IA892 – Análise e Controle de Sistemas Lineares por Desigualdades Matriciais Lineares (LMIs)

Aula 0: Conceitos Básicos

Pedro L. D. Peres & Ricardo C. L. F. Oliveira

Faculdade de Engenharia Elétrica e de Computação
Universidade Estadual de Campinas

2^o Semestre 2022

Tópicos

Matrizes e vetores

$$A_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad ; \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- Transposição:

$$A' = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} ; \; x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} ; \; (A+B)' = A' + B' \; ; \; (AB)' = B'A'$$

$$\begin{bmatrix} X & Y \\ Z & W \end{bmatrix}' = \begin{bmatrix} X' & Z' \\ Y' & W' \end{bmatrix}$$

Matrizes e vetores

- Matriz conjugada $\bar{A}$ $(A \in \mathbb{C}^{m\times n})$

$$A = \begin{bmatrix} 1 & 1+j & 0 \\ -3-3j & j & -1-5j \\ 0 & 4j & 1+j \end{bmatrix} \; ; \; \bar{A} = \begin{bmatrix} 1 & 1-j & 0 \\ -3+3j & -j & -1+5j \\ 0 & -4j & 1-j \end{bmatrix}$$

- Matriz conjugada transposta A^*

$$A^* = \begin{bmatrix} 1 & -3+3j & 0 \\ 1-j & -j & -4j \\ 0 & -1+5j & 1-j \end{bmatrix} \; ; \quad \begin{array}{c} (A+B)^* = A^* + B^* \quad ; \quad (AB)^* = B^*A^* \; ; \\ c \in \mathbb{C} \Rightarrow (cA)^* = \bar{c}A^* \end{array}$$

- Simétrica: $A = A'$ $(a_{ij} = a_{ji})$
- Anti-simétrica: $A = -A'$ $(a_{ij} = -a_{ji})$

- Se $A \in \mathbb{R}^{n\times n}$, então $A + A'$ é simétrica ; $A - A'$ é anti-simétrica

- Se $A \in \mathbb{R}^{m\times n}$, então $A'A$ é simétrica ; AA' é simétrica

Matrizes e vetores

- Considere as matrizes: A $(n \times m)$, B $(m \times r)$, C $(l \times n)$, D $(r \times p)$

Seja a_i a i-ésima coluna de A e b_j a j-ésima linha de B:

$$AB = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_m b_m$$

$$CA = C \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix} = \begin{bmatrix} Ca_1 & Ca_2 & \cdots & Ca_m \end{bmatrix}$$

$$BD = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} D = \begin{bmatrix} b_1 D \\ b_2 D \\ \vdots \\ b_m D \end{bmatrix}$$

$a_i b_i$: matriz n por r (vetor $n \times 1$ multiplicado por vetor $1 \times r$)
$b_i a_i$: só está definido para $n = r$ (escalar)

Matrizes e vetores

- Hermitiana: $A = A^* \ (a_{ij} = \bar{a}_{ji})$
- Anti-hermitiana: $A^* = -A$

- Toda matriz quadrada A pode ser expressa de maneira única como

$$A = X + jY \quad ; \quad X = \frac{1}{2}(A + A^*) \quad ; \quad Y = \frac{1}{2j}(A - A^*) \quad ; \quad X, Y \text{ hermitianas}$$

- Ortogonal: $A \in \mathbb{R}^{n \times n}$, $A'A = AA' = \mathbf{I}$
- Unitária: $A \in \mathbb{C}^{n \times n}$, $A^*A = AA^* = \mathbf{I}$

- **Traço:** soma dos elementos da diagonal de uma matriz quadrada:

$$A_{n \times n} \implies \text{Tr}\,(A) = \sum_{i=1}^{n} a_{ii} \ ; \ \text{Tr}(AB) = \text{Tr}(BA) \ ; \ \text{Tr}(\alpha A + B) = \alpha \text{Tr}(A) + \text{Tr}(B)$$

- $\mathbf{Tr}(MX)$ é uma função linear dos elementos da matriz X.

- Um conjunto de vetores $\{x_1, x_2, \ldots, x_n\}$, $x_i \in \mathbb{R}^n$, é linearmente dependente (LD) se e somente se existem escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$, não todos nulos, tais que $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0$.

- Se a igualdade for verdadeira apenas para $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, diz-se então que o conjunto $\{x_1, x_2, \ldots, x_n\}$ é linearmente independente (LI).

- Se um conjunto de vetores $\{x_1, x_2, \ldots, x_n\}$ é linearmente dependente, existe pelo menos um α_i diferente de zero e (por exemplo, se $\alpha_1 \neq 0$) $x_1 = -(\alpha_2 x_2 + \alpha_3 x_3 \cdots + \alpha_n x_n)/\alpha_1$ isto é, um dos vetores (mas não necessariamente qualquer um) pode ser escrito como uma combinação linear dos demais.

Equivalentemente, os vetores são LI se $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$ implicar que $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2$, ..., $\alpha_n = \beta_n$. Ou, ainda, se nenhum vetor x_i puder ser expresso como combinação linear dos demais.

- Um conjunto de vetores LI do $\mathbb{R}^n$ é uma base se qualquer vetor $x \in \mathbb{R}^n$ puder ser expresso de forma única como uma combinação linear destes vetores.

• Em um espaço de dimensão n, qualquer conjunto de n vetores LI forma uma base

• Quaisquer duas bases de um espaço n-dimensional possuem o mesmo número de elementos.

Se $\{q_1, q_2, \ldots, q_n\}$ forma uma base para o $\mathbb{R}^n$, então qualquer vetor $x \in \mathbb{R}^n$ pode ser escrito de maneira única como

$$x = \alpha_1 q_1 + \alpha_2 q_2 + \cdots + \alpha_n q_n$$

Definindo a matriz quadrada Q $(n \times n)$ $\quad Q \triangleq \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix}$

tem-se $x = Q \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = Q\alpha$ e $\alpha \triangleq \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix}'$ é a representação de x na base $\{q_1, q_2, \ldots, q_n\}$.

• A, $\bar{A}$: matrizes similares • PAP^{-1}, $P^{-1}\bar{A}P$: Transformações de similaridade

- **Range** da matriz $A \in \mathbb{R}^{m \times n}$ é definido como o conjunto de todas as possíveis combinações lineares das colunas de A

$$\mathscr{R}(A) \triangleq \left\{ y = Ax \ : \ \ x \in \mathbb{R}^n \right\} \subseteq \mathbb{R}^m$$

- **Rank** da matriz $A \in \mathbb{R}^{m \times n}$ é definido como a dimensão do range de A (ou, equivalentemente, como o número de colunas LI em A) e denotado $\rho(A)$.

- **Espaço nulo** da matriz A consiste no conjunto de vetores $x \in \mathbb{R}^n$ tais que $Ax = 0$. A dimensão do espaço nulo é chamada de **nulidade** da matriz A e denotada $\nu(A)$.

$$\mathscr{N}(A) \triangleq \left\{ x \in \mathbb{R}^n \ : \ Ax = 0 \right\}$$

Note que $y = Ax$ pode ser escrito $y = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n$ e portanto $x_i \in \mathbb{R}$, $i = 1, \ldots, n$ são as ponderações das colunas de $A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$

- Para $A \in \mathbb{R}^{m\times n}$: $\quad \nu(A) = n - \rho(A)$

  - $$\begin{aligned} \text{rank } A &= \text{número de colunas LI de } A \\ &= \text{número de linhas LI de } A \\ &\leq \min\,(n,m) \end{aligned}$$

- $\mathscr{N}(A)$ e $\mathscr{R}(A)$ são espaços lineares; ($\mathscr{N}(A)$ é um subespaço do $\mathbb{R}^n$ e $\mathscr{R}(A)$ é um subespaço do $\mathbb{R}^m$)

- Se $\nu(A) = 0$, $\mathscr{N}(A) = \{0\}$ e as seguintes afirmações são equivalentes:

  - x pode ser determinado de maneira única de $y = Ax$
  - colunas de A são LI
  - $\det(A'A) \neq 0$

- Se $\nu(A) = k$, $Ax = 0$ possui k soluções LI

- Dada uma matriz $A \in \mathbb{R}^{m \times n}$ e um vetor $y \in \mathbb{R}^{m \times 1}$, existe uma solução $x \in \mathbb{R}^{n \times 1}$ da equação

$$y = Ax$$

se e somente se $y \in \mathscr{R}(A)$ ou, equivalentemente,

$$\rho(A) = \rho(\begin{bmatrix} A & y \end{bmatrix})$$

- Dada uma matriz A, uma solução x de $y = Ax$ existe para todo y se e somente se $\rho(A) = m$ (rank completo de linhas).

- Complemento ortogonal

Dado um conjunto $S \subset \mathbb{R}^n$, o complemento ortogonal de S é dado por

$$S^{\perp} = \left\{ x \in \mathbb{R}^n \ : \ x'y = 0 \ \forall \ y \in S \right\}$$

- Note que $\mathscr{N}(A) = \mathscr{R}(A')^{\perp}$

Determinantes

Notação: $\det(A)$ é o **determinante** da matriz quadrada $A_{n\times n}$, função escalar que pode ser calculada a partir de uma linha arbitrária k da matriz

$$\det(A) = \sum_{j=1}^{n} a_{kj} C_{kj}$$

ou ainda a partir de uma coluna l qualquer

$$\det(A) = \sum_{i=1}^{n} a_{il} C_{il}$$

sendo C_{pq} os **cofatores** dados por

$$C_{pq} = (-1)^{p+q} M_{pq}$$

e M_{pq} os **menores** associados aos elementos a_{pq} da matriz $A_{n\times n}$. O **menor** M_{pq} é o determinante da matriz de dimensão $(n-1)\times(n-1)$ obtida a partir da eliminação da linha p e da coluna q da matriz A.

Determinantes

Determinantes de $A_{n\times n}$, para $n=1,2$ e 3:

$$A = \begin{bmatrix} a_{11} \end{bmatrix} \implies \det(A) = a_{11}$$

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \implies \det(A) = a_{11}a_{22} - a_{12}a_{21}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \implies \begin{array}{l} \det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ -a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{32}a_{23} \end{array}$$

- **Propriedades dos Determinantes**

- Se duas linhas (colunas) são trocadas, troca-se o sinal

- Se um múltiplo de uma linha (coluna) é subtraído de outra linha (coluna) da matriz, o determinante não se altera

- Se uma matriz tem duas linhas (colunas) idênticas o determinante é igual a zero

Determinantes

$$\det\begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix} = -2 \;\; ; \;\; c_2 = c_2 - 2c_1 \;\; \Rightarrow \;\; \det\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} = -2$$

$$c_2 = c_2 + 2c_1 \;\; \Rightarrow \;\; \det\begin{bmatrix} 1 & 4 \\ 0 & -2 \end{bmatrix} = -2 \; ; \; \text{mas} \;\; c_2 = 2c_1 - c_2 \;\; \rightarrow \;\; \det\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = 2$$

- $\det(A') = \det(A) \quad ; \quad \det(A^*) = \det(\bar{A})$
- $\det(AB) = \det(A)\det(B)$
- $\det(\alpha A) = \alpha^n \det(A) \; (A \in \mathbb{R}^{n\times n})$
- $\det\begin{bmatrix} A & B \\ \mathbf{0} & D \end{bmatrix} = \det\begin{bmatrix} A & \mathbf{0} \\ C & D \end{bmatrix} = \det(A)\det(D)$
- Determinante de uma matriz hermitiana é sempre real

$$\det(A) = \det(A^*) = \det(\bar{A}') = \det(\bar{A})$$

Determinantes

- Operações elementares (de linhas, multiplicando à esquerda)
- troca da linha i com a linha j (multiplica o determinante por -1)

$$
\begin{array}{cc}
\begin{bmatrix}
1 & & & \vdots & & & & \vdots & & & \\
& \ddots & & \vdots & & & & \vdots & & & \\
& & 1 & \vdots & & & & \vdots & & & \\
\cdots & \cdots & \cdots & 0 & \cdots & \cdots & \cdots & 1 & \cdots & \cdots & \cdots \\
& & & \vdots & 1 & & & \vdots & & & \\
& & & \vdots & & \ddots & & \vdots & & & \\
& & & \vdots & & & 1 & \vdots & & & \\
\cdots & \cdots & \cdots & 1 & \cdots & \cdots & \cdots & 0 & \cdots & \cdots & \cdots \\
& & & \vdots & & & & \vdots & 1 & & \\
& & & \vdots & & & & \vdots & & \ddots & \\
& & & \vdots & & & & \vdots & & & 1
\end{bmatrix}
&
\begin{array}{l}
\\ \\ \\ \rightarrow \text{linha } i \\ \\ \\ \\ \rightarrow \text{linha } j \\ \\ \\ \\
\end{array}
\\
\begin{array}{ccc}
\downarrow & \qquad\qquad & \downarrow \\
\text{coluna } i & & \text{coluna } j
\end{array}
&
\end{array}
$$

Determinantes

- Exemplos:

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

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{21} & A_{22} \\ A_{11} & A_{12} \end{bmatrix}$$

Determinantes

- multiplicação da linha i por $\alpha > 0$ (multiplica o determinante por α)

$$
\begin{array}{cc}
\begin{bmatrix}
1 & & & \vdots & & & \\
& \ddots & & \vdots & & & \\
& & 1 & \vdots & & & \\
\cdots & \cdots & \cdots & \alpha & \cdots & \cdots & \cdots \\
& & & \vdots & 1 & & \\
& & & \vdots & & \ddots & \\
& & & \vdots & & & 1
\end{bmatrix} & \rightarrow \text{linha } i \\
\qquad\qquad\qquad \downarrow & \\
\qquad\qquad \text{coluna } i &
\end{array}
$$

Determinantes

- linha $j = \alpha$ vezes a linha i + linha j (preserva o determinante)

$$\begin{bmatrix} 1 & \vdots & & & \\ & \ddots & & & \\ & \vdots & 1 & & \\ \cdots & \alpha & \cdots & \ddots & \cdots \\ & \vdots & & & 1 \end{bmatrix} \rightarrow \text{linha } j$$

$\downarrow$

coluna i

Determinantes

- Exemplos

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

$$
\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}
=
\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 7 & 7 & 7 & 7 & 7 & 7 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}
$$

Inversa

Se $A \in \mathbb{R}^{n\times n}$ (matriz quadrada), A é invertível ou não-singular se $\det(A) \neq 0$.

Condições equivalentes:

- as colunas (linhas) de A formam uma base para o $\mathbb{R}^n$
- a equação $y = Ax$ tem uma solução única $x = A^{-1}y$ para todo $y \in \mathbb{R}^n$. Em particular, a única solução de $Ax = 0$ é $x = 0$
- $AA^{-1} = A^{-1}A = \mathbf{I}$ • $\mathscr{N}(A) = \{0\}$ • $\mathscr{R}(A) = \mathbb{R}^n$ • $\det(A'A) = \det(AA') \neq 0$

$A^{-1} = \dfrac{1}{\det(A)}\text{Adj }(A)$, Adj (A): matriz **adjunta** da matriz A, Adj $(A) = [\text{Co }(A)]'$

Co (A): matriz **cofatora** de A, composta pelos cofatores C_{ij} da matriz A

Inversa

- Uma identidade matricial $AB = C$ pode ser particionada de várias maneiras:

$$\left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right] \left[\begin{array}{c|c} B_1 & B_2 \end{array}\right] = \left[\begin{array}{c|c} A_1B_1 & A_1B_2 \\ \hline A_2B_1 & A_2B_2 \end{array}\right] = \left[\begin{array}{c|c} C_1 & C_2 \\ \hline C_3 & C_4 \end{array}\right]$$

$$\left[\begin{array}{c|c} A_1 & A_2 \\ \hline A_3 & A_4 \end{array}\right] \left[\begin{array}{c} B_1 \\ \hline B_2 \end{array}\right] = \left[\begin{array}{c} A_1B_1 + A_2B_2 \\ \hline A_3B_1 + A_4B_2 \end{array}\right] = \left[\begin{array}{c} C_1 \\ \hline C_2 \end{array}\right]$$

$$\left[\begin{array}{c|c} A_1 & A_2 \\ \hline A_3 & A_4 \end{array}\right] \left[\begin{array}{c|c} B_1 & B_2 \\ \hline B_3 & B_4 \end{array}\right] = \left[\begin{array}{c|c} A_1B_1 + A_2B_3 & A_1B_2 + A_2B_4 \\ \hline A_3B_1 + A_4B_3 & A_3B_2 + A_4B_4 \end{array}\right] = \left[\begin{array}{c|c} C_1 & C_2 \\ \hline C_3 & C_4 \end{array}\right]$$

- Note que a inversa de uma matriz ortogonal é igual à sua transposta $A^{-1} = A'$

- Note que a inversa de uma matriz unitária é igual à sua conjugada transposta $A^{-1} = A^*$

- Inversa de uma matriz $A \in \mathbb{R}^{2\times 2}$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Inversa → Fórmulas

Considere a matriz quadrada $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ com A_{11} e A_{22} também quadradas.

- Se A_{11} é não-singular

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ A_{21}A_{11}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} A_{11} & \mathbf{0} \\ \mathbf{0} & \Delta \end{bmatrix} \begin{bmatrix} \mathbf{I} & A_{11}^{-1}A_{12} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

$\Delta \triangleq A_{22} - A_{21}A_{11}^{-1}A_{12}$
A é não-singular sse
Δ é não-singular

- Se A_{22} é não-singular

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & A_{12}A_{22}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \hat{\Delta} & \mathbf{0} \\ \mathbf{0} & A_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ A_{22}^{-1}A_{21} & \mathbf{I} \end{bmatrix}$$

$\hat{\Delta} \triangleq A_{11} - A_{12}A_{22}^{-1}A_{21}$
A é não-singular sse
$\hat{\Delta}$ é não singular

- Δ ($\hat{\Delta}$) é chamado de complemento de Schur de A_{11} (A_{22})

Inversa → Fórmulas

- Se A é não singular

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}\Delta^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}\Delta^{-1} \\ -\Delta^{-1}A_{21}A_{11}^{-1} & \Delta^{-1} \end{bmatrix}$$

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \hat{\Delta}^{-1} & -\hat{\Delta}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}\hat{\Delta}^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}\hat{\Delta}^{-1}A_{12}A_{22}^{-1} \end{bmatrix}$$

$$\Delta \triangleq A_{22} - A_{21}A_{11}^{-1}A_{12} \quad ; \quad \hat{\Delta} \triangleq A_{11} - A_{12}A_{22}^{-1}A_{21}$$

- e portanto (formulas para matriz inversa)

$$(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} = A_{11}^{-1} + A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1}$$

$$(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} = A_{22}^{-1} + A_{22}^{-1}A_{21}(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1}$$

Inversa $\rightarrow$ Fórmulas

- Para A bloco-triangular

$$\begin{bmatrix} A_{11} & \mathbf{0} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & \mathbf{0} \\ -A_{22}^{-1}A_{21}A_{11}^{-1} & A_{22}^{-1} \end{bmatrix} \; ; \; \begin{bmatrix} A_{11} & A_{12} \\ \mathbf{0} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ \mathbf{0} & A_{22}^{-1} \end{bmatrix}$$

- Se A_{11} é não-singular $\det(A) = \det(A_{11})\det(A_{22} - A_{21}A_{11}^{-1}A_{12})$

- Se A_{22} é não-singular $\det(A) = \det(A_{22})\det(A_{11} - A_{12}A_{22}^{-1}A_{21})$

- Para matrizes quaisquer $B \in \mathbb{C}^{m\times n}$ e $C \in \mathbb{C}^{n\times m}$

$$\det\begin{bmatrix} \mathbf{I}_m & B \\ -C & \mathbf{I}_n \end{bmatrix} = \det(\mathbf{I}_n + CB) = \det(\mathbf{I}_m + BC)$$

- Para quaisquer $x, y \in \mathbb{C}^n$ $\det(\mathbf{I}_n + xy^*) = 1 + y^*x$

Produto Interno

- Para dois vetores $x, y \in \mathbb{R}^n$, define-se o produto interno (ou produto escalar) como

$$\left\langle x, y \right\rangle \triangleq x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = x' y$$

Propriedades:

- $\left\langle \alpha x, y \right\rangle = \alpha \left\langle x, y \right\rangle$
- $\left\langle x + y, z \right\rangle = \left\langle x, z \right\rangle + \left\langle y, z \right\rangle$
- $\left\langle x, y \right\rangle = \left\langle y, x \right\rangle$
- $\left\langle x, x \right\rangle = \|x\|^2 \geq 0$
- $\left\langle x, x \right\rangle = 0 \iff x = 0$
- $x' y = 0 \quad \Rightarrow \quad$ vetores ortogonais
- o vetor x' pode ser visto como uma função linear $x' \ : \ \mathbb{R}^n \to \mathbb{R}$

Norma de vetores

Qualquer função real representada por $\|x\|$ pode ser definida como uma **norma** se para qualquer $x \in \mathbb{R}^n$ e para qualquer escalar $\alpha \in \mathbb{R}$

1. $\|x\| \geq 0 \quad ; \quad \|x\| = 0 \iff x = 0$
2. $\|\alpha x\| = |\ \alpha\ |\ \|x\|$
3. $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$ (Desigualdade Triangular)

- **Exemplo**

$$\|x\|_p = \left(\sum_{i=1}^{n} |\ x_i\ |^p \right)^{\frac{1}{p}} \quad ; \quad p \geq 1 \quad p \text{ inteiro}$$

$\|x\|_2 = \sqrt{x'x}$ (norma Euclidiana)

$\|x\|_\infty = \max_i |\ x_i\ |$ (norma infinito)

Norma de vetores

- Para matrizes $X, Y \in \mathbb{R}^{m\times n}$, o produto interno (escalar) é dado por

$$\left\langle X, Y \right\rangle = \mathbf{Tr}(X'Y) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}y_{ij}$$

- note que o produto interno de duas matrizes é o produto interno de vetores do $\mathbb{R}^{nm}$ construídos a partir dos coeficientes das matrizes seguindo uma certa ordem (por exemplo, por linhas)

- A norma de Frobenious de uma matriz $X \in \mathbb{R}^{m\times n}$ é dada por

$$\|X\|_F = \left(\mathbf{Tr}(X'X)\right)^{1/2} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}^2\right)^{1/2}$$

- A norma de Frobenious de uma matriz é a norma euclidiana do vetor obtido a partir dos coeficientes da matriz em uma certa ordem (não é a norma 2 da matriz)

- Note que, para matrizes simétricas $X, Y \in \mathbb{R}^{n\times n}$

$$\left\langle X, Y \right\rangle = \mathbf{Tr}(XY) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}y_{ij} = \sum_{i=1}^{n} x_{ii}y_{ii} + 2\sum_{i<j}^{n}\sum_{j=1}^{n} x_{ij}y_{ij}$$

Norma de Matrizes

$$\|A\|_{a,b} \triangleq \sup_{x \neq 0} \frac{\|Ax\|_a}{\|x\|_b} = \sup_{\|x\|_b = 1} \|Ax\|_a$$

sendo sup o supremo ou o menor limitante superior. A norma $\|A\|_a$ é definida por meio da norma de $\|x\|_b$ (norma induzida). Para a, b diferentes, têm-se diferentes normas de A

$$\|A\|_1 = \max_j \sum_{i=1}^{n} | a_{ij} | \;\; ; \;\; \|A\|_2 = \Big(\lambda_{max}(A^* A)\Big)^{\frac{1}{2}} = \sigma_{max}(A) \;\; ; \;\; \|A\|_\infty = \max_i \sum_{j=1}^{n} | a_{ij} |$$

- Propriedades: $\|Ax\| \leq \|A\| \|x\| \;\; ; \;\; \|A+B\| \leq \|A\| + \|B\| \;\; ; \;\; \|AB\| \leq \|A\| \|B\|$

Decorrência de

$$\|(A+B)x\| = \|Ax + Bx\| \leq \|Ax\| + \|Bx\| \leq \Big(\|A\| + \|B\|\Big) \|x\| \; ;$$
$$\|ABx\| \leq \|A\| \|Bx\| \leq \|A\| \|B\| \|x\|$$

- Exemplo: $A = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$ $\quad \|A\|_1 = 4 \;\; ; \;\; \|A\|_2 = 3.7 \;\; ; \;\; \|A\|_\infty = 5$

Pontos extremos de conjuntos limitados e autovalores

Vetores não nulos $x \in \mathbb{C}^n$ tais que Ax seja um múltiplo escalar de x surgem no contexto da maximização (ou minimização) de uma função quadrática real e simétrica sujeita a uma restrição geométrica

$$\max \; x'Ax \qquad \text{sujeito a} \qquad x \in \mathbb{R}^n \;\; ; \;\; x'x = 1$$

com $A = A' \in \mathbb{R}^{n \times n}$. Introduzindo o lagrangeano $L(x,\lambda) = x'Ax + \lambda(1 - x'x)$ e impondo as condições necessárias de otimalidade tem-se

$$\frac{\nabla L}{\nabla \lambda} = 0 \;\Rightarrow\; x'x = 1 \qquad ; \qquad \frac{\nabla L}{\nabla x} = 0 \;\Rightarrow\; 2(Ax - \lambda x) = 0$$

- Se um vetor $x \in \mathbb{C}^n$ com $x'x = 1$ (e portanto $x \neq 0$) é um ponto extremo de $x'Ax$, necessariamente x satisfaz a equação $Ax = \lambda x$.

Autovalores e autovetores

Um escalar $\lambda \in \mathbb{C}$ é um **autovalor** (valor próprio) de $A \in \mathbb{R}^{n\times n}$ se existe um vetor $x \in \mathbb{C}^n$ não nulo tal que

$$Ax = \lambda x$$

Qualquer vetor $x \in \mathbb{C}$ que satisfaça $Ax = \lambda x$ é chamado de **autovetor** (vetor próprio) de A associado a λ (mais precisamente, esta é a definição para autovetores à direita de A).

- $Ax = \lambda x$ pode ser visto como $(A - \lambda \mathbf{I})x = 0$
- $\exists x \in \mathbb{C}^n, x \neq 0 \ : \ (A - \lambda \mathbf{I})x = 0 \ \Leftrightarrow \ \det(A - \lambda \mathbf{I}) = 0$
- $\Delta(\lambda) \triangleq \det(\lambda \mathbf{I} - A)$: polinômio (mônico) característico de A
- $\Delta(\lambda) = 0$: **equação característica** de A
- Grau de $\Delta(\lambda) = n$ e portanto $A \in \mathbb{R}^{n\times n}$ possui n autovalores.

Autovalores e autovetores

- Exemplo: $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$; $A - \lambda \mathbf{I} = \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix}$

$$\det(A - \lambda \mathbf{I}) = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21})$$

Note que $\lambda \in \mathbb{C}$ é um autovalor de $A \in \mathbb{R}^{n \times n}$ se $\Delta(\lambda) = \det(\lambda \mathbf{I} - A) = 0$

Essa condição é equivalente à existência de $y \in \mathbb{C}$ tal que
$y'A = \lambda y' \quad \Longrightarrow \quad y'(\lambda \mathbf{I} - A) = 0$

- qualquer y que satisfaça a relação acima é chamado de autovetor à esquerda de A (associado ao autovalor λ).

- Se $v \in \mathbb{C}^n$ é um autovetor associado a $\lambda \in \mathbb{C}$, então $\bar{v}$ (complexo conjugado de v) é um autovetor associado a $\bar{\lambda}$.

- Se v é um autovetor de A, a transformação linear A aplicada sobre v produz um escalonamento de λ (na direção v).

Formas Quadráticas (Hermitianas)

São funções de n variáveis complexas da forma

$$\sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij}\bar{x}_i x_j = x^* M_1 x \quad \in \mathbb{R} \ , \quad m_{ij} \in \mathbb{C}$$

Como $x^* M_1 x \in \mathbb{R}$

$$(x^* M_1 x)^* = (x^* M_1^* x) = (x^* M_1 x) \ \ ; \ \ x^* M_1 x - x^* M_1^* x = 0 \quad , \quad x^*(M_1 - M_1^*)x = 0$$

$$x^* M_1 x = x^* \left(\frac{1}{2}(M_1 + M_1^*)\right) x \triangleq x^* M x \quad ; \quad M = \frac{1}{2}(M_1 + M_1^*) \ \ \Rightarrow M = M^*$$

- Toda forma Hermitiana pode ser escrita como $x^* M x$ com $M = M^*$

Formas Quadráticas (Hermitianas)

- **Exemplos**

$$M = \begin{bmatrix} 5 & -1 \\ -1 & 8 \end{bmatrix} \quad ; \quad x'Mx = 5x_1^2 - 2x_1x_2 + 8x_2^2 > 0 \ \ \forall \ x_1, x_2 \neq 0$$

$$\|Bx\|^2 = x'B'Bx \quad ; \quad \sum_{i=2}^{n} (x_{i+1} - x_i)^2 \quad ; \quad \|Fx\|^2 - \|Gx\|^2$$

- Se $x'Ax = x'Bx$ para todo $x \in \mathbb{R}^n$ e as matrizes A e B são simétricas, então $A = B$ (unicidade da forma quadrática)
- $f(x)$ quadrática pode definir conjuntos: $\Big\{x \ : \ f(x) = c\Big\}$; $\Big\{x \ : \ f(x) \leq c\Big\}$

Suponha que a matriz simétrica A foi decomposta na forma $A = Q\Lambda Q'$, com Λ diagonal contendo os autovalores reais de A ordenados $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Então

$$x'Ax = x'Q\Lambda Q'x = (Q'x)'\Lambda(Q'x) = \sum_{i=1}^{n} \lambda_i (q_i'x)^2 \leq \lambda_1 \sum_{i=1}^{n} (q_i'x)^2 = \lambda_1 \|x\|^2$$

Similarmente, $x'Ax \geq \lambda_n \|x\|^2$. Assim, $\quad \lambda_n x'x \leq x'Ax \leq \lambda_1 x'x$

Positividade

- Uma matriz M é **definida positiva** se e somente se $x^* Mx > 0, \ \forall \ x \neq 0, \ x \in \mathbb{C}^n$.

- Uma matriz M é **semidefinida positiva** se e somente se $x^* Mx \geq 0, \ \forall \ x \in \mathbb{C}^n$.

- Uma matriz M é (semi)**definida negativa** se $-M$ é (semi)definida positiva.

- Uma matriz hermitiana $n \times n$ M é definida positiva (semidefinida positiva) se e somente se qualquer das condições seguintes é satisfeita.

- Todos os autovalores de M são positivos (não negativos)

- Todos os menores principais líderes de M são positivos (todos os menores principais de M são não negativos)

- Existe uma matriz não singular $n \times n$ N (uma matriz singular $n \times n$ N ou uma matriz $m \times n$ com $m < n$) tal que $M = N^* N$.

Positividade

Considerando a última condição, se $M = N^* N$, então

$$x^* M x = x^* N^* N x = (Nx)^* (Nx) = \|Nx\|_2^2 \geq 0$$

para qualquer x. Se N é não-singular, $Nx = 0 \Rightarrow x = 0$, e portanto M é definida positiva. Se N é singular, existe $x \neq 0$ tal que $Nx = 0$ e M é semidefinida positiva.

- Menores Principais de uma matriz $M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$

- Menores Principais (determinantes de todas as submatrizes de M cujas diagonais coincidem ou fazem parte da diagonal de M):

$$m_{11} \;,\quad m_{22} \;,\quad m_{33}$$

$$\det\left(\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}\right) \;,\quad \det\left(\begin{bmatrix} m_{11} & m_{13} \\ m_{31} & m_{33} \end{bmatrix}\right) \;,\quad \det\left(\begin{bmatrix} m_{22} & m_{23} \\ m_{32} & m_{33} \end{bmatrix}\right)$$

$$\det(M)$$

Positividade

- Menores Principais Líderes (determinantes das submatrizes de M obtidas ao eliminarem-se as últimas k colunas e k linhas, $k = 2, 1, 0$):

$$m_{11} \quad , \quad \det\left(\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}\right) \quad , \quad \det(M)$$

- todos os menores principais líderes positivos $\implies$ menores principais positivos. No entanto, todos os menores principais líderes não negativos não implica que todos os menores são não negativos. Por exemplo,

$$M = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 2 \end{bmatrix} \quad \begin{array}{l} \Rightarrow \; m_{11} = 0 \; , \; \det \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0 \; , \; \det(M) = 0 \\ \text{mas} \; \Rightarrow \; m_{22} = 0, \; m_{33} = 2, \; \det \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} = -1 \; ; \; \det \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} = 0 \end{array}$$

$M = \begin{bmatrix} 5 & -1 \\ -1 & 8 \end{bmatrix}$ é definida positiva, pois $m_{11} = 5 > 0$; $\det(M) = 39 > 0$; os autovalores de M são $\lambda_1 = 4.697$ e $\lambda_2 = 8.303$ e

$$M = N'N \quad ; \quad N = \begin{bmatrix} 2.2273 & -0.1981 \\ -0.1981 & 2.8215 \end{bmatrix} \quad ; \quad \det(N) = 6.2450 \neq 0$$

Matrizes definidas positivas

- Notaçao: $A > 0 (\geq 0)$ indica que A é (semi)definida positiva.
$A \geq B \Rightarrow A - B \geq 0$.

- Autovalores, traço, determinante e menores principais são positivos

- Considere $A \in \mathbb{R}^{n\times n}$ definida positiva e $T \in \mathbb{R}^{n\times m}$. A matriz $T'AT$ é semi-definida positiva. Como o rank $(T'AT) =$ rank (T), $T'AT$ é definida positiva se e somente se rank $(T) = m$

- Se A e B são definidas positivas, então:

- $A \geq B$ se e somente se $B^{-1} \geq A^{-1}$;

- Se $A \geq B$, então $\det(A) \geq \det(B)$ e $\mathbf{Tr}(A) \geq \mathbf{Tr}(B)$;

- Se $A > 0$ e $B \geq 0$ então $A > B$ se e somente se $|\lambda_{max}(BA^{-1})| < 1$

Matrizes definidas positivas

- A matriz simétrica $H'H$ é sempre semidefinida positiva; será definida positiva se $H'H$ for não singular. O mesmo vale para HH'.

Como $H'H$ e HH' são matrizes simétricas semidefinidas positivas, seus autovalores são reais e não negativos.

Se $H \in \mathbb{R}^{m\times n}$, então $H'H$ tem n autovalores e HH' tem m autovalores. Usando a propriedade de determinantes, para $A \in \mathbb{R}^{m\times n}$ e $B \in \mathbb{R}^{n\times m}$

$$s^n \det(s\mathbf{I}_m - AB) = s^m \det(s\mathbf{I}_n - BA)$$

tem-se

$$\det(s\mathbf{I}_m - HH') = s^{m-n} \det(s\mathbf{I}_n - H'H)$$

e portanto os polinômios característicos de $H'H$ e HH' diferem apenas em s^{m-n}. Como conclusão, $H'H$ e HH' têm os mesmos autovalores positivos mas podem ter um número diferente de autovalores iguais a zero (no máximo igual a $\bar{n} \triangleq \min\{n, m\}$).

Decomposição em Valores Singulares

Seja $H \in \mathbb{R}^{m\times n}$ e defina $M = H'H$ (simétrica, $n \times n$ e semidefinida positiva). Assim, todos os autovalores de M são reais e não negativos (zero ou positivos). Seja r o número de autovalores positivos (portanto, M tem rank r). Os autovalores de $M = H'H$ podem ser arranjados na forma $\lambda_1^2 \geq \lambda_2^2 \geq \cdots \geq \lambda_r^2 > 0 = \lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n$.

$\lambda_i^2,\ i = 1,2,\ldots,n$: autovalores de $H'H$ $\lambda_i,\ i = 1,2,\ldots,\bar{n}$: valores singulares de H

- Exemplo: $H = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 2 & 5 \end{bmatrix}$; $M = H'H = \begin{bmatrix} 2 & 0 & -2 \\ 0 & 8 & 16 \\ -2 & 16 & 34 \end{bmatrix}$

Autovalores de M: 41.6977, 2.3023, 0;
Valores singulares de H: $6.4574 = \sqrt{41.6977}$, $1.5173 = \sqrt{2.3023}$
Autovalores de HH': 41.6977, 2.3023;
Valores singulares de H': $6.4574 = \sqrt{41.6977}$, $1.5173 = \sqrt{2.3023}$

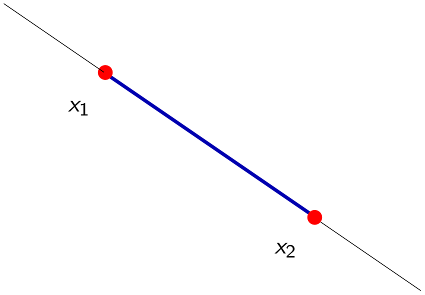
- os autovalores de $H'H$ diferem dos de HH' apenas no número de zeros, e H e H' têm os mesmos valores singulares.

Conjuntos Convexos

Suponha $x_1 \neq x_2 \in \mathbb{R}^n$. Pontos descritos por

$$\alpha x_1 + (1-\alpha)x_2 = x_2 + \alpha(x_1 - x_2)$$

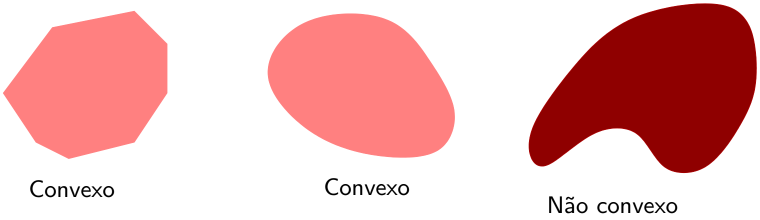
com $\alpha \in \mathbb{R}$ descrevem a reta que passa por x_1 e x_2. Para $0 \leq \alpha \leq 1$, a combinação descreve o segmento de reta entre x_1 e x_2.



Conjuntos Convexos

- Um conjunto $\mathscr{C}$ é **convexo** se o segmento de reta entre dois pontos quaisquer do conjunto estiver em $\mathscr{C}$, isto é, para quaisquer $x_1, x_2 \in \mathscr{C}$ e $\alpha \in [0,1]$,

$$\alpha x_1 + (1-\alpha)x_2 \in \mathscr{C}$$



- Um conjunto $\mathscr{C}$ é um **cone** se para todo $x \in \mathscr{C}$ e $\alpha \geq 0$, $\alpha x \in \mathscr{C}$.

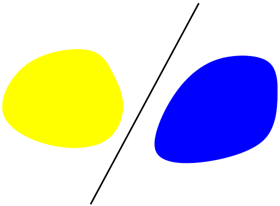
- Um cone convexo é um conjunto $\mathscr{C}$ tal que, para todo $x_1, x_2 \in \mathscr{C}$ e $\alpha_1, \alpha_2 \geq 0$,

$$\alpha_1 x_1 + \alpha_2 x_2 \in \mathscr{C}$$

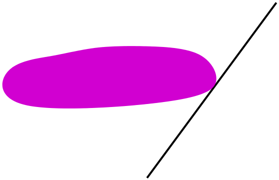
Hiperplanos

- Um hiperplano é um conjunto na forma $\{x \ : \ c'x = b\}$ para $c \neq 0 \in \mathbb{R}^n$ e $b \in \mathbb{R}$. Um hiperplano divide o $\mathbb{R}^n$ em dois semi-espaços, $\{x \ : \ c'x \leq b\}$ e $\{x \ : \ c'x \geq b\}$.

- Hiperplano separador

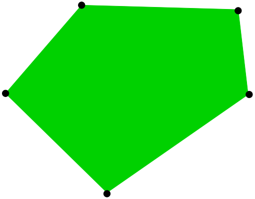


- Hiperplano suporte



Envelope convexo (Convex hull)

• O envelope convexo de um conjunto de N pontos $x_1, x_2, \ldots, x_N$ no $\mathbb{R}^2$ é um poliedro. Um ponto descrito como $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_N x_N$, com $\alpha_i \geq 0$, $i = 1, \ldots, N$ e $\alpha_1 + \alpha_2 + \cdots + \alpha_N = 1$, é chamado de combinação convexa dos pontos $x_1, x_2, \ldots, x_N$. Um conjunto é convexo se e somente se contiver toda combinação convexa de seus pontos.



• Matrizes simétricas $X = X' \in \mathbb{R}^{n \times n}$ definem um espaço vetorial de dimensão $n(n+1)/2$.

• O conjunto das matrizes simétricas semidefinidas positivas é um cone convexo, pois se $\alpha_1, \alpha_2 \geq 0$ e $X_1 > 0$, $X_2 > 0$, então $\alpha_1 X_1 + \alpha_2 X_2$ também é uma matriz simétrica semidefinida positiva.

Funções convexas

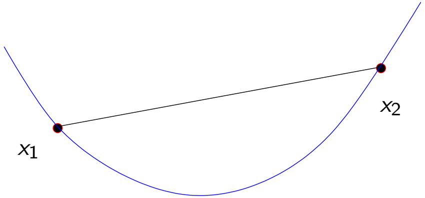
- Uma função $f : \mathbb{R}^n \to \mathbb{R}$ é **linear** se para todo $\alpha_1, \alpha_2 \in \mathbb{R}$ e x_1, x_2 pertencentes ao domínio da função,

$$f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2)$$

- Uma função $f : \mathbb{R}^n \to \mathbb{R}$ é **convexa** se seu domínio for um conjunto convexo e, para quaisquer x_1, x_2 pertencentes ao domínio da função e para $0 \leq \alpha \leq 1$

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$$

- O segmento de reta que une os pontos $(x_1, f(x_1))$ e $(x_2, f(x_2))$ está sempre acima ou sobre o gráfico da função.

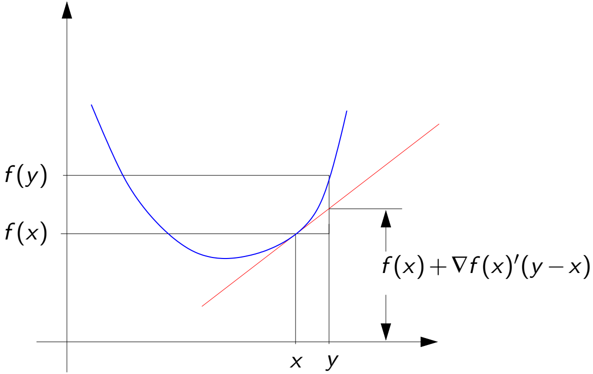


Funções convexas

- Suponha $f(\cdot)$ diferenciável. Então, f é convexa se e somente se seu domínio for um conjunto convexo e

$$f(y) \geq f(x) + \nabla f(x)'(y - x)$$

para todo x, y pertencentes ao domínio de f. $\nabla f(x)$ é o gradiente da função calculado no ponto x, e a função afim em y dada por $f(x) + \nabla f(x)'(y - x)$ é a aproximação de primeira ordem de Taylor da função f na vizinhança de x.



Funções convexas

- Para uma função convexa, a aproximação de primeira ordem está sempre abaixo do valor da função e, se a aproximação de primeira ordem estiver sempre abaixo do valor da função, a função é convexa.

- Se a função for convexa e $\nabla f(x) = 0$, então para todo y pertencente ao domínio da função tem-se $f(y) \geq f(x)$, e portanto x é um mínimo global de f.

- A função f é convexa se e somente se seu domínio for um conjunto convexo e

$$f(y) \geq f(x) + \mu(x)'(y - x)$$

para todo x, y pertencentes ao domínio de f e para todo $\mu(x)$ pertencente ao conjunto de subgradientes da função f calculados no ponto x. Se a função for diferenciável, $\mu(x) = \nabla(x)$ é único. Se f não for diferenciável em x, $\mu(x)$ é a inclinação de todos os hiperplanos suportes da função no ponto x.

- A função f é côncava se $-f$ for convexa. Se f é côncava e diferenciável em x, então $f(y) \leq f(x) + \nabla f(x)'(y - x)$

Complemento de Schur

- Considere a matriz simétrica X particionada

$$X = \begin{bmatrix} A & B \\ B' & C \end{bmatrix}$$

com $A = A' \in \mathbb{R}^{n\times n}$. Se $\det(A) \neq 0$, a matriz $C - B'A^{-1}B$ é o complemento de Schur de X em relação a A.

Note por exemplo que

$$X = \begin{bmatrix} A & B \\ B' & C \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ B'A^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & C - B'A^{-1}B \end{bmatrix} \begin{bmatrix} \mathbf{I} & A^{-1}B \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

e portanto $\det(X) = \det(A)\det(C - B'A^{-1}B)$.

Analogamente, se $\det(C) \neq 0$, $A - BC^{-1}B'$ é o complemento de Schur de X em relação a C e

$$X = \begin{bmatrix} A & B \\ B' & C \end{bmatrix} = \begin{bmatrix} \mathbf{I} & BC^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} A - BC^{-1}B' & \mathbf{0} \\ \mathbf{0} & C \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ C^{-1}B' & \mathbf{I} \end{bmatrix}$$

Complemento de Schur

- O complemento de Schur surge na solução de equações lineares, quando se elimina um bloco de variáveis. Por exemplo, considere o sistema

$$\begin{bmatrix} A & B \\ B' & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Assumindo $\det(A) \neq 0$, da primeira equação tem-se $Ax + By = b_1$, e portanto $x = -A^{-1}By + A^{-1}b_1$. Substituindo na segunda equação, tem-se $(C - B'A^{-1}B)y = b_2 - B'A^{-1}b_1$ e portanto

$$y = (C - B'A^{-1}B)^{-1}(b_2 - B'A^{-1}b_1) = -(C - B'A^{-1}B)^{-1}B'A^{-1}b_1 + (C - B'A^{-1}B)^{-1}b_2$$

Substituindo agora y na primeira equação, tem-se

$$x = (A^{-1} + A^{-1}B(C - B'A^{-1}B)^{-1}B'A^{-1})b_1 - A^{-1}B(C - B'A^{-1}B)^{-1}b_2$$

As expressões para x e y podem também ser obtidas da formula da inversa para uma matriz simétrica particionada em blocos:

$$X^{-1} = \begin{bmatrix} A & B \\ B' & C \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B(C - B'A^{-1}B)^{-1}B'A^{-1} & -A^{-1}B(C - B'A^{-1}B)^{-1} \\ -(C - B'A^{-1}B)^{-1}B'A^{-1} & (C - B'A^{-1}B)^{-1} \end{bmatrix}$$

- o complemento de Schur é a inversa da matriz do bloco (2,2) em X^{-1}

Complemento de Schur

- O complemento de Schur surge também na minimização de uma forma quadrática em relação a algumas das variáveis. Por exemplo,

$$\min_{x} \begin{bmatrix} x \\ y \end{bmatrix}' \begin{bmatrix} A & B \\ B' & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \min_{x}\; x'Ax + x'By + y'B'x + y'Cy = \min_{x}\; x'Ax + 2y'B'x + y'Cy$$

- Supondo $A > 0$ e impondo que a derivada parcial da função em relação a x vale zero, tem-se $2(Ax + By) = 0$, e portanto $x = -A^{-1}By$. Substituindo na função, tem-se

$$\min_{x} \begin{bmatrix} x \\ y \end{bmatrix}' \begin{bmatrix} A & B \\ B' & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = y'(C - B'A^{-1}B)y$$

- Caracterização da positividade de X

- $X > 0$ se e somente se $A > 0$ e $C - B'A^{-1}B > 0$

- Se $A > 0$, $X \geq 0$ se e somente se $C - B'A^{-1}B \geq 0$