Álgebra Linear e Multilinear

André L. F. de Almeida Tarcisio F. Maciel

Universidade Federal do Ceará
Programa de Pós-Graduação em Engenharia de Teleinformática

22/08/2017

Conteúdo

Conteúdo

Revisão de álgebra linear

- Diversos problemas em engenharia recaem em sistemas de equações lineares
- Exemplos
  - Sistemas de forças agindo sobre um corpo (leis de Newton)
  - Sistema de equações para os nós e malhas de um circuito elétrico (leis de Kirchhoff)

Conteúdo

Campo de um espaço vetorial [HJ12]

- O campo escalar subjacente a um espaço vetorial é o conjunto de escalares onde os elementos do vetor são definidos
  - Normalmente o campo dos números reais $\mathbb{R}$ ou complexos $\mathbb{C}$
  - Alternativamente poderia ser o dos racionais, inteiros módulo algum número primo, etc.
- Um campo precisa ser fechado sob duas operações binárias (i.e., que recebem dois operandos)
  - Adição
  - Multiplicação
- As operações precisam ser associativas, comutativas e possuir um elemento neutro
- Elementos inversos precisam existir para todos elemento sob a adição e multiplicação, exceto para a identidade sobre a multiplicação
- Multiplicação precisa ser distributiva sobre adição

Vetores e (sub)espaços vetoriais

- O espaço vetorial linear $\mathbb{R}^n$ é o conjunto de todos os vetores $\mathbf{x}$ de dimensão $n \times 1$ juntamente com as operações de adição de vetores e multiplicação por um escalar [Str88, cap. 2]

- Um **vetor** $\mathbf{x} \in \mathbb{R}^n$ é normalmente representado como

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \text{ com } \mathbf{x}^{\mathrm{T}} = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix} \quad (1)$$

- Embora a maior parte dos conceitos se estenda facilmente para vetores complexos, consideraremos apenas vetores reais (exceto se explicitamente mencionado)

Vetores e (sub)espaços vetoriais

- Há oito propriedades que precisam ser satisfeitas por um espaço vetorial[1] [Str88, Ex. 2.1.5]

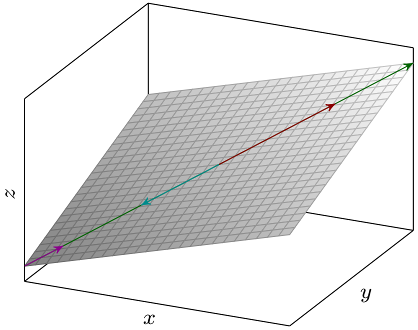
$$\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x} \quad \text{(Comutatividade da adição)} \tag{2a}$$
$$\mathbf{x}+(\mathbf{y}+\mathbf{z})=(\mathbf{x}+\mathbf{y})+\mathbf{z} \quad \text{(Associatividade da adição)} \tag{2b}$$
$$\exists\mathbf{0} \Rightarrow \mathbf{x}+\mathbf{0}=\mathbf{x} \quad \text{(Elemento neutro da adição)} \tag{2c}$$
$$\exists-\mathbf{x} \Rightarrow \mathbf{x}+(-\mathbf{x})=\mathbf{0} \quad \text{(Elemento simétrico)} \tag{2d}$$
$$1\cdot\mathbf{x}=\mathbf{x} \quad \text{(Elemento neutro da multiplicação)} \tag{2e}$$
$$(\alpha_1\cdot\alpha_2)\cdot\mathbf{x}=\alpha_1\cdot(\alpha_2\cdot\mathbf{x}) \quad \text{(Associatividade da multiplicação)} \tag{2f}$$
$$\alpha_1\cdot(\mathbf{x}+\mathbf{y})=\alpha_1\cdot\mathbf{x}+\alpha_1\cdot\mathbf{y} \quad \text{(Distributividade)} \tag{2g}$$
$$(\alpha_1+\alpha_2)\cdot\mathbf{x}=\alpha_1\cdot\mathbf{x}+\alpha_2\cdot\mathbf{x} \quad \text{(Distributividade)} \tag{2h}$$

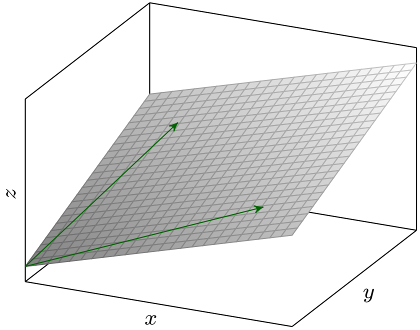
- Um subespaço vetorial linear $\mathbb{S}\subset\mathbb{R}^n$ é um conjunto não-vazio que satisfaz:
  - $\forall\mathbf{x},\mathbf{y}\in\mathbb{S}\Rightarrow\mathbf{x}+\mathbf{y}\in\mathbb{S}$
  - $\forall\mathbf{x}\in\mathbb{S}$ e $\alpha\in\mathbb{R}\Rightarrow\alpha\mathbf{x}\in\mathbb{S}$
- Um subespaço é portanto um subconjunto fechado sob as operações de adição e multiplicação por escalar

[1] Relações similares podem ser definidas para números complexos em $\mathbb{C}^n$

Vetores e (sub)espaços vetoriais

- Considerando o espaço vetorial $\mathbb{R}^3$, alguns exemplos simples de subespaços vetoriais $\mathbb{S} \subset \mathbb{R}^3$ são retas e planos
- Um vetor $\mathbf{v} \in \mathbb{R}^3$ age como suporte de uma reta que representa um subespaço $\mathbb{S}_1 \subset \mathbb{R}^3$
  - Quaisquer vetores $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{S}_1$ encontram-se sobre a reta e, portanto, podem ser escritos como $\mathbf{v}_1 = \alpha\mathbf{v}$ e $\mathbf{v}_2 = \beta\mathbf{v}$, para algum $\alpha, \beta \in \mathbb{R}$
  - Além disso, $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2 = (\alpha + \beta)\mathbf{v}$ está também sobre a reta





Conteúdo

Norma de vetores

- A **norma** de um vetor é uma generalização do conceito de comprimento ou magnitude de um vetor
- A **norma** $\|\mathbf{x}\|$ de $\mathbf{x} \in \mathbb{R}^n$ é uma função $f : \mathbb{R}^n \to \mathbb{R}$ que satisfaz [Che99, pág. 46]:

$$\|\mathbf{x}\| \geq 0 \quad \text{(Não-negatividade)} \tag{3a}$$

$$\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0} \quad \text{(Elemento neutro)} \tag{3b}$$

$$\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|, \forall\alpha \in \mathbb{R} \quad \text{(Escalabilidade)} \tag{3c}$$

$$\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \forall\mathbf{x},\mathbf{y} \in \mathbb{R}^n \quad \text{(Desigualdade triangular)} \tag{3d}$$

- Uma família de normas (norma-p) que atende às propriedades é dada por

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}, \quad p \in \mathbb{N}_+ \tag{4}$$

- Um caso caso de interesse é **norma-1** (ou norma ℓ_1) onde $p = 1$ em (4)
- Outro caso de interesse é **norma euclidiana** (ou norma-2 ou ainda norma ℓ_2) onde $p = 2$ em (4)

Norma euclidiana e produto interno [Rug96, cap. 1]

- A norma euclidiana atende ainda as seguintes propriedades:
  - $|\mathbf{x}^{\mathrm{T}}\mathbf{y}| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2$ (Desigualdade de Cauchy-Schwartz)
  - $\max\limits_{1\leq i\leq n} |x_i| \leq \|\mathbf{x}\|_2 \leq \sqrt{n} \max\limits_{1\leq i\leq n} |x_i|$
- O **produto interno** ou **produto escalar** $\langle \mathbf{x}, \mathbf{y} \rangle$ entre dois vetores $\mathbf{x}$ e $\mathbf{y} \in \mathbb{R}^n$ é escrito como

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^{\mathrm{T}}\mathbf{y} = \sum_{i=1}^{n} x_i y_i = \mathbf{y}^{\mathrm{T}}\mathbf{x} = \langle \mathbf{y}, \mathbf{x} \rangle \tag{5}$$

- A norma euclidiana guarda as seguintes relações com o produto interno

$$\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^{\mathrm{T}}\mathbf{x} = \sum_{i=1}^{n} x_i x_i = \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} |x_i|^2 = \|\mathbf{x}\|_2^2 \tag{6a}$$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^{\mathrm{T}}\mathbf{y} = \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \cos\theta_{\angle^{\mathbf{x}}_{\mathbf{y}}} \tag{6b}$$

- Pare vetores pertencentes a $\mathbb{C}$, $(\cdot)^{\mathrm{T}}$ é substituído por $(\cdot)^{\mathrm{H}}$ que representa o conjugado-transposto de um vetor com a conjugação denotada por $(\cdot)^*$

Conteúdo

Independência linear [Che99, cap. 3]

- Os vetores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in \mathbb{R}^n$ são ditos **linearmente independentes** (L.I.) se e somente se, para $\alpha_1, \alpha_2, \ldots, \alpha_m \in \mathbb{R}$,

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \ldots + \alpha_m\mathbf{x}_m = \mathbf{0} \Leftrightarrow \alpha_1 = \alpha_2 = \ldots = \alpha_m = 0, \tag{7}$$

  caso contrário $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ são ditos **linearmente dependentes** (L.D.)
- Se $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ são L.D., então existe pelo menos um $\alpha_i \neq 0$, tal que é possível escrever

$$\mathbf{x}_i = -\frac{1}{\alpha_i}\left[\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \ldots + \alpha_{i-1}\mathbf{x}_{i-1} + \alpha_{i+1}\mathbf{x}_{i+1} + \ldots + \alpha_m\mathbf{x}_m\right] \tag{8}$$

- A **dimensão** de um (sub)espaço vetorial linear é dado pelo máximo número de vetores L.I. neste (sub)espaço

Bases, representações e ortonormalização

- Um conjunto de n vetores L.I. pertencentes a um (sub)espaço vetorial é chamado uma **base** para esse (sub)espaço
- Dada uma base para um (sub)espaço vetorial, todo vetor desse subespaço pode ser escrito como uma combinação linear única dos vetores que formam a base
- Sejam $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\} \in \mathbb{R}^n$ um conjunto de vetores L.I. que formam uma base para $\mathbb{R}^n$
- Todo vetor $\mathbf{x} \in \mathbb{R}^n$ pode ser representado como

$$\mathbf{x} = \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \ldots + \alpha_n \mathbf{b}_n = \underbrace{\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_n \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}}_{\boldsymbol{\alpha}} = \mathbf{B}\boldsymbol{\alpha} \tag{9}$$

Bases, representações e ortonormalização

- O vetor $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{bmatrix}^{\mathrm{T}}$ é chamado de **representação** do vetor $\mathbf{x}$ na base $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ (ou ainda na base $\mathbf{B}$)
- A cada $\mathbb{R}^n$ é associada uma **base canônica** $\{\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_n\}$ onde $\mathbf{i}_i$ é a i-ésima coluna de uma matriz identidade $\mathbf{I}_n$ de dimensão $n \times n$
- Note que, na base canônica, um vetor $\mathbf{x} \in \mathbb{R}^n$ é representado como

$$\mathbf{x} = x_1\mathbf{i}_1 + x_2\mathbf{i}_2 + \ldots + x_n\mathbf{i}_n = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{10}$$

Vetores normalizados, ortogonais e ortonormais

- Um vetor $\mathbf{x}$ é dito **normalizado** se sua norma euclidiana $\|\mathbf{x}\|_2 = 1$
- Um vetor unitário $\mathbf{u}$ na direção de um vetor $\mathbf{x}$ é obtido normalizando o vetor $\mathbf{x}$ como

$$\mathbf{u}_{\mathbf{x}} = \frac{\mathbf{x}}{\|\mathbf{x}\|_2} \tag{11}$$

- Dois vetores $\mathbf{x}_1$ e $\mathbf{x}_2$ são ditos **ortogonais** se o produto interno entre eles é nulo, i.e., se $\mathbf{x}^{\mathrm{T}}{}_1\mathbf{x}_2 = \mathbf{x}^{\mathrm{T}}{}_2\mathbf{x}_1 = 0$
- Um conjunto de vetores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ é dito **ortonormal** se

$$\mathbf{x}^{\mathrm{T}}{}_i\mathbf{x}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{12}$$

- Um conjunto de vetores L.I. $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\} \in \mathbb{R}^n$ que atende a (12) formam uma **base ortonormal**

Componente ortogonal e paralela

- O comprimento $|P_{\|}(\mathbf{x}_i, \mathbf{x}_j)|$ da componente paralela $P_{\|}(\mathbf{x}_i, \mathbf{x}_j)$ de um vetor $\mathbf{x}_i$ na direção de um vetor $\mathbf{x}_j$ é dado pelo produto escalar do primeiro vetor com o vetor unitário na direção do segundo vetor, i.e.,

$$|P_{\|}(\mathbf{x}_i, \mathbf{x}_j)| = \frac{\mathbf{x}^{\mathrm{T}}{}_i \mathbf{x}_j}{\|\mathbf{x}_j\|_2} = \mathbf{x}^{\mathrm{T}}{}_i \frac{\mathbf{x}_j}{\|\mathbf{x}_j\|_2} = \mathbf{x}^{\mathrm{T}}{}_i \mathbf{u}_j, \text{ onde } \mathbf{u}_j = \frac{\mathbf{x}_j}{\|\mathbf{x}_j\|_2} \tag{13}$$

- Usando (13), a **componente paralela** $P_{\|}(\mathbf{x}_i, \mathbf{x}_j)$ e a **componente ortogonal** $P_{\perp}(\mathbf{x}_i, \mathbf{x}_j)$ de um vetor $\mathbf{x}_i$ em relação a um vetor $\mathbf{x}_j$ são dadas respectivamente por

$$P_{\|}(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}^{\mathrm{T}}{}_i \mathbf{u}_j)\mathbf{u}_j \tag{14a}$$

$$P_{\perp}(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i - P_{\|}(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i - (\mathbf{x}^{\mathrm{T}}{}_i \mathbf{u}_j)\mathbf{u}_j \tag{14b}$$

Processo de ortonormalização de Gram-Schmidt

- O **processo de ortonormalização de Gram-Schmidt** permite construir uma base ortonormal a partir de um conjunto de vetores L.I.
- Segundo esse processo, um conjunto $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\} \in \mathbb{R}^n$ de vetores L.I. produz a base ortonormal $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ como

$$
\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1, & \mathbf{b}_1 &= \mathbf{v}_1 / \|\mathbf{v}_1\|_2 \\
\mathbf{v}_2 &= \mathbf{x}_2 - (\mathbf{x}^{\mathrm{T}}{}_2\mathbf{b}_1)\mathbf{b}_1, & \mathbf{b}_2 &= \mathbf{v}_2 / \|\mathbf{v}_2\|_2 \\
\mathbf{v}_3 &= \mathbf{x}_3 - (\mathbf{x}^{\mathrm{T}}{}_3\mathbf{b}_1)\mathbf{b}_1 - (\mathbf{x}^{\mathrm{T}}{}_3\mathbf{b}_2)\mathbf{b}_2, & \mathbf{b}_3 &= \mathbf{v}_3 / \|\mathbf{v}_3\|_2 \\
\mathbf{v}_4 &= \mathbf{x}_4 - (\mathbf{x}^{\mathrm{T}}{}_4\mathbf{b}_1)\mathbf{b}_1 - (\mathbf{x}^{\mathrm{T}}{}_4\mathbf{b}_2)\mathbf{b}_2 - (\mathbf{x}^{\mathrm{T}}{}_4\mathbf{b}_3)\mathbf{b}_3, & \mathbf{b}_4 &= \mathbf{v}_4 / \|\mathbf{v}_4\|_2 \\
\ldots \\
\mathbf{v}_n &= \mathbf{x}_n - \sum_{i=1}^{n-1} (\mathbf{x}^{\mathrm{T}}{}_n\mathbf{b}_i)\mathbf{b}_i, & \mathbf{b}_n &= \mathbf{v}_n / \|\mathbf{v}_n\|_2
\end{aligned}
$$

Conteúdo

Matrizes

- Uma matriz $\mathbf{A}$ de dimensão $m \times n$ pertencente ao $\mathbb{R}^{m\times n}$ e sua transposta $\mathbf{A}^{\mathrm{T}} \in \mathbb{R}^{n\times m}$ são denotadas por

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \text{ e} \\ \mathbf{A}^{\mathrm{T}} &= \begin{bmatrix} a_{1,1} & a_{2,1} & \dots & a_{m,1} \\ a_{1,2} & a_{2,2} & \dots & a_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \dots & a_{m,n} \end{bmatrix} = \begin{bmatrix} \mathbf{a}^{\mathrm{T}}{}_1 \\ \mathbf{a}^{\mathrm{T}}{}_2 \\ \vdots \\ \mathbf{a}^{\mathrm{T}}{}_n \end{bmatrix},\end{aligned} \tag{15}$$

  onde $\mathbf{a}_i, 1 \leq i \leq n$ denota a i-ésima coluna de $\mathbf{A}$
- Assume-se aqui conhecimento sobre as operações de adição de matrizes, multiplicação por escalar, e multiplicação de matrizes

Algumas operações com matrizes

- Para uma matriz $\mathbf{A} \in \mathbb{C}^{m \times n}$, a **matrix conjugada** $\mathbf{A}^*$ é obtida conjugando cada elemento $a_{i,j}$ de $\mathbf{A}$
- De forma similar, a **matriz conjugada-transposta** $\mathbf{A}^{\mathrm{H}}$ de $\mathbf{A}$ é obtida conjugando a matriz transposta $\mathbf{A}^{\mathrm{T}}$ de $\mathbf{A}$, i.e., $\mathbf{A}^{\mathrm{H}} = \left(\mathbf{A}^{\mathrm{T}}\right)^*$
- A **matriz inversa** $\mathbf{A}^{-1}$ de uma matriz $\mathbf{A}$ de dimensão $n \times n$ é a matriz que satisfaz $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$
- Algumas propriedades relevantes envolvendo matrizes são [PP08]:

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \tag{16a}$$

$$(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} \tag{16b}$$

$$(\mathbf{AB})^{\mathrm{H}} = \mathbf{B}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}} \tag{16c}$$

$$\left(\mathbf{A}^{\mathrm{T}}\right)^{-1} = \left(\mathbf{A}^{-1}\right)^{\mathrm{T}} \tag{16d}$$

$$\left(\mathbf{A}^{\mathrm{H}}\right)^{-1} = \left(\mathbf{A}^{-1}\right)^{\mathrm{H}} \tag{16e}$$

$$(\mathbf{A} + \mathbf{B})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}} \tag{16f}$$

$$(\mathbf{A} + \mathbf{B})^{\mathrm{H}} = \mathbf{A}^{\mathrm{H}} + \mathbf{B}^{\mathrm{H}} \tag{16g}$$

Traço de uma matriz

- O **traço** $\mathrm{tr}\left(\mathbf{A}\right)$ de uma matriz $\mathbf{A}$ de dimensão $n \times n$ é definido como a soma dos elementos da diagonal da mesma, i.e.,

$$\mathrm{tr}\left(\mathbf{A}\right) = \sum_{i=1}^{n} a_{i,i} \tag{17}$$

- Entre as propriedades do $\mathrm{tr}\left(\cdot\right)$ temos [PP08]:

$$\mathrm{tr}\left(\mathbf{A}\right) = \mathrm{tr}\left(\mathbf{A}^{\mathrm{T}}\right) \tag{18a}$$

$$\mathrm{tr}\left(\mathbf{AB}\right) = \mathrm{tr}\left(\mathbf{BA}\right) \tag{18b}$$

$$\mathrm{tr}\left(\mathbf{ABC}\right) = \mathrm{tr}\left(\mathbf{CAB}\right) = \mathrm{tr}\left(\mathbf{BCA}\right) \tag{18c}$$

$$\mathrm{tr}\left(\alpha\mathbf{A} + \beta\mathbf{B}\right) = \alpha\,\mathrm{tr}\left(\mathbf{A}\right) + \beta\,\mathrm{tr}\left(\mathbf{B}\right) \tag{18d}$$

Determinante uma matriz

- O **determinante** $\det\left(\mathbf{A}\right)$ de uma matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$ pode ser escrito através da expansão de Laplace sobre uma linha ou uma coluna da matriz como

$$\det\left(\mathbf{A}\right) = \sum_{i=1}^{n} a_{i,j} c_{i,j} = \sum_{j=1}^{n} a_{i,j} c_{i,j}, \tag{19}$$

  onde $c_{i,j}$ é o **cofator** associado a $a_{i,j}$, o qual é dado por

$$c_{i,j} = (-1)^{i+j} \det\left(\tilde{\mathbf{A}}_{i,j}\right), \tag{20}$$

  onde a matriz $\tilde{\mathbf{A}}_{i,j} \in \mathbb{R}^{(n-1)\times(n-1)}$ é formada a partir de $\mathbf{A}$ pela exclusão de sua i-ésima linha e j-ésima coluna
- A **matriz adjunta** $\operatorname{adj}\left(\mathbf{A}\right)$ de $\mathbf{A}$ é a matriz transposta dos cofatores de $\mathbf{A}$
- Uma matriz $\mathbf{A}$ é dita **não-singular** e possui inversa $\mathbf{A}^{-1}$ se $\det\left(\mathbf{A}\right) \neq 0$
- A matriz inversa pode ser calculada como

$$\mathbf{A}^{-1} = \frac{1}{\det\left(\mathbf{A}\right)} \operatorname{adj}\left(\mathbf{A}\right) \tag{21}$$

Exemplo para inversa de uma matriz

- Se a matriz $\mathbf{A}$ é dada por

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$$

  temos que

$$\begin{aligned} \mathbf{A}^{-1} &= \frac{1}{\det(\mathbf{A})} \operatorname{adj}(\mathbf{A}) = \frac{1}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}} \begin{bmatrix} c_{1,1} & c_{2,1} \\ c_{1,2} & c_{2,2} \end{bmatrix} \\ &= \frac{1}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}} \begin{bmatrix} a_{2,2} & -a_{1,2} \\ -a_{2,1} & a_{1,1} \end{bmatrix} \end{aligned}$$

Conteúdo

Transformação linear

- Um função $f(\cdot) : \mathbb{R}^n \to \mathbb{R}^m$ é um **operador linear** se e somente se

$$f(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) = \alpha_1 f(\mathbf{x}_1) + \alpha_2 f(\mathbf{x}_2), \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n, \alpha_1, \alpha_2 \in \mathbb{R} \tag{22}$$

- Sejam $\mathbb{X} \subset \mathbb{R}^n$ e $\mathbb{Y} \subset \mathbb{R}^m$ são dois espaços vetoriais e sejam:
  - $L(\cdot) : \mathbb{X} \to \mathbb{Y}$ um operador linear
  - $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ um conjunto de n vetores L.I. em $\mathbb{X}$
  - $\mathbf{y}_i = L(\mathbf{x}_i), i = 1, 2, \ldots, n$ um conjunto de n vetores em $\mathbb{Y}$
- Então, podemos afirmar que:
  - O operador linear $L(\cdot)$ é unicamente determinado pelos n pares $(\mathbf{x}_i, \mathbf{y}_i), i = 1, 2, \ldots, n$
  - Se $\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ e $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m\}$ são bases para $\mathbb{X}$ e $\mathbb{Y}$, respectivamente, então o operador linear $L(\cdot)$ pode ser representado por uma matriz $\mathbf{T}$ de dimensão $m \times n$
  - A i-ésima coluna da matriz $\mathbf{T}$ é a representação de $\mathbf{y}_i$ na base $\mathbf{P} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \ldots & \mathbf{p}_n \end{bmatrix}$

Transformações lineares: mudança de base

- Exemplos de transformações lineares comuns no $\mathbb{R}^2$
  - **Mudança de escala**: $\mathbf{T} = \alpha\mathbf{I}$ (mesma escala nos eixos x e y) ou $\mathbf{T} = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix}$ (escalas diferentes para os eixos x e y)
  - **Rotação**: $\mathbf{T} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$
- Outra transformação linear de interesse refere-se à mudança de base em um espaço vetorial
- Se $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$ são respectivamente as representações de $\mathbf{x} \in \mathbb{X} \subset \mathbb{R}^n$ nas bases $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix}$ e $\mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_n \end{bmatrix}$, temos

$$\mathbf{x} = \underbrace{\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}}_{\boldsymbol{\alpha}} = \underbrace{\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_n \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}}_{\boldsymbol{\beta}} \tag{23}$$

Transformações lineares: mudança de base

- Se $\mathbf{p}_i$ é a representação de $\mathbf{a}_i$ na base $\mathbf{B}$, então temos que

$$\mathbf{a}_i = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_n \end{bmatrix} \mathbf{p}_i = \mathbf{B}\mathbf{p}_i, \quad i = 1, 2, \dots, n \tag{24}$$

- Juntando as n equações acima em expressão matricial obtemos

$$\mathbf{A} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_n \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \dots & \mathbf{p}_n \end{bmatrix} = \mathbf{B}\mathbf{P} \tag{25}$$

- Substituindo (25) em (23), obtemos

$$\mathbf{A}\boldsymbol{\alpha} = \mathbf{B}\boldsymbol{\beta} \Rightarrow \mathbf{B}\mathbf{P}\boldsymbol{\alpha} = \mathbf{B}\boldsymbol{\beta} \Rightarrow \mathbf{B}^{-1}\mathbf{B}\mathbf{P}\boldsymbol{\alpha} = \mathbf{B}^{-1}\mathbf{B}\boldsymbol{\beta} \Rightarrow \boxed{\mathbf{P}\boldsymbol{\alpha} = \boldsymbol{\beta}}$$

- Logo, temos que $\mathbf{P}$ é a transformação linear que leva a representação $\boldsymbol{\alpha}$ de $\mathbf{x}$ na base $\mathbf{A}$ para sua representação $\boldsymbol{\beta}$ na base $\mathbf{B}$

Transformações lineares: mudança de base

- Um desenvolvimento análogo ao anterior provê a transformação linear $\mathbf{Q}$ que leva a representação $\boldsymbol{\beta}$ de $\mathbf{x}$ na base $\mathbf{B}$ para sua representação $\boldsymbol{\alpha}$ na base $\mathbf{A}$
- Se $\mathbf{P}$ e $\mathbf{Q}$ são conhecidas, então para um vetor $\mathbf{x}$ qualquer com representações $\boldsymbol{\alpha}$ na base $\mathbf{A}$ e $\boldsymbol{\beta}$ na base $\mathbf{B}$ temos

$$\boldsymbol{\beta} = \mathbf{P}\boldsymbol{\alpha} \text{ e } \boldsymbol{\alpha} = \mathbf{Q}\boldsymbol{\beta} \Rightarrow \boldsymbol{\beta} = \mathbf{P}\mathbf{Q}\boldsymbol{\beta} \Rightarrow \mathbf{Q}\mathbf{P} = \mathbf{I} \Rightarrow \boxed{\mathbf{P}^{-1} = \mathbf{Q}}$$

Transformações lineares: transformações de similaridade

- Considere que:
  - $\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_n \end{bmatrix}$ e $\bar{\mathbf{U}} = \begin{bmatrix} \bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2, \ldots, \bar{\mathbf{u}}_n \end{bmatrix}$ são duas bases para um subespaço vetorial $\mathbb{X}$
  - $L(\cdot)$ é um operador linear tal que $\mathbf{y}_i = L(\mathbf{u}_i)$ e $\bar{\mathbf{y}}_i = L(\bar{\mathbf{u}}_i), i = 1, 2, \ldots, n$
  - $\mathbf{V} = \begin{bmatrix} \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \end{bmatrix}$ e $\bar{\mathbf{V}} = \begin{bmatrix} \bar{\mathbf{v}}_1, \bar{\mathbf{v}}_2, \ldots, \bar{\mathbf{v}}_n \end{bmatrix}$ são duas uma base para o espaço gerado por $\mathbf{y}_i$
  - $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \end{bmatrix}$ e $\bar{\mathbf{A}} = \begin{bmatrix} \bar{\mathbf{a}}_1, \bar{\mathbf{a}}_2, \ldots, \bar{\mathbf{a}}_n \end{bmatrix}$ são as representação de $\mathbf{y}_i$ e $\bar{\mathbf{y}}_i$ nas bases $\mathbf{V}$ e $\bar{\mathbf{V}}$, respectivamente
- Como $\mathbf{U}$ e $\bar{\mathbf{U}}$ são base para $\mathbb{X}$, um vetor $\mathbf{x} \in \mathbb{X}$ pode ser escrito como combinação linear das colunas de $\mathbf{U}$ ou de $\bar{\mathbf{U}}$ e como $L(\mathbf{x})$ é um operador linear temos

$$L(\mathbf{x}) = L\left(\sum_{i=1}^{n} \alpha_i \mathbf{u}_i\right) = \sum_{i=1}^{n} \alpha_i L\left(\mathbf{u}_i\right) = \sum_{i=1}^{n} \alpha_i \mathbf{y}_i \tag{26a}$$

$$L(\mathbf{x}) = L\left(\sum_{i=1}^{n} \bar{\alpha}_i \bar{\mathbf{u}}_i\right) = \sum_{i=1}^{n} \bar{\alpha}_i L\left(\bar{\mathbf{u}}_i\right) = \sum_{i=1}^{n} \bar{\alpha}_i \bar{\mathbf{y}}_i \tag{26b}$$

Transformações lineares: transformações de similaridade

- Podemos escrever ainda que

$$
\begin{aligned}
L\left(\begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_n\end{bmatrix}\right) &= \begin{bmatrix}\mathbf{y}_1 & \mathbf{y}_2 & \dots & \mathbf{y}_n\end{bmatrix} \\
&= \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n\end{bmatrix}\begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n\end{bmatrix} = \mathbf{VA} \qquad (27a) \\
L\left(\begin{bmatrix}\bar{\mathbf{u}}_1 & \bar{\mathbf{u}}_2 & \dots & \bar{\mathbf{u}}_n\end{bmatrix}\right) &= \begin{bmatrix}\bar{\mathbf{y}}_1 & \bar{\mathbf{y}}_2 & \dots & \bar{\mathbf{y}}_n\end{bmatrix} \\
&= \begin{bmatrix}\bar{\mathbf{v}}_1 & \bar{\mathbf{v}}_2 & \dots & \bar{\mathbf{v}}_n\end{bmatrix}\begin{bmatrix}\bar{\mathbf{a}}_1 & \bar{\mathbf{a}}_2 & \dots & \bar{\mathbf{a}}_n\end{bmatrix} = \bar{\mathbf{V}}\bar{\mathbf{A}} \qquad (27b)
\end{aligned}
$$

- Usando (26) e (27), para um vetor $\mathbf{x}$ com representação $\boldsymbol{\alpha}$ na base $\mathbf{U}$ e um vetor $\mathbf{y} = L(\mathbf{x})$ com representação $\boldsymbol{\beta}$ na base $\mathbf{V}$ temos

$$
\begin{aligned}
&\mathbf{y} = L(\mathbf{x}) \Rightarrow \mathbf{V}\boldsymbol{\beta} = L(\mathbf{U}\boldsymbol{\alpha}) \Rightarrow \mathbf{V}\boldsymbol{\beta} = \sum_{i=1}^{n} \alpha_i L\left(\mathbf{u}_i\right) \Rightarrow \mathbf{V}\boldsymbol{\beta} = \sum_{i=1}^{n} \alpha_i \mathbf{y}_i \\
&\Rightarrow \mathbf{V}\boldsymbol{\beta} = \mathbf{VA}\boldsymbol{\alpha} \Rightarrow \boxed{\boldsymbol{\beta} = \mathbf{A}\boldsymbol{\alpha}}
\end{aligned} \qquad (28)
$$

- Analogamente para um vetor $\mathbf{x}$ com representação $\bar{\boldsymbol{\alpha}}$ na base $\bar{\mathbf{U}}$ e um vetor $\bar{\mathbf{y}} = L(\bar{\mathbf{x}})$ com representação $\bar{\boldsymbol{\beta}}$ na base $\bar{\mathbf{V}}$ obtemos

$$
\boxed{\bar{\boldsymbol{\beta}} = \bar{\mathbf{A}}\bar{\boldsymbol{\alpha}}} \qquad (29)
$$

Transformações lineares: transformações de similaridade

- Seja $\mathbf{P}$ a transformação linear que leva a representação $\boldsymbol{\alpha}$ de $\mathbf{x}$ da base $\mathbf{U}$ para $\bar{\boldsymbol{\alpha}}$ na base $\bar{\mathbf{U}}$ e seja $\mathbf{Q} = \mathbf{P}^{-1}$ a transformação linear que traz a representação $\boldsymbol{\alpha}$ de $\mathbf{x}$ na base $\bar{\mathbf{U}}$ de volta para a base $\mathbf{U}$
- Como há um representação única para o operador linear $L(\mathbf{x})$ e $\mathbf{y} = L(\mathbf{x})$, as transformações $\mathbf{P}$ e $\mathbf{Q} = \mathbf{P}^{-1}$ também realizam as mudança de base de $\mathbf{V}$ para $\bar{\mathbf{V}}$ e vice-versa para um vetor $\mathbf{y}$

$$\mathbf{x} \xrightarrow{\;L\;} \mathbf{y} = L(\mathbf{x})$$

$$\begin{tikzcd}
\mathbf{U}\boldsymbol{\alpha} \arrow[r, "\mathbf{A}"] \arrow[d, bend right, "\mathbf{P}"'] & \boldsymbol{\beta} = \mathbf{A}\boldsymbol{\alpha} \arrow[d, bend right, "\mathbf{P}"'] \\
\bar{\mathbf{U}}\bar{\boldsymbol{\alpha}} \arrow[r, "\bar{\mathbf{A}}"] \arrow[u, bend right, "\mathbf{P}^{-1}"'] & \bar{\boldsymbol{\beta}} = \bar{\mathbf{A}}\bar{\boldsymbol{\alpha}} \arrow[u, bend right, "\mathbf{P}^{-1}"']
\end{tikzcd}$$

- Com o auxílio do diagrama ao lado e usando (28) e (29) temos que

$$\begin{aligned}
\bar{\boldsymbol{\beta}} = \bar{\mathbf{A}}\bar{\boldsymbol{\alpha}} &\Rightarrow \bar{\mathbf{A}}\bar{\boldsymbol{\alpha}} = \mathbf{P}\boldsymbol{\beta} \\
&\Rightarrow \bar{\mathbf{A}}\bar{\boldsymbol{\alpha}} = \mathbf{P}\mathbf{A}\boldsymbol{\alpha} \\
&\Rightarrow \bar{\mathbf{A}}\bar{\boldsymbol{\alpha}} = \mathbf{P}\mathbf{A}\mathbf{P}^{-1}\bar{\boldsymbol{\alpha}} \\
&\Rightarrow \boxed{\bar{\mathbf{A}} = \mathbf{P}\mathbf{A}\mathbf{P}^{-1}} \\
&\Rightarrow \boxed{\mathbf{A} = \mathbf{P}^{-1}\bar{\mathbf{A}}\mathbf{P}}
\end{aligned} \tag{30}$$

Transformações lineares: transformações de similaridade

- As transformações em $\mathbf{P}\bar{\mathbf{A}}\mathbf{P}^{-1}$ e $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ mostradas em (30) são chamadas **transformações de similaridade**
- Matrizes $\mathbf{A}$ e $\bar{\mathbf{A}}$ que se relacionam conforme (30) são ditas **matrizes similares**
- Em particular, as matrizes $\mathbf{A}$ e $\bar{\mathbf{A}}$ são representações de um mesmo operador linear $L(\cdot)$ em duas bases distintas
- Todas as representações de um mesmo operador linear são similares
- Como um operador linear em uma dada base pode ser representado por uma matriz $\mathbf{A}$, essa matriz pode ser vista como o operador linear propriamente dito
- Sendo o operador linear $L(\cdot) : \mathbb{R}^n \to \mathbb{R}^n$ descrito pela matriz $\mathbf{A} \in \mathbb{R}^{n\times n}$, temos que

$$\begin{aligned} \mathbf{y} = L(\mathbf{x}) = \mathbf{A}\left(\sum_{i=1}^{n} \alpha_i \mathbf{u}_i\right) = \sum_{i=1}^{n} \alpha_i \mathbf{A}\mathbf{u}_i = \sum_{i=1}^{n} \alpha_i \mathbf{y}_i \\ \Rightarrow \mathbf{y}_i = \mathbf{A}\mathbf{u}_i, i = 1, 2, \ldots, n \end{aligned} \tag{31}$$

Conteúdo

Sistema de equações lineares

- Considere um sistema de m equações lineares e n variáveis $x_1, x_2, \ldots, x_n$,

$$\begin{aligned} a_{1,1} \cdot x_1 + a_{1,2} \cdot x_2 + \ldots + a_{1,n} \cdot x_n &= y_1 \\ a_{2,1} \cdot x_1 + a_{2,2} \cdot x_2 + \ldots + a_{2,n} \cdot x_n &= y_2 \\ \ldots \\ a_{m,1} \cdot x_1 + a_{m,2} \cdot x_2 + \ldots + a_{m,n} \cdot x_n &= y_m \end{aligned} \tag{32}$$

  onde os coeficientes $a_{i,j}$ e y_i são dados e $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, n$
- Usando notação matricial, podemos definir

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \ldots & a_{m,n} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{e} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, \tag{33}$$

  e reescrever (32) como

$$\mathbf{A}\mathbf{x} = \mathbf{y} \tag{34}$$

Posto de uma matriz

- O **posto** ou **rank** de uma matriz $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix}$ é denotado por $\operatorname{rank}(\mathbf{A})$ e pode ser definido como o número de colunas $\mathbf{a}_i$ de $\mathbf{A}$ que são L.I.
- O posto de uma matriz corresponde portanto à dimensão do espaço vetorial gerado pelas colunas da matriz
- Dada uma matriz $\mathbf{A}$ de dimensão $m \times n$, temos que

$$\operatorname{rank}(\mathbf{A}) = \operatorname{rank}\left(\mathbf{A}^{\mathrm{T}}\right) \Rightarrow \operatorname{rank}(\mathbf{A}) \leq \min\{m, n\} \tag{35}$$

  e portanto os espaço vetoriais gerados pelas colunas e pelas linhas de $\mathbf{A}$ têm a mesma dimensão
- Para $\mathbf{A}$ com dimensão $m \times n$ e $\mathbf{B}$ com dimensão $n \times p$, a **desigualdade de Sylvester** estabelece que

$$\operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{B}) - n \leq \operatorname{rank}(\mathbf{AB}) \leq \min\{\operatorname{rank}(\mathbf{A}), \operatorname{rank}(\mathbf{B})\} \tag{36}$$

Subespaços fundamentais: espaço coluna

- O sistema (32) possui solução se $\mathbf{y}$ pode ser escrito como combinação linear das colunas de $\mathbf{A}$
- Nesse caso, cada vetor $\mathbf{x}$ (se existir) que satisfaz (32) é uma representação de $\mathbf{y}$ no subespaço gerado pelas colunas de $\mathbf{A}$
- Considere que $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_r\}$ formam uma base para o subespaço gerado pelas colunas de $\mathbf{A}$, onde $r = \operatorname{rank}(\mathbf{A})$
- O sistema (32) possui solução se e somente se $\mathbf{y}$ pertence ao subespaço $\mathcal{R}(\mathbf{A})$ gerado pela base $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_r\}$, ou seja, o subespaço gerado pelas colunas de $\mathbf{A}$
- O subespaço $\mathcal{R}(\mathbf{A})$ de $\mathbf{A} \in \mathbb{R}^{m\times n}$ é chamado **espaço coluna** ou **espaço *range*** de $\mathbf{A}$

$$\mathcal{R}(\mathbf{A}) \triangleq \{\mathbf{y} = \mathbf{A}\mathbf{x} | \mathbf{x} \in \mathbb{R}^n\} \subset \mathbb{R}^m \tag{37}$$

- O $\operatorname{rank}(\mathbf{A})$ é a dimensão do espaço coluna $\mathcal{R}(\mathbf{A})$
- Um matriz $\mathbf{A}$ possui (pseudo-)inversa quando o seu $\operatorname{rank}(\mathbf{A})$ é máximo, i.e., $\operatorname{rank}(\mathbf{A}) = \min\{m, n\}$
- Para $\mathbf{A} \in \mathbb{R}^{m\times p}$ e $\mathbf{B} \in \mathbb{R}^{m\times q}$, temos que

$$\mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B}) = \mathcal{R}\left(\begin{bmatrix}\mathbf{A} & \mathbf{B}\end{bmatrix}\right) \tag{38}$$

Subespaços fundamentais: espaço nulo

- Quando $\mathbf{y} = \mathbf{0}$ em (32) temos $\mathbf{Ax} = \mathbf{0}$
- O conjunto de soluções de $\mathbf{Ax} = \mathbf{0}$ define por si só um subespaço vetorial: o **espaço nulo** $\mathcal{N}(\mathbf{A})$ de $\mathbf{A}$

$$\mathcal{N}(\mathbf{A}) \triangleq \{\mathbf{x} \in \mathbb{R}^n | \mathbf{Ax} = 0\} \subset \mathbb{R}^n \tag{39}$$

- A dimensão do espaço nulo $\mathcal{N}(\mathbf{A})$ é chamada de **nulidade** de $\mathbf{A}$ e é denotada por $\mathrm{nullity}(\mathbf{A})$
- Para $\mathbf{A} \in \mathbb{R}^{p\times n}$ e $\mathbf{B} \in \mathbb{R}^{q\times n}$, temos que

$$\mathcal{N}(\mathbf{A}) + \mathcal{N}(\mathbf{B}) = \mathcal{N}\left(\begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix}\right) \tag{40}$$

- Note que o espaço coluna e o espaço nulo de uma matriz $\mathbf{A}$ são duais tal que o $\mathrm{rank}(\mathbf{A})$ é o dual da $\mathrm{nullity}(\mathbf{A})$
- De fato, para $\mathbf{A}$ com dimensão $m \times n$ temos que

$$\mathrm{rank}(\mathbf{A}) + \mathrm{nullity}(\mathbf{A}) = n \qquad \text{(Teorema do posto-nulidade)} \tag{41}$$

Subespaços fundamentais: espaço linha e espaço nulo à esquerda

- Associados a uma matriz $\mathbf{A}$ e seus subespaços $\mathcal{R}(\mathbf{A})$ e $\mathcal{N}(\mathbf{A})$ temos ainda dois outros subespaços:
  - O **espaço linha** de $\mathbf{A}$, denotado por $\mathcal{R}(\mathbf{A}^{\mathrm{T}})$, que corresponde ao espaço coluna de $\mathbf{A}^{\mathrm{T}}$ e é gerado pelas linhas de $\mathbf{A}$
  - O **espaço nulo à esquerda** de $\mathbf{A}$, denotado $\mathcal{N}(\mathbf{A}^{\mathrm{T}})$, que corresponde ao espaço nulo de $\mathbf{A}^{\mathrm{T}}$ e é o espaço que contém todos os vetores $\mathbf{z}$ que satisfazem $\mathbf{A}^{\mathrm{T}}\mathbf{z} = 0$
- De maneira similar à anterior, temos para $\mathbf{A}$ com dimensão $m \times n$ que

$$\operatorname{rank}(\mathbf{A}^{\mathrm{T}}) + \operatorname{nullity}(\mathbf{A}^{\mathrm{T}}) = m \qquad \text{(Teorema do posto-nulidade)} \tag{42}$$

- Além disso, temos ainda de acordo com (35) que

$$\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{A}^{\mathrm{T}}) = r \leq \min\{m, n\} \tag{43}$$

Caracterização do posto de uma matriz [HJ12]

- Na caracterização do posto de uma matriz $\mathbf{A} \in \mathbb{R}^{m \times n}$ são equivalentes
  1. $\operatorname{rank}(\mathbf{A}) = r$
  2. r, e não mais que r, linhas de $\mathbf{A}$ são linearmente independentes
  3. r, e não mais que r, colunas de $\mathbf{A}$ são linearmente independentes
  4. Alguma submatriz $r \times r$ de $\mathbf{A}$ tem determinante não-nulo e toda submatriz $(r+1) \times (r+1)$ de $\mathbf{A}$ tem determinante nulo
  5. $\dim(\mathcal{R}(\mathbf{A})) = r$
  6. Há r, e não mais que r, vetores $\mathbf{b}_1, \ldots, \mathbf{b}_r$ tais que o sistema linear $\mathbf{A}\mathbf{x} = \mathbf{b}_j$ é consistente para $j = 1, \ldots, r$
  7. $r = n - \dim(\mathcal{N}(\mathbf{A}))$ (teorema do posto-nulidade)
  8. $r = \min\left\{p : \mathbf{A} = \mathbf{X}\mathbf{Y}^{\mathrm{T}}\right\}$ para algum $\mathbf{X} \in \mathbb{R}^{m \times p}$, $\mathbf{Y} \in \mathbb{R}^{n \times p}$
  9. $r = \min\left\{p : \mathbf{A} = \mathbf{x}_1\mathbf{y}^{\mathrm{T}}{}_1 + \ldots + \mathbf{x}_p\mathbf{y}^{\mathrm{T}}{}_p\right\}$ para algum $\mathbf{x}_1, \ldots, \mathbf{x}_p \in \mathbb{R}^m$, $\mathbf{y}_1, \ldots, \mathbf{y}_p \in \mathbb{R}^n$

Desigualdades do posto de uma matriz [HJ12]

1. Se $\mathbf{A} \in \mathbb{R}^{m \times n}$, então $\operatorname{rank}(\mathbf{A}) \leq \min\{m, n\}$
2. Desigualdade de Sylvester: se $\mathbf{A} \in \mathbb{R}^{m \times k}$ e $\mathbf{B} \in \mathbb{R}^{k \times n}$, então $(\operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{B})) - k \leq \operatorname{rank}(\mathbf{AB}) \leq \min\{\operatorname{rank}(\mathbf{A}), \operatorname{rank}(\mathbf{B})\}$ ()
3. Desigualdade soma-posto: se $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times k}$ então $|\operatorname{rank}(\mathbf{A}) - \operatorname{rank}(\mathbf{B})| \leq \operatorname{rank}(\mathbf{A} + \mathbf{B}) \leq \operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{B})$ com igualdade na segunda desigualdade se e somente se $\mathcal{R}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B}) = \emptyset$
4. Desigualdade de Frobenius: se $\mathbf{A} \in \mathbb{R}^{m \times k}$, $\mathbf{B} \in \mathbb{R}^{k \times p}$ e $\mathbf{C} \in \mathbb{R}^{p \times n}$, então $\operatorname{rank}(\mathbf{AB}) + \operatorname{rank}(\mathbf{BC}) \leq \operatorname{rank}(\mathbf{B}) + \operatorname{rank}(\mathbf{ABC})$ com igualdade se e somente se $\exists \mathbf{X}, \mathbf{Y} : \mathbf{B} = \mathbf{BCX} + \mathbf{YAB}$

Igualdades do posto de uma matriz [HJ12]

1. Se $\mathbf{A} \in \mathbb{C}^{m\times n}$, então $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}\left(\mathbf{A}^{\mathrm{T}}\right) = \operatorname{rank}(\mathbf{A}^{*}) = \operatorname{rank}\left(\mathbf{A}^{\mathrm{H}}\right)$
2. Se $\mathbf{A} \in \mathbb{R}^{m\times m}$ e $\mathbf{C} \in \mathbb{R}^{n\times n}$ são matrizes não-singulares e $\mathbf{B} \in \mathbb{R}^{m\times n}$, então $\operatorname{rank}(\mathbf{AB}) = \operatorname{rank}(\mathbf{B}) = \operatorname{rank}(\mathbf{BC}) = \operatorname{rank}(\mathbf{ABC})$, ou seja, multiplicações à direita ou esquerda por matrizes não-singulares não afetam o posto
3. Se $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ então $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{B})$ se e somente se $\exists \mathbf{X} \in \mathbb{R}^{m\times m}, \mathbf{Y} \in \mathbb{R}^{n\times n}$ não-singulares tais que $\mathbf{B} = \mathbf{XAY}$
4. Se $\mathbf{A} \in \mathbb{C}^{m\times n}$ então $\operatorname{rank}\left(\mathbf{A}^{\mathrm{H}}\mathbf{A}\right) = \operatorname{rank}(\mathbf{A})$
5. Fatorização de posto completo: se $\mathbf{A} \in \mathbb{R}^{m\times n}$, então $\operatorname{rank}(\mathbf{A}) = k$ se e somente se $\mathbf{A} = \mathbf{X}\mathbf{Y}^{\mathrm{T}}$ onde $\mathbf{X} \in \mathbb{R}^{m\times k}$ e $\mathbf{Y} \in \mathbb{R}^{n\times k}$ têm colunas independentes cada
6. Se $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{X} \in \mathbb{R}^{n\times k}$, e $\mathbf{Y} \in \mathbb{R}^{m\times k}$ e $\mathbf{W} = \mathbf{Y}^{\mathrm{T}}\mathbf{AX}$ é não-singular, então $\operatorname{rank}\left(\mathbf{A} - \mathbf{AXW}^{-1}\mathbf{Y}^{\mathrm{T}}\mathbf{A}\right) = \operatorname{rank}(\mathbf{A}) - \operatorname{rank}\left(\mathbf{AXW}^{-1}\mathbf{Y}^{\mathrm{T}}\mathbf{A}\right)$

Conteúdo

Definições básicas

- Sejam uma matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$, um vetor $\mathbf{v} \in \mathbb{R}^{n}$ e um escalar $\lambda \in \mathbb{R}$ tais que

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \tag{44}$$

- A transformação linear expressa por $\mathbf{A}$ aplicada ao vetor $\mathbf{v}$ resulta em uma versão escalonada de $\mathbf{v}$, i.e., $\lambda\mathbf{v}$
- Nesse caso, o escalar λ é denominado um autovalor de $\mathbf{A}$ e $\mathbf{v}$ é chamado o autovetor de $\mathbf{A}$ associado ao autovalor λ
- Autovalores e autovetores encontram diversas aplicações em engenharia
  - Esforços e direções principais em mecânica dos materiais
  - Estudos de momentos de inércia
  - Frequências naturais e modos de vibração
  - Alocação ótima de potência em comunicações co-canal
  - Formação de feixe em sistemas de comunicação com antenas inteligentes

Equação característica

- Sendo $\mathbf{I}$ uma matriz identidade e $\mathbf{0}$ um vetor de zeros, reescrevemos (44) como

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0} \tag{45}$$

- Se a matriz $(\mathbf{A} - \lambda\mathbf{I})$ não é singular ➜ $\exists\ (\mathbf{A} - \lambda\mathbf{I})^{-1}$ ➜ $\mathbf{v} = (\mathbf{A} - \lambda\mathbf{I})^{-1}\mathbf{0}$ (solução trivial)
- Se a matriz $(\mathbf{A} - \lambda\mathbf{I})$ é singular, temos que

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0, \tag{46}$$

- Expandindo (46) usando as regras para cálculo de determinantes ➜ equação polinomial de grau n
- De fato, (46) é chamada de equação característica da matriz $\mathbf{A}$ e suas raízes $\lambda_1, \lambda_2, \ldots, \lambda_n$ são os autovalores de $\mathbf{A}$
- O autovetor $\mathbf{v}_i$ associado a $\lambda_i,\ i = 1, 2, \ldots, n$, pode ser determinado substituindo-se λ_i em (44), ou seja,

$$\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i, \quad i = 1, 2, \ldots, n \tag{47}$$

- Por convenção, assume-se que $|\lambda_{\max}| = |\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n| = |\lambda_{\min}|$ e que $\|\mathbf{v}_i\|_2 = 1, \quad i = 1, 2, \ldots, n$

Propriedades de autovalores e autovetores

Propriedade 1

Se $\lambda_1, \lambda_2, \ldots, \lambda_n$ são os autovalores da matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$, então os autovalores de $\mathbf{A}^k$, $k > 0$, são $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$.

Prova.

$$\mathbf{A}^k \mathbf{v}_i = \mathbf{A}^{k-1} \mathbf{A} \mathbf{v}_i = \lambda_i \mathbf{A}^{k-1} \mathbf{v}_i = \lambda_i^2 \mathbf{A}^{k-2} \mathbf{v}_i = \ldots = \lambda_i^{k-1} \mathbf{A} \mathbf{v}_i = \lambda_i^k \mathbf{v}_i. \qquad \square$$

- Logo, todo autovetor $\mathbf{v}_i$ de $\mathbf{A}$ é autovetor de $\mathbf{A}^k$.

Propriedade 2

Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ os autovetores da matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$, correspondentes a autovalores distintos $\lambda_1, \lambda_2, \ldots, \lambda_n$, então $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente independentes.

Propriedades de autovalores e autovetores

Prova.

- Se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, então existe $\alpha_i, i = 1, 2, \ldots, n$, não todos nulos, tal que $\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = 0$, que multiplicado repetidamente por $\mathbf{A}$ leva ao conjunto de n equações

$$\sum_{i=1}^{n} \alpha_i \lambda_i^k \mathbf{v}_i = 0, \quad k = 1, 2, \ldots, n, \tag{48}$$

o qual pode ser reescrito matricialmente como

$$\begin{bmatrix} \alpha_1 \mathbf{v}_1 & \alpha_2 \mathbf{v}_2 & \alpha_3 \mathbf{v}_3 & \ldots & \alpha_n \mathbf{v}_n \end{bmatrix} \underbrace{\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \ldots & \lambda_n^{n-1} \end{bmatrix}}_{\mathbf{S}} = \mathbf{0} \tag{49}$$

□

Propriedades de autovalores e autovetores

Prova.

- A matriz $\mathbf{S}$ em (49) é chamada matriz de Vandermonde e para λ_i distintos é não-singular. Logo

$$\begin{bmatrix}\alpha_1\mathbf{v}_1 & \alpha_2\mathbf{v}_2 & \dots & \alpha_n\mathbf{v}_n\end{bmatrix}\mathbf{S} = \mathbf{0} \Rightarrow \begin{bmatrix}\alpha_1\mathbf{v}_1 & \alpha_2\mathbf{v}_2 & \dots & \alpha_n\mathbf{v}_n\end{bmatrix} = \mathbf{0}\mathbf{S}^{-1} = \mathbf{0} \tag{50}$$

- Logo, α_i precisam ser todos nulos, já que $\mathbf{v}_i$ são não-nulos e, portanto, $\mathbf{v}_i$ são L.I.

$\square$

Propriedade 3

Sejam $\lambda_1, \lambda_2, \dots, \lambda_n$ os autovalores da matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\mathbf{A} = \mathbf{A}^{\mathrm{H}}$, então $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ e $\lambda_i \geq 0$.

Prova.

Temos que $\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i \Rightarrow \mathbf{v}_i^{\mathrm{H}}\mathbf{A}^{\mathrm{H}} = \lambda_i^*\mathbf{v}_i^{\mathrm{H}} \Rightarrow \mathbf{v}_i^{\mathrm{H}}\mathbf{A} = \lambda_i^*\mathbf{v}_i^{\mathrm{H}}$. Logo
$\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i \Rightarrow \mathbf{v}_i^{\mathrm{H}}\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i^{\mathrm{H}}\mathbf{v}_i \geq 0 \Rightarrow \lambda_i^*\mathbf{v}_i^{\mathrm{H}}\mathbf{v}_i = \lambda_i\mathbf{v}_i^{\mathrm{H}}\mathbf{v}_i \geq 0 \Rightarrow \lambda_i = \lambda_i^* \geq 0$ $\square$

Propriedades de autovalores e autovetores

Propriedade 4

Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ *os autovetores da matriz* $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\mathbf{A} = \mathbf{A}^{\mathrm{H}}$, *correspondentes a autovalores distintos* $\lambda_1, \lambda_2, \ldots, \lambda_n$, *então* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ *são ortogonais.*

Prova.

Temos que $\mathbf{A}\mathbf{v}_j = \lambda_j \mathbf{v}_j \Rightarrow \mathbf{v}_j^{\mathrm{H}}\mathbf{A}^{\mathrm{H}} = \lambda_j^* \mathbf{v}_j^{\mathrm{H}} \Rightarrow \mathbf{v}_j^{\mathrm{H}}\mathbf{A} = \lambda_j \mathbf{v}_j^{\mathrm{H}}$. Logo
$\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i \Rightarrow \mathbf{v}_j^{\mathrm{H}}\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_j^{\mathrm{H}}\mathbf{v}_i \Rightarrow \lambda_j \mathbf{v}_j^{\mathrm{H}}\mathbf{v}_i = \lambda_i \mathbf{v}_j^{\mathrm{H}}\mathbf{v}_i \Rightarrow \mathbf{v}_j^{\mathrm{H}}\mathbf{v}_i = 0$, pois $\lambda_i \neq \lambda_j$ $\square$

- Os autovetores de uma matriz $\mathbf{A}$, são ortonormais se associados a autovalores distintos, i.e.,

$$\mathbf{v}^{\mathrm{T}}{}_i \mathbf{v}_i = 1, \quad \forall i, \quad i = 1, 2, \ldots, n \tag{51a}$$

$$\mathbf{v}^{\mathrm{T}}{}_i \mathbf{v}_j = 0, \quad \forall i \neq j, \quad i, j = 1, 2, \ldots, n \tag{51b}$$

- Formam uma base para o espaço coluna gerado pela matriz, i.e., qualquer vetor $\mathbf{x}$ nesse espaço pode ser escrito como uma combinação linear

$$\mathbf{x} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n \tag{52}$$

onde $\alpha_i, \ i = 1, 2, \ldots, n$, são constantes reais tais $\alpha_i = 0, \forall i \Leftrightarrow \mathbf{x} = \mathbf{0}$

Propriedades de autovalores e autovetores

Propriedade 5

Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ *os autovetores da matriz* $\mathbf{A}$*, correspondentes a autovalores distintos* $\lambda_1, \lambda_2, \ldots, \lambda_n$*, então* $\mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \mathbf{\Lambda}$, $\mathbf{A} \in \mathbb{R}^{n\times n}$ *e* $\mathbf{V}^{\mathrm{H}}\mathbf{A}\mathbf{V} = \mathbf{\Lambda}$, $\mathbf{A} \in \mathbb{C}^{n\times n}$ *onde* $\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{bmatrix}$ *e* $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$.

Prova.

A prova segue diretamente das propriedades de independência linear ($\exists \mathbf{V}^{-1}$) e de ortogonalidade ($\mathbf{V}^{\mathrm{H}}\mathbf{V} = \mathbf{I}$). □

Propriedade 6

Sejam $\lambda_1, \lambda_2, \ldots, \lambda_n$ *os autovalores da matriz* $\mathbf{A}$*, então o traço* $\mathrm{tr}\,(\mathbf{A})$ *de* $\mathbf{A}$ *é igual à soma dos autovalores de* $\mathbf{A}$.

Prova.

Temos que $\mathrm{tr}\,(\mathbf{A}) = \sum_{i=1}^{n} a_{i,i} = \mathrm{tr}\,\left(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}\right) = \mathrm{tr}\,\left(\mathbf{\Lambda}\mathbf{V}^{-1}\mathbf{V}\right) = \mathrm{tr}\,(\mathbf{\Lambda}) = \sum_{i=1}^{n} \lambda_i$. □

Propriedades de autovalores e autovetores

Propriedade 7

Sejam $\lambda_1, \lambda_2, \ldots, \lambda_n$ os autovalores da matriz $\mathbf{A}$, então o determinante $\det(\mathbf{A})$ de $\mathbf{A}$ é igual ao produto dos autovalores de $\mathbf{A}$.

Prova.

Temos que

$$\det(\mathbf{A}) = \det\left(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}\right) = \det(\mathbf{V})\det(\mathbf{\Lambda})\det\left(\mathbf{V}^{-1}\right) = \det(\mathbf{V})\det(\mathbf{\Lambda})\det(\mathbf{V})^{-1} = \det(\mathbf{\Lambda}) = \prod_{i=1}^{n}\lambda_i. \qquad \square$$

Transformação de similaridade

- Transformações de similaridade são úteis para transformar uma matriz associada a um problema em uma forma similar e de mais simples manipulação
- Se agruparmos as n equações em (47) em uma única equação podemos escrever

$$\mathbf{A}\underbrace{\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n\end{bmatrix}}_{\mathbf{V}} = \underbrace{\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n\end{bmatrix}}_{\mathbf{V}}\underbrace{\begin{bmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \lambda_n\end{bmatrix}}_{\mathbf{\Lambda}} \tag{53}$$

$$\mathbf{AV} = \mathbf{V\Lambda} \Rightarrow \boxed{\mathbf{A} = \mathbf{V\Lambda V}^{-1}}$$

- As matrizes $\mathbf{\Lambda}$ e $\mathbf{V}$ são as matrizes dos autovalores e autovetores de $\mathbf{A}$, respectivamente

Transformação de similaridade

- Duas matrizes $\mathbf{A}$ e $\mathbf{B}$ de dimensão $n \times n$ são ditas **similares** se existe uma transformação $\mathbf{T}$ tal que

$$\mathbf{A} = \mathbf{T}\mathbf{B}\mathbf{T}^{-1} \Rightarrow \mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} \tag{54}$$

- Matrizes similares possuem os mesmos autovalores e encontram várias aplicações em engenharia
- A partir de (53) e (54), podemos observar que

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1} \text{ e } \boldsymbol{\Lambda} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}, \tag{55}$$

  de modo que a matriz $\mathbf{A}$ é similar à matriz diagonal $\boldsymbol{\Lambda}$
- Observe ainda que há uma relação entre as inversas de matrizes similares dada por

$$\mathbf{A}^{-1} = \left(\mathbf{V}^{-1}\right)^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{V}^{-1} = \mathbf{V}\boldsymbol{\Lambda}^{-1}\mathbf{V}^{-1}, \tag{56}$$

  de modo que a inversa de $\mathbf{A}$ pode ser facilmente obtida se $\boldsymbol{\Lambda}$ e $\mathbf{V}$ forem conhecidas, pois $\boldsymbol{\Lambda}$ é diagonal
- Note ainda que $\det\left(\mathbf{A}^{-1}\right) = \det\left(\boldsymbol{\Lambda}^{-1}\right) = \prod_{i=1}^{n} \lambda_i^{-1}$

Forma de Jordan

- ...

Conteúdo

Polinômios de matrizes quadradas

- Se $\mathbf{A}$ é uma matriz quadrada e k é um número não-negativo, então

$$\mathbf{A}^k = \underbrace{\mathbf{A} \cdot \mathbf{A} \cdot \mathbf{A} \cdots \mathbf{A}}_{k \text{ vezes}} \quad \text{e} \quad \mathbf{A}^0 = \mathbf{I} \tag{57}$$

- Se $f(\lambda)$ é um polinômio qualquer e, por exemplo, $f_1(\lambda) = \lambda^3 + 2\lambda^2 - 6$ e $f_2(\lambda) = (\lambda + 2)(4\lambda - 3)$ então

$$f_1(\mathbf{A}) = \mathbf{A}^3 + 2\mathbf{A}^2 - 6\mathbf{I} \quad \text{e} \quad f_2(\mathbf{A}) = (\mathbf{A} + 2\mathbf{I})(4\mathbf{A} - 3\mathbf{I}) \tag{58}$$

- Em outros termos, polinômios podem ser aplicados diretamente a matrizes constituindo assim uma classe de funções de matrizes

Teorema de Cayley-Hamilton

- Se $\mathbf{A}$ é uma matriz quadrada, então seu polinômio característico $\Delta(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A})$ é dado por
$$\Delta(\lambda) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \alpha_{n-2}\lambda^{n-2} + \ldots + \alpha_1\lambda + \alpha_0 \tag{59}$$
- O teorema de Cayley-Hamilton estabelece que a função polinomial (59) aplicada à matriz $\mathbf{A}$ é identicamente nula, i.e.,
$$\Delta(\mathbf{A}) = \mathbf{A}^n + \alpha_{n-1}\mathbf{A}^{n-1} + \alpha_{n-2}\mathbf{A}^{n-2} + \ldots + \alpha_1\mathbf{A} + \alpha_0\mathbf{I} = \mathbf{0}, \tag{60}$$
- Logo, $\mathbf{A}$ satisfaz sua própria equação característica
- O teorema de Cayley-Hamilton é útil para calcular potências de $\mathbf{A}^k$, $k > n$ em função de $\mathbf{A}^n, \ldots, \mathbf{A}$, como por exemplo
$$\begin{aligned} \mathbf{A}^{n+1} &= \mathbf{A} \cdot (\mathbf{A}^n + \alpha_{n-1}\mathbf{A}^{n-1} + \alpha_{n-2}\mathbf{A}^{n-2} + \ldots + \alpha_1\mathbf{A} + \alpha_0\mathbf{I}) = \mathbf{0} \Rightarrow \\ \mathbf{A}^{n+1} &= -\alpha_{n-1}\mathbf{A}^n - \alpha_{n-2}\mathbf{A}^{n-1} - \ldots - \alpha_1\mathbf{A}^2 - \alpha_0\mathbf{A} \end{aligned} \tag{61}$$

Teorema de Cayley-Hamilton

- De acordo com (60), $\mathbf{A}^n$ pode ser escrita como combinação linear de $\{\mathbf{I}, \mathbf{A}, \ldots, \mathbf{A}^{n-1}\}$
- De acordo com (61), $\mathbf{A}^{n+1}$ pode ser escrita como combinação linear de $\{\mathbf{A}, \mathbf{A}^2, \ldots, \mathbf{A}^n\}$
- Logo, para qualquer função polinomial $f(\lambda)$ temos que $f(\mathbf{A})$ pode ser escrita como

$$f(\mathbf{A}) = \beta_0\mathbf{I} + \beta_1\mathbf{A} + \ldots + \beta_{n-2}\mathbf{A}^{n-2} + \beta_{n-1}\mathbf{A}^{n-1} \quad (62)$$

  para algum conjunto de valores $\beta_0, \beta_1, \ldots, \beta_{n-1}$
- Se $\mathbf{A} = \mathbf{T}\bar{\mathbf{A}}\mathbf{T}^{-1}$ então $\mathbf{A}^k = \mathbf{T}\bar{\mathbf{A}}^k\mathbf{T}^{-1}$ e, portanto $f(\mathbf{A}) = \mathbf{T}f(\bar{\mathbf{A}})\mathbf{T}^{-1}$, pois

$$\begin{aligned} f(\mathbf{A}) &= f(\mathbf{T}\bar{\mathbf{A}}\mathbf{T}^{-1}) \\ &= \beta_0\mathbf{T}\mathbf{T}^{-1} + \beta_1\mathbf{T}\bar{\mathbf{A}}\mathbf{T}^{-1} + \ldots + \beta_{n-2}\mathbf{T}\bar{\mathbf{A}}^{n-2}\mathbf{T}^{-1} + \beta_{n-1}\mathbf{T}\bar{\mathbf{A}}^{n-1}\mathbf{T}^{-1} \\ &= \mathbf{T}(\beta_0\mathbf{I} + \beta_1\mathbf{A} + \ldots + \beta_{n-2}\mathbf{A}^{n-2} + \beta_{n-1}\mathbf{A}^{n-1})\mathbf{T}^{-1} \\ &= \mathbf{T}f(\bar{\mathbf{A}})\mathbf{T}^{-1} \end{aligned}$$

Conteúdo

Norma induzida

- A norma induzida $\|\mathbf{A}\|$ de uma matriz $\mathbf{A}$ pode ser definida através do problema de otimização

$$\|\mathbf{A}\| = \max\left\{\|\mathbf{A}\mathbf{x}\|_2\right\} = \max\left\{\left(\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x}\right)^{\frac{1}{2}}\right\}, \tag{63a}$$

$$\text{subject to: } \|\mathbf{x}\|_2 = 1 \tag{63b}$$

- Dado que o $\mathcal{R}\left(\mathbf{A}\right) = \mathcal{R}\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)$ e que $\|\mathbf{x}\|_2 = 1$ conforme (63b), verificamos que a solução de (63) corresponde à raiz quadrada do maior autovalor de $\mathbf{A}^{\mathrm{T}}\mathbf{A}$
- A norma induzida acima pode ser definida em termos de outras normas que não a norma euclidiana
- Apenas a norma induzida pela norma euclidiana é também chamada norma espectral

Norma induzida

- As normas induzidas para matrizes satisfazem as mesmas propriedades que as normas de vetores, tais como:

$$\|\mathbf{A}\| = \|\mathbf{A}^{\mathrm{T}}\| \geq 0 \qquad \text{(Não-negatividade)} \qquad (64a)$$

$$\|\mathbf{A}\mathbf{B}\| \leq \|\mathbf{A}\|\|\mathbf{B}\| \qquad \text{(Desigualdade de Cauchy-Schwarz)} \qquad (64b)$$

$$\|\mathbf{A}\| = 0 \Leftrightarrow \mathbf{A} = \mathbf{0} \qquad \text{(Elemento neutro)} \qquad (64c)$$

$$\|\alpha\mathbf{A}\| = |\alpha|\|\mathbf{A}\|, \forall \alpha \in \mathbb{R} \qquad \text{(Escalabilidade)} \qquad (64d)$$

$$\|\mathbf{A}+\mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|, \forall \mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n} \qquad \text{(Desigualdade triangular)} \qquad (64e)$$

$$\max_{\substack{1\leq i\leq m\\1\leq j\leq n}} |a_{i,j}| \leq \|\mathbf{A}\| \leq \sqrt{mn} \max_{\substack{1\leq i\leq m\\1\leq j\leq n}} |a_{i,j}| \qquad (64f)$$

- Para matrizes complexas, as matrizes transpostas são substituídas por matrizes hermitianas

Conteúdo

Cálculo com matrizes

- As regras de cálculo aplicadas a matrizes devem ser consistentes com as regras de cálculo utilizando escalares
- Os resultados sugerem apenas a reorganização dos termos entre formas escalares e matriciais
- Este princípio conduz à conclusão de que a derivada e a integral de matrizes pode ser definida elemento-a-elemento, i.e., o elemento i, j das matrizes

$$\int_0^t \mathbf{A}(\tau)d\tau \quad \text{e} \quad \frac{d\mathbf{A}(t)}{dt} \quad \text{são} \quad \int_0^t a_{i,j}(\tau)d\tau \quad \text{e} \quad \frac{da_{i,j}(t)}{dt}, \tag{65}$$

  respectivamente
- O teorema fundamental do cálculo aplicado a matrizes leva a

$$\frac{d}{dt}\left(\int_0^t \mathbf{A}(\tau)d\tau\right) = \mathbf{A}(t) \tag{66}$$

Cálculo com matrizes

- Similarmente, a regra do produto para matrizes $\mathbf{A}(t)$ e $\mathbf{B}(t)$ resulta em

$$\frac{d}{dt}\left(\mathbf{A}(t)\mathbf{B}(t)\right) = \dot{\mathbf{A}}(t)\mathbf{B}(t) + \mathbf{A}(t)\dot{\mathbf{B}}(t) \tag{67}$$

- No entanto, é importante observar que

$$\frac{d\mathbf{A}^2(t)}{dt} = \dot{\mathbf{A}}(t)\mathbf{A}(t) + \mathbf{A}(t)\dot{\mathbf{A}}(t) \tag{68}$$

- Nesse contexto, um relação importante é a desigualdade triangular

$$\|\int_{t_0}^{t} \mathbf{x}(\tau)d\tau\| \leq |\int_{t_0}^{t} \|\mathbf{x}(\tau)\|d\tau| \tag{69}$$

onde para $t \geq t_0$ o módulo pode ser desconsiderado

Conteúdo

Forma quadrática e matriz positiva (semi)definida

- Seja $\mathbf{Q}$ uma matriz pertencente a $\mathbb{R}^{n\times n}$ e $\mathbf{x}$ um vetor pertencente a $\mathbb{R}^n$
- O produto $\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x}$ é chamado forma quadrática em $\mathbf{x}$
- A matriz $\mathbf{Q}$ pode ser considerada simétrica (i.e., $\mathbf{Q} = \mathbf{Q}^{\mathrm{T}}$) no estudo de formas quadráticas pois

$$\mathbf{x}^{\mathrm{T}}\left(\mathbf{Q} + \mathbf{Q}^{\mathrm{T}}\right)\mathbf{x} = \mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} + \mathbf{x}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\mathbf{x} = 2\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} \tag{70}$$

  e a forma quadrática não muda ao se substituir $\mathbf{Q}$ por $\dfrac{\mathbf{Q} + \mathbf{Q}^{\mathrm{T}}}{2}$
- Um matriz $\mathbf{Q}$ é chamada:
  - Positiva definida se $\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} > 0, \forall \mathbf{x} \neq \mathbf{0}$
  - Positiva semidefinida se $\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} \geq 0, \forall \mathbf{x}$
- Um matriz $\mathbf{Q}$ é negativa definida ou semidefinida se $-\mathbf{Q}$ é positiva definida ou semidefinida, respectivamente

Forma quadrática e matriz positiva (semi)definida

- A notação $\mathbf{Q} \succ 0$ e $\mathbf{Q} \succeq 0$ é normalmente utilizada para indicar que $\mathbf{Q}$ é positiva definida ou semidefinida, respectivamente
- A notação $\mathbf{Q}_1 \succ \mathbf{Q}_2$ e $\mathbf{Q}_1 \succeq \mathbf{Q}_2$ implicam $\mathbf{Q}_1 - \mathbf{Q}_2 \succ 0$ e $\mathbf{Q}_1 - \mathbf{Q}_2 \succeq 0$, respectivamente
- Todos os autovalores de uma matriz simétrica são reais
    - Todos os autovalores de $\mathbf{Q}$ são reais e positivos se $\mathbf{Q} \succ 0$
    - Todos os autovalores de $\mathbf{Q}$ são reais e não-negativos se $\mathbf{Q} \succeq 0$
- Algumas relações importantes para formas quadráticas são:

$$\lambda_{\min}\mathbf{x}^{\mathrm{T}}\mathbf{x} \leq \mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} \leq \lambda_{\max}\mathbf{x}^{\mathrm{T}}\mathbf{x}, \qquad \text{(Desigualdade de Rayleigh-Ritz)} \tag{71a}$$

$$\|\mathbf{Q}\| \leq \mathrm{tr}\left(\mathbf{Q}\right) \leq n\|\mathbf{Q}\| \tag{71b}$$

- Para matrizes e vetores complexos, a operação $(\cdot)^{\mathrm{T}}$ é substituída por $(\cdot)^{\mathrm{H}}$ e a forma quadrática em $\mathbf{x}$ torna-se $\mathbf{x}^{\mathrm{H}}\mathbf{Q}\mathbf{x}$

Menores principais dominantes

- Um menor principal de uma matriz $\mathbf{Q}$ simétrica em $\mathbb{R}^{n\times n}$ é o determinante da submatriz formada pela remoção de r linhas e colunas de mesmo índice da matriz
- Para a mesma matriz $\mathbf{Q}$, o p-ésimo menor principal dominante de $\mathbf{Q}$ é o determinante da submatriz $\mathbf{Q}_p$ superior esquerda compreendendo os elementos $q_{i,j}$ de $\mathbf{Q}$ para $i, j = 1, 2, \ldots, p$, de modo que o 1º, 2º e 3º menores principais dominantes $\mathbf{Q} \in \mathbb{R}^{n\times n}$, $n \geq 3$, são

$$\det\left(\mathbf{Q}_1\right) = \det\left(\left[q_{1,1}\right]\right), \quad \det\left(\mathbf{Q}_2\right) = \det\left(\begin{bmatrix} q_{1,1} & q_{1,2} \\ q_{2,1} & q_{2,2} \end{bmatrix}\right), \text{ e}$$
$$\det\left(\mathbf{Q}_3\right) = \det\left(\begin{bmatrix} q_{1,1} & q_{1,2} & q_{1,3} \\ q_{2,1} & q_{2,2} & q_{2,3} \\ q_{3,1} & q_{3,2} & q_{3,3} \end{bmatrix}\right) \tag{72}$$

- A matriz $\mathbf{Q}$ é positiva definida se e somente se todos os seus menores principais dominantes são positivos, i.e., se $\det\left(\mathbf{Q}_p\right) > 0, \forall p = 1, 2, \ldots, n$
- A matriz $\mathbf{Q}$ é positiva semidefinida se e somente se todos os seus menores principais dominantes são não-negativos

Conteúdo

Conteúdo

Definição da decomposição LU

- A decomposição matricial LU fatora uma matriz $\mathbf{A}$ em um produto entre uma matriz triangular inferior $\mathbf{L}$ e uma matriz triangular superior $\mathbf{U}$ tal que

$$\underbrace{\begin{bmatrix} \bullet & \bullet & \bullet & \dots & \bullet \\ \bullet & \bullet & \bullet & \dots & \bullet \\ \bullet & \bullet & \bullet & \dots & \bullet \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \bullet & \bullet & \bullet & \dots & \bullet \end{bmatrix}}_{\mathbf{A}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ * & 1 & 0 & \dots & 0 \\ * & * & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \dots & 1 \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} \circ & \circ & \circ & \dots & \circ \\ 0 & \circ & \circ & \dots & \circ \\ 0 & 0 & \circ & \dots & \circ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \circ \end{bmatrix}}_{\mathbf{U}} \tag{73}$$

$$\mathbf{A} = \mathbf{LU}$$

- A matriz $\mathbf{L}$, cf. (73), tem ainda a característica particular de que todos os elementos de sua diagonal são unitários
- A decomposição LU é muito aplicada na solução de sistemas lineares da forma

$$\mathbf{AX} = \mathbf{B} \tag{74}$$

onde $\mathbf{A}$, $\mathbf{B}$, e $\mathbf{X}$ são matrizes $n \times n$ com $\mathbf{A}$ e $\mathbf{B}$ conhecidas e $\mathbf{X}$ incógnita

- O sistema (74) corresponde a n sistemas $\mathbf{A}\mathbf{x}_i = \mathbf{b}_i, i = 1, 2, \dots, n$, onde $\mathbf{x}_i$ e $\mathbf{b_i}$ correspondem às colunas de $\mathbf{X}$ e $\mathbf{B}$, respectivamente

Conteúdo

Decomposição LU

- A decomposição LU pode ser obtida usando a eliminação de Gauss
- O método de eliminação de Gauss converte um sistema de n equações da forma $\mathbf{Ax} = \mathbf{b}$ em sistema triangular superior em $n-1$ passos

$$
\begin{gathered}
\underbrace{\begin{bmatrix} a_{1,1}^{(0)} & a_{1,2}^{(0)} & a_{1,3}^{(0)} & \dots & a_{1,n}^{(0)} \\ a_{2,1}^{(0)} & a_{2,2}^{(0)} & a_{2,3}^{(0)} & \dots & a_{2,n}^{(0)} \\ a_{3,1}^{(0)} & a_{3,2}^{(0)} & a_{3,3}^{(0)} & \dots & a_{3,n}^{(0)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1}^{(0)} & a_{n,2}^{(0)} & a_{n,3}^{(0)} & \dots & a_{n,n}^{(0)} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} b_1^{(0)} \\ b_2^{(0)} \\ b_3^{(0)} \\ \vdots \\ b_n \end{bmatrix}}_{\mathbf{b}} \xrightarrow{\text{Eliminação de Gauss}} \\
\underbrace{\begin{bmatrix} a_{1,1}^{(0)} & a_{1,2}^{(0)} & a_{1,3}^{(0)} & \dots & a_{1,n}^{(0)} \\ 0 & a_{2,2}^{(1)} & a_{2,3}^{(1)} & \dots & a_{2,n}^{(1)} \\ 0 & 0 & a_{3,3}^{(2)} & \dots & a_{3,n}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{n,n}^{(n-1)} \end{bmatrix}}_{\mathbf{A}'} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} b_1^{(0)} \\ b_2^{(1)} \\ b_3^{(2)} \\ \vdots \\ b_n^{(n-1)} \end{bmatrix}}_{\mathbf{b}'}
\end{gathered} \tag{75}
$$

Decomposição LU

- Como vimos anteriormente, no método de eliminação de Gauss, os coeficientes $a_{i,j}^{(k)}$ são obtidos em passos
- No passo k, $k = 1, 2, \ldots, n-1$, os coeficientes $a_{i,j}^{(k)}$ de cada linha i, $k < i \leq n$, são obtidos subtraindo da linha i da matriz o produto de $m_{i,k} = \dfrac{a_{i,k-1}^{(k-1)}}{a_{k-1,k-1}^{(k-1)}}$ pela linha k da matriz $\mathbf{A}$ do passo $k-1$, ou seja,

$$a_{i,j}^{(k)} = a_{i,j}^{(k-1)} - m_{i,k} a_{k-1,j}^{(k-1)} \tag{76}$$

Decomposição LU

- Aplicando esse procedimento sucessivamente, temos que

$$
\begin{bmatrix}
a_{1,1}^{(0)} & a_{1,2}^{(0)} & a_{1,3}^{(0)} & \cdots & a_{1,n}^{(0)} \\
a_{2,1}^{(0)} & a_{2,2}^{(0)} & a_{2,3}^{(0)} & \cdots & a_{2,n}^{(0)} \\
a_{3,1}^{(0)} & a_{3,2}^{(0)} & a_{3,3}^{(0)} & \cdots & a_{3,n}^{(0)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a_{n,1}^{(0)} & a_{n,2}^{(0)} & a_{n,3}^{(0)} & \cdots & a_{n,n}^{(0)}
\end{bmatrix}
\xrightarrow{\text{Passo 1}}
\begin{bmatrix}
a_{1,1}^{(0)} & a_{1,2}^{(0)} & a_{1,3}^{(0)} & \cdots & a_{1,n}^{(0)} \\
0 & a_{2,2}^{(1)} & a_{2,3}^{(1)} & \cdots & a_{2,n}^{(1)} \\
0 & a_{3,2}^{(1)} & a_{3,3}^{(1)} & \cdots & a_{3,n}^{(1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & a_{n,2}^{(1)} & a_{n,3}^{(1)} & \cdots & a_{n,n}^{(1)}
\end{bmatrix}
\xrightarrow{\text{Passo 2}}
$$

$$
\begin{bmatrix}
a_{1,1}^{(0)} & a_{1,2}^{(0)} & a_{1,3}^{(0)} & \cdots & a_{1,n}^{(0)} \\
0 & a_{2,2}^{(1)} & a_{2,3}^{(1)} & \cdots & a_{2,n}^{(1)} \\
0 & 0 & a_{3,3}^{(2)} & \cdots & a_{3,n}^{(2)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & a_{n,3}^{(2)} & \cdots & a_{n,n}^{(2)}
\end{bmatrix}
\xrightarrow{\text{Passo 3}} \cdots \xrightarrow{\text{Passo n-1}}
\begin{bmatrix}
a_{1,1}^{(0)} & a_{1,2}^{(0)} & a_{1,3}^{(0)} & \cdots & a_{1,n}^{(0)} \\
0 & a_{2,2}^{(1)} & a_{2,3}^{(1)} & \cdots & a_{2,n}^{(1)} \\
0 & 0 & a_{3,3}^{(2)} & \cdots & a_{3,n}^{(2)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & a_{n,n}^{(n-1)}
\end{bmatrix}
$$

Decomposição LU

- De fato, a aplicação do método da eliminação de Gauss já fornece a decomposição LU da matriz $\mathbf{A}$
- As matrizes $\mathbf{L}$ e $\mathbf{U}$ são escritas como

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ l_{2,1} & 1 & 0 & \dots & 0 & 0 \\ l_{3,1} & l_{3,2} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ l_{n-1,1} & l_{n-1,2} & l_{n-1,3} & \dots & 1 & 0 \\ l_{n,1} & l_{n,2} & l_{n,3} & \dots & l_{n,n-1} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ m_{2,1} & 1 & 0 & \dots & 0 & 0 \\ m_{3,1} & m_{3,2} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ m_{n-1,1} & m_{n-1,2} & m_{n-1,3} & \dots & 1 & 0 \\ m_{n,1} & m_{n,2} & m_{n,3} & \dots & m_{n,n-1} & 1 \end{bmatrix} \quad \text{(77a)}$$

$$\mathbf{U} = \begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} & \dots & u_{1,n-1} & u_{1,n} \\ 0 & u_{2,2} & u_{2,3} & \dots & u_{2,n-1} & u_{2,n} \\ 0 & 0 & u_{3,3} & \dots & u_{3,n-1} & u_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & u_{n-1,n-1} & u_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 & u_{n,n} \end{bmatrix} = \begin{bmatrix} a_{1,1}^{(0)} & a_{1,2}^{(0)} & a_{1,3}^{(0)} & \dots & a_{1,n-1}^{(0)} & a_{1,n}^{(0)} \\ 0 & a_{2,2}^{(1)} & a_{2,3}^{(1)} & \dots & a_{2,n-1}^{(1)} & a_{2,n}^{(1)} \\ 0 & 0 & a_{3,3}^{(2)} & \dots & a_{3,n-1}^{(2)} & a_{3,n}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a_{n-1,n-1}^{(n-2)} & a_{n-1,n}^{(n-2)} \\ 0 & 0 & 0 & \dots & 0 & a_{n,n}^{(n-1)} \end{bmatrix} \quad \text{(77b)}$$

Decomposição LU

- Considerando que a decomposição LU de $\mathbf{A}$ é calculada linha por linha de cima para baixo (omitindo, portanto, os índices (k) dos $n-1$ passos) e utilizando as equações 76 e 77, podemos sintetizar as equações da decomposição LU como

$$u_{1,j} = a_{1,j} \qquad j = 1, 2, \ldots, n \tag{78a}$$

$$m_{i,1} = \frac{a_{i,1}}{u_{1,1}} \qquad i = 2, \ldots, n \tag{78b}$$

$$u_{i,j} = a_{i,j} - \sum_{s=1}^{i-1} m_{i,s} u_{s,j} \qquad j = i, i+1, \ldots, n,\ i \geq 2 \tag{78c}$$

$$m_{i,j} = \frac{1}{u_{j,j}} \left(a_{i,j} - \sum_{s=1}^{j-1} m_{i,s} u_{s,j} \right) \qquad i = j+1, j+2, \ldots, n, j \geq 2 \tag{78d}$$

Conteúdo

Método de Doolittle

- O método de Doolittle utiliza a decomposição LU para solucionar sistemas de equações lineares
- Considere o sistema

$$\mathbf{A}\mathbf{x} = \mathbf{b} \Rightarrow \mathbf{L}\underbrace{\mathbf{U}\mathbf{x}}_{\mathbf{y}} = \mathbf{b} \Rightarrow \begin{cases} \mathbf{L}\mathbf{y} = \mathbf{b} \\ \mathbf{U}\mathbf{x} = \mathbf{y} \end{cases} \tag{79}$$

- Como $\mathbf{L}$ é triangular inferior, o sistema $\mathbf{L}\mathbf{y} = \mathbf{b}$ pode ser solucionado primeiro utilizando o algoritmo de substituição progressiva discutido anteriormente de tal modo que

$$y_1 = b_1 \tag{80a}$$

$$y_i = b_i - \sum_{j=1}^{i-1} m_{i,j} y_j, \ i = 2, 3, \ldots, n \tag{80b}$$

Método de Doolittle para solução de sistemas usando a decomposição LU

- Uma vez que $\mathbf{y}$ já foi determinado no passo anterior e dado que $\mathbf{U}$ é triangular inferior, o sistema $\mathbf{U}\mathbf{x} = \mathbf{y}$ pode ser facilmente solucionado utilizando o algoritmo de substituição progressiva discutido anteriormente de tal modo que

$$x_n = \frac{y_n}{u_{n,n}} \tag{81a}$$

$$x_i = \frac{1}{u_{i,i}} \left(y_i - \sum_{j=i+1}^{n} u_{i,j} y_j \right), \; i = n-1, n-2, \ldots, 1 \tag{81b}$$

Conteúdo

Inversão de matrizes pelo método de Doolittle

- Como mencionado anteriormente, um sistema da forma $\mathbf{AX} = \mathbf{B}$ pode ser visto como um conjunto de n sistemas da forma $\mathbf{Ax_i} = \mathbf{b}_i$, onde $\mathbf{x}_i$ e $\mathbf{b}_i$ são as colunas de $\mathbf{X}$ e $\mathbf{B}$
- Para o caso particular em que $\mathbf{B} = \mathbf{I}$, a solução $\mathbf{X}$ do sistema, se existir, corresponde à inversa $\mathbf{A}^{-1}$ de $\mathbf{A}$, pois $\mathbf{AX} = \mathbf{I}$
- Aplicando uma única vez a decomposição LU à matriz $\mathbf{A}$ obtemos $\mathbf{L}$ e $\mathbf{U}$ e aplicando o método de Doolittle a cada sistema $\mathbf{Ax_i} = \mathbf{b}_i$ obtemos então sucessivamente as colunas de $\mathbf{A}^{-1}$
- Uma condição necessária para a existência de $\mathbf{X} = \mathbf{A}^{-1}$ é que $\det(\mathbf{A}) \neq 0$
- Uma vez que o determinante de uma matriz triangular inferior ou superior é igual ao produto dos elementos de sua diagonal, temos que

$$\det(\mathbf{L}) = 1, \text{ e que } \det(\mathbf{U}) = u_{1,1} \cdot u_{2,2} \cdot u_{3,3} \cdots u_{n,n} \tag{82}$$

- Como o determinante do produto de matrizes é igual ao produto dos determinantes temos que

$$\mathbf{A} = \mathbf{LU} \Rightarrow \det(\mathbf{A}) = \det(\mathbf{L})\det(\mathbf{U}) = u_{1,1} \cdot u_{2,2} \cdot u_{3,3} \cdots u_{n,n} \tag{83}$$

Conteúdo

Conteúdo

Definição da decomposição de Cholesky

- A decomposição de Cholesky fatora uma matriz simétrica, positiva definida $\mathbf{A}$ em um produto de uma matriz triangular inferior $\mathbf{L}$ por sua transposta $\mathbf{L}^T$, i.e.,

$$\mathbf{A} = \mathbf{L}\mathbf{L}^T \Rightarrow$$

$$\begin{bmatrix} a_{1,1} & a_{2,1} & \dots & a_{n,1} \\ a_{2,1} & a_{2,2} & \dots & a_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{bmatrix} = \begin{bmatrix} l_{1,1} & 0 & \dots & 0 \\ l_{2,1} & l_{2,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n,1} & l_{n,2} & \dots & l_{n,n} \end{bmatrix} \begin{bmatrix} l_{1,1} & l_{2,1} & \dots & l_{n,1} \\ 0 & l_{2,2} & \dots & l_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & l_{n,n} \end{bmatrix} \tag{84}$$

- Uma matriz é denominada simétrica se $\mathbf{A} = \mathbf{A}^T$
- Uma matriz é denominada positiva definida se para todo vetor não-nulo $\mathbf{x} \neq \mathbf{0}$, o produto $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$
- Os autovalores de uma matriz positiva definida são todos positivos, pois

$$\mathbf{x}^T\mathbf{A}\mathbf{x} > 0 \Rightarrow \underbrace{\mathbf{x}^T\mathbf{V}}_{\tilde{\mathbf{x}}^{\mathbf{T}}}\boldsymbol{\Lambda}\underbrace{\mathbf{V}^T\mathbf{x}}_{\tilde{\mathbf{x}}} > 0 \Rightarrow \tilde{\mathbf{x}}^{\mathbf{T}}\boldsymbol{\Lambda}\tilde{\mathbf{x}}^{\mathbf{T}} \Rightarrow [\tilde{x}_1 \quad \tilde{x}_2 \quad \dots \quad \tilde{x}_n] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_n \end{bmatrix} > 0$$

$$\Rightarrow \lambda_1\tilde{x}_1^2 \cdot \lambda_2\tilde{x}_2^2 \cdots \lambda_n\tilde{x}_n^2 > 0 \Rightarrow \lambda_i > 0, \forall i = 1, 2, \dots, n$$

Conteúdo

Método de Cholesky

- O método de Cholesky é um método direto para a determinação de $\mathbf{L}$
- O método consiste em equacionar diretamente os coeficientes $l_{i,j}$ de $\mathbf{L}$, o que leva ao seguinte conjunto de equações

$$l_{1,1} = \sqrt{a_{1,1}} \tag{85a}$$

$$l_{i,1} = \frac{a_{i,1}}{l_{1,1}} \qquad i = 2, \ldots, n \tag{85b}$$

$$l_{i,i} = \sqrt{a_{i,i} - \sum_{j=1}^{i-1} l_{i,j}^2} \qquad i = 2, \ldots, n \tag{85c}$$

$$l_{i,j} = \frac{1}{l_{j,j}} \sqrt{a_{i,j} - \sum_{s=1}^{j-1} l_{j,s} l_{i,s}} \qquad i = j+1, j+2, \ldots, n, j \geq 2 \tag{85d}$$

Conteúdo

Decomposição QR [GS08, cap. 4]

- O método da fatoração QR faz uso de transformações de Householder e de transformações de similaridade para decompor uma matriz $\mathbf{A}$ em um produto de uma matriz ortogonal $\mathbf{Q}$ ($\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$), por uma matriz triangular superior $\mathbf{R}$
- A fatoração QR se inicia com a matriz $\mathbf{A}^{(1)} = \mathbf{A}$ cujos autovalores devem ser determinados
- A matriz $\mathbf{A}^{(1)}$ é fatorada como
$$\mathbf{A}^{(1)} = \mathbf{Q}^{(1)}\mathbf{R}^{(1)}, \tag{86}$$
onde $\mathbf{Q}^{(1)}$ é uma matriz ortogonal, i.e., $\mathbf{Q}^{(1)}\left(\mathbf{Q}^{(1)}\right)^{\mathrm{T}} = \mathbf{I}$, mas a matriz $\mathbf{R}^{(1)}$ não é ainda uma matriz triangular superior
- A matriz $\mathbf{A}^{(2)}$ é obtida multiplicando $\mathbf{R}^{(1)}$ à direita por $\mathbf{Q}^{(1)}$, i.e.,
$$\mathbf{A}^{(2)} = \mathbf{R}^{(1)}\mathbf{Q}^{(1)} \tag{87}$$
- Usando (86), temos que $\mathbf{R}^{(1)} = \left(\mathbf{Q}^{(1)}\right)^{-1}\mathbf{A}^{(1)}$, a qual substituida em (87) resulta em
$$\mathbf{A}^{(2)} = \left(\mathbf{Q}^{(1)}\right)^{-1}\mathbf{A}^{(1)}\mathbf{Q}^{(1)} \tag{88}$$
- Logo, $\mathbf{A}^{(1)}$ e $\mathbf{A}^{(2)}$ são matrizes similares, i.e., possuem os mesmos autovalores

Matriz de Householder

- Para obter $\mathbf{Q}$ e $\mathbf{R}$, o método de decomposição QR utiliza matrizes de transformação de Householder
- Dado um vetor

$$\mathbf{v} = \mathbf{c} + \|\mathbf{c}\|_2 \mathbf{e}, \tag{89}$$

  a matriz de transformação de Householder $\mathbf{H}$ associada ao vetor $\mathbf{v}$ é definida como

$$\mathbf{H} = \mathbf{I} - 2\frac{\mathbf{v}\mathbf{v}^{\mathrm{T}}}{\mathbf{v}^{\mathrm{T}}\mathbf{v}}, \tag{90}$$

  onde $\mathbf{I}$ é a matriz identidade e $\mathbf{e}$ é um vetor com uma única componente igual a ± 1 e todas as demais componentes igual a zero.
- Por conveniência, costuma-se definir ainda o vetor $\mathbf{e}$ como $\pm\mathbf{i}_i$, i.e., em termos da i-ésima coluna $\mathbf{I}$
- Note que a matriz de transformação de Householder é ortogonal, i.e., $\mathbf{H}^{\mathrm{T}}\mathbf{H} = \mathbf{I}$

Algoritmo de fatoração QR

- O **passo 1** do algoritmo QR consiste em determinar $\mathbf{Q}^{(1)}$ e $\mathbf{R}^{(1)}$
- Para tanto, os vetores $\mathbf{c}^{(1)}$ e $\mathbf{e}^{(1)}$ geradores da matriz de Householder $\mathbf{H}^{(1)}$ são definidos como

$$\mathbf{c}^{(1)} = \mathbf{a}_1, \text{ onde } \mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{bmatrix}, \text{ e}$$
$$\mathbf{e}^{(1)} = \begin{cases} \mathbf{i}_1, & a_{1,1} \geq 0 \\ -\mathbf{i}_1, & a_{1,1} < 0 \end{cases}, \text{ onde } \mathbf{I} = \begin{bmatrix} \mathbf{i}_1 & \mathbf{i}_2 & \dots & \mathbf{i}_n \end{bmatrix} \tag{91}$$

- Usando (91) em (90), as matrizes $\mathbf{Q}^{(1)}$ e $\mathbf{R}^{(1)}$ são definidas como

$$\mathbf{Q}^{(1)} = \mathbf{H}^{(1)} \quad \text{e} \quad \mathbf{R}^{(1)} = \mathbf{H}^{(1)} \mathbf{A}^{(1)} \tag{92}$$

Algoritmo de fatoração QR

- A matriz $\mathbf{R}^{(1)}$ obtida no **passo 1** é da forma

$$\mathbf{R}^{(1)} = \begin{bmatrix} \mathbf{r}_{1,1}^{(1)} & \mathbf{r}_{1,2}^{(1)} & \mathbf{r}_{1,3}^{(1)} & \dots & \mathbf{r}_{1,n}^{(1)} \\ 0 & \mathbf{r}_{2,2}^{(1)} & \mathbf{r}_{2,3}^{(1)} & \dots & \mathbf{r}_{2,n}^{(1)} \\ 0 & \mathbf{r}_{3,2}^{(1)} & \mathbf{r}_{3,3}^{(1)} & \dots & \mathbf{r}_{3,n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & \mathbf{r}_{n,2}^{(1)} & \mathbf{r}_{n,3}^{(1)} & \dots & \mathbf{r}_{n,n}^{(1)} \end{bmatrix} \tag{93}$$

- No **passo 2** do algoritmo QR, aproximadamente o mesmo processo do **passo 1** é repetido
- No entanto, os vetores $\mathbf{c}^{(2)}$ e $\mathbf{e}^{(2)}$ geradores da matriz de Householder $\mathbf{H}^{(2)}$ são definidos como

$$\mathbf{c}^{(2)} = \begin{bmatrix} 0 & r_{2,2}^{(1)} & r_{3,2}^{(1)} & \dots & r_{n,2}^{(1)} \end{bmatrix}^T, \quad \text{e} \quad \mathbf{e}^{(2)} = \begin{cases} \mathbf{i}_2, & r_{2,2}^{(1)} \geq 0 \\ -\mathbf{i}_2, & r_{2,2}^{(1)} < 0 \end{cases} \tag{94}$$

Algoritmo de fatoração QR

- Usando (94) e (51), as matrizes $\mathbf{Q}^{(2)}$ e $\mathbf{R}^{(2)}$ são obtidas como

$$\mathbf{Q}^{(2)} = \mathbf{Q}^{(1)}\mathbf{H}^{(2)} \quad \text{e} \quad \mathbf{R}^{(2)} = \mathbf{H}^{(2)}\mathbf{R}^{(1)} \tag{95}$$

onde a matriz $\mathbf{R}^{(2)}$ tem a forma

$$\mathbf{R}^{(2)} = \begin{bmatrix} \mathbf{r}_{1,1}^{(2)} & \mathbf{r}_{1,2}^{(2)} & \mathbf{r}_{1,3}^{(2)} & \dots & \mathbf{r}_{1,n}^{(2)} \\ 0 & \mathbf{r}_{2,2}^{(2)} & \mathbf{r}_{2,3}^{(2)} & \dots & \mathbf{r}_{2,n}^{(2)} \\ 0 & 0 & \mathbf{r}_{3,3}^{(2)} & \dots & \mathbf{r}_{3,n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \mathbf{r}_{n,3}^{(2)} & \dots & \mathbf{r}_{n,n}^{(2)} \end{bmatrix} \tag{96}$$

- No **passo 3** aproximadamente o mesmo processo do **passo 2** é repetido, exceto que os vetores $\mathbf{c}^{(3)}$ e $\mathbf{e}^{(3)}$ geradores $\mathbf{H}^{(3)}$ são dados por

$$\mathbf{c}^{(3)} = \begin{bmatrix} 0 & 0 & r_{3,3}^{(2)} & \dots & r_{n,3}^{(2)} \end{bmatrix}^T, \quad \text{e} \quad \mathbf{e}^{(3)} = \begin{cases} \mathbf{i}_3, & r_{3,3}^{(2)} \geq 0 \\ -\mathbf{i}_3, & r_{3,3}^{(2)} < 0 \end{cases} \tag{97}$$

Algoritmo de fatoração QR

- Usando (97) e (90), as matrizes $\mathbf{Q}^{(3)}$ e $\mathbf{R}^{(3)}$ são obtidas como

$$\mathbf{Q}^{(3)} = \mathbf{Q}^{(2)}\mathbf{H}^{(3)} \quad \text{e} \quad \mathbf{R}^{(3)} = \mathbf{H}^{(3)}\mathbf{R}^{(2)} \tag{98}$$

- Em cada passo, o processo descrito anteriormente é repetido
- Observe ainda que a cada passo i, os elementos da i-ésima coluna da matriz $\mathbf{R}^{(i)}$ são zerados
- De forma geral, as iterações descritas anteriormente podem ser escritas como

$$\mathbf{Q}^{(i)} = \mathbf{Q}^{(i-1)}\mathbf{H}^{(i)} \quad \text{e} \quad \mathbf{R}^{(i)} = \mathbf{H}^{(i)}\mathbf{R}^{(i-1)} \tag{99}$$

  onde $\mathbf{Q}^{(0)} = \mathbf{I}$ e $\mathbf{R}^{(0)} = \mathbf{A}$
- Um total de $n-1$ passos é realizado até que a matriz $\mathbf{R}^{(n-1)} = \mathbf{H}^{(n-1)}\mathbf{R}^{(n-2)}$ obtida é triangular superior
- Esse processo pode ser repetido até que a matriz $\mathbf{A}^{(n)} = \mathbf{R}^{(n-1)}\mathbf{Q}^{(n-1)}$ seja triangular. Nesse caso, os autovalores de $\mathbf{A}$ serão os elementos da diagonal de $\mathbf{A}^{(n)}$

Conteúdo

Decomposição em valores singulares

- A escrever...

Conteúdo

Bibliografia

[Che99] C.-T. Chen, *Linear system theory and design*, 3rd ed. Oxford Univeristy Press, 1999.

[GS08] A. Gilat and V. Subramanian, *Métodos numéricos para engenheiros e cientistas: uma introdução com aplicações usando o MATLAB*, 1st ed. Bookman, 2008.

[HJ12] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge University Press, 2012.

[PP08] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Technical University of Denmark, Oct. 2008, version 20081110. [Online]. Available: http://www2.imm.dtu.dk/pubdb/p.php?3274

[Rug96] W. J. Rugh, *Linear system theory*, 2nd ed., ser. Information and systems sciences, T. Kailath, Ed. Prentice Hall, 1996.

[Str88] G. Strang, *Linear algebra and its applications*, 3rd ed. Harcourt College Publishers, 1988.