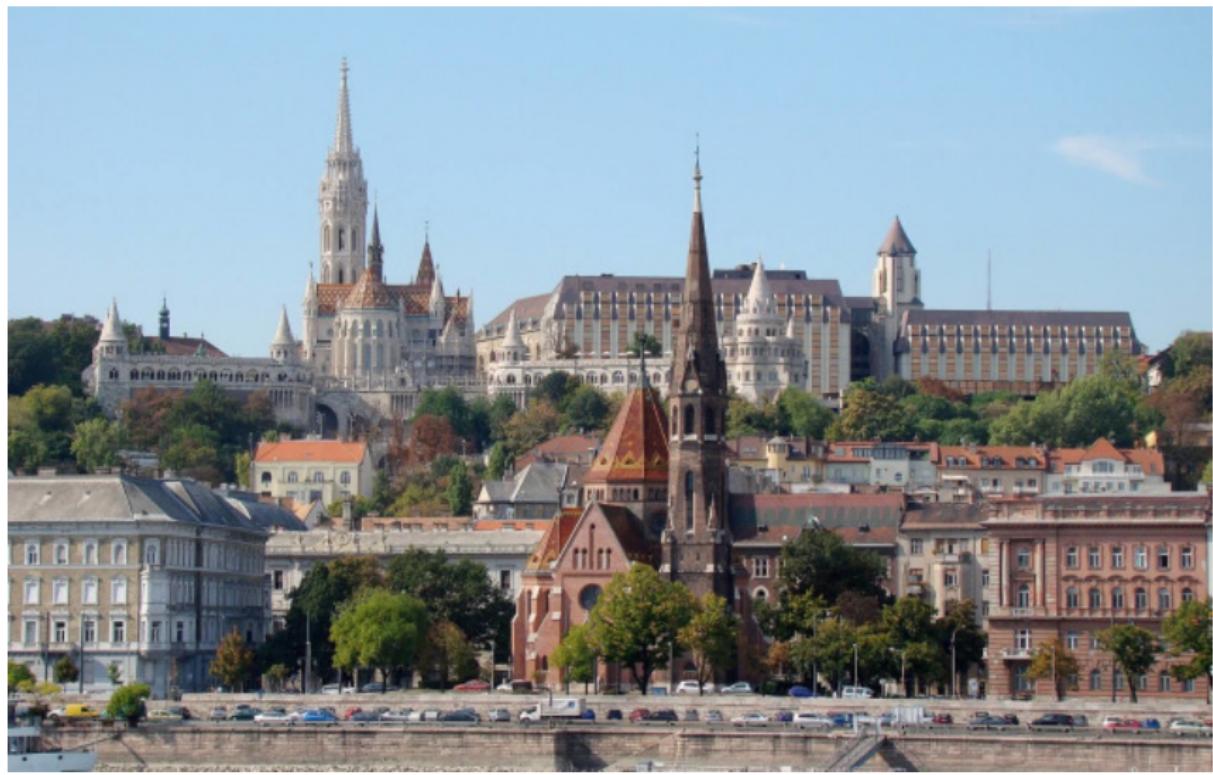
$$\min_{A_k \in \mathbb{R}^{m \times n} \mid \text{posto}(A_k) = k} \|A - A_k\|_F^2$$

onde $1 \leq k \leq r$.

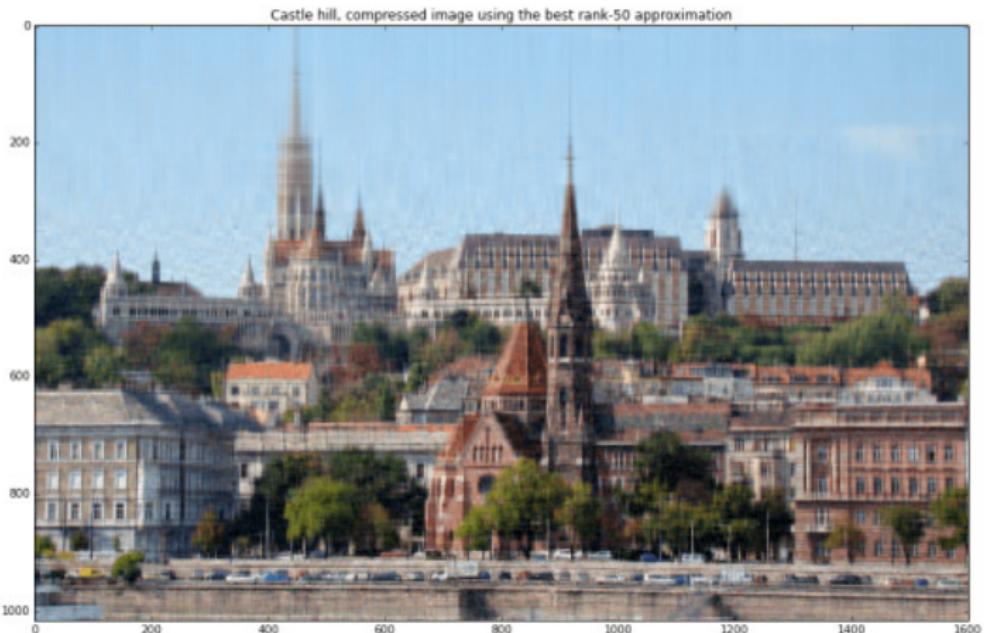
- A norma de Frobenius é unitariamente invariante:
 $\|Y\|_F = \|QYR\|_F$ para $Q \in \mathbb{R}^{m \times m}$, $R \in \mathbb{R}^{n \times n}$ ortogonais.
- Então $\|A - A_k\|_F^2 = \|U^T(A - A_k)V\|_F^2 = \|\Sigma - Z\|_F^2$.
- A solução ótima tem que ser com uma Z diagonal e com a diagonal nula da $k + 1$ posição em diante.
- Então a função objetivo ótima é: $\sum_{i=k+1}^r \sigma_i^2$.

Low rank approximation, full image

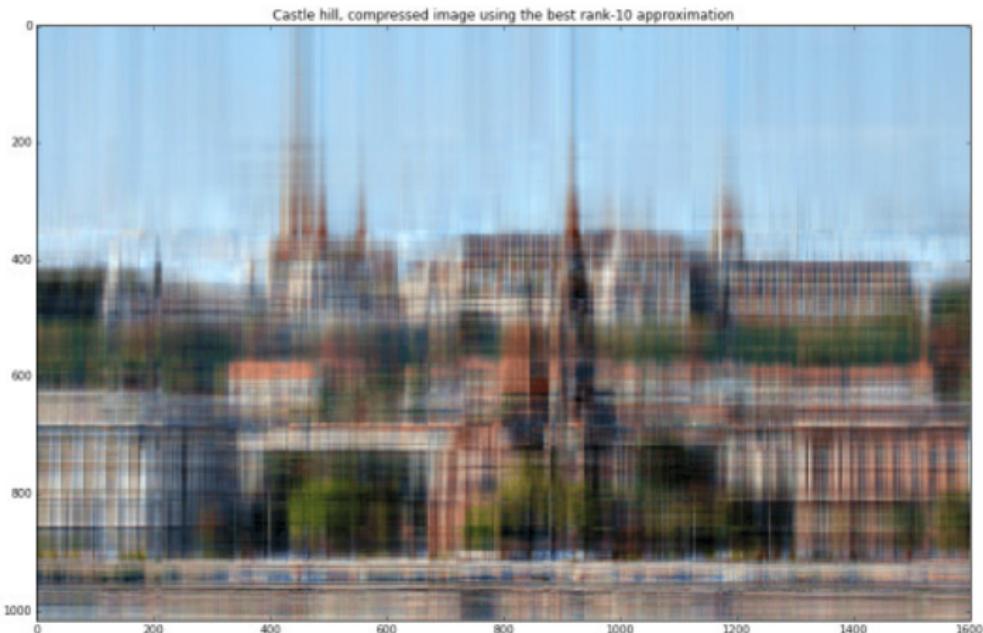
<https://www.balabit.com/blog/image-compression-using-singular-value-decomposition/>



Low rank approximation, $k = 50$



Low rank approximation, $k = 10$



A não defectível, diagonalizável

$$A \in \mathbb{R}^{n \times n}$$

$\{u_1, \dots, u_n\}$ autovetores de A são l.i., $u_j \in \mathbb{C}^n, j = 1, \dots, n$

Imagem da transformação linear Au_j . Para $j = 1, \dots, n : Au_j = \lambda_j u_j$.
Então

$$A[u_1 \ u_2 \ \dots \ u_n] = AU = [u_1 \ u_2 \ \dots \ u_n]\Lambda = U\Lambda.$$

$$A = U\Lambda U^{-1}$$

Λ é uma matriz diagonal com os autovalores de A . (A é similar a uma matriz diagonal Λ , espectro de A é o espectro de Λ)

$$A = \begin{pmatrix} 1.2 & 0.4 \\ 0.6 & 1 \end{pmatrix}$$

--> [R,D] = spec(A)

D =

$$\begin{matrix} 1.6 & 0. \\ 0. & 0.6 \end{matrix}$$

R =

$$\begin{matrix} 0.7071068 & -0.5547002 \\ 0.7071068 & 0.8320503 \end{matrix}$$

--> A - R*D*inv(R)

ans =

$$\begin{matrix} 0. & 0. \\ 0. & 0. \end{matrix}$$

Teorema

Seja $A \in \mathcal{S}^n$ uma matriz real simétrica, $\lambda_i : i = 1, \dots, k \leq n$ seus autovalores distintos. Sejam $\mu_i, v_i : i = 1, \dots, k$, respectivamente as multiplicidades algébrica e geométrica dos autovalores e ϕ_i o autoespaço associado à λ_i .

Então, para todo $i = 1, \dots, k$:

- $\lambda_i \in \mathbb{R}$
- $\phi_i \perp \phi_j, i \neq j$
- $v_i = \mu_i$.

Corolário - diagonalização unitária (matrizes normais $AA^T = A^TA$)

$A \in \mathcal{S}_n$ admite uma decomposição espectral

$$A = X \Lambda X^T = \sum_{i=1}^n \lambda_i x_i x_i^T$$

onde $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ é a matriz diagonal real com os n autovalores de A (contando as multiplicidades das raízes do polinômio característico) e $\{x_1, \dots, x_n\}$ são autovetores de A , formando uma base ortonormal X para \mathbb{R}^n .

Consequência direta do terceiro ponto do Teorema acima.

Toda matriz real simétrica é ortogonalmente similar a uma matriz diagonal real.

$A \in \mathbb{R}^{n \times n}$

- Toda matriz quadrada admite uma fatoração de Schur, isto é, é similar a uma triangular superior: $A = XUX^{-1}$.
- As matrizes não defectíveis são similares a uma matriz diagonal: $A = X\Lambda X^{-1}$.
- As matrizes normais (simétricas são um caso particular) admitem uma diagonalização unitária: $A = X\Lambda X^T$.

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

- Toda matriz $A \in \mathbb{R}^{m \times n}$ ($m \geq n$) admite uma decomposição em valores singulares, $A = U\Sigma V^T$ onde $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ são ortogonais e $\Sigma \in \mathbb{R}^{m \times n}$ é diagonal.

Lemma

Seja (λ, u) um autopar para $Q \in \mathbb{R}^{n \times n}$ ($Qu = \lambda u$) e $P^k(Q)$ e $P^k(\lambda)$ os polinômios de Q e de λ respectivamente definidos como:

$$P^k(Q) = \sum_{i=0}^k a_i Q^i = a_k Q^k + a_{k-1} Q^{k-1} + \cdots + a_1 Q + a_0 I_n$$

$$P^k(\lambda) = \sum_{i=0}^k a_i \lambda^i = a_k \lambda^k + a_{k-1} \lambda^{k-1} + \cdots + a_1 \lambda + a_0.$$

Então, $P^k(\lambda)$ e u formam um autopar para a matriz $P^k(Q)$, isto é,

$$P^k(Q)u = P^k(\lambda)u.$$

Prova

$$Qu = \lambda u$$

$$Q^2 u = Q(Qu) = Q(\lambda u) = \lambda Qu = \lambda^2 u$$

$$Q^3 u = Q(Q^2 u) = Q(\lambda^2 u) = \lambda^2 Qu = \lambda^3 u$$

$$\vdots$$

$$\vdots$$

$$\vdots$$

$$Q^k u = Q(Q^{k-1} u) = Q(\lambda^{k-1} u) = \lambda^{k-1} Qu = \lambda^k u$$

$$P^k(Q)u = \sum_{i=0}^k a_i Q^i u = \sum_{i=0}^k a_i \lambda^i u = \left(\sum_{i=0}^k a_i \lambda^i\right) u = P^k(\lambda)u$$

Consequências

- ① Eigenvalue shift rule: se os autovalores de $A \in \mathbb{R}^{n \times n}$ são $\lambda_i : i = 1, \dots, n$, então os autovalores da matriz $A + \mu I_n$ são $\lambda_i(A + \mu I_n) = \lambda_i + \mu$, $i = 1, \dots, n$.
- ② Decomposição espectral de $P^k(Q)$ para Q que admita diagonalização na forma $Q = U\Lambda U^{-1}$

$$P^k(Q) = \sum_{i=0}^k a_i Q^i = U P(\Lambda) U^{-1}$$

$$P(\Lambda) = \text{diag}(P^k(\lambda_1), \dots, P^k(\lambda_n))$$

Teorema

- Seja Q uma matriz $n \times n$ que admite uma fatoração diagonal da forma $Q = U\Lambda U^{-1}$ onde Λ é uma matriz diagonal com os autovalores de Q e U uma matriz com os correspondentes autovetores.
- Seja $P^k(t) : t \in \mathbb{R}$ o polinômio $P^k(t) = \sum_{i=0}^k a_i t^i$ de grau k .

Então:

$$P^k(Q) = \sum_{i=0}^k a_i Q^i = UP(\Lambda)U^{-1}$$

onde $P(\Lambda) \in \mathbb{R}^{n \times n}$ é a matriz diagonal $\text{diag}(P^k(\lambda_1), \dots, P^k(\lambda_n))$.

Prova

$$Q = U\Lambda U^{-1}$$

$$Q^2 = U\Lambda U^{-1} U\Lambda U^{-1} = U\Lambda^2 U^{-1}$$

$$Q^k = U\Lambda^{k-1} U^{-1} U\Lambda U^{-1} = U\Lambda^k U^{-1}$$

$$\begin{aligned} P^k(Q) &= \sum_{i=0}^k a_i Q^i = \sum_{i=0}^k a_i U\Lambda^i U^{-1} \\ &= \sum_{i=0}^k a_i U\Lambda^i U^{-1} = U(a_k\Lambda^k + a_{k-1}\Lambda^{k-1} + \cdots + a_1\Lambda + a_0 I_n)U^{-1} \\ &= UP(\Lambda)U^{-1} \end{aligned}$$

$$\lambda_{\max}(A) = \lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A) = \lambda_{\min}(A)$$

Teorema - Coeficientes de Rayleigh

$$\lambda_{\min}(A) \leq \frac{x^T A x}{x^T x} \leq \lambda_{\max}(A), \quad x \neq 0_n$$

$$\lambda_{\min}(A) = \min\{x^T A x : \|x\|_2 = 1\}$$

$$\lambda_{\max}(A) = \max\{x^T A x : \|x\|_2 = 1\}$$

Consequência da decomposição espectral de matrizes simétricas:
 $A = U \Lambda U^T = \sum_{i=1}^n \lambda_i u_i u_i^T$ para U ortogonal e da invariância da norma Euclideana a transformações ortogonais: $\|Ux\|_2 = \|x\|_2$. Os valores máximos e mínimos são atingidos para u_1, u_n associados a

$$\lambda_1, \lambda_n.$$

$A \in \mathbb{R}^{n \times n}$

- semipositiva definida: $x^T A x \geq 0$ para qualquer $x \in \mathbb{R}^n, x \neq 0$
($A \succeq 0$)
- positiva definida: $x^T A x > 0$ para qualquer x não nulo. ($A \succ 0$)

Se forem também simétricas, indicamos $A \in \mathcal{S}_+^n$, $A \in \mathcal{S}_{++}^n$.

Fatos

- Se $A \succ 0$ ($A \succeq 0$) então qualquer submatriz principal $A_{\mathcal{I}}$ de A (onde \mathcal{I} é um subconjunto dos índices $\{1, \dots, n\}$) é tal que $A_{\mathcal{I}} \succ 0$ ($A_{\mathcal{I}} \succeq 0$)
- $A \succ 0$ ($A \succeq 0$) $\iff \lambda_i(A) > 0 : i = 1, \dots, n$
($\lambda_i(A) \geq 0 : i = 1, \dots, n$)
- $\lambda_k(A + B) \geq \lambda_k(A)$ para $A \in \mathcal{S}^n, B \in \mathcal{S}_+^n : k = 1, \dots, n$
(autovalores não podem decrescer com a soma de uma matriz semipostiva).

Teorema

Dada $A \in \mathcal{S}^n$, $B \in \mathbb{R}^{n \times m}$ e considere o produto $C = B^T AB \in \mathcal{S}^m$.

- a) Se $A \succeq 0$, então $C \succeq 0$.
- b) Se $A \succ 0$, então $C \succ 0$ se e somente se $\text{posto}(B) = m$
- c) Se B é quadrada e inversível, então $A \succ 0$ (resp. $A \succeq 0$) se e somente se $C \succ 0$ ($C \succeq 0$)

Prova

- a) $x^T C x = x^T B^T A B x = z^T A z \geq 0$.
- b) Veja que $A \succ 0$, $C \succ 0 \iff Bx \neq 0$ para qualquer $x \neq 0$, cc C seria singular. Então $\dim(\mathcal{N}(B)) = 0$ e $\dim(\mathcal{N}(B)) + \text{posto}(B) = m$ e então $\text{posto}(B) = m$.

Quando B é quadrada e admite inversa, $C = B^T AB$ é chamada transformação de congruência, A e C são ditas congruentes.

Corolário

Dada $A \in \mathbb{R}^{m \times n}$

- a) Se $A^T A \succ 0 \iff A$ possui posto coluna completo, i.e,
 $\text{posto}(A) = n$.
- b) Se $AA^T \succ 0 \iff A$ possui posto linha completo, i.e,
 $\text{posto}(A) = m$.
- c) $AA^T \succeq 0, A^T A \succeq 0$

Prova

A matriz I é positiva definida, logo $C = A^T IA$ e aplica-se o Teorema anterior.

Obs: C é positiva definida se e somente se for congruente à matriz identidade (Existe uma matriz A quadrada inversível tal que $C = A^T IA$).

Dada $A \in \mathcal{S}^n$:

① $A \succeq 0 \iff \exists B \succeq 0 : A = B^2$.

② $A \succ 0 \iff \exists B \succ 0 : A = B^2$

Prova: Partindo de $A \in \mathcal{S}^n \rightarrow A = U\Lambda U^T$.

(\rightarrow) Se $A \in \mathcal{S}_+^n$ então $\Lambda_{ii} \geq 0, i = 1, \dots, n$.

Definindo $B = U\Lambda^{\frac{1}{2}}U^T$ temos $B^2 = U\Lambda^{\frac{1}{2}}U^T U\Lambda^{\frac{1}{2}}U^T = U\Lambda U = A$.

(\leftarrow) Por outro lado, se para alguma matriz simétrica B vale $A = B^T B = B^2$, então $A \succeq 0$.

$B = A^{\frac{1}{2}}$ é a matriz raiz-quadrada de A .

Pode-se provar que $B : A = BB$ é única.

Repetindo o processo acima com $B = \Lambda^{\frac{1}{2}}U^T$, concluímos que:

• $A \succeq 0 \iff \exists B : A = B^T B$.

• $A \succ 0 \iff \exists B$ não singular tal que $A = B^T B$.

Congruência para A simétrica

- $A \succeq 0 \iff \exists B : A = B^T B.$
- $A \succ 0 \iff \exists B$ não singular tal que $A = B^T B.$

Recordando que qualquer B quadrada admite uma fatoração $B = QR$ tal que Q é ortogonal e R é triangular superior com o mesmo posto de B .

Teorema de Cholesky

- Então $A = B^T B = R^T Q^T QR = R^T R$, isto é, qualquer matriz semipositiva definida admite uma fatoração $A = R^T R$ onde R é triangular superior.
- Se $A \succ 0$, a diagonal de R é positiva e esta forma $A = R^T R$ é chamada fatoração de Cholesky de A . A fatoração pode ser feita de forma que a diagonal de R seja não negativa. É única, neste caso.

Matrizes positivas definidas são intimamente relacionadas a objetos geométricos chamados elipsóides: $\varepsilon = \{x \in \mathbb{R}^n : x^T P^{-1} x \leq 1\}$ onde $P \succ 0$ ($\iff P^{-1} \succ 0$).

- Autovetores $u_i : i = 1, \dots, n$ e autovalores $\lambda_i : i = 1, \dots, n$ de P definem a orientação e a forma do elipsóide.
- Os autovetores u_i de P determinam a orientação dos eixos principais do elipsóide e $\sqrt{\lambda_i}$ seu comprimento.
- Como existe uma decomposição de Cholesky para $P \succ 0$ e $P^{-1} \succ 0$, $x^T P^{-1} x = x^T A^T A x = \|Ax\|_2^2$.
- Logo o elipsóide pode ser caracterizado como $\varepsilon = \{x \in \mathbb{R}^n : \|Ax\|_2 \leq 1\}$, onde A é a triangular superior obtida na decomposição de Cholesky de P^{-1} .

Positividade de matrizes bloco diagonais

Dadas $A \in \mathcal{S}^n$, $B \in \mathcal{S}^m$ e a matriz bloco diagonal

$$M = \begin{pmatrix} A & 0_{n,m} \\ 0_{m,n} & B \end{pmatrix} \text{ temos que}$$

$$M \succ 0 \ (M \succeq 0) \iff A \succ 0, B \succ 0 \ (A \succeq 0, B \succeq 0).$$

Positividade de matrizes bloco simétricas - Complemento de Schur

Considere $A \in \mathcal{S}^n$, $B \in \mathcal{S}_{++}^m$, $X \in \mathbb{R}^{n \times m}$ e a matriz simétrica

$M = \begin{pmatrix} A & X \\ X^T & B \end{pmatrix}$. Defina o complemento de Schur de A em M como $S := A - XB^{-1}X^T$.

$$\text{Então } M \succ 0 \ (M \succeq 0) \iff S \succ 0 \ (S \succeq 0).$$

- Defina a matriz triangular inferior **não singular**

$$C = \begin{pmatrix} I_n & 0_{n,m} \\ -B^{-1}X^T & I_m \end{pmatrix}.$$

- Considere a transformação de congruência em M :

$$C^T MC = \begin{pmatrix} S & 0_{n,m} \\ 0_{m,n} & B \end{pmatrix}$$

- Vimos que $M \succ 0$ ($M \succeq 0$) $\iff C^T MC \succ 0$ ($C^T MC \succeq 0$).
Como $C^T MC$ é bloco diagonal e $B \succ 0$, concluímos que
 $M \succ 0$ ($M \succeq 0$) $\iff S \succ 0$ ($S \succeq 0$).

Para uma função $f \in \mathbb{C}^2$:

- Operador $\nabla = (\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})^T$.
- Vetor gradiente $\nabla f = (\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n})^T$
- Matriz Hessiana

$$H(x) = \nabla(\nabla f^T) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

A matriz Hessiana é simétrica, uma vez que $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$.

$$\nabla(x^T) = \nabla(x_1, x_2, \dots, x_n)$$

$$\nabla x^T = \begin{pmatrix} \frac{\partial x_1}{\partial x_1} & \frac{\partial x_2}{\partial x_1} & \cdots & \frac{\partial x_n}{\partial x_1} \\ \frac{\partial x_1}{\partial x_2} & \frac{\partial x_2}{\partial x_2} & \cdots & \frac{\partial x_n}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_1}{\partial x_n} & \frac{\partial x_2}{\partial x_n} & \cdots & \frac{\partial x_n}{\partial x_n} \end{pmatrix} = (e_1, e_2, \dots, e_n) = I_n$$

Exemplo - Função de Rosenbrock

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

$$\nabla f(x) = \begin{pmatrix} -400x_1(x_2 - x_1^2) - 2(1 - x_1) \\ 200(x_2 - x_1^2) \end{pmatrix}$$

$$\nabla^2 f(x) = \begin{pmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}$$

Definição

Dado um ponto $\hat{x} \in \mathbb{R}^n$ e uma direção $s \in \mathbb{R}^n$, o conjunto afim

$$\{x \in \mathbb{R}^n : x = x(\alpha) = \hat{x} + \alpha s\}$$

é chamado de linha.

Desejamos conhecer a inclinação $\frac{df}{d\alpha}$ e a curvatura $\frac{d^2f}{d\alpha^2}$ ao longo da linha, dado que dispomos de \hat{x} , ∇f e da direção s .

Pela regra da cadeia temos:

$$\frac{d}{d\alpha} = \sum_{i=1}^n \frac{dx_i(\alpha)}{d\alpha} \frac{\partial}{\partial x_i} = s^T \nabla$$

$$\frac{df}{d\alpha} = \sum_{i=1}^n \frac{dx_i(\alpha)}{d\alpha} \frac{\partial f}{\partial x_i} = s^T \nabla f = \nabla f^T s$$

$$\frac{d^2f}{d\alpha^2} = \frac{d}{d\alpha} \left(\frac{df}{d\alpha} \right) = s^T \nabla (\nabla f^T s) = s^T (\nabla^2 f) s$$

Exemplo - Função de Rosenbrock

$$\hat{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad s = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\nabla f(\hat{x}) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \quad s^T \nabla f(\hat{x}) = -2$$

$$\nabla^2 f(\hat{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 200 \end{pmatrix}, \quad s^T \nabla^2 f(\hat{x}) s = 2$$

- $f : \mathbb{R}^m \rightarrow \mathbb{R}$ função diferenciável das variáveis $z_i : i = 1, \dots, m$
- Cada variável z_i é uma função diferenciável de n variáveis x_1, \dots, x_n , isto é, $z_i = g_i(x)$ ou $z : \mathbb{R}^n \rightarrow \mathbb{R}^m$. **Matrix Jacobiana** de g :

$$J_{g(x)} = \begin{pmatrix} \frac{\partial g_1(x)}{\partial x_1} & \dots & \frac{\partial g_m(x)}{\partial x_1} \\ \vdots & \vdots & \vdots \\ \frac{\partial g_1(x)}{\partial x_j} & \dots & \frac{\partial g_m(x)}{\partial x_j} \\ \vdots & \vdots & \vdots \\ \frac{\partial g_1(x)}{\partial x_n} & \dots & \frac{\partial g_m(x)}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \nabla g_1(x) & \dots & \nabla g_m(x) \end{pmatrix}$$

- Vamos considerar a função composta $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ definida como $\phi = f(g(x))$. **Seu gradiente** $\nabla \phi(x) = [J_{g(x)}] \nabla f_g \in \mathbb{R}^n$ é o vetor **n dimensional** cuja j -ésima coordenada é dada por:

$$[\nabla \phi(x)]_j = \sum_{i=1}^m \frac{\partial f(g)}{\partial dg_i} \frac{\partial dg_i}{\partial dx_j}$$

Matriz G é simétrica, $G \in \mathcal{S}^n$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ constantes.

$$q(x) = \frac{1}{2}x^T Gx + b^T x + c \quad \text{ou}$$

$$q(x) = \frac{1}{2}(x - \hat{x})^T G(x - \hat{x}) + \hat{c}$$

onde $G\hat{x} = -b$ e $\hat{c} = c - \frac{1}{2}\hat{x}^T G\hat{x}$ e $\hat{x} \in \mathbb{R}^n$ é um ponto qualquer.

Para u, v funções de x , **pela regra da cadeia**

$$\nabla(u^T v) = \nabla u^T v + \nabla v^T u.$$

Logo,

$$\nabla q = \frac{1}{2}((\nabla x^T)Gx + (\nabla(Gx)^T)x) + b = \frac{1}{2}(IG + G^T)x + b = Gx + b$$

- Matriz Hessiana $\nabla^2 q = G$ é uma matriz constante.
- Gradiente $\nabla q = Gx + b$ é uma função afim.
- Consequentemente, dados \hat{x}, \bar{x} e os gradientes $\hat{q} = \nabla q(\hat{x}), \bar{q} = \nabla q(\bar{x})$ de q avaliados nestes pontos, temos que $\hat{q} - \bar{q} = G(\hat{x} - \bar{x})$, isto é, a Hessiana mapeia diferença entre pontos em diferenças entre gradientes.

- Considerando $f \in \mathbb{C}^\infty$, função de uma variável α :

$$f(\alpha) = f(0) + \alpha f'(0) + \frac{1}{2} \alpha^2 f''(0) + \dots$$

- Considerando $f(x(\alpha))$, onde $x = \hat{x} + \alpha s$, função de n variáveis:

$$\begin{aligned} f(\hat{x} + \alpha s) &= f(\hat{x}) + \alpha s^T \nabla f(\hat{x}) + \frac{1}{2} (\alpha s)^T \nabla^2 f(\hat{x}) (\alpha s) + \dots \\ &= f(\hat{x}) + \alpha s^T \nabla f(\hat{x}) + \frac{1}{2} \alpha^2 s^T \nabla^2 f(\hat{x}) s + \dots &= \\ f(\hat{x} + h) &= f(\hat{x}) + h^T \nabla f(\hat{x}) + \frac{1}{2} h^T \nabla^2 f(\hat{x}) h + \dots & \text{e então} \end{aligned}$$

$\nabla f(\hat{x} + h) = \nabla f(\hat{x}) + [\nabla^2 f(\hat{x})]h + \dots$ é uma expansão em série para o gradiente.

Aproximação quadrática nas vizinhanças de \hat{x}

Quando $s \rightarrow 0, h \rightarrow 0$, a função afim $\nabla f(\hat{x}) + [\nabla^2 f(\hat{x})]h$ é uma razoável aproximação para o gradiente nas vizinhanças de \hat{x} .

Dada uma direção $s \in \mathbb{R}^n$ e $\alpha \rightarrow 0$

$$f(x_0 + \alpha s) \approx f(x_0) + \alpha f'(x_0) = f(x_0) + \alpha s^T \nabla f(x_0)$$

$$\lim_{\alpha \rightarrow 0} \frac{f(x_0 + \alpha s) - f(x_0)}{\alpha} = s^T \nabla f(x_0)$$

- $s^T \nabla f(x_0)$: derivada direcional, taxa de variação de f ao longo de s .
- Se $\alpha > 0$ e $s^T \nabla f(x_0) > 0$, a função cresce ao longo de s . Se $s^T \nabla f(x_0) < 0$, a função decresce ao longo de s .
- Se $s^T \nabla f(x_0) = 0$, a aproximação da função até a primeira ordem permanece constante ao longo da direção. Isto é, a direção é tangente à curva de nível da função.

- Conjuntos: convexo, afim
- Operações que preservam convexidade
- Desigualdades generalizadas
- Separação de conjuntos convexos
- Hiperplanos suporte
- Cones duais, polar

Aberto

\mathcal{X} é aberto se para todo ponto $x \in \mathcal{X}$ for possível inserir no conjunto uma bola de raio ϵ (pequeno o suficiente) centrada em x . Isto é, \mathcal{X} é aberto se $\mathcal{B}_\epsilon = \{y \in \mathbb{R}^n : \|y - x\|_2 < \epsilon\} \subseteq \mathcal{X}$.

Fechado

\mathcal{X} é fechado se seu complemento $\mathbb{R}^n \setminus \mathcal{X}$ é aberto.

Interior

$int(\mathcal{X}) = \{x \in \mathcal{X} : \exists \epsilon > 0 \text{ tal que } \mathcal{B}_\epsilon \subseteq \mathcal{X}\}.$

Fecho

O fecho $\overline{\mathcal{X}}$ de \mathcal{X} é o conjunto de pontos limites de sequências no conjunto, isto é,

$$\overline{\mathcal{X}} = \{z \in \mathbb{R}^n : z = \lim_{k \rightarrow \infty} x^k, x^k \in \mathcal{X}, \forall k\}$$

Ou, é o menor conjunto fechado que contém \mathcal{X} .

Fronteira - $bd(\mathcal{X})$

$$bd(\mathcal{X}) = \overline{\mathcal{X}} \setminus int(\mathcal{X})$$

- O conjunto \mathcal{X} é aberto se $\mathcal{X} = \text{int}(\mathcal{X})$.
- Um conjunto aberto não contém nenhum ponto de sua fronteira.
- Um conjunto é fechado se contém todos os pontos em sua fronteira.
- A união e interseção de conjuntos abertos (fechados) resulta em conjuntos abertos (fechados).

Limitado

Um conjunto \mathcal{X} é limitado se é contido em uma bola de raio finito, isto é, se existem $r \in \mathbb{R}, x \in \mathbb{R}^n$ tais que

$$\mathcal{X} \subseteq \{z \in \mathbb{R}^n : \|z - x\|_2 < r\}.$$

Compacto

Um conjunto \mathcal{X} é compacto se for limitado e fechado.

Dado um conjunto $\mathcal{P} = \{x^1, \dots, x^m\}$ de m pontos do \mathbb{R}^n :

Combinação linear

É o subespaço vetorial

$$\mathcal{S} = \{x \in \mathbb{R}^n : x = \sum_{i=1}^m \lambda_i x^i, \lambda_i \in \mathbb{R}, i = 1, \dots, m\}.$$

Combinação afim

A envoltória afim de \mathcal{P} é o conjunto

$$aff(\mathcal{P}) = \{x \in \mathbb{R}^n : x = \sum_{i=1}^m \lambda_i x^i, \sum_{i=1}^m \lambda_i = 1\}.$$

$aff(\mathcal{P})$ é o menor conjunto afim contendo todos os pontos de \mathcal{P} .

Combinação convexa

$$co(\mathcal{P}) = \{x \in \mathbb{R}^n : x = \sum_{i=1}^m \lambda_i x^i, \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0 : i = 1, \dots, m\}.$$

Combinação cônica

$$conic(\mathcal{P}) = \{x \in \mathbb{R}^n : x = \sum_{i=1}^m \lambda_i x^i, \lambda_i \geq 0 : i = 1, \dots, m\}.$$

Um conjunto $\mathcal{C} \subseteq \mathbb{R}^n$ é convexo se contém o segmento da linha que conecta quaisquer dois pontos $x^1, x^2 \in \mathcal{C}$.

Isto é, é convexo se para $x^1, x^2 \in \mathcal{C}$ e $\lambda \in [0, 1]$ então $\{x \in \mathbb{R}^n : x = \lambda x^1 + (1 - \lambda)x^2\} \subseteq \mathcal{C}$.

- A dimensão d de um conjunto convexo é a dimensão de sua envoltória afim.
- Pode ocorrer da dimensão d do conjunto ser inferior à do espaço \mathbb{R}^n no qual está imerso.
- Neste caso $d < n$, o conjunto convexo não possui um interior regular, no sentido da definição anterior. Ele possui um *interior relativo* à sua envoltória afim.

O interior relativo de um conjunto convexo \mathcal{C} , $\text{relint}(\mathcal{C})$, é o interior de \mathcal{C} relativo à sua envoltória afim.

Isto é, um ponto $x \in \text{relint}(\mathcal{C})$ se existe uma bola d -dimensional de raio positivo, centrada em x , totalmente contida em $\text{aff}(\mathcal{C})$.

O $\text{reint}(\mathcal{C})$ e $\text{int}(\mathcal{C})$ são coincidentes se \mathcal{C} for um conjunto de dimensão completa, isto é, se a dimensão de sua envoltória afim for a dimensão do espaço na qual está imerso.

Cone

Um conjunto \mathcal{C} é um cone se satisfaz a seguinte propriedade: para qualquer $x \in \mathcal{C}$ e $\lambda \geq 0$ então $\lambda x \in \mathcal{C}$.

Nem todo cone é conexo. Por exemplo, $\{(x, y) \in \mathbb{R}^2 : y = |x|\}$ é um cone não convexo.

Cone convexo

Um conjunto \mathcal{C} é um cone convexo se é um conjunto convexo e um cone.

\mathcal{C} é um conjunto estritamente convexo se para $x^1, x^2 \in \mathcal{C}$ e $\lambda \in (0, 1)$ então $x = \lambda x^1 + (1 - \lambda)x^2 \in \text{relint}(\mathcal{C})$.

- Interseção
- Transformação afim
- Projeção
- Transformação perspectiva, transformação linear-fracionária (não veremos)

Dados m conjuntos convexos C_1, \dots, C_m , sua interseção $C = \bigcap_{i=1}^m C_i$ é um conjunto convexo.

Prova simples, pode ser obtida aplicando-se a definição de convexidade.

Em particular, o resultado vale para uma interseção de infinitos conjuntos convexos.

Dada uma transformação afim $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ (isto é, que admita uma representação do tipo $f(x) = Ax + b$ onde $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$) e um conjunto convexo $C \subseteq \mathbb{R}^n$, a **imagem de C por f**

$$f(C) = \{f(x) : x \in C\}$$

é um conjunto convexo.

Inversa da imagem

Similarmente, se $f : \mathbb{R}^k \rightarrow \mathbb{R}^n$ é uma função afim, a inversa da imagem de C por f , $f^{-1} = \{x \in \mathbb{R}^k : f(x) \in C\}$, é convexa.

- **Projeção de um conjunto convexo em um subespaço é convexo.**
A projeção no subespaço pode ser escrita como uma transformação linear, logo a projeção de um convexo é convexa.
- Para C_1, C_2 convexos:
 - Sua soma, $(C_1 + C_2) := \{x + y : x \in C_1, y \in C_2\}$, é convexa
 - Produto cartesiano, $C_1 \times C_2 := \{(x_1, x_2) : x_1 \in C_1, x_2 \in C_2\}$, é convexo.
- Soma parcial de $C_1, C_2 \in \mathbb{R}^n \times \mathbb{R}^m$,
 $S = \{(x, y_1 + y_2) : (x, y_1) \in C_1, (x, y_2) \in C_2\}$
($x \in \mathbb{R}^n, y_1, y_2 \in \mathbb{R}^m$) é convexa.

O conjunto $H = \{x \in \mathbb{R}^n : a^T x = b\}$ é chamado de hiperplano suporte de um conjunto convexo $C \in \mathbb{R}^n$ em um ponto $z \in bd(C)$ se $z \in H$ e $C \in H_-$ onde H_- é a monotonização inferior de H , $H_- := \{x \in \mathbb{R}^n : a^T x \leq b\}$.

Teorema

Todo conjunto convexo C admite um hiperplano suporte para qualquer $z \in bd(C)$, isto é para qualquer ponto em sua fronteira.

Separação

Dados dois conjuntos convexos $C_1, C_2 \subseteq \mathbb{R}^n$, dizemos que um hiperplano H **separa** C_1 de C_2 se $C_1 \subseteq H_-$ e $C_2 \subseteq H_+$, onde $H_+ := \{x \in \mathbb{R}^n : a^T x \geq b\}$ é a monotonização superior de H .

Separação estrita

Dados dois conjuntos convexos $C_1, C_2 \subseteq \mathbb{R}^n$, dizemos que um hiperplano H **estritamente separa** C_1 de C_2 se $C_1 \subseteq H_{--}$ e $C_2 \subseteq H_{++}$, onde $H_{--} := \{x \in \mathbb{R}^n : a^T x < b\}$ e $H_{++} := \{x \in \mathbb{R}^n : a^T x > b\}$.

Teorema de Separação por Hiperplanos

Quaisquer dois conjuntos convexos C_1, C_2 disjuntos ($C_1 \cap C_2 = \emptyset$) admitem um hiperplano separador H . Além diso, se C_1 é fechado e limitado (logo, compacto) e C_2 é fechado, então C_1, C_2 podem ser separados de forma estrita.

Definição

Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é convexa se:

- $\text{dom } f$ é um conjunto convexo e
- para todo $x, y \in \text{dom } f$ e $\lambda \in [0, 1]$

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad (2)$$

- f é côncava se $-f$ é convexa.
- É estritamente convexa se a desigualdade (2) vale na forma estrita (folgada, $<$) para qualquer par $x \neq y$ e $\lambda \in (0, 1)$.

Definição

Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é fortemente convexa se existe $m > 0$ tal que $\hat{f}(x) := f(x) - \frac{m}{2}\|x\|_2^2$ é convexa, ou seja se:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \frac{m}{2}\lambda(1 - \lambda)\|x - y\|_2^2 \quad (3)$$

- Convexidade forte implica em convexidade estrita: tomando $x \neq y, \lambda \in (0, 1)$, o último termo do lado direito de (3) é diferente de zero, dado que $m > 0$.

Teorema

Se $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é convexa:

- Então f é uma função contínua em $\text{int}(\text{dom } f)$.
- Além disto, f é Liptschitz em todo subconjunto compacto $\mathcal{X} \subseteq \text{int}(\text{dom } f)$, isto é, existe $M > 0$ tal que

$$|f(x) - f(y)| \leq M\|x - y\|_2, \quad \forall x, y \in \mathcal{X}$$

$$f(x) = \begin{cases} x^2 & \text{se } x \in (-1, 1] \\ 2 & \text{se } x = -1 \\ +\infty & \text{c.contrário} \end{cases}$$

- Para o caso acima $\text{dom } f = [-1, 1]$, f é contínua em $\text{int}(\text{dom } f)$ mas há uma descontinuidade em $x = -1 \in \text{bd}(\text{dom } f)$
- Se ao invés de $f(-1) = 2$ tivéssemos $f(-1) = k < 1$, a função seria convexa ?

Uma função $f : C \rightarrow (-\infty, +\infty]$ ($C \subseteq \mathbb{R}^n$) é convexa se e somente se $epi(f)$ é um subconjunto convexo do \mathbb{R}^{n+1} .

Se $f : C \rightarrow (-\infty, +\infty]$ é convexa (estritamente) então os conjuntos de níveis $S_\alpha = \{x \in \mathbb{R}^n : f(x) \leq \alpha\}$ são conjuntos convexos (estritamente convexos) para qualquer $\alpha \in \mathbb{R}$.

O sentido inverso da última afirmativa é verdadeiro ? Isto é, toda função com sublevel sets convexos é convexa ? Considere $f(x) = \ln(x)$.

Função quasi-convexa

Uma função f tal que os conjuntos de nível S_α são convexos para qualquer $\alpha \in \mathbb{R}$ é chamada quasi-convexa (quasi-côncava se $-f$ é quasi-convexa).

Objetivo

Criar um *Cálculo de Funções Convexas*, visando caracterizar convexidade de funções mais complicadas por meio de operações que preservam convexidade, aplicadas sobre funções base, mais simples, convexas.

Como ?

- Caracterizando estas operações
- Identificando formas alternativas de caracterização de convexidade, uma vez que o uso da definição pode ser muito mais complicado.

Seja $f_i : i = 1, \dots, m$ funções convexas e $\alpha_i \geq 0 > i = 1, \dots, m$.
Então, $f(x) = \sum_{i=1}^m \alpha_i f_i(x)$ é convexa em $\bigcap_{i=1}^m \text{dom } f_i(x)$.

$$\begin{aligned} f(\lambda x + (1 - \lambda)y) &= \sum_{i=1}^m \alpha_i f_i(\lambda x + (1 - \lambda)y) \\ &\leq \sum_{i=1}^m \alpha_i (\lambda f_i(x) + (1 - \lambda)f_i(y)) \\ &= \lambda \sum_{i=1}^m \alpha_i f_i(x) + (1 - \lambda) \sum_{i=1}^m \alpha_i f_i(y) \\ &= \lambda f(x) + (1 - \lambda) f(y) \end{aligned}$$

Exemplo: função entropia negativa

$f(x) = -\sum_{i=1}^n x_i \log(x_i)$ é convexa em $\text{dom } f(x) = \mathbb{R}_{++}^n$.

- f, g convexas, $\rightarrow f + g$ é convexa.
- f convexa, g estritamente convexa, $\rightarrow f + g$ é estritamente convexa.
- f convexa, g fortemente convexa, $\rightarrow f + g$ é fortemente convexa.

No último caso:

$g(x)$ fortemente convexa $\exists m > 0 : g(x) - \frac{m}{2}\|x\|_2^2$ é convexa

$h(x) = f(x) + g(x) - \frac{m}{2}\|x\|_2^2$ é convexa, soma de duas convexas

Logo $f(x) + g(x)$ é fortemente convexa

Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa.

Defina $g(x) = f(Ax + b)$, $g : \mathbb{R}^m \rightarrow \mathbb{R}$, onde $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$.

Então $g(x)$ é convexa em $\text{dom } g = \{x \in \mathbb{R}^m : Ax + b \in \text{dom } f\}$.

Exemplo

$f(z) = -\log(z)$ é convexa em $\text{dom } f = \mathbb{R}_{++}$.

Logo, $g(x) = -\log(ax + b)$ é convexa em $\text{dom } g = \{x : ax + b > 0\}$.

Caracterização de 1a. ordem

- Se f é diferenciável (isto é, $\text{dom } f$ é aberto e ∇f existe em todo o domínio),

então f é convexa se e somente se, a proximação de primeira ordem de f em qualquer ponto de seu domínio sub-estima f em todo o domínio. Isto é:

para quaisquer $x, y \in \text{dom } f$, deve valer:

$$f(y) \geq f(x) + \nabla f(x)^T(y - x) \quad (4)$$

- Caracterização de convexidade estrita de f diferenciável se (4) é satisfeita folgada para todo $y \neq x$.

É convexa \rightarrow (4)

$$f(x + \lambda(y - x)) \leq (1 - \lambda)f(x) + \lambda f(y) \quad f \text{ é cvx}$$

$$\frac{f(x + \lambda(y - x)) - f(x)}{\lambda} \leq f(y) - f(x)$$

$$(\text{quando } \lambda \rightarrow 0) \rightarrow \nabla f(x)^T(y - x) \leq f(y) - f(x)$$

É convexa \leftarrow (4)

Diante de (4), tomando $\lambda \in [0, 1]$, $x, y \in \text{dom } f$ e definindo
 $z = \lambda x + (1 - \lambda)y$:

$$f(x) \geq f(z) + \nabla f(z)^T(x - z) \quad \textcolor{red}{\times \lambda}$$

$$f(y) \geq f(z) + \nabla f(z)^T(y - z) \quad \textcolor{red}{\times (1 - \lambda)}$$

$$\begin{aligned} \lambda f(x) + (1 - \lambda)f(y) &\geq f(z) + \nabla f(z)^T(\lambda x - \lambda z + (1 - \lambda)y - (1 - \lambda)z) \\ &= f(z) + \nabla f(z)^T(0) = f(z) \end{aligned}$$

O gradiente de uma função convexa em um ponto $x \in \mathbb{R}^n \setminus \{0\}$ divide o espaço em dois semi-espacos, a partir de x :

- $H_{++} = \{y \in \mathbb{R}^n : \nabla f(x)^T(y - x) > 0\}$
- $H_- = \{y \in \mathbb{R}^n : \nabla f(x)^T(y - x) \leq 0\}$

todo ponto em $y \in H_{++}$ satisfaz $y : f(y) > f(x)$.

Talvez a mais empregada caracterização de convexidade.

Se $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é duas vezes diferenciável, então f é convexa **se e somente se sua matriz Hessiana é positiva semidefinida em todo o domínio**, i.e., $\nabla^2 f \succeq 0$ ($\nabla^2 f(x) \in \mathcal{S}_+^n$) $\forall x \in \text{dom } f$.

(\rightarrow): se é convexa, então $\nabla^2 f(x) \succeq 0$

Tomando $x_0 \in \text{dom } f$ (que é aberto, uma vez que f é diferenciável) e $v \in \mathbb{R}^n$ uma direção. Então existe $\lambda > 0$ tão pequeno quanto necessário tal que $z = x_0 + \lambda v \in \text{dom } f$.

$$f(x_0) + \lambda \nabla f(x_0)^T v + \frac{1}{2} \lambda^2 v^T \nabla^2 f(x_0) v + O(\lambda^3) = f(z) \quad \text{Taylor } x_0$$

$$\frac{1}{2} \lambda^2 v^T \nabla^2 f(x_0) v + O(\lambda^3) = f(z) - f(x_0) - \lambda \nabla f(x_0)^T v \geq 0 \quad f \text{ é cvx}$$

$$\frac{1}{2} v^T \nabla^2 f(x_0) v + \frac{O(\lambda^3)}{\lambda^2} \geq 0 \quad \div \lambda^2$$

Quando $\lambda \rightarrow 0$, $\frac{1}{2} v^T \nabla^2 f(x_0) v \geq 0$.

(\leftarrow): se $\nabla^2 f(x) \succeq 0$ em todo domínio, então f é cvx.

Versão da expansão em Série de Taylor em torno de x :

$$f(y) = f(x) + \nabla f(x)^T(y - x) + \frac{1}{2}(y - x)^T \nabla^2 f(z)(y - x)$$

onde a Hessiana é avaliada no ponto desconhecido $z \in \theta x + (1 - \theta)y$ para algum $\theta \in [0, 1]$.

Então se $\nabla^2 f(x) \succeq 0$

$$f(y) - f(x) - \nabla f(x)^T(y - x) = \frac{1}{2}(y - x)^T \nabla^2 f(z)(y - x) \geq 0$$

e f é convexa.

- Por argumento similar pode-se provar que f é fortemente convexa se e somente se $\nabla^2 f(x) \succeq ml$ para algum $m > 0$ e para todo $x \in \text{dom } f$.
- Também vale que se $\nabla^2 f \succ 0$ para todo $x \in \text{dom } f$ então f é estritamente convexa. **Atenção: o reverso não é verdadeiro.**
Contra-exemplo: $f(x) = x^4$ já que $f''(x) = 12x^2$ e em $x = 0$ a Hessiana é nula.

$$f_1(x) = 4x_1^2 + 2x_2^2 + 3x_1x_2 + 4x_1 + 5x_2 + 2 \times 10^5$$

$$f_2(x) = 4x_1^2 - 2x_2^2 + 3x_1x_2 + 4x_1 + 5x_2 + 2 \times 10^5$$

$$\nabla^2 f_1 = \begin{pmatrix} 8 & 3 \\ 3 & 4 \end{pmatrix} \quad \nabla^2 f_2 = \begin{pmatrix} 8 & 3 \\ 3 & -4 \end{pmatrix}$$

$$\lambda(\nabla^2 f_1) = \{2.39; 9.6\} \quad \lambda(\nabla^2 f_2) = \{-4.71; 8.71\}$$

$$f(x, y) : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$$

$$f(x, y) = \begin{cases} \frac{x^T x}{y} & y > 0 \\ \infty & \text{c.c.} \end{cases} \quad \text{dom } f = \{(x, y) \in \mathbb{R}^{n+1} : y > 0\}$$

$\text{dom } f$ é convexo, então vamos investigar a Hessiana de f no interior de $\text{dom } f$.

$$\nabla^2 f(x, y) = \frac{2}{y^3} \begin{pmatrix} y^2 I & -yx \\ -yx^T & x^T x \end{pmatrix}$$

Tomando um $(z, t)^T \in \mathbb{R}^{n+1}$ temos (bastaria tomar $(z, t)^T \in \mathbb{R}^n \times \mathbb{R}_+$):

$$\begin{pmatrix} z^T & t \end{pmatrix} \begin{pmatrix} y^2 I & -yx \\ -yx^T & x^T x \end{pmatrix} \begin{pmatrix} z \\ t \end{pmatrix} = \|yz - tx\|_2^2 \geq 0$$

Logo a Hessiana é positiva definida semidefinida.

Uma das mais poderosas maneiras de provar convexidade

Uma função f é convexa se e somente se sua restrição a uma linha é convexa.

Por sua restrição à uma linha consideramos a função

$$g(t) = f(x_0 + tv)$$

do escalar $t \in \mathbb{R}$, onde $x_0 \in \mathbb{R}^n$ é um ponto e $v \in \mathbb{R}^n$ é uma direção.

Então: f é convexa se e somente se $g(t) : \mathbb{R} \rightarrow \mathbb{R}$ é convexa para todo $x_0 \in \mathbb{R}^n, v \in \mathbb{R}^n$

$$f(X) = -\log \det(X) : X \in \mathcal{S}_{++}^n$$

$$\text{dom } f = \{X \in \mathcal{S}^n : z^T X z > 0 \text{ para qualquer } z \in \mathbb{R}^n \setminus \{0\}\}$$

Tomando matrizes $X_0 \in \mathcal{S}_{++}^n$, $V \in \mathcal{S}^n$, consideramos a função de $t \in \mathbb{R}$,

$$g(t) = -\log \det(X_0 + tV)$$

Lembrando que X_0 pode ser fatorada (matriz raiz quadrada)

$$X_0 = X_0^{\frac{1}{2}} X_0^{\frac{1}{2}} \text{ e } X_0^{-\frac{1}{2}} \text{ existe.}$$

$$\begin{aligned}
 \det(X_0 + tV) &= \det(X_0^{\frac{1}{2}} X_0^{\frac{1}{2}} + tV) \\
 &= \det(X_0^{\frac{1}{2}} (I + tX_0^{-\frac{1}{2}} V X_0^{-\frac{1}{2}}) X_0^{\frac{1}{2}}) \\
 &= \det(X_0^{\frac{1}{2}}) \det(I + tX_0^{-\frac{1}{2}} V X_0^{-\frac{1}{2}}) \det(X_0^{\frac{1}{2}}) \\
 &= \det(X_0) \det(I + tX_0^{-\frac{1}{2}} V X_0^{-\frac{1}{2}}) \\
 &= \det(X_0) \prod_{i=1}^n (1 + t\lambda_i(Z))
 \end{aligned}$$

onde $\lambda(Z)$ é o espectro da matriz simétrica $Z = X_0^{-\frac{1}{2}} V X_0^{-\frac{1}{2}}$.

$$g(t) = -\log \det(X_0 + tV) = -\log \det(X_0) - \sum_{i=1}^n \log(1 + t\lambda_i(Z))$$

constante + soma de funções convexas = convexa

$f_\alpha(x)$ é uma função convexa indexada pelo parâmetro $\alpha \in \mathcal{A}$, onde \mathcal{A} é um conjunto arbitrário.

Função supremo ponto-a-ponto

$$f(x) = \sup_{\alpha \in \mathcal{A}} f_\alpha(x)$$

é convexa no domínio $\{x \in \bigcap_{\alpha \in \mathcal{A}} \text{dom } f_\alpha\} \cap \{x : f(x) < \infty\}$

(obs: sempre que \mathcal{A} for compacto, sup pode ser substituído por max.)

$$f(x) = \max\{f_1(x), f_2(x) : x \in \text{dom } f_1 \cap \text{dom } f_2\}$$

$$\begin{aligned}f(\lambda x + (1 - \lambda)y) &= \max\{f_1(\lambda x + (1 - \lambda)y), f_2(\lambda x + (1 - \lambda)y)\} \\&\leq \max\{\lambda f_1(x) + (1 - \lambda)f_1(y), \lambda f_2(x) + (1 - \lambda)f_2(y)\} \\&\leq \lambda \max\{f_1(x), f_2(x)\} + (1 - \lambda) \max\{f_1(y), f_2(y)\} \\&= \lambda f(x) + (1 - \lambda)f(y)\end{aligned}$$

Função norma dual

Dada uma norma vetorial $\|\cdot\|$,

$f(x) = \|x\|^* := \max_{y \in \mathbb{R}^n: \|y\| \leq 1} y^T x$ é convexa no \mathbb{R}^n .

(máximo de infinitas funções lineares)

Máximo valor singular

Dada $X \in \mathbb{R}^{n \times m}$,

$f(X) = \sigma_{\max}(X) = \max_{v: \|v\|_2=1} \|Xv\|_2$ é convexa no $\mathbb{R}^{n \times m}$

(máximo de infinitas funções convexas: composições da norma Euclideana com a transformação afim $X \rightarrow Xv$)

(Decomposição SVD de $A = U\Sigma V^T$, $AV = U\Sigma$)

Se f é uma função convexa em (x, z) e Z é um subconjunto não vazio e convexo, então a função:

$$g(x) = \inf_{z \in Z} f(x, z)$$

é convexa, dado que $g(x) > -\infty$ para todo x .

Suponha que para A, C simétricas, quadradas e $C \succ 0$,
 $f(x, z) = x^T Ax + 2x^T Bz + z^T Cz$ seja convexa.

$$f \text{ é cvx} \iff U = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \succeq 0 \iff S_A = A - BC^{-1}B^T \succeq 0$$

Definindo $g(x) = \min_z f(x, z)$, temos que
 $g(x) = x^T (A - BC^{-1}B^T)x$.

Isto pode ser obtido impondo CNPO:

$$\nabla f(x, z) = 0 \rightarrow \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = 0$$

obtendo z em função de x e substituindo: $g(x) = f(x, z^*(x))$

Dadas as funções $h(x) : \mathbb{R}^k \rightarrow \mathbb{R}$, $g(x) : \mathbb{R}^n \rightarrow \mathbb{R}^k$.

A composição destas funções é a função $f := h \circ g = h(g(x))$

- Nem sempre a composição de funções convexas gera uma função convexa.
- Veremos alguns casos onde a composição de funções convexas preserva a convexidade.

Dadas as funções $h(x) : \mathbb{R} \rightarrow \mathbb{R}$, $g(x) : \mathbb{R} \rightarrow \mathbb{R}$, g, h convexas e h não decrescente. Então $f = h \circ g$ é convexa.

Podemos nos restringir ao caso onde $n = k = 1$, tendo em vista que g é convexa se e somente se sua restrição a uma linha é convexa.

Dadas as funções $h(x) : \mathbb{R} \rightarrow \mathbb{R}$, $g(x) : \mathbb{R} \rightarrow \mathbb{R}$, onde
 $\text{dom } h = \text{dom } g = \mathbb{R}$

Assumindo que h, g sejam duas vezes diferenciáveis, temos:

$$f' = \frac{dh}{dg} \frac{dg}{dx}$$

$$f'' = \frac{d^2h}{dg^2} \left(\frac{dg}{dx} \right)^2 + \frac{dh}{dg} \frac{d^2g}{dx^2}$$

$$f'' = h''(g(x))(g'(x))^2 + h'(g(x))g''(x)$$

$$f'' = h''(g(x))(g'(x))^2 + h'(g(x))g''(x) \quad (5)$$

f é cvx se e somente se $f'' \geq 0, \forall x$.

De (5), temos:

- ① f é cvx se h é convexa e não decrescente e g é cvx.
- ② f é cvx se h é convexa e não crescente e g é ccv.
- ③ f é ccv se h é côncava e não decrescente e g é ccv.
- ④ f é ccv se h é côncava e não crescente e g é cvx.

- ① f é cvx se \bar{h} é convexa e não decrescente e g é cvx.
- ② f é cvx se \bar{h} é convexa e não crescente e g é ccv.
- ③ f é ccv se \bar{h} é côncava e não decrescente e g é ccv.
- ④ f é ccv se \bar{h} é côncava e não crescente e g é cvx.

- \bar{h} denota a estensão de h que atribui o valor ∞ ($-\infty$) para pontos fora de $\text{dom } h$ para h convexa (h côncava).
- A única diferença entre estes resultados e o caso diferenciável é que necessitamos que \bar{f} seja não crescente ou não decrescente em todo \mathbb{R} . Exemplo: dizer que \bar{h} é não descrescente para h convexa significa que para $x < y$ implica que $\bar{h}(x) \leq \bar{h}(y)$. Logo, se $y \in \text{dom } h$ então $x \in \text{dom } h$, ou seja, o domínio de h se estende indefinidamente na direção negativa.

- $g(x)$ é cvx, então $e^{g(x)}$ é cvx.
- $g(x)$ é ccv e positiva, então $\log(g(x))$ é ccv, $\frac{1}{g(x)}$ é cvx.
- $g(x)$ é cvx e não negativa e $p \geq 1$, então $(g(x))^p$ é cvx.
- $g(x)$ é cvx, então $-\log(-g(x))$ é cvx em $\{x : g(x) < 0\}$.
- $g_i(x) : i = 1, \dots, k$ são cvxs então $\ln(\sum_{i=1}^k e^{g_i(x)})$ é cvx.