Apostila de Álgebra Multilinear

Sumário

1 Álgebra Linear **2**

1.1 Revisão Matemática 2

1.2 Cálculo Matricial e Transformadas 6

1.2.1 Transformada de Laplace 7

1.2.2 Transformada Z 7

1.3 Complemento Ortogonal e Estrutura das Transformações Lineares 8

1.4 Subespaços Vetoriais Fundamentais 12

1.5 Decomposição de Transformações Lineares 15

1.6 Teoremas 17

1.7 Exercícios 26

2 Autoanálise de Sistemas **28**

2.1 Autoanálise de sistemas 28

2.1.1 Matriz de Auto-Corelação 28

2.1.2 O problema do autovalor 29

2.2 Propriedades de autovalores e autovetores associados a uma matriz Hermitiana definida não-negativa 29

2.3 Modelagem de rank baixo 36

2.4 Autofiltros 39

2.5 Cálculo de autovalores 40

2.6 Cálculo de Autovalores 40

2.7 Exercícios 42

3 Decomposição em Valores Singulares **43**

3.1 Decomposição em Valor Singular (SVD) 43

Capítulo 1

Álgebra Linear

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

1.1 Revisão Matemática

Vetores

- Vetores L.I (linearmente independentes): $\alpha\mathbf{x} + \beta\mathbf{y} = 0 \longrightarrow \alpha, \beta = 0$
- Base de espaço vetorial (E.V.): número de vetores L.I.
- Conjunto Varredura (span): é um E.V. com todos vetores formados por combinações lineares de n vetores específicos. ex:$\mathbb{R}^3 \subset \mathbb{R}^n$, x
- Norma Euclidiana:$\|\mathbf{z}\| = \left\{ \begin{array}{c} \sqrt{\mathbf{x}\mathbf{x}^\mathsf{T}} \\ \sqrt{\mathbf{x}\mathbf{x}^\mathsf{T}} \end{array} \right\} = \left[\sum_{j=1}^{n} |x_j|^2 \right]^{1/2}$,$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1}$
- $max_{1\leq i\leq n}|x_i| \leq \|\mathbf{x}\| \leq \sqrt{n}\, max_{1\leq i\leq n}|x_i|$
- $\|\mathbf{x}\| \geq 0$ OBS: Conjuntos compelxos dado por: $\mathbf{x}^*$ ou $\bar{\mathbf{x}}$
- $\|\mathbf{x}\| = 0$ se e somente se $\mathbf{x} = 0$
- $\|\alpha\mathbf{x}\| = |\alpha|\, \|\mathbf{x}\|$
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ Desigualdade Triangular.
- $\left\|\mathbf{x}^\mathsf{T}\mathbf{y}\right\| \leq \|\mathbf{x}\|\, \|\mathbf{y}\|$ Desigualdade de Cauchy-Schwarz.

Matrizes

- Matriz Nula: $\mathbf{0_{mxn}}$
- Matriz Identidade: $\mathbf{I}(m = n)$ tal que $a_{mn} = 1$ para $m = n$, caso contrário $a_{mn} = 0$
  Desta forma: $\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$
- Multiplicação de matrizes: Não comutativa. $\mathbf{AB} \neq \mathbf{BA}$
- Matriz Transposta: $(\mathbf{AB})^\intercal = \mathbf{B}^\intercal \mathbf{A}^\intercal$
- Traço de uma matriz quadrada: Soma da sua diagonal principal. $Tr(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}; Tr(\mathbf{AB}) = Tr(\mathbf{BA})$
- Determinante: função de valor escalar para uma matriz. $det(\mathbf{A})$
- $det(A)$ para expansão de Laplace: $det(\mathbf{A}) = \sum_{j=1}^{n} a_{ij} c_{ij}$; Onde c_{ij} é o cofator, calculado por: $c_{ij} = (-1)^{i+j} det[\mathbf{A_{ij}}]_{(n-1)\times(n-1)}$
- OBS:
  1. uso recursivo (det de escalar é o próprio escalar);
  2. det é a soma dos produtos de entradas (componentes) das matrizes;
  3. det é função das entradas da matriz contínua diferenciável.
- $det(\mathbf{AB}) = det(\mathbf{A})det(\mathbf{B}) = det(\mathbf{BA}), (\mathbf{A}, \mathbf{B}\ \ n \times n)$
- $\exists \mathbf{A^{-1}}$ se e somente se $det(\mathbf{A}) \neq 0, \mathbf{A^{-1}} = \frac{adj\mathbf{A}}{det(\mathbf{A})}, adj\mathbf{A} = [c_{ji}]$ Transposta da matriz de cofatores.
- $(\mathbf{AB})^{-1} = \mathbf{B^{-1}A^{-1}}, \mathbf{A}, \mathbf{B}\ \ n \times n$
- Equação de autovalores: $\mathbf{Ap} = \lambda\mathbf{p}$; Onde $\mathbf{p}$ é denominado autovetor e λ um autovalor.
- se λ e $\mathbf{p}$ são complexos, então: $\mathbf{Ap}^* = \lambda^*\mathbf{p}^*$
- Autovetor à esquerda: $\mathbf{qA} = \lambda\mathbf{q}$ para algum λ.
- Equação característica de $\mathbf{A}$:$det(\mathbf{sI_n} - \mathbf{A}) = 0$, raízes do polinômio característico S é função contínua dos coeficientes do polinômio. S é função contínua das entradas da matriz $\mathbf{A}$.
- $det(\mathbf{A}) = \lambda_1\lambda_2 \ldots \lambda_n = \prod_{i=1}^{n} \lambda_i$
- Teorema de Caley-Hamilton:
  Se $det(\mathbf{sI_n} - \mathbf{A}) = s^n + a_{n-1}s^{n-1} + \ldots + a_0$
  Então,$\mathbf{A}^{n+k}, k \in \mathbb{Z}$ pode ser escrito como combinação linear de $\mathbf{I_n}, \mathbf{A}, \mathbf{A^2}, \ldots, \mathbf{A^{n-1}}$.

- Transformação de similaridade: $\mathbf{B} = \mathbf{T^{-1}AT}, \mathbf{A} e \mathbf{T}$ são $n \times n$
$det(\mathbf{B}) = det(\mathbf{T^{-1}AT}) = det(\mathbf{T^{-1}})det(\mathbf{T})det(\mathbf{A}) = det(\mathbf{A})$.
Se $\mathbf{A}$ tem $\lambda's$ distintos e $\mathbf{T}$ é formado pelos autovetores de $\mathbf{A}$: Então $\mathbf{B} = \mathbf{\Lambda}$. Onde $\mathbf{\Lambda}$ é uma matriz diagonal contendo todos os autovalores em sua diagonal principal.

- Exemplos:
$\mathbf{A} = \begin{bmatrix} 0 & 2 \\ 2 & -2 \end{bmatrix} \rightarrow$ polinômio característico:
$det(\lambda\mathbf{I} - \mathbf{A}) = [\lambda - (-1 - j\sqrt{3})][\lambda - (-1 + j\sqrt{3})]$
$\lambda_1 : (-1 + j\sqrt{3})$;
$\lambda_2 : (-1 - j\sqrt{3})$
$$\mathbf{Ap} = \lambda\mathbf{p} \therefore \left\{ \begin{array}{ll} \lambda_1 : & \mathbf{p_1} = \begin{pmatrix} 2 \\ 1 - j\sqrt{3} \end{pmatrix} \\ \lambda_2 : & \mathbf{p_2} = \begin{pmatrix} 2 \\ 1 + j\sqrt{3} \end{pmatrix} \end{array} \right\} \rightarrow \mathbf{T} = (\mathbf{p_1 p_2})$$
$$\rightarrow \mathbf{\Lambda} = \mathbf{T^{-1}AT} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Transformações Lineares (TL)

- $\mathbf{A}_{m\times n}\mathbf{x} = \mathbf{b}_{m\times 1}$

- Espaço range (ou imagem) de A: sub-espaço de $\mathbb{R}^m$ varrido pelas colunas de A.

- Espaço nulo de A: Espaço gerados por vetores de $\mathbf{x}$ tal que $\mathbf{Ax} = \mathbf{0}$.

- Transformações Lineares têm soluções se e somente se $\mathbf{b} \in$ espaço range de $\mathbf{A}$: Se $m = n$, $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

- Rank ou posto de $\mathbf{A}$ é a dimensão do espaço range de $\mathbf{A}$: número de colunas ou de linhas (menor número) linearmente independentes de $\mathbf{A}$.

- Se $\mathbf{A}_{m\times n}$ e $\mathbf{B}_{n\times p} \rightarrow rank(\mathbf{A}) + rank(\mathbf{B}) - n \leq rank(\mathbf{AB}) \leq min\{rank(\mathbf{A}), rank(\mathbf{B})\}$

- Partições de vetores e matrizes: Operações conformáveis.
  - $\begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ 0 & \mathbf{A}_4 \end{bmatrix} + \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{B}_3 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 + \mathbf{B}_1 & \mathbf{A}_1 + \mathbf{B}_2 \\ \mathbf{B}_3 & \mathbf{A}_4 \end{bmatrix}$
  $\begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ 0 & \mathbf{A}_4 \end{bmatrix} \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{B}_3 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1\mathbf{B}_1 + \mathbf{A}_2\mathbf{B}_3 & \mathbf{A}_1\mathbf{B}_2 \\ \mathbf{A}_4\mathbf{B}_3 & 0 \end{bmatrix}$
  - $\mathbf{x}_{n\times 1}$ e $\mathbf{A}_{m\times n}$: $\begin{bmatrix} \mathbf{a}_1^\mathsf{T} \\ \mathbf{a}_2^\mathsf{T} \\ \vdots \\ \mathbf{a}_m^\mathsf{T} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{a}_1^\mathsf{T}\mathbf{x} \\ \mathbf{a}_2^\mathsf{T}\mathbf{x} \\ \vdots \\ \mathbf{a}_m^\mathsf{T}\mathbf{x} \end{bmatrix}$, $\mathbf{a}$ é coluna.
  - $\mathbf{z}_{m\times 1}$ e $\mathbf{A}_{m\times n}$: $\mathbf{z}^\mathsf{T} \left[\ \mathbf{a}_1^\mathsf{T}, \mathbf{a}_2^\mathsf{T}, \cdots \mathbf{a}_m^\mathsf{T} \ \right] = \left[\ \mathbf{z}^\mathsf{T}\mathbf{a}_1^\mathsf{T}, \mathbf{z}^\mathsf{T}\mathbf{a}_2^\mathsf{T}, \cdots \mathbf{z}^\mathsf{T}\mathbf{a}_m^\mathsf{T} \ \right]$

- Partição quadrada com blocos diagonais : $det \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} \end{bmatrix} = det[\mathbf{A}_{11}]det[\mathbf{A}_{22}]$

- Norma induzida ou norma espectral de $\mathbf{A}_{m\times n}$: baseado no problema de maximização com restrição.
  $\|\mathbf{A}\| = max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\|$ = maior autovalor de $\mathbf{A}^\mathsf{T}\mathbf{A}$ ou $\mathbf{A}\mathbf{A}^\mathsf{T} ex(1.11)$.
  exemplo: $\|\mathbf{A}\| = max_{\sqrt{x_1^2+x_2^2}=1} = \left[\sqrt{(\lambda_1 x_1 + x_2)^2 + \lambda_2^2 x_2}\right]$ Calculando λ_s' de $\mathbf{A}^\mathsf{T}\mathbf{A}$:
  $det(\lambda\mathbf{I} - \mathbf{A}^\mathsf{T}\mathbf{A}) = det\begin{bmatrix} (\lambda-\lambda_1) & -\lambda_1 \\ -\lambda_1 & \lambda-\lambda_2^2-1 \end{bmatrix} = \lambda^2 - (1+\lambda_1^2+\lambda_2^2)\lambda + \lambda_1^2\lambda_2^2$
  $\lambda = \frac{1+\lambda_1^2\lambda_2^2+\pm\sqrt{(1+\lambda_1^2\lambda_2^2)-4\lambda_1^2\lambda_2^2}}{2}$;
  $\|\mathbf{A}\| = \frac{\sqrt{(\lambda_1+\lambda_2)^2+1}+\sqrt{(\lambda_1-\lambda_2)}}{2}$

- Resultados da norma induzida:
  - $\|\mathbf{A}\| = \|\mathbf{A}^\mathsf{T}\|$
  - $\|\mathbf{A}\mathbf{x}\| \leq \|\mathbf{A}\|\|\mathbf{x}\|$
  - $\|\mathbf{A}\mathbf{B}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|$
  - $max_{i,j}|a_{ij}| \leq \|\mathbf{A}\| \leq \sqrt{nm}\, max_{i,j}|ai,j|$

- $(\mathbf{A}^*)^\mathsf{T} = \mathbf{A}^H$

Formas Quadráticas

- $\left.\begin{matrix} \mathbf{Q} & \to & n\times n \\ \mathbf{x} & \to & n\times 1 \end{matrix}\right\}$ Entradas reais $\to \mathbf{x}^\mathsf{T}\mathbf{Q}\mathbf{x}$; norma quadrática em $\mathbf{x}$.

- Supondo: $\mathbf{Q} = \mathbf{Q}^\mathsf{T}$ (simétrica) $\to \mathbf{A}^\mathsf{T}(\mathbf{Q}+\mathbf{Q}^\mathsf{T})\mathbf{x} = 2\mathbf{x}^\mathsf{T}\mathbf{Q}\mathbf{x}$; Logo: forma quadrática não muda se $\mathbf{Q} \to \frac{\mathbf{Q}+\mathbf{Q}^\mathsf{T}}{2}$ (em geral).

- Definição:
  - Se $\mathbf{x}^\mathsf{T}\mathbf{Q}\mathbf{x} \geq 0 \to \mathbf{Q}$ é semidefinida positiva ($\mathbf{Q} \succcurlyeq 0$)
  - Se $\mathbf{x}^\mathsf{T}\mathbf{Q}\mathbf{x} > 0 \to \mathbf{Q}$ é definida positiva ($\mathbf{Q} \succ 0$)
  - Se $\mathbf{x}^\mathsf{T}\mathbf{Q}\mathbf{x} = 0 \to \mathbf{x} = 0$

- $\mathbf{Q}_a \succeq \mathbf{Q}_b \to [\mathbf{Q}_a - \mathbf{Q}_b]$ é semidefinida positiva.

- $\lambda' s$ de $\mathbf{Q}$ simétrica ($\mathbf{Q} = \mathbf{Q}^H$) são reais. Prova:
  Temos que $\mathbf{A}\mathbf{p} = \lambda\mathbf{p} \to \mathbf{p}^H\mathbf{A}^H = \lambda^*\mathbf{p}^H$. Logo, multiplicando a esquerda por $\mathbf{p}^H$ temos: $\mathbf{p}^H\mathbf{A}\mathbf{p} = \lambda\mathbf{p}^H\mathbf{p} \geq 0$. Substituindo $\mathbf{p}^H\mathbf{A}$ por $\lambda^*\mathbf{p}^H$ temos $\lambda^*\mathbf{p}^H\mathbf{p} = \lambda\mathbf{p}^H\mathbf{p} \geq 0 \to \lambda = \lambda^* \geq 0$.

- Desigualde de Rayleigh-Ritz : $\lambda_{min}||\mathbf{x}||^2 \leq \mathbf{x}^\mathsf{T}\mathbf{Q}\mathbf{x} \leq \lambda_{max}||\mathbf{x}||^2$

- Se :$\mathbf{Q} \succeq 0 \to ||\mathbf{Q}|| = \lambda_{max}$ e $||\mathbf{Q}|| \leq Tr(\mathbf{Q}) \leq n||\mathbf{Q}||$

- Teste sobre $\mathbf{Q} \succeq 0$,$\mathbf{Q} = \mathbf{Q}^\mathsf{T}$: Baseado no sinal algébrco de determinantes de submatrizes de $\mathbf{Q}$.
  $\mathbf{Q} = (q_{ij})_{n\times n}$, $p = 1$,$\cdots$, n e $1 \leq i_1 < i_2 < \cdots < i_p \leq n$

Escalares :$\mathbf{Q}(i_1,\cdots,i_p) = det \begin{bmatrix} p_{i_1i_1} & pi_1i_2 & \cdots & p_{i_1i_p} \\ p_{i_2i_1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ p_{i_pi_1} & \cdots & \cdots & p_{i_pi_p} \end{bmatrix}$

Teorema:
$\mathbf{Q} \succ 0 \Longleftrightarrow \mathbf{Q}(1,2,3,\cdots,p), \succ 0,\ p = 1,2,3,\cdots,n.$
$\mathbf{Q} \prec 0 \Longleftrightarrow (-1)^p\mathbf{Q}(1,2,3,\cdots,p), \succ 0,\ p = 1,2,3,\cdots,n.$

- Se $\mathbf{Q}$ é complexa e hermitiana, $\mathbf{Q} = \mathbf{Q}^H$, logo a forma quadrática $\mathbf{x}^H\mathbf{Q}\mathbf{x}$ é real.

1.2 Cálculo Matricial e Transformadas

- Entradas são funções do tempo, em geral contínuas e com derivadas contínuas, definidas em $[t_0, t_1]$, $t_0 < t_1$ ou em $(-\infty, \infty)$.

- $\|\mathbf{x}(t)\| = \sqrt{\mathbf{x}^\mathsf{T}(t)\mathbf{x}(t)}$

- $\mathbf{A}(t)_{nxn}$ é inversível $\forall t$ se $\mathbf{A}^{-1}\ \exists\ \forall t$

- Valem as outras definições e resultados para matrizes.

- Exemplo:

$$\frac{d}{dt}[\mathbf{A}(t)\mathbf{B}(t)] = \dot{\mathbf{A}}(t)\mathbf{B}(t) + \mathbf{A}(t)\dot{\mathbf{B}}(t) \tag{1.1}$$

$$\frac{d}{dt}\int_0^t \mathbf{A}(\sigma)d\sigma = \mathbf{A}(t) \tag{1.2}$$

$$\frac{d}{dt}\int_{f(t)}^{g(t)} \mathbf{A}(t,\sigma)d\sigma = \mathbf{A}(t,g(t))g(t) - \mathbf{A}(t,f(t))f(t) + \int_{f(t)}^{g(t)} \frac{\partial}{\partial t}\mathbf{A}(t,\sigma)d\sigma \tag{1.3}$$

$$\left\|\int_{t_0}^t \mathbf{x}(\sigma)d\sigma\right\| \leq \left|\int_{t_0}^t \|\mathbf{x}(\sigma)\|d\sigma\right|,\ t \geq t_0 \tag{1.4}$$

Atenção:

$$\frac{d}{dt}\mathbf{A}(t)\mathbf{A}(t) = \dot{\mathbf{A}}(t)\mathbf{A}(t) + \mathbf{A}(t)\dot{\mathbf{A}}(t) \neq 2\mathbf{A}(t)\dot{\mathbf{A}}(t) \tag{1.5}$$

- Convergência: Importante para discussão sobre existência e unicidade de soluções para equações de estado lineares.

- Sequência infinita de vetores coluna $(n \times 1)$:

$$\{\mathbf{x}_k\}_{k=0}^{\infty} = \{\mathbf{x}_0, \mathbf{x}_1, ...\}$$

- Limite da sequência: $\|\hat{\mathbf{x}} - \mathbf{x}_k\| < \varepsilon, k > K(\varepsilon)$, converge para $\hat{\mathbf{x}}$ se

$$\lim_{\mathbf{x}\to\infty} \mathbf{x}_k = \hat{\mathbf{x}} \tag{1.6}$$

- Sequência de funções vetoriais: Há convergência uniforme em $[t_0, t_1]$ se
  - $\|\hat{\mathbf{x}}(t_a) - \mathbf{R}_k(t_a)\| < \varepsilon$ para $k > \mathbf{K}(\varepsilon), \forall t_a \in [t_0, t_1], \varepsilon > 0$

  - $\mathbf{S}_k(t) = \sum_{j=0}^{\infty} \mathbf{x}_j(t)$ e $\lim_{k\to\infty} \|\hat{\mathbf{x}}(t) - \mathbf{S}_k(t)\|$

- Então $\sum_{j=0}^{\infty} \mathbf{x}_j(t)$ converge uniformemente para $\hat{\mathbf{x}}(t)$ em $[t_0, t_1]$ se $\mathbf{S}_k(t)$ converge uniformemente para $\hat{\mathbf{x}}(t)$ em $[t_0, t_1]$.

Teorema 1. *Se $\sum_{j=0}^{\infty} \mathbf{x}_j(t)$, com $\mathbf{x}_j(t)$ contínuo em $[t_0, t_1]$, converge uniformemente para $\mathbf{x}_j(t)$ logo $\mathbf{x}_j(t)$ é contínuo $\forall \in [t_0, t_1]$.*

Teorema 2. *Se $\sum_{j=0}^{\infty} \mathbf{x}_j(t)$, com $\mathbf{x}_j(t)$ continuamente diferenciável em $[t_0, t_1]$, converge uniformemente para $\hat{\mathbf{x}}(t)$ se $\sum_{j=0}^{\infty} \frac{d}{dt}\mathbf{x}_j(t)$ converge uniformemente em $[t_0, t_1] \to$ ela converge para $\frac{d}{dt}\hat{\mathbf{x}}(t)$.*

- Observação: $\sum_{j=0}^{\infty} \mathbf{x}_j(t)$ converge absolutamente se $\sum_{j=0}^{\infty} \|\mathbf{x}_j\|$ converge em $[t_0, t_1]$.

- Propriedade chave: ordenamento da série é indiferente em relação à convergência.

1.2.1 Transformada de Laplace

Definição 1. *Seja $\mathbf{f}(t) : \mathbb{R}^{N\times N} \to \mathbb{R}^{N\times N}$ uma função matricial. A sua transformada de Laplace é definida como*

$$\mathbf{F}(s) = \int_0^{\infty} \mathbf{f}(t) e^{-st} dt = \mathcal{L}[\mathbf{f}(t)] \tag{1.7}$$

em que $s = a + jb \in \mathbb{C}$, $a, b \in \mathbb{R}$.

Algumas propriedades:

1. Diferenciação: $\mathcal{L}[\dot{\mathbf{f}}(t)] = s\mathcal{L}[\mathbf{f}(t)] - \mathbf{f}(0)$
2. Integração: $\mathcal{L}\left[\int_0^{\infty} \mathbf{f}(\sigma) d\sigma\right] = \frac{1}{s}\mathcal{L}[\mathbf{f}(t)]$
3. Convolução: $\mathcal{L}\left[\int_0^{\infty} \mathbf{F}(t-\sigma)\mathbf{G}(\sigma) d\sigma\right] = \mathcal{L}[\mathbf{F}(t)]\mathcal{L}[\mathbf{G}(t)]$
4. Teorema do Valor Inicial: $\lim_{t\to 0} \mathbf{F}(t) = \lim_{s\to\infty} s\mathbf{F}(s)$
5. Teorema do Valor Final: $\lim_{t\to\infty} \mathbf{F}(t) = \lim_{s\to 0} s\mathbf{F}(s)$
6. $\mathbf{F}(s)$ polinomial $\to \mathbf{F}^{-1}(s)$ não é sempre polinomial.

1.2.2 Transformada Z

Definição 2. *Seja $\mathbf{u}(n) : \mathbb{R}^{N\times N} \to \mathbb{R}^{N\times N}$ uma função matricial. A sua transformada Z é definida como*

$$\mathbf{U}(z) = \sum_{n=-\infty}^{\infty} \mathbf{u}(n) z^{-n} = \mathcal{Z}[\mathbf{u}(n)] \tag{1.8}$$

em que $z \in \mathbb{C}$.

- Utilizada para frequências temporais escalares, vetoriais e matriciais.

- Tipos de entradas: Sequências na forma de somas com termos do tipo $k^r\lambda^k$, $k = 0, 1, 2, \ldots$, com versões deslocadas destas sequências, $\lambda \in \mathbb{C}$, $r \in \mathbb{Z}$, $r > 0$.

- Fazendo: $r = \lambda = 0 \rightarrow$ Delta de Kronecker: $\delta(k) = 1$ se $k = 0$ e $\delta(k) = 0$, caso contrário.
- Mudança de notação temporal para vetores e matrizes: sub-índice$\rightarrow$ argumento: $\mathbf{x}_k \rightarrow \mathbf{x}(k)$.
- Tipos de $\mathbf{F}(z)$ mais usuais: Funções racionais (razões de polinômios).
- $\mathbf{Z}^{-1}$: $\mathbf{F}(k) = \mathbf{Z}^{-1}\{\mathbf{F}(z)\}$, e região de convergência (R.C.)

Algumas propriedades:

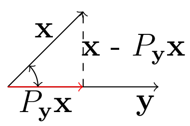
1. $\mathbf{Z}\{\mathbf{F}(k-1)\} = z^{-1}\mathbf{Z}\{\mathbf{F}(k)\}$
2. $\mathbf{Z}\{\mathbf{F}(k+1)\} = z\mathbf{Z}\{\mathbf{F}(k)\} - z\mathbf{F}(0)$
3. $\mathbf{Z}\{\sum_{j=0}^{k} \mathbf{F}(k-j)\mathbf{H}(j)\} = \mathbf{Z}\{\mathbf{F}(k)\}\mathbf{Z}\{\mathbf{H}(k)\}$
4. Teorema do Valor Inicial: $\mathbf{F}(0) = \lim_{z\to\infty} \mathbf{F}(z)$
5. Teorema do Valor Final: $\lim_{k\to\infty} \mathbf{F}(k) = \lim_{z\to 1}(z-1)\mathbf{F}(z)$

1.3 Complemento Ortogonal e Estrutura das Transformações Lineares

Projeção Ortogonal de x em y

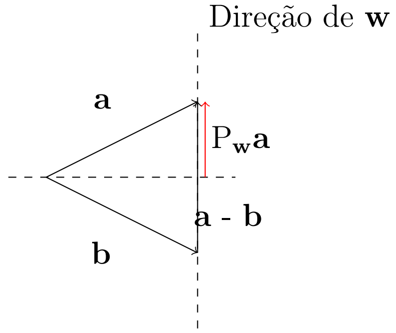
$$P_{\mathbf{y}}\mathbf{x} = \left[\frac{\mathbf{y}\,\mathbf{y}^{\mathsf{T}}}{\|\mathbf{y}\|^2}\right]\mathbf{x} \tag{1.9}$$

Na definição, $\mathbf{y}\,\mathbf{y}^{\mathsf{T}}$ é uma matriz cujas colunas são todas linearmente dependentes entre si.



Projeção Ortogonal em Subespaço Se $\mathbf{S}$ é um subespaço de $\mathbb{R}^n$ varrido por um conjunto ortonormal $\{\mu_1, \ldots, \mu_m\}$, então $P_{\mathbf{S}}\mathbf{x} \triangleq \sum_{i=1}^{m} P_{\mu_i}\mathbf{x}$.

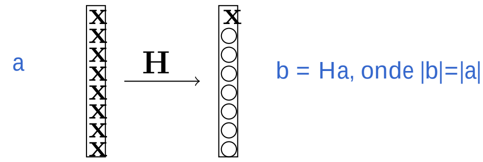
Transformação de Householder Sejam $\mathbf{a}$ e $\mathbf{b}$ vetores com magnitudes iguais: $\|a\| = \|b\|$



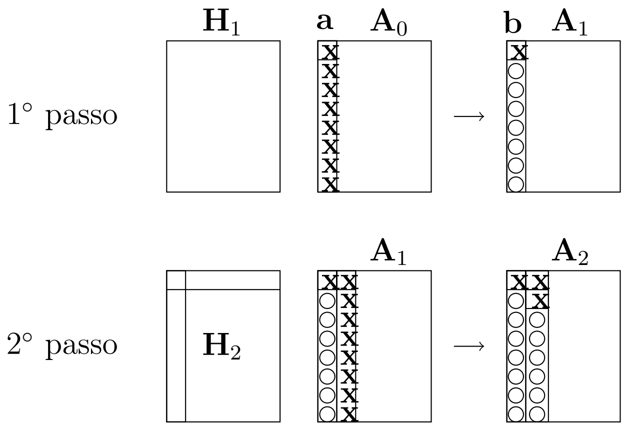
Objetivo: Girar $\mathbf{a}$ para fornecer $\mathbf{b}$: $\mathbf{b} = \mathrm{H}\mathbf{a}$. Seja $\mathbf{w} = \gamma(\mathbf{a}-\mathbf{b})$ o vetor de Householder. Fazendo $\|b\| = \|a\|$, tem-se $\mathbf{a} - \mathbf{b} = 2\,\mathrm{P}_{\mathbf{w}}\mathbf{a} \;\therefore\; \mathbf{b} = \mathbf{a} - 2\,\mathrm{P}_{\mathbf{w}}\mathbf{a}$. Therefore,

$$\mathbf{b} = \underbrace{\left[\mathrm{I} - \frac{2\mathbf{w}\,\mathbf{w}^{\mathsf{T}}}{\|\mathbf{w}\|^2}\right]}_{\text{Matriz de Householder, }\mathbf{H}} \mathbf{a}, \tag{1.10}$$

Aplicação: Transformar um vetor cheio em um vetor com um só elemento não-nulo.



Aplicação: Transformar uma matriz cheia em uma matriz triangular.



Complemento Ortogonal de um Subespaço

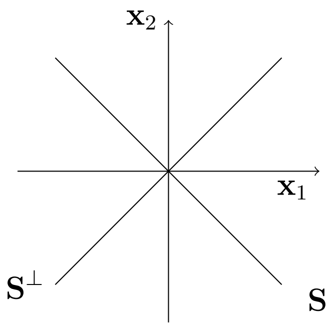
$$\mathbf{S}^{\perp} \triangleq \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \perp \mathbf{S}\} \tag{1.11}$$

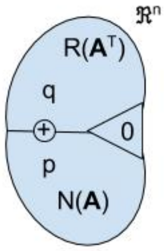
Observação Note $\mathbf{S}^{\perp}$ também é um subespaço de $\mathbb{R}^n$ e $\mathbf{S} \perp \mathbf{S}^{\perp}$.

Os subespaços $\mathbf{S}$ e $\mathbf{S}^{\perp}$ decompõem $\mathbb{R}^n$ no sentido de que qualquer $\mathbf{x} \in \mathbb{R}^n$ pode ser unicamente expresso por

$$\mathbf{x} = \mathbf{P}_S\mathbf{x} + \mathbf{P}_{S^{\perp}}\mathbf{x} \triangleq \mathbf{y} + \mathbf{z},\; \mathbf{y} \perp \mathbf{z} \tag{1.12}$$

Diz-se ainda que $\mathbb{R}^n$ é a soma direta de $\mathbf{S}$ e $\mathbf{S}^{\perp}$: $\mathbf{S} \oplus \mathbf{S}^{\perp}$, onde $\mathbf{S} = \{\mathbf{x} \in \mathbb{R}^2 \| \, \mathbf{x}_1 + \mathbf{x}_2 = 0\}$ e $\mathbf{S}^{\perp} = \{\mathbf{x} \in \mathbb{R}^2 \| \, \mathbf{x} \perp \mathbf{S}\}$.





Representação Pictográfica

Observação: Não se trata do diagrama de Venn.

Estrutura das Transformações Lineares

Transformação Linear: a multiplicação de um vetor $\mathbf{x}$ no $\mathbb{R}^n$ por uma matriz $\mathbf{A}$ produz o vetor:

$$\mathbf{y} = \mathbf{Ax} = \begin{bmatrix} \alpha_1^\mathsf{T} \\ \alpha_2^\mathsf{T} \\ \vdots \\ \alpha_m^\mathsf{T} \end{bmatrix}_{m\times n} \mathbf{x} = \begin{bmatrix} \alpha_1^\mathsf{T}\mathbf{x} \\ \alpha_2^\mathsf{T}\mathbf{x} \\ \vdots \\ \alpha_m^\mathsf{T}\mathbf{x} \end{bmatrix}_{m\times 1}$$

onde $\mathbf{y}$ está no $\mathbf{R}^m$, sendo esta operação conhecida como transformação linear do $\mathbb{R}^n$ no $\mathbb{R}^m$, representada por $\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^m$.

Formas de representação de matrizes particionadas em vetores

$$\mathbf{A} = \begin{bmatrix} \alpha_1^\mathsf{T} \\ \alpha_2^\mathsf{T} \\ \vdots \\ \alpha_m^\mathsf{T} \end{bmatrix} \qquad \mathbf{A}^\mathsf{T} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_m \end{bmatrix}$$

Outra maneira de representar a mesma transformação

$$\mathbf{y} = \mathbf{Ax} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} x_1\alpha_1 & x_2\alpha_2 & \cdots & x_m\alpha_m \end{bmatrix}_{m\times 1}$$

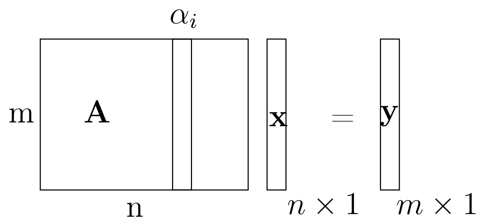
logo, $\mathbf{Ax}$ é uma combinação linear das n colunas de $\mathbf{A}$.

A matriz $\mathbf{A}^\mathsf{T}$ é também uma transformação linear $\mathbf{A}^\mathsf{T} : \mathbb{R}^m \to \mathbb{R}^n$.

Definição 3. *O espaço range (ou imagem, varrido) da matriz* $\mathbf{A}$ *é o subespaço de* $\mathbb{R}^m$ *definido por*

$$R(\mathbf{A}) \triangleq \{\mathbf{y} \in \mathbb{R}^m \| \mathbf{y} = \mathbf{Ax},\ \textit{para algum}\, \mathbf{x} \in \mathbb{R}^n\} \tag{1.13}$$

Observação Voltando ao fato que $\mathbf{AX}$ é uma combinação linear das colunas de $\mathbf{A}$:



Em geral, o subespaço range de $\mathbf{A}$: $R(\mathbf{A}) \subset \mathbb{R}^m$, ou seja, depende do número de vetores α_i de $\mathbf{A}$ que são linearmente independentes (LI). Se $\exists m(m = n)$ vetores $\alpha_i, i = 1, \ldots, m$ LI, então $R(\mathbf{A}) = \mathbb{R}^m$. Assim, todo espaço de ~~$\mathbf{y}$~~ é varrido. Agora, se $m > n$, mesmo que todos α_i sejam LI, $R(\mathbf{A})$ será sempre uma parte de $\mathbb{R}^m$, pois há n colunas LI de $\mathbf{A}$ e são m linhas de $\mathbf{A}$. Por outro lado, se $m < n$, haverá $(n - m)$ vetores α_i que são LI e $R(\mathbf{A}) = \mathbb{R}^m$.

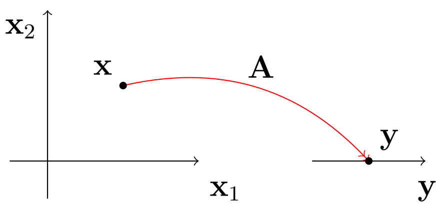
Exemplos

01. $\mathbf{A} = \begin{bmatrix} 1 & 2 \end{bmatrix}$, $m = 1$, $n = 2$ m<n

$\mathbf{A} : \mathbb{R}^2 \to \mathbb{R}$

$R(\mathbf{A}) \equiv \mathbb{R}$

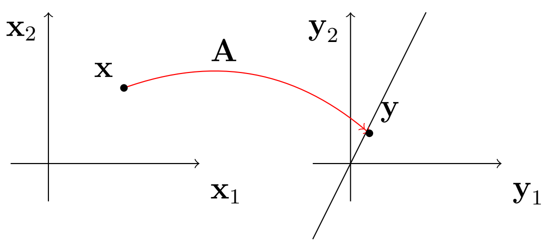
$$\mathbf{y} = \mathbf{Ax} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 + 2x_2$$



02. $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$, $m = 2$, $n = 2$ m=n com 1 vetor LI (matriz degenerativa)

$\mathbf{A} : \mathbb{R}^2 \to R(\mathbf{A}) \subset \mathbb{R}^2$

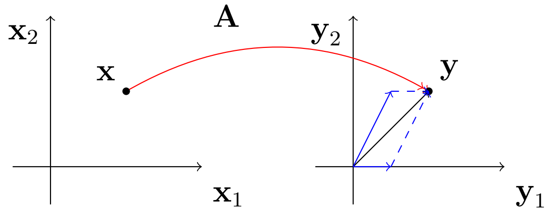
$$\mathbf{y} = \mathbf{Ax} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ 2(x_1 + x_2) \end{bmatrix}$$



03. $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $m = 2$, $n = 2$

m=n com 2 vetor LI (full-rank)

$\mathbf{y} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathbf{x}_1 + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathbf{x}_2$

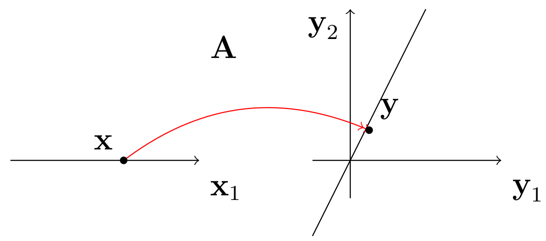


04. $\mathbf{A} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $m = 2$, $n = 1$

m>n

$\mathbf{y} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathbf{x}_1$

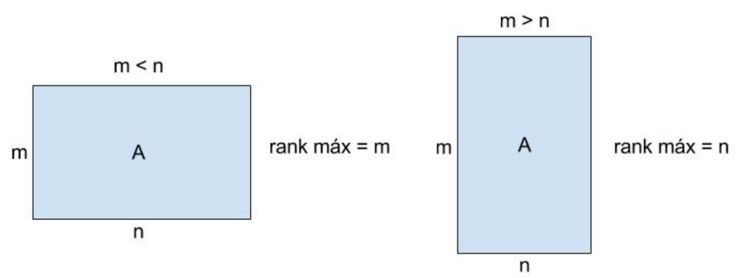
$R(\mathbf{A}) = \{\mathbf{y} \in \mathbb{R}^2 \| \mathbf{y}_1 = \mathbf{y}_2\}$



1.4 Subespaços Vetoriais Fundamentais

Definição 4. *O rank (ou posto) de* $\boldsymbol{A}$ *é a dimensão do* $R(\boldsymbol{A})$ *(do espaço range), isto é, é igual ao número de de colunas (ou de linhas) L.I. de* $\boldsymbol{A}$*.*

- ~~O máximo rank de~~ $\cancel{\mathbf{A}\ (m \times n)}$ ~~é~~ $\cancel{m}$~~, que é a dimensão de~~ $\cancel{\mathbf{y} = \mathbf{A}\mathbf{x}}$~~.~~
- A matriz $\mathbf{A}$ é dita de *rank completo*, se o rank($\mathbf{A}$) for igual à menor dimensão da matrix $\mathbf{A}$.



* **Lembrete**: Espaço range: $R(\mathbf{A}) \triangleq \{y \in \mathbb{R}^m \mid \mathbf{y} = \mathbf{A}\mathbf{x}$, p/ algum $\mathbf{x} \in \mathbb{R}^n\}$

Definição 5. *O espaço nulo de* $\boldsymbol{A}$ *é o sub-espaço de* $\mathbb{R}^n$ *definido por:*

$$N(\boldsymbol{A}) \triangleq \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{Ax} = \boldsymbol{0}\}$$ A is a m-by-n matrix

Logo, $\exists\ \mathbf{a'_k} \in \mathbb{R}^n \mid \mathbf{a'_k}.\mathbf{x} = \mathbf{0}$, $\forall\ k \in [1, m]$, onde $\mathbf{a'_k}$ representa as linhas de $\mathbf{A}$.

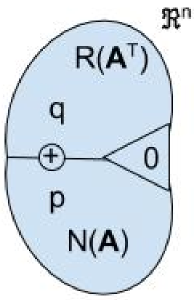
Assim, $\mathrm{N}(\mathbf{A})$ é constituído de todos os vetores $\mathbf{x}$ que são ortogonais às linhas de $\mathbf{A}$, ou às colunas de $\mathbf{A}^\top$. Consequentemente, como:

$$R(\mathbf{A}^\top) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{A}^\top\mathbf{y},\ \text{p/ algum } \mathbf{y} \in \mathbb{R}^m\}$$

então: $\mathrm{N}(\mathbf{A}) = \mathrm{R}(\mathbf{A}^\top)^\perp$, pois $\mathbf{x} \perp \mathbf{a'_i}$, $i \in [1, m]$.
onde: $\mathrm{R}(\mathbf{A}^\top) \oplus \mathrm{R}(\mathbf{A}^\top)^\perp = \mathbb{R}^n$.

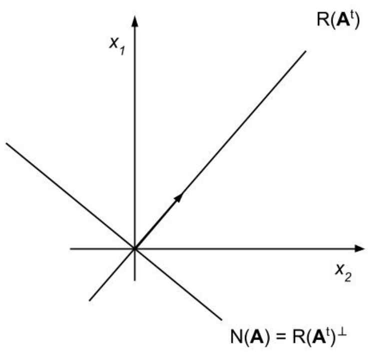
Propriedade: A nulidade de $\mathbf{A}$ é a dimensão de $\mathrm{N}(\mathbf{A})$ e o $\mathrm{rank}(\mathbf{A}^\top)$ + $\mathrm{nulidade}(\mathbf{A})$ = n, pois $\mathbb{R}^n = \mathrm{R}(\mathbf{A}^\top) \oplus \mathrm{R}(\mathbf{A}^\top)^\perp$.

Se $\exists\{\mathbf{a_k}^\perp \in \mathbb{R}^n \mid \mathbf{a_k}^\perp \perp \mathbf{x}\}$ e $\exists\{\mathbf{a_i}' \in \mathbb{R}^n \mid \mathbf{x} = \sum_{i=1}^m \mathbf{y_i}\mathbf{a_i}'\}$, então $\{\mathbf{a_k}^\perp \perp \mathbf{a_i}'\}$ e como $\mathrm{R}(\mathbf{A}^\top) \oplus \mathrm{R}(\mathbf{A}^\top)^\perp = \mathbb{R}^n$, então $\mathrm{R}(\mathbf{A}^\top)^\perp = N(\mathbf{A})$ $\therefore \dim \mathrm{R}(\mathbf{A}^\top) + \dim \mathrm{R}(\mathbf{A}^\top)^\perp = n \Longrightarrow$ Se $\dim \mathrm{R}(\mathbf{A}^\top) = q$ e $\dim \mathrm{N}(\mathbf{A}) = p \Longrightarrow p + q = n$.



Exemplo: $\mathbf{A} = \begin{bmatrix} 1 & 1 \end{bmatrix}$ $\therefore$ $\mathbf{Ax} = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$ $\therefore$ $x_1 + x_2 = 0$.

$\mathbf{A}^\top = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\therefore$ $R(\mathbf{A}^\top) \perp R(\mathbf{A}^\top)^\perp = N(\mathbf{A})$



Teorema 3. *Rank(*$\boldsymbol{A}$*) + nulidade (*$\boldsymbol{A}$*) = n.*

Seja $\mathbf{A}_{m\times n}$ e uma base para o espaço nulo de A:

$$\text{base de } N_x(\mathbf{A}) : \{\mu_\mathbf{1}, \mu_\mathbf{2}, ..., \mu_\mathbf{p}\} \Longrightarrow \text{nulidade de } \mathbf{A} = p$$

Então, uma base para $N_x(\mathbf{A})^\perp$ é: $\{\boldsymbol{\mu}_{\mathbf{p+1}}, \boldsymbol{\mu}_{\mathbf{p+2}}, ..., \boldsymbol{\mu}_{\mathbf{n}}\}$

Note que $\mathbb{R}^n = N_x(\mathbf{A}) \oplus N_x(\mathbf{A})^\perp$, logo $\{\boldsymbol{\mu}_{\mathbf{1}}, \boldsymbol{\mu}_{\mathbf{2}}, ..., \boldsymbol{\mu}_{\mathbf{n}}\}$ forma uma base para $\mathbb{R}^n$, tal que é possível escrever: $\mathbf{x} = \mathbf{x_N} + \mathbf{x_{N^\perp}}$.

Na verdade, o objetivo é mostrar que rank(A)=rank(A^T)

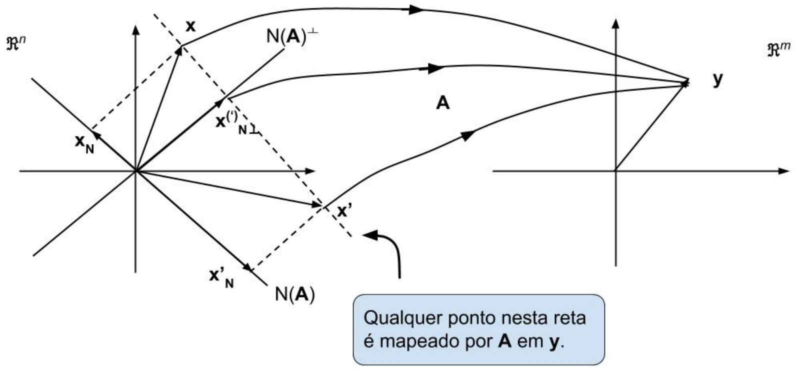
Objetivo: Mostrar que $\{\mathbf{A}\boldsymbol{\mu}_{\mathbf{p+1}}, \mathbf{A}\boldsymbol{\mu}_{\mathbf{p+2}}, ..., \mathbf{A}\boldsymbol{\mu}_{\mathbf{n}}\}$ é uma base para $\mathrm{R}(\mathbf{A})$.

Seja a transformação: $\mathbf{y} = \mathbf{Ax} = \mathbf{A}(\mathbf{x_N} + \mathbf{x_{N^\perp}}) = \mathbf{Ax_N} + \mathbf{Ax_{N^\perp}} = \mathbf{0} + \mathbf{Ax_{N^\perp}}$

A é uma matrix m-por-n

Dimensionalidades:

*dim(null(A)) = p
*dim(range(A^T)) = n-p
*null(A)=range(A^T)^\perp



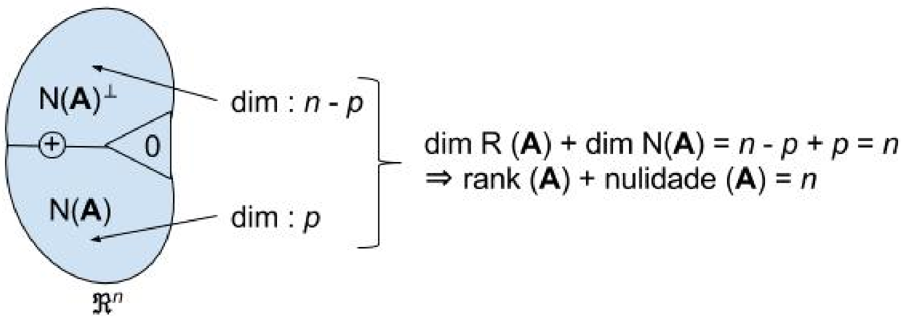
Fazendo o caminho de volta: de y para x

Figura 1.1: Ilustração da decomposição e sua transformação.

Como: $\mathbf{x}_{N^\perp} = \alpha_{p+1}\boldsymbol{\mu}_{\mathbf{p+1}} + ... + \alpha_n\boldsymbol{\mu}_{\mathbf{n}} \Longrightarrow dim\ N(\mathbf{A}^\perp) = n - p$,=dim R(A^T) e $\mathbf{y} = \mathbf{y_{N_y}} + \mathbf{y_{N_y^\perp}}$, onde $\mathbf{A}^\mathsf{T}\mathbf{y_{N_y}} = \mathbf{0}$ e $\mathbf{A}^\mathsf{T}\mathbf{y_{N_y^\perp}} \in \mathbb{R}^n \Longrightarrow \mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{A}^\mathsf{T}\mathbf{y_{N_y^\perp}} = \mathbf{A}^\mathsf{T}\mathbf{Ax_{N_x^\perp}} \therefore$ $\mathbf{y_{N_y^\perp}} = \mathbf{Ax_{N_x^\perp}}$.

Então: $\mathbf{y_{N_y^\perp}} = \alpha_{p+1}\mathbf{A}\boldsymbol{\mu}_{\mathbf{p+1}} + ... + \alpha_n\mathbf{A}\boldsymbol{\mu}_{\mathbf{n}} \Longrightarrow dim\ = (n-p)$ e $\mathbf{y_{N_y^\perp}} \in R(\mathbf{A})$.

Logo a dim $\mathrm{R}(\mathbf{A})$ é a mesma dimensão do complemento ortogonal do espaço nulo de A: $\mathrm{N}(\mathbf{A})^\perp \Longrightarrow$ rank $(\mathbf{A})$ = rank $(\mathbf{A}^\mathsf{T})$



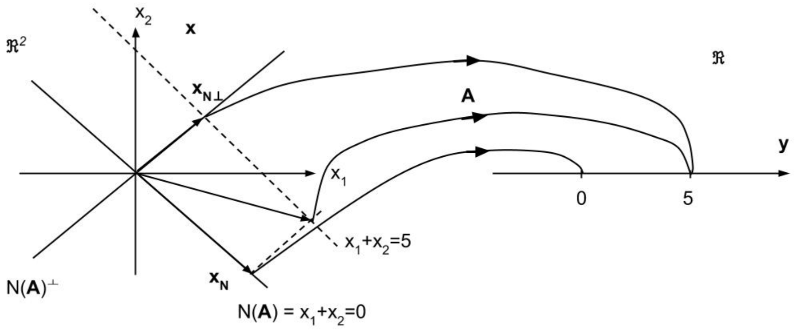
Exemplo: $\mathbf{Ax} = \mathbf{0}$; $A = \begin{bmatrix} 1 & 1 \end{bmatrix}$, $n = 2$, $x_1 + x_2 = 5$.

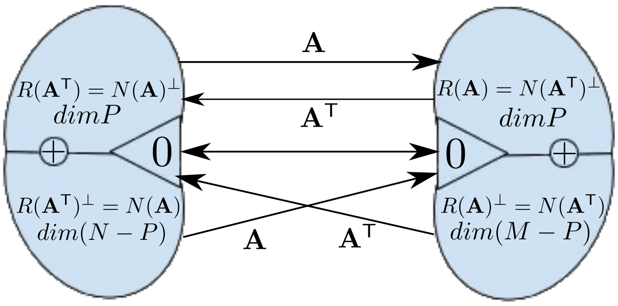
Lema: Seja a transformação linear:

$$\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^m$$

$$\text{então: } \mathrm{R}(\mathbf{A}) = \mathrm{R}(\mathbf{AA}^\mathsf{T})$$

ou seja, a independência linear em $\mathbf{A}$ é mantida em $(\mathbf{AA}^\mathsf{T})$.





$$\mathbb{R}^N = N(\mathbf{A}) \bigoplus R(\mathbf{A}^\mathsf{T}) \qquad \mathbb{R}^M = N(\mathbf{A}^T) \bigoplus R(\mathbf{A})$$

$$\mathbf{x} = \mathbf{x}_N + \mathbf{x}_{N^\perp} \qquad \mathbf{y} = \mathbf{y}_N + \mathbf{y}_{N^\perp}$$

Figura 1.2: Figura

Demonstração. 1. Se $\mathbf{y} \in R(\mathbf{A}\mathbf{A}^\mathsf{T})$, então: $\mathbf{y} = \mathbf{A}(\mathbf{A}^\mathsf{T}\mathbf{z})$, p/ algum $\mathbf{z} \in \mathbb{R}^m$. Então $\exists\ \mathbf{z} \in \mathbb{R}^m$, tal que $\mathbf{A}^\mathsf{T}\mathbf{z} = \mathbf{x} \in \mathbb{R}^n$, então $\mathbf{y} = \mathbf{A}\mathbf{x}$, e concluímos que $\mathbf{y} \in R(\mathbf{A})$; portanto: se $\mathbf{y} \in R(\mathbf{A}\mathbf{A}^\mathsf{T}) \implies \mathbf{y} \in R(\mathbf{A})$. (*)

2. Se $\mathbf{y} \in R(\mathbf{A})$, então: $\mathbf{y} = \mathbf{A}\mathbf{x}$, p/ algum $\mathbf{x} \in \mathbb{R}^n$ Agora, $\mathbf{x}$ pode ser decomposto em $\mathbf{x} = \mathbf{x_1} + \mathbf{x_2}$, onde:
   $\mathbf{x_1} \in R(\mathbf{A}^\mathsf{T}) \implies \mathbf{x_1} = \mathbf{A}^\mathsf{T}\mathbf{z}$, p/ algum $\mathbf{z} \in \mathbb{R}^m$
   $\mathbf{x_2} \in R(\mathbf{A}^\mathsf{T})^\perp = N(\mathbf{A}) \implies \mathbf{A}\mathbf{x_2} = \mathbf{0}$ Então temos: $\mathbf{y} = \mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{x_1} + \mathbf{x_2}) = \mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{z} + \mathbf{A}\mathbf{x_2} = \mathbf{A}\mathbf{A}$
   p/ algum $\mathbf{z} \in \mathbb{R}^m$, e portanto, $\mathbf{y} \in R(\mathbf{A}\mathbf{A}^\mathsf{T})$.
   Assim, se $\mathbf{y} \in R(\mathbf{A}) \implies \mathbf{y} \in R(\mathbf{A}\mathbf{A}^\mathsf{T})$ (**)

Finalmente, considerando (*) e (**): $R(\mathbf{A}) = R(\mathbf{A}\mathbf{A}^\mathsf{T}) \therefore$ dim $R(\mathbf{A}) =$ dim $R(\mathbf{A}\mathbf{A}^\mathsf{T}) \therefore$ rank $(\mathbf{A}) =$ rank $(\mathbf{A}\mathbf{A}^\mathsf{T})$ □

1.5 Decomposição de Transformações Lineares

Decomposição de transformação linear: $\mathbf{A} : \mathbb{R}^N \to \mathbb{R}^M$, $\mathbf{A} \in \mathbb{R}^{M \times N}$

- Caso $M < N$: Temos rank$(\mathbf{A}) = M$. Além disso, temos que $p = M$, pois rank$(\mathbf{A}) =$ dim$R(\mathbf{A}) = p = M$.

- Caso $M > N$: Temos rank$(\mathbf{A}) = N$, pois rank$(\mathbf{A}) =$ rank$(\mathbf{A}^\mathsf{T}) = p = N$

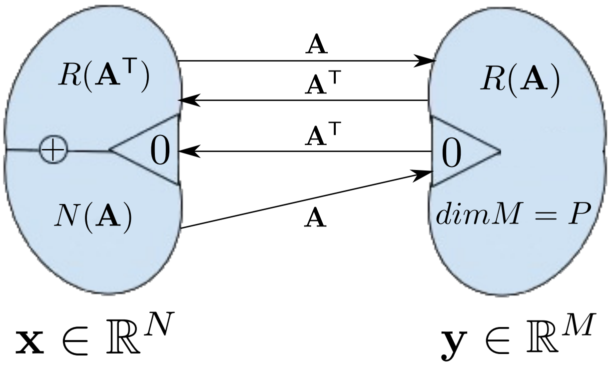


Figura 1.3: Figura caso1

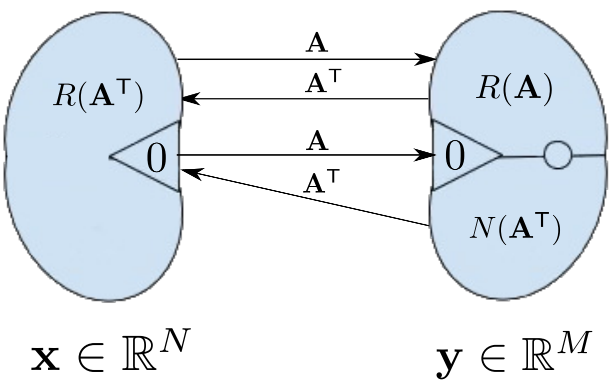


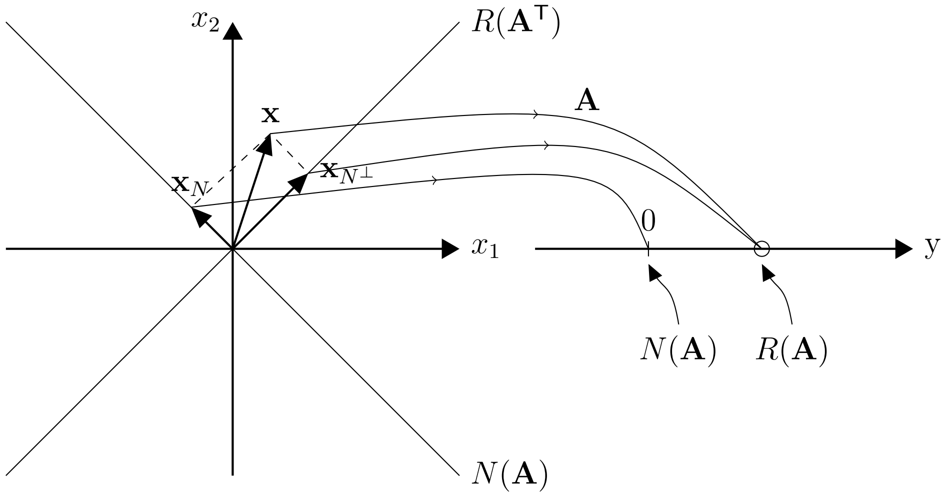
Figura 1.4: Figura caso2

Exemplo: Seja a transformação linear $y = \mathbf{A}\mathbf{x}$, em que $\mathbf{A} = \begin{bmatrix} 1 & 1 \end{bmatrix}$ e $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\mathsf{T}$. Assim

$$y = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 + x_2 \tag{1.14}$$

Portanto

- $N(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^2 \,|\, \mathbf{A}\mathbf{x} = \mathbf{0}\} = \{\mathbf{x} \in \mathbb{R}^2 \,|\, x_1 + x_2 = 0\}$
- $R(\mathbf{A}) = \{y \in \mathbb{R} \,|\, y = \mathbf{A}\mathbf{x},\, \forall \mathbf{x} \in \mathbb{R}^2\}$
- $N(\mathbf{A}^\mathsf{T}) = \{y \in \mathbb{R} \,|\, \mathbf{A}^\mathsf{T} y = 0\}$
- $R(\cancel{\mathbf{A}}) = \{\mathbf{x} \in \mathbb{R}^2 \,|\, \mathbf{x} = \mathbf{A}^\mathsf{T} y,\, \forall y \in \mathbb{R}\} = \{\mathbf{x} \in \mathbb{R}^2 \,|\, x_1 = y,\, x_2 = y, \forall y \in \mathbb{R}\}$



1.6 Teoremas

Teorema a

A equação linear $\mathbf{A}\mathbf{x} = \mathbf{b}$ tem solução (pode ser 1 ou mais) sss $\mathbf{b} \in R(\mathbf{A}) = R(\mathbf{A}\mathbf{A}^\top)$, ou o que da no mesmo: $\mathbf{b} \perp R(\mathbf{A}\mathbf{A}^\top)$. Prova diagramática:

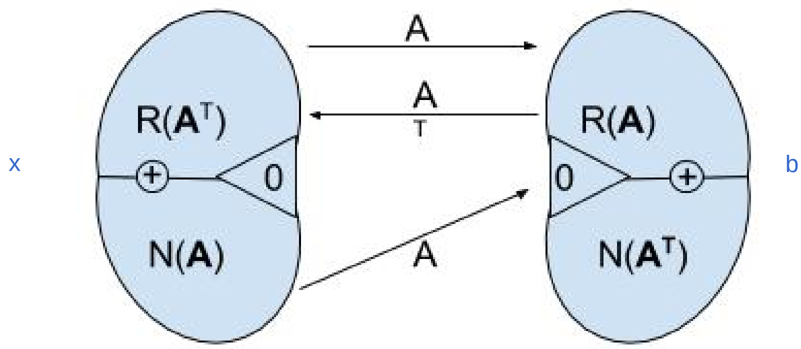


Figura 1.5: Temos que $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{b} \in \mathbb{R}^m$.

Obs: Este problema só tem solução se $\mathbf{b}$ não tem qualquer componente em $N(\mathbf{A}^\top)$

Teorema b

Há no máximo uma solução sss: $N(A) = N(AA^{\top}) = 0$, ou o que da no mesmo: $rank(A^{\top}) = rank(A) = n$.

Obs: Pode ser que nem tenha solução, basta que $\mathbf{b}$ tenha componentes em $N(A^{\top}) = R(A)^{\perp}$.

Prova diagramática:

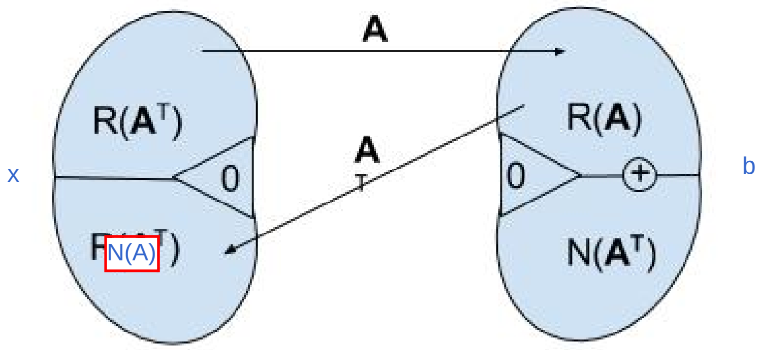


Figura 1.6: Temos que $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{b} \in \mathbb{R}^m$.

Teorema c

$\exists$ pelo menos uma solução para cada $\mathbf{b} \in \mathbb{R}^{\top}$ sss $R(\mathbf{A}) = R(\mathbf{A}\mathbf{A}^{\top})(N(\mathbf{A}^{\top}) = N(\mathbf{A}^{\top}\mathbf{A}) = \{0\})$. Prova diagramática:

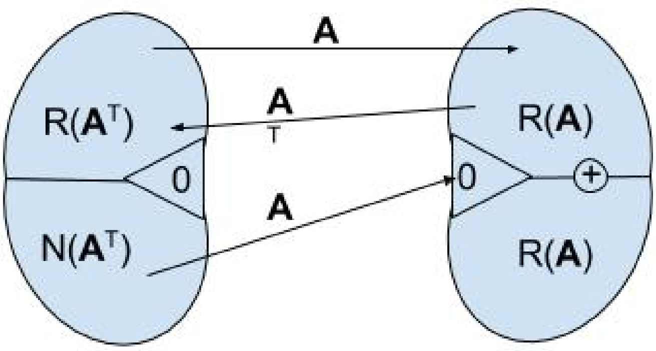


Figura 1.7: Temos que $\mathbf{b} \in \mathbb{R}^m$.

Obs: Este fato está ligado à questão da estimação de estado.

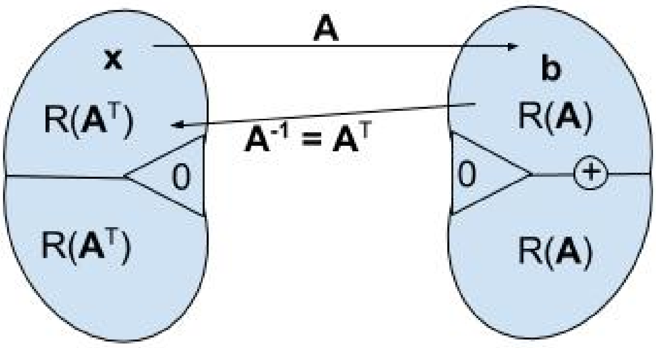


Figura 1.8: Temos que $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{b} \in \mathbb{R}^m$.

Teorema d

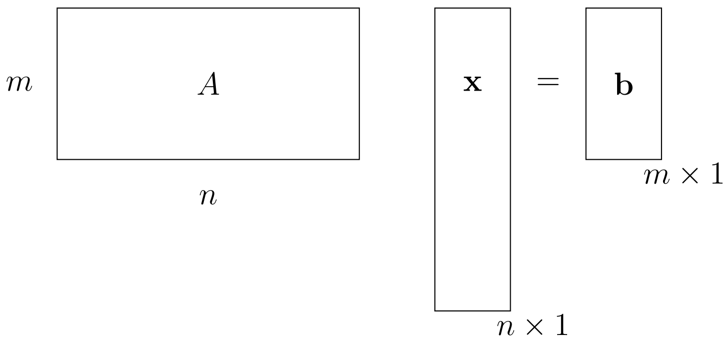
$\exists$ uma e somente uma solução para cada $\mathbf{b}$ sss $\mathbf{A}$ for quadrada e não singular, caso em que a solução é $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Prova:

Obs: Neste caso : $rank(\mathbf{A}) = rank(\mathbf{A}^\top) = m = n$ pois $dimR(\mathbf{A}^\top) = p$ e $dimR(\mathbf{A}) = p$.

PS: $\mathbf{b}$ não tem componente em $N(\mathbf{A}^\top)$, $\mathbf{x}$ não tem componente em $N(\mathbf{A})$.

Teorema e

Seja A uma matriz $m \times n$; $m < n$, com rank completo m, ou seja:



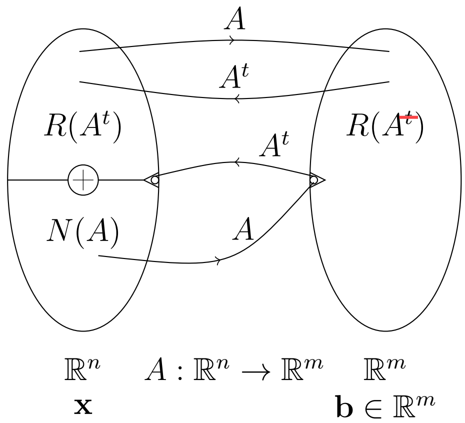
então:

a) $A\mathbf{x} = \mathbf{b}$ tem pelo menos uma solução para qualquer $\mathbf{b} \in \mathbb{R}^m$;

b) AA^t é não-singular, onde a matriz $[A^t(AA^t)^{-1}]$ é a pseudo-inversa.

Prova:

a)



logo: $\mathbf{b} \in R(A)$ (sempre há uma solução)

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

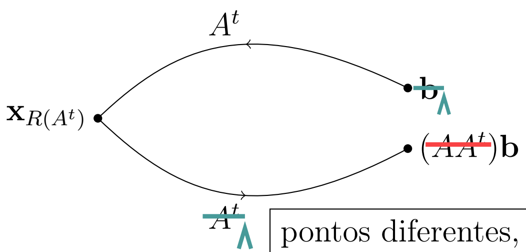
$dim(R(A)) + nulidade(A^t) = m$

como $dimR(A) = m$, $nulidade(A^t) = 0$ e $N(A^t) = \emptyset$

b) $rank(A) = rank(AA^t) = m \rightarrow AA^t$ é uma matriz $m \times m$ de rank m, logo é não-singular!

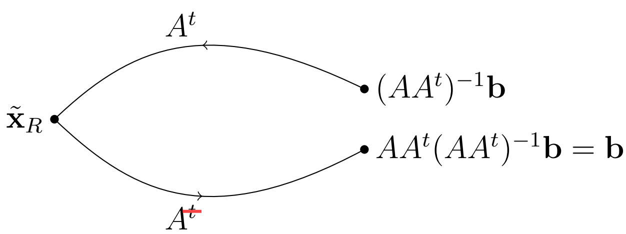
façamos: $\mathbf{x} \in \mathbb{R}^n, \mathbf{x} = \mathbf{x}_{R(A^t)} + \mathbf{x}_{N(A)} \therefore$
$A\mathbf{x} = A[\mathbf{x}_{R(A^t)} + \mathbf{x}_{N(A)}] = A\mathbf{x}_{R(A^t)} + A\mathbf{x}_{N(A)} = A\mathbf{x}_{R(A^t)} = \mathbf{b}$

1o passo



pontos diferentes, ou seja:
no caso A é não-unitária,
pois os pontos não coincidem
$(A^t \neq A^{-1})$.

2o passo



$$
\begin{tikzcd}
\tilde{\mathbf{x}}_R \arrow[r, bend right, "A^{\cancel{t}}"'] & \mathbf{b} \\
& (AA^t)^{-1}\mathbf{b} \arrow[ul, bend right, "A^t"']
\end{tikzcd}
$$

3o passo

$$
\begin{tikzcd}
& \mathbf{z} \arrow[dl, bend right, "A^t"'] \\
\tilde{\mathbf{x}}_R \arrow[r, bend right, "A"'] & \mathbf{b} \arrow[u, bend left, "(AA^t)^{-1}"]
\end{tikzcd}
$$

$\mathbf{x}_{R(\cancel{A})} = \tilde{\mathbf{x}}_R$

$\tilde{\mathbf{x}}_R = A^t z \dashrightarrow \boxed{\tilde{\mathbf{x}}_R = A^t(AA^t)^{-1}\mathbf{b}}$

$A\tilde{\mathbf{x}}_R = \mathbf{b} \dashrightarrow AA^t\mathbf{z} = \mathbf{b} \therefore \mathbf{z} = (AA^t)^{-1}\mathbf{b}$

$$
\begin{tikzcd}
& (AA^t)^{-1}\mathbf{b} \arrow[dl, bend right, "A^t"'] \\
\tilde{\mathbf{x}}_R \arrow[r, bend right, "A"'] & \mathbf{b} \arrow[u, bend left, "(AA^t)^{-1}"]
\end{tikzcd}
\equiv
\begin{tikzcd}
\tilde{\mathbf{x}}_R \arrow[r, bend right, "A : \mathbb{R}^m \to \mathbb{R}^m"'] & \mathbf{b} \arrow[l, bend right, "A^t(AA^t)^{-1}"']
\end{tikzcd}
$$

no entanto $\mathbf{x}_{R(A^t)}$ em geral não é igual a $\tilde{\mathbf{x}}_R$, pois $A \neq A^t(AA^t)^{-1}$.

então: partimos da imagem e voltamos a ela. Teremos que ter a pseudo-inversa. Assim, AA^t é não-singular.

Pode-se dizer ainda que "achamos"uma transformação linear tal que saindo de $\tilde{\mathbf{x}}_R$ voltamos a ela em $\mathbb{R}^n$:

$$
\begin{tikzcd}
\tilde{\mathbf{x}}_R \arrow[r, bend right, "A"'] & \mathbf{b} \arrow[l, bend right, "A^t(AA^t)^{-1}"']
\end{tikzcd}
$$

c) O vetor $\tilde{\mathbf{x}}_R = \mathbf{x}^* = A^t(AA^t)^{-1}\mathbf{b}$ é a solução de norma mínima de $A\mathbf{x} = \mathbf{b}$, no sentido de $||\mathbf{x}^*|| < ||\mathbf{x}||$ para qualquer outra solução $\mathbf{x}$.

Prova:

i) $\mathbf{x}^*$ é uma solução, pois: $A\mathbf{x}^* = (AA^t)(AA^t)^{-1}\mathbf{b} = \mathbf{b}$.

ii) $\mathbf{x}^*$ é solução de norma mínima porque qualquer $\mathbf{x}$ dado por $\mathbf{x} = \mathbf{x}^* + \mathbf{x}_{N(A)}(\mathbf{x}^* \in R(A^t) \perp \mathbf{x}_{N(A)})$ é solução.

Mas $\exists$ um único $\mathbf{x}^* \in R(A^t)$ e que e dado por $\mathbf{x}^* = A^t(AA^t)^{-1}\mathbf{b}$, e pelo teorema de Pitágoras: $||\mathbf{x}||^2 = ||\mathbf{x}^*||^2 + ||\mathbf{x}_{N(A)}||^2$. Logo: $||\mathbf{x}|| > ||\mathbf{x}^*||$ para qualquer $\mathbf{x}_{N(A)}$ não-nulo.

Como obter $\mathbf{x}_{N(A)}$? Só é possível colocando alguma informação sobre ele em edição a $A\mathbf{x}_{N(A)} = \emptyset$.

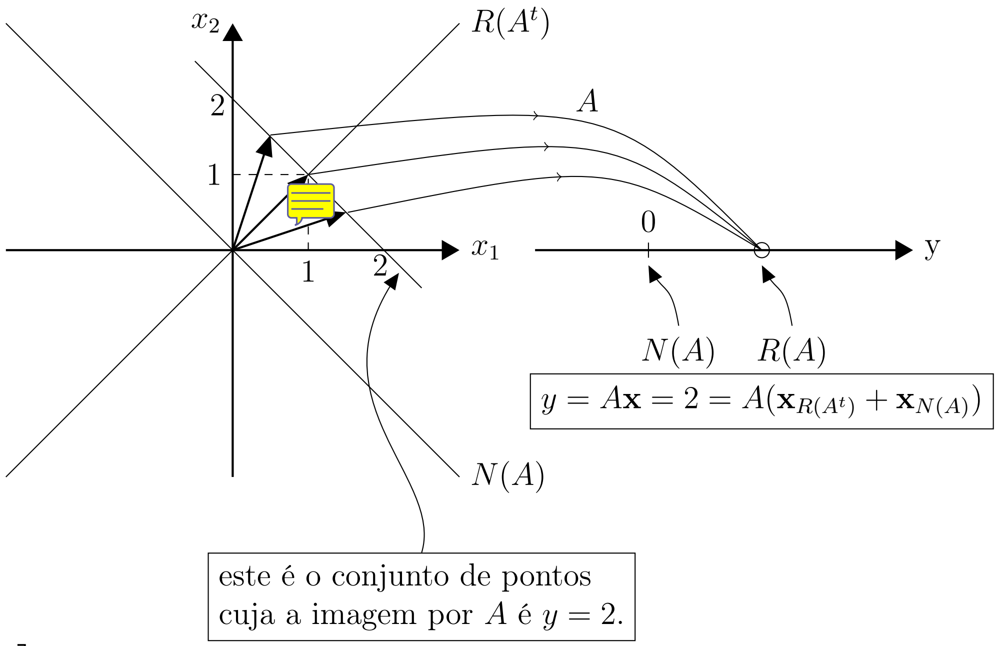
Por exemplo: $||\mathbf{x}_{N(A)}||^2 = k^2$.

Ex: $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \mathbf{x}_{N1} \\ \mathbf{x}_{N2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \therefore \begin{bmatrix} a \\ c \end{bmatrix} \mathbf{x}_{N1} + \begin{bmatrix} b \\ d \end{bmatrix} \mathbf{x}_{N2} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \therefore$

$\mathbf{x}_{N1} = -(\frac{b}{a})\mathbf{x}_{N2} = -(\frac{d}{c})\mathbf{x}_{N2} \therefore$

$(ad - bc) = det(A) = 0 \therefore \mathbf{x}_{N2}^2 = \frac{k^2}{(\frac{b}{a})^2+1} \therefore \mathbf{x}_{N1}^2 = \frac{k^2}{(\frac{a}{b})^2+1} \therefore \mathbf{x}_{N1}^2 + \mathbf{x}_{N2}^2 = k^2$

Exemplo: $A\mathbf{x} = \mathbf{y}$, $A = \begin{bmatrix} 1 & 1 \end{bmatrix}$



fazendo $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow A\mathbf{x} = \mathbf{y} = y = 2$.

Aplicando a pseudo-inversa:

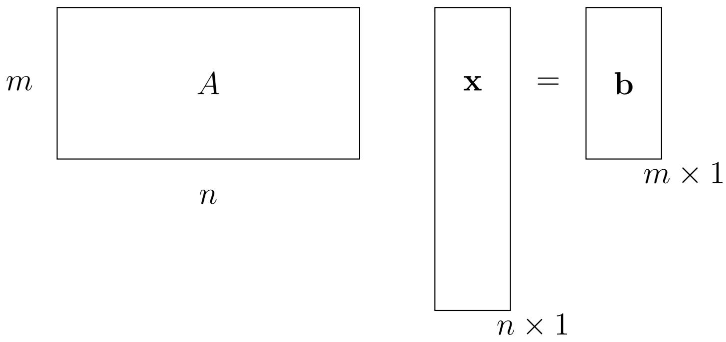
$\mathbf{x}^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \left\{ \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}^{-1} \mathbf{y} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} 2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, que é o vetor de menor norma para $y = 2$ e $A = \begin{bmatrix} 1 & 1 \end{bmatrix}$.

Outro exemplo: $N(A) = \{\mathbf{x} \in \mathbb{R}^2 | x_1 + x_2 = 0\}$, $A = \ 1 \quad 1 \ $, e $R(A^t) = \{\mathbf{x} \in \mathbb{R} | x_1 - x_2 = 0\}$.

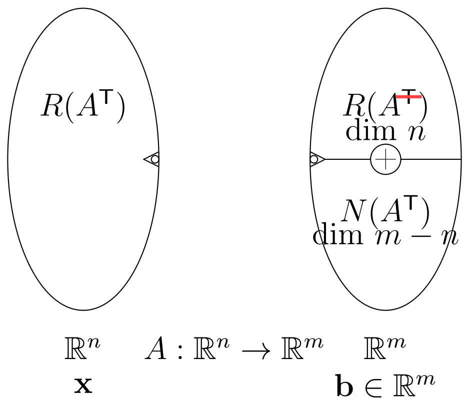
Além disso: $A\mathbf{x} = 2 \therefore \mathbf{x}_1 + \mathbf{x}_2 = 2$. Se $\mathbf{x}_2 = \frac{3}{2}$, então: $\mathbf{x}_1 = \frac{1}{2}$. Logo: $\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \frac{1}{2} + \frac{3}{2} = 2$

Teorema f

Seja $\mathbf{A}$ uma matriz $m \times n$, $n < m$, com rank completo n.



a) $\mathbf{Ax} = \mathbf{b}$ tem no máximo uma solução (e pode ter nenhuma).



Prova:

é dado $\mathbf{b} \in \mathbf{R}^m : \mathbf{b} = \mathbf{b}_{\mathbf{R(A)}} + \mathbf{b}_{\mathbf{N(A^\mathsf{T})}} = \mathbf{b}_{\mathbf{R(A)}} + \mathbf{b}_{\mathbf{N(A)}^\perp}$

Se $\mathbf{b} = \mathbf{b}_{\mathbf{R(A)}}$ ($\mathbf{b}_{\mathbf{N(A^\mathsf{T})}} = \mathbf{0}$), $\exists$ 1 só solução; porém se $\mathbf{b}_{\mathbf{N(A^\mathsf{T})}} \neq \mathbf{0}$, não existe solução (não $\exists \mathbf{x}$ levado a $\mathbf{b} = \mathbf{b}_{\mathbf{R(A)}} + \mathbf{b}_{\mathbf{N(A)}^\perp}$).

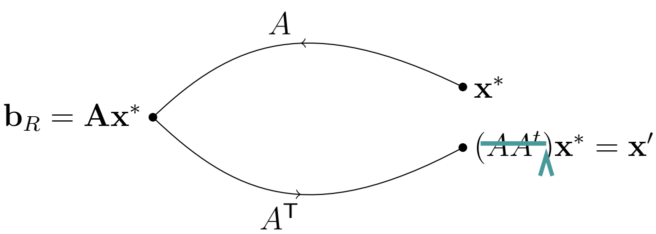
b) $\mathbf{A}^\mathsf{T}\mathbf{A}$ é não singular (a matriz $[(\mathbf{A}^\mathsf{T}\mathbf{A}^{-1})\mathbf{A}^\mathsf{T}]$ é a psedo-inversa).

Prova: conforme exercício (b) passado (p.18)

$rank(\mathbf{A}^\mathsf{T}) = rank(\mathbf{A}) = rank(\mathbf{A}^\mathsf{T}\mathbf{A}) \to \mathbf{A}^\mathsf{T}\mathbf{A}$ é $n \times n$ com rank n, logo é não singular. Logo, $\mathbf{R}(\mathbf{A}^\mathsf{T}) = \mathbf{R}(\mathbf{A}^\mathsf{T}\mathbf{A})$: Espaço de procura de $\mathbf{x}_{\mathbf{R(A^\mathsf{T})}}$ é o mesmo de $\mathbf{x}_{\mathbf{R(A^\mathsf{T}A)}}$.

Se deseja obter $\mathbf{b}_1 = \mathbf{Ax}, \mathbf{x} \in \mathbb{R}^n, \in \mathbf{R(A)}$ tal que seja o mais próximo de $\mathbf{b}$:

1º passo : Sabemos que $\mathbf{R}(\mathbf{A}^\mathsf{T}) = \mathbf{R}(\mathbf{A}^\mathsf{T}\mathbf{A})$. $\mathbf{Ax} = \mathbf{b}$???



$\boxed{\text{Em geral } (\mathbf{A}^\mathsf{T}\mathbf{A}) \text{ é não unitária.}}$

2º passo :

$$\begin{tikzcd} \mathbf{x}^* \arrow[r, bend right, "\mathbf{A}"'] & \tilde{b} = \mathbf{b}_{R(A)} \arrow[l, bend right, "(\mathbf{A}^\intercal\mathbf{A})^{-1}\mathbf{A}^\intercal"'] \end{tikzcd}$$

Ideia Geométrica :

c) Nos caso em que $\mathbf{Ax} = \mathbf{b}$ não há solução, e a solução : $\mathbf{x}^* = (\mathbf{A}^\intercal\mathbf{A})^{-1}\mathbf{b}$ {estado estimado para um psedo $\mathbf{b}$ (assumido como verdadeiro)} é a melhor aproximação de uma solução no sentido que $||\mathbf{Ax}^* - \mathbf{b}|| < ||\mathbf{Ax}' - \mathbf{b}||, \forall \text{outros} \mathbf{x}' \in \mathbb{R}^n$

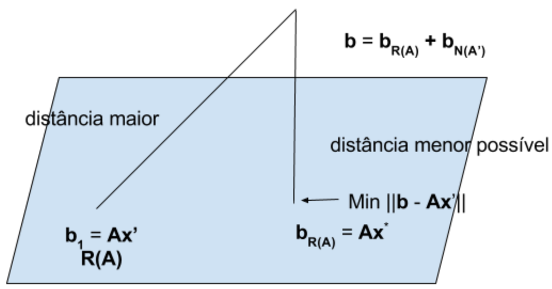
prova :

Solução da pseudo-inversa: $\mathbf{Ax} = \mathbf{b}$

a) $m < n : \mathbf{x}^* = \mathbf{A}^\intercal(\mathbf{AA}^\intercal)^{-1}\mathbf{b}$

b) $n < m : \mathbf{x}^* = (\mathbf{AA}^\intercal)^{-1}\mathbf{A}^\intercal\mathbf{b}$

c) $n \cancel{<} m : \mathbf{x}^* = \mathbf{A}^{-1}\mathbf{b}$



Quando **Ax** = **b**: não há solução!

Teorema g

$$\mathbf{x}^* \xrightarrow{A} \mathbf{b}_{R(A)} = A\mathbf{x}^*\} \in R(A)$$

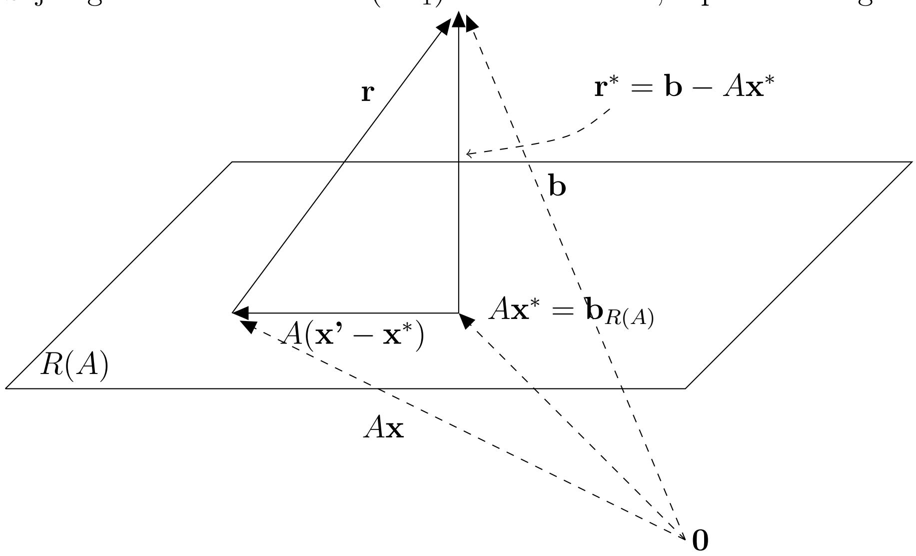
Seja o vetor resíduo $\mathbf{r}_1$:
$\mathbf{r}_1 = (A\mathbf{x}' - \mathbf{b})$ pode ser reescrito na forma:

$$\mathbf{r}_1 = (A\mathbf{x}' - \mathbf{b}) = \underbrace{(A\mathbf{x}^* - \mathbf{b})}_{-b_{N(A^t)}} + \underbrace{A(\mathbf{x}' - \mathbf{x}^*)}_{\in R(A)}$$

como $(\mathbf{b}_{N(A^t)}) \in N(A^t)$ e $[A(\mathbf{x}' - \mathbf{x}^*)] \in R(A)$, logo são ortogonais.
$\mathbf{b} = \mathbf{b}_{R(A)} + \mathbf{b}_{N(A^t)}$, onde $A\mathbf{x}^* = \mathbf{b}_{R(A)}$.
Assim, pode-se obter $\mathbf{b}_{R(A)}$ pela pseudo-inversa:

$$\begin{tikzcd} \mathbf{x}^* \arrow[r, bend right, "A"'] & \mathbf{b}_{R(A)} \arrow[l, bend right, "(A^tA)^{-1}A^t"'] \end{tikzcd} \quad \mathbf{b}_{R(A)} = A\mathbf{x}^* = A\underbrace{\left[(A^tA)^{-1}A^t\right]}_{pseudo-inversa}\mathbf{b}$$

Seja agora o vetor resíduo $(-\mathbf{r}_1) = \mathbf{r} = \mathbf{b} - A\mathbf{x}'$, representado geometricamente:



1.7 Exercícios

Ex. 1 — Demonstre as propriedades da Transformada de Laplace na Seção 1.2.1.

Ex. 2 — Demonstre as propriedades da Transformada de Z na Seção 1.2.2.

Ex. 3 — Se a matriz $\mathbf{A}$ $(n \times n)$ tem autovalores $\lambda_1, \ldots, \lambda_n$, quais são os autovalores de

a) $\mathbf{A}^k$, onde k é um inteiro positivo,

b) $\mathbf{A}^{-1}$, assumindo a existência da inversa,

c) $\mathbf{A}^T$,

d) $\mathbf{A}^H$,

e) $\alpha\mathbf{A}^k$, onde α é um número real,

f) $\mathbf{A}^T\mathbf{A}$?

Ex. 4 — a) Prove uma condição necessária e suficiente para nilpotência em termos de autovalores.

b) Mostre que os autovalores de uma matriz simétrica são reais.

c) Prove que os autovalores de uma matriz triangular superior são suas diagonais.

Ex. 5 — Para um matriz $\mathbf{A}(m \times n)$, prove que a norma espectral é dada por:

$$||\mathbf{A}|| = \max_{x \neq 0} \frac{||\mathbf{Ax}||}{||\mathbf{x}||}$$

Concluir que para qualquer vetor $\mathbf{x}$ $(n \times 1)$:

$$\|\mathbf{Ax}\| \leq \|\mathbf{A}\|\|\mathbf{x}\|$$

Ex. 6 — Usando o resultado da questão anterior, prove que para "conformable matrices" $\mathbf{A}$ e $\mathbf{B}$,

$$\|\mathbf{AB}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|$$

Se $\mathbf{A}$ é invertível, mostre que:

$$\|\mathbf{A}^{-1}\| \geq \frac{1}{\|\mathbf{A}\|}.$$

Ex. 7 — Se $\mathbf{Q}$ é uma matriz simétrica de ordem $n \times n$, prove que a norma espectral é dada por,

$$\|\mathbf{Q}\| = \max_{\|\mathbf{x}\|=1} |\mathbf{x}^{\mathsf{T}}\mathbf{Q}\mathbf{x}| = \max_{1 \leq i \leq n} |\lambda_i|$$

onde $\lambda_1, ..., \lambda_n$ são os autovalores de $\mathbf{Q}$.

Ex. 8 — Mostre que a norma espectral de uma matriz $m \times n$ $\mathbf{A}$ é dada por

$$\|\mathbf{A}\| = \left[\max_{\|\mathbf{x}\|=1} \mathbf{x}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}\mathbf{A}\mathbf{x}\right]^{\frac{1}{2}}$$

Obtenha da desigualdade de Rayleigh-Ritz que $\|\mathbf{A}\|$ é dado pela raiz quadrada não-negativa do maior autovalor de $\mathbf{A}^{\mathsf{T}}\mathbf{A}$.

Ex. 9 — Se $\mathbf{A}$ é uma matriz inversível $n \times n$, prove que $\|\mathbf{A}^{-1}\| \leq \dfrac{\|\mathbf{A}^{n-1}\|}{\|\det \mathbf{A}\|}$.

Ex. 10 — Se $A(t)$ é uma matrix $n \times n$ continuamente diferenciável e invertível em cada ponto t, mostre que:

$$\frac{d}{dt}A^{-1}(t) = -A^{-1}(t)\dot{A}(t)A^{-1}(t)$$

Ex. 11 — Mostre que $\mathbf{x} - P_{\mathbf{y}}\mathbf{x} \perp \mathbf{y}$.

Ex. 12 — Mostre que a matriz de Householder $\mathbf{H}$, definida na Equação (1.10), é ortogonal, ou seja, $\mathbf{H}^{\mathsf{T}}\mathbf{H} = \mathbf{I}$.

Ex. 13 — Mostre que se $\mathbf{b} = \mathbf{Ha}$, então $\|a\| = \|b\|$.

Ex. 14 — Mostre que o vetor transformado $\mathbf{b}$ é idêntico ao vetor original $\mathbf{a}$ em todas as componentes para as quais o vetor de Householder $\mathbf{w}$ é nulo.

Ex. 15 — Mostrar: $\mathbf{N}(\mathbf{A}) = \mathbf{N}(\mathbf{A}\mathbf{A}^{\mathsf{T}}) = \mathbf{R}(\mathbf{A}^{\mathsf{T}})^{\perp} = \mathbf{R}(\mathbf{A}^{\mathsf{T}}\mathbf{A})^{\perp}$

Ex. 16 — Mostrar a partir do lema: $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}\mathbf{A}^{\mathsf{T}})$, que: $\text{rank}(\mathbf{A}^{\mathsf{T}}\mathbf{A}) = \text{rank}(\mathbf{A}^{\mathsf{T}})$ Obs.: este resultado permite determinar o $\text{rank}(\mathbf{A})$ a partir de $(\mathbf{A}\mathbf{A}^{\mathsf{T}})$ ou $(\mathbf{A}^{\mathsf{T}}\mathbf{A})$.

Ex. 17 — Mostrar nulidade $(\mathbf{A})$ = nulidade $(\mathbf{A}^{\mathsf{T}}\mathbf{A})$.

Ex. 18 — Dúvida: nulidade $(\mathbf{A})$ = nulidade $(\mathbf{A}^{\mathsf{T}})$?

Ex. 19 — Calcular os subespaços fundamentais da transformação linear $\mathbf{A} = \begin{bmatrix} 1 & 1 \end{bmatrix}^{\mathsf{T}}$.

Ex. 20 — A equação $\mathbf{Ax} = \mathbf{b}$ tem pelo menos uma solução se e somente se

$$(\mathbf{I}_m - \mathbf{A}\mathbf{A}^{\dagger})\mathbf{b} = \mathbf{0},$$

em que $(\cdot)^{\dagger}$ denota a pseudo-inversa. Quando esta condição é satisfeita, a solução geral para $\mathbf{x}$ é dada por:

$$\mathbf{x} = \mathbf{A}^{\dagger}\mathbf{b} + (\mathbf{I}_n - \mathbf{A}^{\dagger}\mathbf{A})\xi,$$

em que ξ é um vetor arbitrário com dimensão n. Analise os casos:

1. $m \geq n$ e $\text{rank}(\mathbf{A}) = n$
2. $n \geq m$ e $\text{rank}(\mathbf{A}) = m$

Capítulo 2

Autoanálise de Sistemas

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.1 Autoanálise de sistemas

2.1.1 Matriz de Auto-Corelação

- Vetor de observação: $\boldsymbol{\mu}(n) = [\mu(n)\, \mu(n-1)\, \ldots\, \mu(n-M+1)]^{\mathsf{T}}$
- Matriz de auto-correlação $\mathbf{R}$ de série temporal $\mu(n)\, \mu(n-1)\, \ldots\, \mu(n-M+1)$, tem como entradas as médias dos produtos formados a partir de $\boldsymbol{\mu}(n)$ consigo mesmo: $\mathbf{R} = \mathbb{E}\left[\boldsymbol{\mu}(n)\boldsymbol{\mu}^{\mathsf{H}}(n)\right] = \mathbb{E}\left[([\mu^*(n)\boldsymbol{\mu}(n)]\, [\mu^*(n)\boldsymbol{\mu}(n-1)]\, \ldots\, [\mu^*(n-M+1)\boldsymbol{\mu}(n)])\right]$
- Fazendo: $r(n,k) = r(k) = \mathbb{E}\left[\mu(n)\, \mu^*(n)\right]$, então:

$$R = \begin{pmatrix} r(0) & r(1) & \cdots & r(M-1) \\ r(-1) & r(0) & r(1) & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ r(-M+1) & r(-M+2) & \cdots & r(0) \end{pmatrix}_{M\times M} = (\mu(m)\mu^*(M+i-j))_{M\times M}$$

Obs. É suposto em $r(k)$ que a diferença entre os instantes no tempo é importante.

Propriedades da matriz $\mathbf{R}$:

a) $\mathbf{R}$ é Hermitiana: $\mathbf{B}^{\mathsf{H}} = \mathbf{R}$, isto é: $r(-k) = r^*(k)$ e no caso real: $\mathbf{R}^{\mathsf{T}}\mathbf{R}$ (matriz simétrica)

b) $\mathbf{R}$ é toeplitz: diagonais com entradas iguais, consequência da independência de $r(n,k)$ de n. Se $\mathbf{R}$ é toeplitz, então $r(n,k) = r(k)$

c) $\mathbf{R}$ é sempre def. não-negativa, e em geral def. positiva: Sejam $\mathbf{x}_{M\times 1} \in \mathbb{C}^M$ e $\mathbf{y}(n) = \mathbf{x}^{\mathsf{H}}\boldsymbol{\mu}(n) \Rightarrow \mathbf{y}^*(n) = \boldsymbol{\mu}^{\mathsf{H}}(n)\mathbf{x}$.

- $\mathbb{E}\left[\mathbf{y}(n)\mathbf{y}^*(n)\right] = \mathbb{E}\left[\|\mathbf{y}(n)\|^2\right] = \mathbf{x}^{\mathsf{H}}\mathbf{R}\mathbf{x}$, forma quadrática ou hermitiana.
- Como $\mathbb{E}\left[\|\mathbf{y}(n)\|^2\right] \leq 0 \Rightarrow \mathbf{x}^{\mathsf{H}}\mathbf{R}\mathbf{x} \geq 0$, $\mathbf{R}$ é semi-definida positiva.
- Se: $\mathbf{x}^{\mathsf{H}}\mathbf{R}\mathbf{x} > 0$, $\mathbf{R}$ é def. positiva e não singular.

2.1.2 O problema do autovalor

- Encontrar $\mathbf{q}_{M\times 1}$ tal que $\mathbf{R}\mathbf{q} = \lambda\mathbf{q}$ em que λ é constante, isso define uma transformação invariante na direção dentro do espaço M-dimensional.
- Logo: $(\mathbf{R} - \lambda\mathbf{I})\mathbf{q} = \mathbf{0}$, $(\mathbf{R} - \lambda\mathbf{I})$ é uma matriz singular. Para que seja garantido que $\mathbf{q} \neq \mathbf{0}$: $det(\mathbf{R} - \lambda\mathbf{I}) = 0$.
- Solução: polinômio com M raízes distintas ou não (degeneração).
- Seja λ_i o i-ésimo autovalor de $\mathbf{R}$, e que $\mathbf{q}_i$ é o autovetor associado, tal que:

$$\mathbf{R}\mathbf{q}_i = \lambda_i \mathrm{q}_i$$

Um $\mathrm{q}_i \Rightarrow$ um λ_i, porém! Um λ_i é associado a vários $\mathbf{q}_i$

ex.: $(a\mathbf{q})$ é também autovetor p/ $\lambda_i \neq 0$.

ex.: $\mathbf{R} = \sigma^2\mathbf{I}$ (ruído branco) $\Rightarrow \lambda_i = \sigma^2, i = 1, \ldots, M$.

ex.: senóide complexa:

$$R = \begin{pmatrix} 1 & e^{j\omega} & \cdots & e^{j(M-1)\omega} \\ e^{j\omega} & 1 & \cdots & e^{j(M-2)\omega} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(M-1)\omega} & e^{j(M-2)\omega} & \cdots & 1 \end{pmatrix}$$

Um autovetor possível é: $\mathbf{q} = [1, e^{j\omega}, \cdots, e^{j(M-1)\omega}]^{\mathsf{T}}$, $rank(\mathbf{R}) = 1, \lambda_1 = M, \lambda_i = 0, i = 2, \ldots, M$.

2.2 Propriedades de autovalores e autovetores associados a uma matriz Hermitiana definida não-negativa

Propriedade 1

Se $\lambda_1, \ldots, \lambda_M$ são autovalores de $\mathbf{R}$, então $\lambda_1^k, \ldots, \lambda_M^k$ são autovalores de $\mathbf{R}^k$.

Demonstração. Seja a equação que relaciona os autovetores aos seus autovalores

$$\mathbf{R}\mathbf{q} = \lambda\mathbf{q} \tag{2.1}$$

Multiplicando-se pela esquerda esta equação por $\mathbf{R}^{k-1}$ obtemos:

$$\begin{aligned}
\mathbf{R}^{k-1}\mathbf{R}\mathbf{q} &= \lambda\mathbf{R}^{k-1}\mathbf{q} \\
\mathbf{R}^{k}\mathbf{q} &= \lambda\left(\mathbf{R}^{k-1}\mathbf{q}\right) \\
\mathbf{R}^{k}\mathbf{q} &= \lambda^{2}\left(\mathbf{R}^{k-2}\mathbf{q}\right) \\
&\vdots \\
\mathbf{R}^{k}\mathbf{q} &= \lambda^{k}\mathbf{q}
\end{aligned}$$

□

Propriedade 2

Sejam $\mathbf{q}_i$ e λ_i, $i \in [1, M]$ autovetores e autovalores distintos de $\mathbf{R}$. Então $\mathbf{q}_i$, $i = 1, \ldots, M$ são L.I.

Demonstração. (Por contradição) Suponhamos que os vetores $\mathbf{q}_i$, $i = 1, \ldots, M$, sejam linearmente dependentes, ou seja:

$$\sum_{i=1}^{M} v_i \mathbf{q}_i = \mathbf{0} \tag{2.2}$$

para v_i não-nulos. Multiplicando a equação acima por $\mathbf{R}^{k-1}$ e utilizando a Propriedade 1 temos:

$$\sum_{i=1}^{M} v_i \mathbf{R}^{k-1}\mathbf{q}_i = \sum_{i=1}^{M} v_i \lambda_i^{k-1}\mathbf{q}_i = \mathbf{0}, \quad k = 1, \ldots, M \tag{2.3}$$

Podemos reescrever a equação acima como:

$$[v_1\mathbf{q}_1, v_2\mathbf{q}_2, \ldots, v_M\mathbf{q}_M]_{M\times M}\,\mathbf{S} = \mathbf{0}_M, \tag{2.4}$$

em que

$$\mathbf{S} = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{M-1} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{M-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_M & \lambda_M^2 & \ldots & \lambda_M^{M-1} \end{bmatrix}_{M\times M}$$

é uma matriz de Vandermonde. Mostra-se que $\mathbf{S}$ é não-singular se λ_i são distintos $\forall i \in [1, M]$. Logo podemos multiplicar (2.4) por $\mathbf{S}^{-1}$ pela direita:

$$[v_1\mathbf{q}_1, v_2\mathbf{q}_2, \ldots, v_M\mathbf{q}_M]\underbrace{\mathbf{S}\mathbf{S}^{-1}}_{\mathbf{I}} = \mathbf{0} \therefore [v_1\mathbf{q}_1, v_2\mathbf{q}_2, \ldots, v_M\mathbf{q}_M] = \mathbf{0}$$

ou ainda: $v_i\mathbf{q}_i = \mathbf{0}$ para $i = 1, \ldots, M$. Como $\mathbf{q}_i \neq \mathbf{0}$, então $v_i = 0\,\forall i \in [1, M]$. Isto contradiz a suposição inicial (2.2). Portanto, os autovetores $\mathbf{q}_i$ são L.I., formando uma base de representação de um vetor no espaço $\mathbb{R}^N$ ou $\mathbb{C}^N$, ou seja: $\mathbf{w} = \sum_{i=1}^{M} v_i\mathbf{q}_i$, em que os v_i são constantes. Logo:

$$\mathbf{R}\mathbf{w} = \sum_{i=1}^{M} v_i\mathbf{R}\mathbf{q}_i = \underbrace{\sum_{i=1}^{M} v_i\lambda_i\mathbf{q}_i}_{\text{outro vetor no mesmo espaço}}$$

□

Propriedade 3

Se $\lambda_1, \ldots, \lambda_M$ são autovalores de $\mathbf{R} \in \mathbb{C}^{M\times M}$, então $\lambda_m \in \mathbb{R}$ para $m = 1, \ldots, M$.

Demonstração. Temos que $\mathbf{R}\mathbf{q}_i = \lambda_i \mathbf{q}_i$, $i = 1, \ldots, M$. Portanto $\mathbf{q}_i^{\mathsf{H}}\mathbf{R}\mathbf{q}_i = \lambda_i \mathbf{q}_i^{\mathsf{H}}\mathbf{q}_i$. Isolando λ_i:

$$\lambda_i = \frac{\mathbf{q}_i^{\mathsf{H}}\mathbf{R}\mathbf{q}_i}{\mathbf{q}_i^{\mathsf{H}}\mathbf{q}_i} \quad \text{(coeficiente ou quociente de Rayleigh do vetor } \mathbf{q}_i\text{)}$$

Sendo $\mathbf{q}_i^{\mathsf{H}}\mathbf{q}_i = \|\mathbf{q}_i\|^2 > 0$ e $\mathbf{q}_i^{\mathsf{H}}\mathbf{R}\mathbf{q}_i > 0 \,\forall i \in [1, M]$ pois $\mathbf{R}$ é definida positiva. Assim, $\lambda_i > 0 \,\forall i \in [1, M]$. □

Propriedade 4

Sejam $\mathbf{q}_i$ e λ_i, $i \in [1, M]$ autovetores e autovalores de $\mathbf{R}$. Logo $\mathbf{q}_i$ e $\mathbf{q}_j$ s]ao mutuamente ortogonais, $i \neq j$, $i, j \in [1, M]$.

Demonstração. Considere

$$\mathbf{R}\mathbf{q}_i = \lambda_i \mathbf{q}_i \tag{2.5}$$

$$\mathbf{R}\mathbf{q}_j = \lambda_j \mathbf{q}_j \tag{2.6}$$

com $i \neq j$. Aplicando-se o operador hermitiano na Eq. (2.6): $(\mathbf{R}\mathbf{q}_j)^{\mathsf{H}} = (\lambda_j \mathbf{q}_j)^{\mathsf{H}} \therefore \mathbf{q}_j^{\mathsf{H}}\mathbf{R}^{\mathsf{H}} = \lambda_j^* \mathbf{q}_j^{\mathsf{H}}$. Porém $\mathbf{R}^{\mathsf{H}} = \mathbf{R}$ e $\lambda_j^* = \lambda_j$. Portanto

$$\mathbf{q}_j^{\mathsf{H}}\mathbf{R} = \lambda_j \mathbf{q}_j^{\mathsf{H}}. \tag{2.7}$$

Agora vamos multiplicar a Equação (2.5) por $\mathbf{q}_j^{\mathsf{H}}$ pela esquerda:

$$\mathbf{q}_j\mathbf{R}\mathbf{q}_i = \lambda_i \mathbf{q}_j^{\mathsf{H}}\mathbf{q}_i$$

Substituindo (2.7) na equação acima: $\lambda_j \left(\mathbf{q}_j^{\mathsf{H}}\mathbf{q}_i\right) = \lambda_i \left(\mathbf{q}_j^{\mathsf{H}}\mathbf{q}_i\right) \therefore (\lambda_j - \lambda_i)(\mathbf{q}_j^{\mathsf{H}}\mathbf{q}_i) = 0$. Portanto, se $\lambda_i \neq \lambda_j$, então $\mathbf{q}_j^{\mathsf{H}}\mathbf{q}_i = 0$, $i \neq j$. □

Propriedade 5: Transformação de Similaridade Unitária

Sejam $\mathbf{q}_1, \ldots, \mathbf{q}_M$ autovetores de $\lambda_1, \ldots, \lambda_M$ em $\mathbf{R} \in \mathbb{C}^{M\times M}$. Seja $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_M] \in \mathbb{C}^{M\times M}$ tal que $\mathbf{q}_i^{\mathsf{H}}\mathbf{q}_j = \delta_{ij}$. Seja $\mathbf{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_M)$. Então $\mathbf{R}$ pode ser diagonalizada por $\mathbf{Q}$: $\mathbf{\Lambda} = \mathbf{Q}^{\mathsf{H}}\mathbf{R}\mathbf{Q}$.

Demonstração. Podemos expressar o conjunto de sistemas lineares $\mathbf{R}\mathbf{q}_m = \lambda_m \mathbf{q}_m$, $m = 1, \ldots, M$ sucintamente:

$$\mathbf{R}\mathbf{Q} = \mathbf{Q}\mathbf{\Lambda} \tag{2.8}$$

Multiplicando a equação acima por $\mathbf{Q}^{\mathsf{H}}$ pela esquerda:

$$\mathbf{Q}^{\mathsf{H}}\mathbf{R}\mathbf{Q} = \underbrace{\mathbf{Q}^{\mathsf{H}}\mathbf{Q}}_{\mathbf{I}}\mathbf{\Lambda} = \mathbf{\Lambda}$$

Podemos observar que $\mathbf{Q}^{\mathsf{H}}\mathbf{Q} = \mathbf{I} \therefore \mathbf{Q}^{\mathsf{H}}\mathbf{Q}\mathbf{Q}^{-1} = \mathbf{Q}^{-1} \therefore \mathbf{Q}^{\mathsf{H}} = \mathbf{Q}^{-1}$ (matriz unitária). □

Resultado: $\mathbf{Q}^{\mathsf{H}}\mathbf{R}\mathbf{Q} = \mathbf{\Lambda} \to \mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathsf{H}}$. A matriz $\mathbf{Q}$ é formada por vetores ortonormais e $\mathbf{\Lambda}$ é uma matriz diagonal com os autovalores de $\mathbf{R}$. Dizemos que a matriz $\mathbf{R}$ é *similiar* à matriz diagonal $\mathbf{\Lambda}$.

Propriedade 6: Teorema de Mercer ou Teorema Espectral

A partir de $\mathbf{R} = \mathbf{Q\Lambda Q}^{\mathsf{H}} = \sum_{i=1}^{M} \lambda_i \mathbf{q}_i \mathbf{q}_i^{\mathsf{H}} = \sum_{i=1}^{M} \lambda_i \mathbf{P}_i$, em que $\mathbf{P}_i = \mathbf{q}_i \mathbf{q}_i^{\mathsf{H}}$ é uma matriz de rank unitário. Resultados sobre $\mathbf{P}_i$:

- $\mathbf{P}_i \mathbf{P}_i = \mathbf{q}_i \underbrace{\mathbf{q}_i^{\mathsf{H}} \mathbf{q}_i}_{1} \mathbf{q}_i^{\mathsf{H}} = \mathbf{q}_i \mathbf{q}_i^{\mathsf{H}} = \mathbf{P}_i$

- $\mathbf{P}_i^{\mathsf{H}} = \left(\mathbf{q}_i \mathbf{q}_i^{\mathsf{H}}\right)^{\mathsf{H}} = \mathbf{q}_i \mathbf{q}_i^{\mathsf{H}} = \mathbf{P}_i$

- $\mathbf{P}_i \mathbf{x} = \mathbf{q}_i \mathbf{q}_i^{\mathsf{H}} \mathbf{x} = \left[\sum_{k=1}^{M} (x_k q_{ik}^*)\right] \mathbf{q}_i$, ou seja, $\mathbf{P}_i$ projeta um vetor $\mathbf{x}$ numa única direção $\mathbf{q}_i$ (autovetor). Logo, $\mathbf{P}_i$ é uma matriz de projeção de rank unitário.

Propriedade 7

Seja $\mathbf{R} = \mathbf{Q\Lambda Q}^{\mathsf{H}}$ a autodecomposição de $\mathbf{R}$. Logo, $\sum_{i=1}^{M} \lambda_i = \operatorname{tr}(\mathbf{R}) = \operatorname{tr}(\mathbf{\Lambda})$.

Demonstração. Podemos expressar a matriz de autovalores como $\Lambda = \mathbf{Q}^{\mathsf{H}}\mathbf{RQ}$. O seu traço é dado por: $\operatorname{tr}(\Lambda) = \operatorname{tr}(\mathbf{Q}^{\mathsf{H}}\mathbf{RQ})$. Sabendo que $\operatorname{tr}(\mathbf{A}_{M\times N}\mathbf{B}_{N\times M}) = \operatorname{tr}(\mathbf{B}_{N\times M}\mathbf{A}_{M\times N})$, temos:

$$\operatorname{tr}(\underbrace{\mathbf{Q}^{\mathsf{H}}}_{\mathbf{A}} \underbrace{\mathbf{RQ}}_{\mathbf{B}}) = \operatorname{tr}(\mathbf{R} \underbrace{\mathbf{QQ}^{\mathsf{H}}}_{\mathbf{I}}) = \operatorname{tr}(\mathbf{R}) \therefore \operatorname{tr}(\mathbf{R}) = \operatorname{tr}(\Lambda) = \sum_{i=1}^{M} \lambda_i, \ \forall \mathbf{R}$$

□

Propriedade 8

A matriz $\mathbf{R}$ é mal-condicionada se a razão entre o seu maior e menor autovalor for elevada.

Demonstração. Resultado da Teoria da Pertubação: Considere o sistema linear $\mathbf{A}\omega = \mathbf{d}$, em que $\mathbf{A}$ e $\mathbf{d}$ dependem de dados e ω é um filtro linear. Se $\mathbf{A}$ e $\mathbf{d}$ são pertubados por $\delta\mathbf{A}$ e $\delta\mathbf{d}$ e ambos $\|\delta\mathbf{A}\|/\|\mathbf{A}\|$ e $\|\delta\mathbf{d}\|/\|\mathbf{d}\|$ são da ordem de $\varepsilon << 1$, logo:

$$\frac{\delta\mathbf{d}}{\omega} \leq \varepsilon\chi(\mathbf{A})$$

em que $\chi(\mathbf{A})$ é a condição numérica de $\mathbf{A}$ relativa a $\mathbf{A}^{-1}$. Este valor descreve a má condição numérica de $\mathbf{A}$ quantitativamente.

O valor $\chi(\mathbf{A})$ é definido como $\chi(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$, em que $\|\mathbf{A}\|$ mede a magnitude de $\mathbf{A}$ em algum sentido, satisfazendo um campo algébrico:

1. $\|\mathbf{A}\| \geq 0$, $\|\mathbf{A}\| = 0$ se e somente se $\mathbf{A} = \mathbf{0}$;

2. $\|c\mathbf{A}\| = |c|\|\mathbf{A}\|$, $c \in \mathbb{R}$;

3. Desigualdade triangular: $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$

4. Desigualdade de Cauchy-Schwarz: $\|\mathbf{AB}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|$

Algumas definições de $\|\mathbf{A}\|$:

1. Norma espectral: $\|\mathbf{A}\|_S = \left[\text{maior autovalor de } (\mathbf{A}^{\mathsf{H}}\mathbf{A})\right]^{\frac{1}{2}}$

2. Norma de Frobenius: $\|\mathbf{A}\|_F = \left[\sum_{i=1}^{I}\sum_{j=1}^{J}|a_{ij}|^2\right]^{\frac{1}{2}}$

Se $\mathbf{R} = \mathbf{A}^{\mathsf{H}}\mathbf{A}$, então $\mathbf{R}^{\mathsf{H}} = \mathbf{R}$, ou seja, a matriz $\mathbf{R}$ é hermitiana e definida não-negativa. Assim, os autovalores de $\mathbf{R} = \mathbf{A}^{\mathsf{H}}\mathbf{A}$ são reais e não-negativos.

O autovalor de $\mathbf{A}^{\mathsf{H}}\mathbf{A}$ é o coeficiente de Rayleigh do autovetor correspondente:

$$\|\mathbf{A}\|_S^2 = \max \frac{\mathbf{x}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{A}\mathbf{x}}{\mathbf{x}^{\mathsf{H}}\mathbf{x}} = \max \frac{\|\mathbf{A}\mathbf{x}\|^2}{\|\mathbf{x}\|} \therefore \|\mathbf{A}\|_S = \max \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} \tag{2.9}$$

Interpretação: a normal espectral de $\|\mathbf{A}\|$ mede o maior valor pelo qual qualquer vetor é amplificado pela matriz multiplicante, e o vetor que é amplificado ao máximo é o autovetor que corresponde ao maior autovalor de $\mathbf{A}^{\mathsf{H}}\mathbf{A}$. Logo, a norma espectral de $\mathbf{R}$ é $\|\mathbf{R}\|_S = \lambda_{\max}$. A norma espectral de $\mathbf{R}^{-1}$ é :

$$\|\mathbf{R}^{-1}\|_S = \left\|(\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathsf{H}})^{-1}\right\|_S = \|\mathbf{Q}\mathbf{\Lambda}^{-1}\mathbf{Q}\|_S = 1/\lambda_{\min} \tag{2.10}$$

Dessa forma, o condicionamento de $\mathbf{R}$ é dado por

$$\chi(\mathbf{R}) = \|\mathbf{R}\|\|\mathbf{R}^{-1}\| = \frac{\lambda_{\max}}{\lambda_{\min}} \geq 1$$

O valor $\chi(\mathbf{X})$ também é conhecido como fator de espalhamento da matriz. Se $\chi(\mathbf{X})$ é muito grande, então $\mathbf{R}^{-1}$ tem elementos muito grandes. Portanto em problemas cuja solução utiliza $\mathbf{R}^{-1}$ podem enfrentar problemas numéricos. $\square$

Propriedade 9: Teorema Minimax

Seja $\mathbf{R_{MxM}}$ onde $\lambda_1, \lambda_2, \ldots, \lambda_M$ são arranjados em ordem decrescente:

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_M$$

Logo:

$$\lambda_k = \min_{\dim(\mathcal{J})=k} \max_{\substack{x\in\mathcal{J}\\ \mathbf{x}\neq\mathbf{0}}} \frac{\mathbf{x}^{\mathsf{H}}\mathbf{R}\mathbf{x}}{\mathbf{x}^{\mathsf{H}}\mathbf{x}}, \quad k = 1, \ldots, M$$

onde $\mathcal{J}$ é um subespaço do vetor de espaço de todos os vetores complexos Mx1; $\dim(\mathcal{J})$ denota a dimensão do subespaço $\mathcal{J}$; e $\mathbf{x} \in \mathcal{J}$ significa que o vetor $\mathbf{x}$ (não-nulo) varia sobre o subespaço $\mathcal{J}$.

Discussão: Seja $\mathbb{C}^M = \{\mathbf{y}\}, \quad \mathbf{y} = \sum_{i=1}^{M} a_i\mathbf{q_i}, \quad \mathbf{q_i}^{\mathsf{H}}\mathbf{q_i} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

Seja o subespaço $\mathcal{J}$ de $\mathbb{C}^M$, com k vetores de base $\mathbf{q_1}, \ldots, \mathbf{q_k}, \quad k \leq M$.

Se $\mathbf{x} \in \mathcal{J} \rightarrow \mathbf{x} = \sum_{i=1}^{k} a_i\mathbf{q_i}, \quad k \leq M$, e $\dim(\mathbf{x}) = M$, pois $\mathbf{x}$ tem componentes nulas em certas direções.

Demonstração. i) Do teorema espectral, sabe-se que: $\mathbf{R} = \sum_{i=1}^{M} \lambda_i\mathbf{q_i}\mathbf{q_i}^{\mathsf{H}}, \quad \mathbf{q_i}$ ortonormal.

Seja $\mathbf{x}$ restrito ao subespaço $\mathcal{J}$ de dimensão k.

Seja o quociente de Rayleigh associado a $\mathbf{x}$:

$$\frac{\mathbf{x^H R x}}{\mathbf{x^H x}} = \frac{\mathbf{x^H Q \Lambda Q^H x}}{\sum_{i=1}^{k} |a_i|^2} = \frac{\sum_{i=1}^{k} a_i^* \mathbf{q_i}^{\mathsf{H}} \sum_{l=1}^{M} \lambda_l \mathbf{q_l} \mathbf{q_l}^{\mathsf{H}} \sum_{m=1}^{k} a_m \mathbf{q_m}}{\sum_{i=1}^{k} |a_i|^2}$$
$$= \frac{\sum_{i=1}^{k} \sum_{l=1}^{M} \sum_{m=1}^{k} \lambda_l a_i^* a_m (\mathbf{q_i}^{\mathsf{H}} \mathbf{q_l})(\mathbf{q_l}^{\mathsf{H}} \mathbf{q_m})}{\sum_{i=1}^{k} |a_i|^2} = \left.\frac{\sum_{i=1}^{k} \lambda_i |a_i|^2}{\sum_{i=1}^{k} |a_i|^2}\right\} \text{média ponderada por } \lambda_i\text{'s}$$

Por suposição: $\lambda_1 \leq \lambda_2 \leq, \ldots, \leq \lambda_k$, logo o valor máximo de $\left[\frac{\sum_{i=1}^{k} \lambda_i |a_i|^2}{\sum_{i=1}^{k} |a_i|^2}\right]$ é menor ou igual a λ_k, ou seja, fazendo $\lambda_i = \lambda_k$, $\forall i, i = 1, \ldots, k$, temos:

$$\frac{\sum_{i=1}^{k} \lambda_i |a_i|^2}{\sum_{i=1}^{k} |a_i|^2} \leq \frac{\lambda_k \sum_{i=1}^{k} |a_i|^2}{\sum_{i=1}^{k} |a_i|^2} = \lambda_k, \quad \text{p/ } \mathbf{x} \neq \mathbf{0}, \mathbf{x} \in \mathcal{J}.$$

Logo, teremos que:

$$\max_{\substack{x \in \mathcal{J} \\ \mathbf{x} \neq \mathbf{0}}} \frac{\mathbf{x^H R x}}{\mathbf{x^H x}} \leq \lambda_k, \quad \text{p/ qualquer subespaço } \mathcal{J} \text{ de dimensão } k$$

E ainda, o menor dentre os máximos de $\frac{\mathbf{x^H R x}}{\mathbf{x^H x}} \leq \lambda_k$, ou seja:

$$\min_{\dim(\mathcal{J}) = k} \max_{\substack{x \in \mathcal{J} \\ \mathbf{x} \neq \mathbf{0}}} \frac{\mathbf{x^H R x}}{\mathbf{x^H x}} \leq \lambda_k, \quad k \in [1, M]$$

ii) Agora, seja $\mathbf{x}$ restrito ao subespaço $\mathcal{J}'$ de dimensão $M - k + 1$:

$$\mathbf{x} = \sum_{i=k}^{M} b_i \mathbf{q_i}$$

onde $\mathbf{x}$ tem componentes nulas nas direções $\mathbf{q_i}$, $i \in [1, k-1]$.

Considere ainda encontrar $\mathbf{x}$ de tal forma que:

$$\mathbf{x} = \sum_{i=k}^{M} b_i \mathbf{q_i} = \sum_{j=1}^{k} a_j \mathbf{q_{i_j}} \tag{2.11}$$

onde $\mathbf{q_{i_1}}$, $\mathbf{q_{i_2}}, \ldots$, $\mathbf{q_{i_k}}$ são autovalores que varrem o subespaço $\mathcal{J}$ de dimensão k; e $\{i_1, \ldots, i_k\} \subset \{1, 2, \ldots, M\}$.

O sistema de M equações homogêneas (equação 2.11) tem $(M+1)$ variáveis e terá sempre uma solução não-trivial. Além disso, todos os autovetores são LI. Por isto, é possível afirmar que há no mínimo um vetor $\mathbf{x} = \sum_{j=1}^{k} a_j \mathbf{q_{i_j}}$ que é comum aos subespaços $\mathcal{J}'$ e $\mathcal{J}$. Portanto, o quociente de Rayleigh relativo a $\mathbf{x}$ em $\mathcal{J}'$ é, por associação ao caso anterior:

$$\frac{\mathbf{x^H R x}}{\mathbf{x^H x}} = \frac{\sum_{i=k}^{M} \lambda_i |b_i|^2}{\sum_{i=k}^{M} |b_i|^2}$$

Supondo: $\lambda_k \leq \lambda_{k+1} \leq \ldots \leq \lambda_M$ e como $\mathbf{x} \in \mathcal{J}$, teremos:

$$\lambda_k \leq \frac{\sum_{i=k}^{M} \lambda_i |b_i|^2}{\sum_{i=k}^{M} |b_i|^2} \leq \underbrace{\frac{\sum_{i=k}^{M} \lambda_M |b_i|^2}{\sum_{i=k}^{M} |b_i|^2}}_{\max_{\substack{x \in \mathcal{J} \\ \mathbf{x} \neq \mathbf{0}}} \frac{\mathbf{x^H R x}}{\mathbf{x^H x}}} = \lambda_M$$

Logo, temos que:

$$\min_{\dim(\mathcal{J})\,=\,k}\ \max_{\substack{x\in\mathcal{J}\\ \mathbf{x}\neq\mathbf{0}}} \frac{\mathbf{x}^{\mathsf{H}}\mathbf{R}\mathbf{x}}{\mathbf{x}^{\mathsf{H}}\mathbf{x}} \geq \lambda_k \tag{2.12}$$

E ainda que:

$$\min_{\dim(\mathcal{J})\,=\,k}\ \max_{\substack{x\in\mathcal{J}\\ \mathbf{x}\neq\mathbf{0}}} \frac{\mathbf{x}^{\mathsf{H}}\mathbf{R}\mathbf{x}}{\mathbf{x}^{\mathsf{H}}\mathbf{x}} \leq \lambda_k \tag{2.13}$$

Consequentemente, para que sejam satisfeitas as equações 2.12 e 2.13, temos:

$$\min_{\dim(\mathcal{J})\,=\,k}\ \max_{\substack{x\in\mathcal{J}\\ \mathbf{x}\neq\mathbf{0}}} \frac{\mathbf{x}^{\mathsf{H}}\mathbf{R}\mathbf{x}}{\mathbf{x}^{\mathsf{H}}\mathbf{x}} = \lambda_k$$

□

Consequências do teorema minimax:

1. λ_k é determinado sem conhecer a autoestrutura de $\mathbf{R}$ e é baseado sobre qualquer $\mathbf{x} \in \mathcal{J}$.
2. Os $\mathbf{q}_i's$, $\quad i = 1, \ldots, M$, formam uma base para $\mathbb{C}^M$ mais eficiente no sentido de energia.
3. os $\lambda_i's$ são certas energias do vetor de entrada $\mu(\mathbf{n})$

Ainda há uma forma alternativa para o teorema minimax, chamada de teorema maximin, onde:

$$\lambda_k = \max_{\dim(\mathcal{J})\,=\,M-k+1}\ \min_{\substack{x\in\mathcal{J}'\\ \mathbf{x}\neq\mathbf{0}}} \frac{\mathbf{x}^{\mathsf{H}}\mathbf{R}\mathbf{x}}{\mathbf{x}^{\mathsf{H}}\mathbf{x}}$$

Casos especiais:

1. $k = M: \quad \mathcal{J} = \mathbb{C}^M \rightarrow \lambda_M = \max_{\substack{x\in\mathbb{C}^M\\ \mathbf{x}\neq\mathbf{0}}} \frac{\mathbf{x}^{\mathsf{H}}\mathbf{R}\mathbf{x}}{\mathbf{x}^{\mathsf{H}}\mathbf{x}}$
2. $k = 1: \quad \mathcal{J}' = \mathbb{C}^M \rightarrow \lambda_1 = \min_{\substack{x\in\mathbb{C}^M\\ \mathbf{x}\neq\mathbf{0}}} \frac{\mathbf{x}^{\mathsf{H}}\mathbf{R}\mathbf{x}}{\mathbf{x}^{\mathsf{H}}\mathbf{x}}$

Propriedade 10: Expansão de Karhunen-Loève (versão discreta)

O vetor de dados $\mu(\mathbf{n}) = [\mu(n)\ \mu(n-1)\ \ldots\ \mu(n-M+1)]^{\mathsf{T}}$, onde $\mu(\mathbf{n})$ é um processo estacionário no sentido amplo e no tempo discreto n, de média zero, com matriz de autocorrelação $\mathbf{R}$; $\mathbf{q_i} \leftrightarrow \lambda_i$ são autovetores (autovalores) associados a $\mathbf{R}$.

Logo: $\quad \underbrace{\mu(\mathbf{n}) = \sum_{i=1}^{M} C_i(n)\mathbf{q_i}}_{\text{síntese}} \quad , \quad \underbrace{C_i(n) = \mathbf{q_i}^{\mathsf{H}}\mu(\mathbf{n})}_{\text{análise}}, \quad i = 1, \ldots, M$

As equações de síntese e análise representam, respectivamente, construção e a decomposição de um sinal no espaço M-dimensional.

Pode-se calcular a energia de cada componente do espaço M-dimensional:

$$\mathbb{E}\{\underbrace{C_i(n)C_j^*(n)}_{\text{descorrelacionados}}\} = \mathbb{E}\{\mathbf{q_i}^{\mathsf{H}}\mu(\mathbf{n})\mathbf{q_j}\mu^*(\mathbf{n})\} = \mathbf{q_i}^{\mathsf{H}}\underbrace{\mathbb{E}\{\mu(\mathbf{n})\mu^{\mathsf{H}}(\mathbf{n})\}}_{\mathbf{R}}\mathbf{q_j}$$

$$= \mathbf{q_i}^{\mathsf{H}}\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{\mathsf{H}}\mathbf{q_j} = \begin{cases} \lambda_i, & i = j \\ 0, & i \neq j \end{cases}$$

Assim, λ_i é a energia de $\mu(\mathbf{n})$ ao longo de $\mathbf{q_i}$.
E a energia do sinal $\mu(\mathbf{n})$ é dada por:

$$\mathbb{E}\{\|\mu(\mathbf{n})\|^2\} = \sum_{i=1}^{M} \mathbb{E}\{|C_i(n)|^2\} = \sum_{i=1}^{M} \lambda_i$$

2.3 Modelagem de rank baixo

Modelagem de Baixo Rank

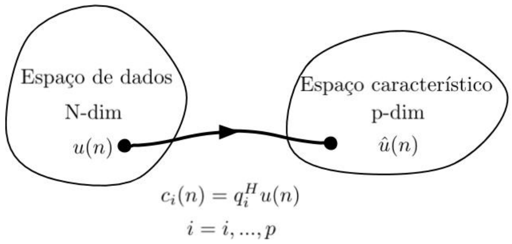
- Seleção de características importantes (essenciais) de um processo baseado num espaço de dados originais.
- **Objetivo desejável**

  Realização de uma transformação sobre o vetor de dados representando-o em sua essência, ou seja, por um número reduzido de suas características efetivas, retendo a maioria das informações intrínsicas dos dados originais (redução de ordem).

Problema de seleção-característica: Abordagem para expansão de Korhunen-Loève

- $u(n)$: Vetor de dados, no espaço M-dimensional, de uma realização particular de um processo estacionário no sentido amplo.
- **Questão**: Como transmitir $\mathbf{u}(n)$ num canal ruidoso baseado em "$\mathbf{p}$"parâmetros, $\mathbf{p} < \mathbf{M}$?
- $\mathbf{u}(n)$ expandido na base q_i, $i = 1, ..., M$, com autovalores associados λ de $\mathbf{R}$ de $\mathbf{u}(n)$, onde: $\lambda_i \neq \lambda_j, i \neq j$ e $\lambda_1 > \lambda_2 > ... > \lambda_i > \lambda_M$.
- $\mathbf{u}(n) \to$ representação exata: $\mathbf{u}(n) = \sum_{i=1}^{M} c_i(n)q_i$ (expansão de Karhunen-Loève). Rank $(\mathbf{u}(n)) = \mathbf{M}$.
- **Suposição**: Conhecimento a priori de $(M - p)$ autovalores $\lambda_{p+1}, ..., \lambda_M$ pequenos. Assim se tem os p^T significativos componentes de $\mathbf{u}(n)$, truncando a expansão.
- **Reconstrução aproximada de** $\mathbf{u}(n)$: $\hat{\mathbf{u}}(n) = \sum_{i=1}^{p} C_i(n)q_i, p < \mathbf{M}$, rank $(\hat{\mathbf{u}}(n)) = p$.
- $\mathbf{c}(n) = [c_1(n)c_2(n)...c_p(n)]^T$ representação de $\mathbf{u}(n)$ com rank reduzido.

  $\mathbf{c}_i(n) = q_i^H u(n)$ $i = i, ..., p$ decomposição em sub-espaço.



- **Erro instantânero produzido no processo de reconstrução de baixo rank**

$$\mathbf{e}(n) = \mathbf{u}(n) - \hat{\mathbf{u}}(n) = \sum_{i=p+1}^{M} \mathbf{c}_i(n) p_i \tag{2.14}$$

- **Erro quadrático médio**

$$E(n) = E\{\|\mathbf{e}(n)^2\|\} = E\{\mathbf{e}^H(n)\mathbf{e}(n)\} == \sum_{i=p+1}^{\mathbf{M}} \sum_{j=p+1}^{\mathbf{M}} E\{\mathbf{c}_i^*(n)\mathbf{c}_j(n)\} q_i^H q_j \tag{2.15}$$

com $1, i == j$ e $0, i \neq j$

$$E(n) = \sum_{i=p+1}^{M} E\{|c_i(n)|^2 = \sum_{i=p+1}^{M} \lambda i \tag{2.16}$$

logo, escolhido os menores autovalores, a reconstrução dos dados pode ser boa.

Uma aplicação de modelagem de Baixo-rank.

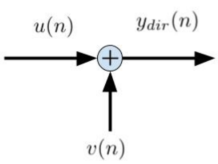
- **Problema**: Transmissão de dados em canal de comunicação ruidoso.
- **Modelo do sinal**

$$\mathbf{y}(n) = \mathbf{u}(n) + \mathbf{v}(n) \tag{2.17}$$

$\mathbf{v(n)} \rightarrow$ Problema.

- **Resultados estatísticos**:

$$E\{\mathbf{u}(n)\mathbf{v}^H(n)\} = 0 E\{\mathbf{v}(n)\mathbf{v}^H(n)\} = \sigma^2 \mathbf{I} \tag{2.18}$$



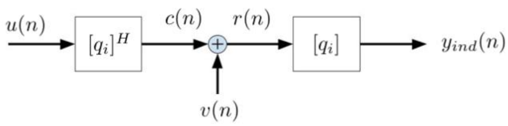
- **Reconstrução dos dados por dois métodos**
  Avaliação do erro de transmissão cometido.

- **i**) Método direto

$$E_{dir} = E\{|\mathbf{y}_{dir}(n) - \mathbf{u}(n)|^2\} = E\{|\mathbf{v}_m|^2\} = \sum_{i=1}^{M} E\{|\mathbf{v}_m|^2\} = \mathbf{M}\sigma^2 \tag{2.19}$$

- **ii**) Método indireto

  Processo de filtragem antes de transmitir e de recuperação por filtragem dos dados transformados recebidos.



- Seja

$$\mathbf{c}(n) = [q_1 q_2 ... q_p]^H \mathbf{u}(n) \tag{2.20}$$

  e o sinal recebido:

$$\mathbf{r}(n) = \mathbf{c}(n) + \mathbf{v}(n) \tag{2.21}$$

  .

- O sinal reconstruído é:

$$\mathbf{y}_{ind}(n) = [q_1 q_2 ... q_p]\mathbf{r}(n) \therefore \tag{2.22}$$

$$\mathbf{y}_{ind}(n) = [q_1 q_2 ... q_p]\mathbf{c}(n) + [q_1 q_2 ... q_p]\mathbf{v}(n)] \tag{2.23}$$

- **O erro de reconstrução total instantâneo**:

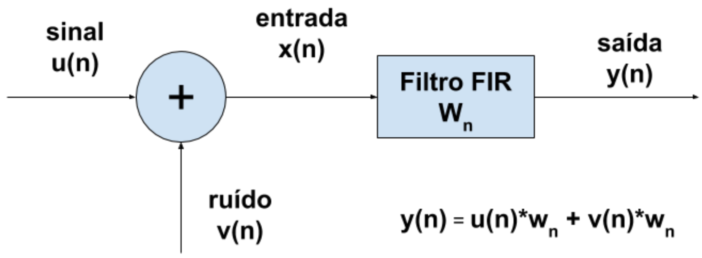
$$E_{dir} = E\{|\mathbf{y}_{dir}(n) - \mathbf{u}(n)|^2\} = \sum_{i=p+1}^{M} \lambda_i + p\sigma^2 \tag{2.24}$$

  sendo λ_i o erro devido ao baixo rank e $p\sigma^2$ o erro devido ao ruído do canal.

- Comparando os dois métodos, a modelagem para baixo rank **pode** oferecer uma vantagem: o erro indireto pode pode ser menor que o direto se:

$$\sum_{i=p+1}^{M} \lambda_i < (\mathbf{M} - p)\sigma^2 \tag{2.25}$$

- **Consequência**: É possível diminuir o erro com uma sub-modelagem do problema, sem associar perfeitamente com a física que gera os dados.



2.4 Autofiltros

- problema: otimização da relação sinal/ruído -> maximização.
- aplicações: comunicações, áreas ruidosas visando o cancelamento de ruído em geral.
- resultado: otimização do filtro <-> problema de autovalor.

Considere o sinal em ambiente ruidoso:

- $u(n)$: Processo estacionário no sentido amplo, média nula e matriz de autocorrelação $\mathbf{R}$.
- $v(n)$: Processo estacionário no sentido amplo, média nula, branco e potência σ^2.
- $\mathbb{E}\{u(n)v^*(m)\} = 0, \forall(n, m)$
- potência de $u(n) * w(n)$: $P_0 = \mathbf{w}^{\mathsf{H}}\mathbf{R}\mathbf{w}$ (mostrar!)
- potência de $v(n) * w(n)$: $N_0 = \mathbf{w}^{\mathsf{H}}(\sigma^2\mathbf{I})\mathbf{w} = \sigma^2||\mathbf{w}||^2$ (mostrar!)

  $(SNR)_0 = \dfrac{P_0}{N_0} = \dfrac{\mathbf{w}^{\mathsf{H}}\mathbf{R}\mathbf{w}}{\sigma^2\mathbf{w}^{\mathsf{H}}\mathbf{w}}$ -> coeficiente de Rayleigh
- Problema de otimização com restrição: achar $\mathbf{w}$ tal que $(SNR)_0$ seja máximo com $\mathbf{w}^{\mathsf{H}}\mathbf{w} = 1$ (restrição).
- Como $\left(\dfrac{P_0}{N_0}\right)$, a menos de $\left(\dfrac{1}{\sigma^2}\right)$, é o coeficiente de Rayleigh, este problema de otimização é semelhante ao teorema Minimax.

- Supondo $K = M$ no teorema Minimax: $\lambda_M = \max\limits_{\mathbf{x}\in\mathbb{J},\mathbf{x}\neq\mathbf{0}} \dfrac{\mathbf{x}^\mathsf{H}\mathbf{R}\mathbf{x}}{\mathbf{x}^\mathsf{H}\mathbf{x}} = \lambda_{\max}$

- Fazendo $\mathbf{x} = \mathbf{w} \to \max(SNR)_0 = \dfrac{\lambda_{\max}}{\sigma_2}$

- Como consequência: $\mathbf{W}_0 = \mathbf{q}_{\max}$, $\mathbf{q}_{\max} \leftrightarrow \lambda_{\max}$, $\mathbf{W}_0$ é o autofiltro máximo.

- Comentários:

  - O filtro FIR que maximiza a SNR é obtido da estrutura de autodecomposição de $\mathbf{R}$ de $u(n)$, fornecendo o autovetor associado ao maior autovalor, que é a reposta ao impulso do filtro.
  - O filtro ótimo é a contrapartida estocástica do filtro casado (determinístico), em ambiente ruidoso.

2.5 Cálculo de autovalores

2.6 Cálculo de Autovalores

- Vários programas: matlab®, mathematica®...

- Livro de referência: Handbook of Automatic Computation, V02: Linear Algebra, Wilkison & Reinsh (eds), 1971.

- Outras fontes apra matrizes reais: em C, em Fortran.

- Extensão de rotinas de autovalores reais para hermitianas:
  $\mathbf{A} = \mathbf{A}_r + j\mathbf{A}_i$
  $\mathbf{q} = \mathbf{q}_r + j\mathbf{q}_i$
  $\to \mathbf{A}\mathbf{q} = \lambda\mathbf{q} \to (\mathbf{A}_r + j\mathbf{A}_i)(\mathbf{q}_r + j\mathbf{q}_i) = \lambda(\mathbf{q}_r + j\mathbf{q}_i)$,
  ou ainda, (*): $\begin{pmatrix} \mathbf{A}_r & -\mathbf{A}_i \\ \mathbf{A}_r & \mathbf{A}_r \end{pmatrix}$
  $\begin{pmatrix} \mathbf{q}_r \\ \mathbf{q}_i \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{q}_r \\ \mathbf{q}_i \end{pmatrix}, \lambda \in \mathbb{R}$
  Como $\mathbf{A}$ é hermitiano: $\mathbf{A}^\mathsf{H} = \mathbf{A} \to \mathbf{A}_r^\mathsf{T} = \mathbf{A}_r, \mathbf{A}_i^\mathsf{T} = -\mathbf{A}_i$, logo $\begin{pmatrix} \mathbf{A}_r & \mathbf{A}_i \\ \mathbf{A}_i & \mathbf{A}_r \end{pmatrix}_{2M\times 2M}$ é uma matriz simétrica e real. Como o autovalor $\begin{pmatrix} -\mathbf{q}_i \\ \mathbf{q}_r \end{pmatrix}$ satisfaz (*), logo os (2M) autovalores λ's são duplamente repetidos: $\lambda_1, \lambda_1, \lambda_2, \lambda_2, \ldots, \lambda_M, \lambda_M$, e portanto:

  - λ's tem multiplicidade 2;
  - Autovetores para cada λ com multiplicidade 2: $\mathbf{q}_A = \mathbf{q}_r + j\mathbf{q}_i$; $\mathbf{q}_B = j\mathbf{q}_A$; $\angle\mathbf{q}_B = \angle\mathbf{q}_A + 90$

- Estratégias para calcular matrizes de autovalores

(i) diagonização (somente para matrizes Hermitianas): diagonalizar $\mathbf{A}$ com aplicação sucessiva de transformação de similaridade unitária.

(ii) triangularização (é igual para qualquer matriz quadrada)

- ∗ aplicação iterativa: algoritmo QL (usa matriz triangular inferior)
- ∗ procedimento iterativo na fatoração da matriz Hermitiana $M \times M$: $\mathbf{A}_n = \mathbf{Q}_n\mathbf{L}_n$, em que $\mathbf{Q}_n$ é matriz unitária, $\mathbf{L}_n$ é matriz triangular superior e n é o número de iterações.
  Em $(n+1)$: $\mathbf{A}_{n+1} = \mathbf{Q}_n\mathbf{L}_n$; como $\mathbf{Q}_n^{-1} = \mathbf{Q}_n^{\mathsf{H}} \rightarrow \mathbf{L}_n = \mathbf{Q}_n^{-1}\mathbf{A}_n = \mathbf{Q}_n^{\mathsf{H}}\mathbf{A}_n \therefore$ $\mathbf{A}_{n+1} = \mathbf{Q}_n^{\mathsf{H}}\mathbf{A}_n\mathbf{Q}_n$.
  Inicialização: $\mathbf{A}_0 = \mathbf{A}$

2.7 Exercícios

Ex. 21 — Calcule $\mathbf{R}^{-1}$ em função de λ_i e $\mathbf{P_i}$, onde λ_i representa os autovalores da matriz hermitiana $\mathbf{R}$ e $\mathbf{P_i}$ representa as respectivas matrizes de projeção de rank 1.

Ex. 22 — Prove:

$$\|\mathbf{u}(n)\|^2 = \sum_{i=1}^{M} |c_i(n)|^2 = \sum_{i=1}^{M} \lambda_i$$

Ex. 23 — Mostrar que a Potência de $\mathbf{u}(n) * \mathbf{w}_n$:

$$P_0 = \mathbf{w}^H R \mathbf{w}$$

Ex. 24 — Mostrar que a potência de $v(n) * \mathbf{w}_n$:

$$N_0 = \mathbf{w}^H (\sigma^2 I)\mathbf{w} = \sigma^2 \|\mathbf{w}\|$$

Capítulo 3

Decomposição em Valores Singulares

3.1 Decomposição em Valor Singular (SVD)

- método aplicável a matrizes quadradas e retangulares, reais e complexas;
- conveniente para superar problemas numéricos em soluções de problemas lineares;
- aplicável diretamente à matrizes de dados (decomposição em tempo real de aproximação das matrizes estatísticas);

Existem 2 formas diferentes das equações normais para calcular a solução dos mínimos quadrados lineares, matematicamente equivalentes

1. $\hat{\mathbf{w}} = \phi^{-1}\mathbf{z}$
2. $\hat{\mathbf{w}} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{d}$

O estimador $\hat{\mathbf{w}}$ pode ser escrito como $\hat{\mathbf{w}} = \mathbf{A}^\dagger\mathbf{d}$, onde $\mathbf{A}^\dagger = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ é a matriz pseudo-inversa ou inversa generalizada de Moore-Paramore da matriz $\mathbf{A}$.

Se $\mathbf{A}^\dagger$ for linearmente dependente em colunas, nulidade($\mathbf{A} \neq 0$) e a solução mínima quadrática para $\hat{\mathbf{w}}$ não pode ser obtia através de $\mathbf{A}^\dagger$ (não há soluçã única). O SVD é empregado para resolver essa situação.

O SVD também é empregado para encontrar $\hat{\mathbf{w}}$ com $\mathbf{A}^\dagger$, pois apresenta melhor precisão que com ϕ.

Teorema de Decomposição do Valor Singular

Procedimento algorítmico para fornecer informaçã quantitativa sobre a estrutura matricial de um sistema de equações lineares.

$$\mathbf{A}_{k\times m}\hat{\mathbf{w}}_{m\times 1} = \mathbf{d}_{k\times 1}$$

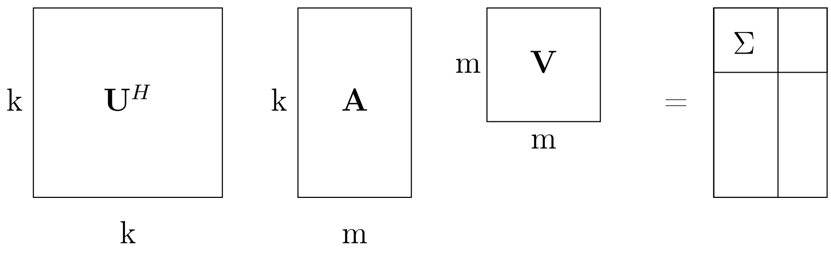
onde $\hat{\mathbf{w}}$ é a estimativa de um vetor de parâmetros desconhecidos.

Sejam $\mathbf{U}$ e $\mathbf{V}$ duas matrizes unitárias tal que

$$\mathbf{U}^H_{k\times k}\mathbf{A}_{k\times m}\mathbf{V}_{m\times m} = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \quad , \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_w \end{bmatrix}$$

onde $\sigma_1 \geq 0, \sigma_2 \geq 0, \cdots, \sigma_w \geq 0$ e $\mathbf{0} \equiv$ matriz nula.

Representação Gráfica



Observação:

$$\mathbf{U}^H\mathbf{AV} = \Sigma' \quad \therefore \quad \mathbf{UU}^H\mathbf{AV} = \mathbf{U}\Sigma' \quad \therefore \quad \underbrace{(\mathbf{U}^H\mathbf{U})^H}_{\mathbf{I}}\mathbf{AV} = \mathbf{U}\Sigma' \quad \therefore \quad \mathbf{AV} = \mathbf{U}\Sigma' \quad \therefore \quad \mathbf{AVV}^H = \mathbf{U}\Sigma'\mathbf{V}^H \quad \therefore \quad \mathbf{A}\underbrace{(\mathbf{V}^H\mathbf{V})^H}_{I} = \mathbf{U}\Sigma'\mathbf{V}^H \quad \therefore \quad \mathbf{A} = \mathbf{U}\Sigma'\mathbf{V}^H$$

Pode ser mostrado que $\mathbf{w} = \text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^H)$

Análise da relação entre as dimensões k e m das matrizes:

Um sistema é dito sobredeterminado se $k > m$. $(\mathbf{A}^H\mathbf{A})_{m\times m}$ é hermitiana e definida não negativa: autovalores positivos ($\lambda_i > 0$). Fazendo $\lambda_i = \sigma_i^2, \quad i = 1, \cdots, m$, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_w > 0$ e $\sigma_{w+1} = \sigma_{w+2} = \cdots = \sigma_m = 0$, $1 \leq w \leq m$. Assim, $\text{rank}(\mathbf{A}^H\mathbf{A}) = \text{rank}(\mathbf{A})$, e por isto há w autovalores não nulos.

Considere $\mathbf{v}_1, \cdots, \mathbf{v}_m$ os autovetores de $(\mathbf{A}^H\mathbf{A})$, com $\lambda_i = \sigma_i^2, \quad i = 1, \cdots, m$. Definindo a matriz $\mathbf{V}_{m\times m} = [\mathbf{v}_1, \cdots, \mathbf{v}_m]$, temos que a autodecomposição de $\mathbf{A}^H\mathbf{A}$ é

$$\begin{bmatrix} \lambda_1 & 0 & \dots & 0 & \cdots & 0 \\ 0 & \lambda_2 & \dots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \dots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma_m^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} = \mathbf{V}^H(\mathbf{A}^H\mathbf{A})\mathbf{V} = \begin{bmatrix} \Sigma^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}_{m\times m}$$

Particionando $\mathbf{V}$ temos $\mathbf{V} = [\mathbf{V}_1, \mathbf{V}_2]$, onde $\mathbf{V}_1 = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_w]$ e $\mathbf{V}_2 = [\mathbf{v}_{w+1}, \mathbf{v}_{w+2}, \cdots, \mathbf{v}_m]$. Logo,

$$\mathbf{V}_1^H \mathbf{V}_2 = \begin{bmatrix} \mathbf{v}_1^H \\ \mathbf{v}_2^H \\ \vdots \\ \mathbf{v}_w^H \end{bmatrix}_{w\times m} \begin{bmatrix} \mathbf{v}_{w+1} & \mathbf{v}_{w+2} & \cdots & \mathbf{v}_m \end{bmatrix}_{w\times m} = \mathbf{0}_{w\times(m-w)}$$

Dado que $\Sigma^2_{w\times w} = \Sigma \cdot \Sigma$,

$$[\mathbf{V}_1, \mathbf{V}_2]^h \left(\mathbf{A}^H\mathbf{A}\right)[\mathbf{V}_1, \mathbf{V}_2] = \begin{bmatrix} \mathbf{V}_1^H \\ \mathbf{V}_2^H \end{bmatrix} \left(\mathbf{A}^H\mathbf{A}\right)[\mathbf{V}_1, \mathbf{V}_2] = \begin{bmatrix} \mathbf{V}_1^H\mathbf{A}^H\mathbf{A} \\ \mathbf{V}_2^H\mathbf{A}^H\mathbf{A} \end{bmatrix}_{m\times m} [\mathbf{V}_1, \mathbf{V}_2]_{m\times m} =$$

$$\begin{bmatrix} \mathbf{V}_1^H\mathbf{A}^H\mathbf{A}\mathbf{V}_1 & \mathbf{V}_1^H\mathbf{A}^H\mathbf{A}\mathbf{V}_2 \\ \mathbf{V}_2^H\mathbf{A}^H\mathbf{A}\mathbf{V}_1 & \mathbf{V}_2^H\mathbf{A}^H\mathbf{A}\mathbf{V}_2 \end{bmatrix} \text{ onde } \begin{bmatrix} \overbrace{\mathbf{V}_1^H\mathbf{A}^H\mathbf{A}\mathbf{V}_1}^{\Sigma^2} & \overbrace{\mathbf{V}_1^H\mathbf{A}^H\mathbf{A}\mathbf{V}_2}^{\mathbf{0}} \\ \underbrace{\mathbf{V}_2^H\mathbf{A}^H\mathbf{A}\mathbf{V}_1}_{\mathbf{0}} & \underbrace{\mathbf{V}_2^H\mathbf{A}^H\mathbf{A}\mathbf{V}_2}_{\mathbf{0}} \end{bmatrix}$$

ou seja,

$$\mathbf{V}_1^H\mathbf{A}^H\mathbf{A}\mathbf{V}_1 = \Sigma^2 \quad \therefore \quad \Sigma^{-1}(\mathbf{V}_1^H\mathbf{A}^H\mathbf{A}\mathbf{V}_1)\Sigma{-1} = \mathbf{I}_{w\times w}$$

$$\mathbf{V}_2^H\mathbf{A}^H\mathbf{A}\mathbf{V}_2 = \mathbf{0} \quad \therefore \quad (\mathbf{A}\mathbf{V}_2)^H(\mathbf{A}\mathbf{V}_2) = \mathbf{0} \quad \therefore \quad \mathbf{A}\mathbf{V}_2 = \mathbf{0}$$

Seja $\mathbf{U}_1 = \mathbf{A}\mathbf{V}_1\Sigma^{-1} \rightarrow \mathbf{U}_1^H\mathbf{U}_1 = (\mathbf{A}\mathbf{V}_1\Sigma^{-1})^H\mathbf{A}\mathbf{V}_1\Sigma^{-1} = \Sigma^{-1}\underbrace{\mathbf{V}_1^H\mathbf{A}^H\mathbf{A}\mathbf{U}_1}_{\Sigma\cdot\Sigma}\Sigma^{-1} = \mathbf{I}$

Logo, as colunas de $\mathbf{U}_1$ são ortonormais entre si.

Seja $\mathbf{U}_2$ uma matriz $(k \times (k-w))$ tal que $\mathbf{U}_{k\times k} = [\mathbf{U}_1, \mathbf{U}_2]$ seja unitária.

Logo, $[\mathbf{U}_1, \mathbf{U}_2]^H [\mathbf{U}_1, \mathbf{U}_2] = \mathbf{I} \therefore \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix} [\mathbf{U}_1, \mathbf{U}_2] = \begin{bmatrix} (\mathbf{U}_1^H\mathbf{U}_1) & (\mathbf{U}_1^H\mathbf{U}_2) \\ (\mathbf{U}_2^H\mathbf{U}_1) & (\mathbf{U}_2^H\mathbf{U}_2) \end{bmatrix}$.

Desta forma, $\mathbf{U}_2^H\mathbf{U}_1 = \mathbf{U}_1^H\mathbf{U}_2 = \mathbf{0}$ e $\mathbf{U}_2^H\mathbf{U}_2 = I$. Ou seja, $\mathbf{U}_2$ também é unitária.

- Sistema Subdeterminado: $k < m$
  k linhas (número de equações) $< m$ colunas (número de variáveis).

  - Seja a matriz $k \times k$ $(\mathbf{A}\mathbf{A}^H)$ hermitiana e definida não negativa ($\lambda's$ reais).
  - λ's de $\mathbf{A}\mathbf{A}^H = \lambda_i = \sigma_i^2, i = 1, 2, 3 \cdots, k; \sigma_1 \geq \sigma_2 \geq \cdots \sigma_w > 0$ e $\sigma_{w+1} = \sigma_{w+2} = \cdots = \sigma_k = 0, 1 \leq w \leq k$.
  - autovetores de $\mathbf{A}\mathbf{A}^H$: $\mathbf{u}_i, i = 1, 2 \cdots, k$ associados aos autovalores: σ_i^2.

– Simétrica Unitária: $\mathbf{U}_{m\times k} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k] \rightarrow \mathbf{U}^H_{k\times m}\mathbf{U}_{m\times k} = \begin{bmatrix} \mathbf{u}_1^H \\ \mathbf{u}_2^H \\ \vdots \\ \mathbf{u}_k^H \end{bmatrix} [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k] =$ $\mathbf{I}_{k\times k}$

– Logo pode-se mostrar que a autodecomposição de $(\mathbf{A}\mathbf{A}^H)$(rank w) é :

$$\mathbf{U}^H(\mathbf{A}\mathbf{A}^H)\mathbf{U} = \begin{bmatrix} \mathbf{\Sigma^2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{3.1}$$

– Seja $\mathbf{U} = [\mathbf{U}_1\mathbf{U}_2], \mathbf{U}_1 = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_w]_{m\times w}\mathbf{U}_2 = [\mathbf{u}_{w+1}, \mathbf{u}_{w+2}, \cdots, \mathbf{u}_k]_{m\times(k-w)}$e portanto: $\mathbf{U}_1^H\mathbf{U}_2 = \mathbf{0}_{w\times(k-w)}$

– De (1) temos: $[\mathbf{U}_1, \mathbf{U}_2]^H(\mathbf{A}\mathbf{A}^H)[\mathbf{U}_1, \mathbf{U}_2] \rightarrow \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix} \mathbf{A}\mathbf{A}^H[\mathbf{U}_1\mathbf{U}_2] = \begin{bmatrix} \mathbf{U}_1^H\mathbf{A} \\ \mathbf{U}_2^H\mathbf{A} \end{bmatrix}$ $[\mathbf{A}^H\mathbf{U}_1\mathbf{A}^H\mathbf{U}_2] = \begin{bmatrix} (\mathbf{U}_1^H\mathbf{A}\mathbf{A}^H\mathbf{U}_1) & (\mathbf{U}_1^H\mathbf{A}\mathbf{A}^H\mathbf{U}_2) \\ (\mathbf{U}_2^H\mathbf{A}\mathbf{A}^H\mathbf{U}_1) & (\mathbf{U}_2^H\mathbf{A}\mathbf{A}^H\mathbf{U}_2) \end{bmatrix}$

– Comparando:
  * $\mathbf{U}_1^H\mathbf{A}\mathbf{A}^H\mathbf{U}_1 = \mathbf{\Sigma}^2 \rightarrow \mathbf{\Sigma}^{-1}\mathbf{U}_1^H\mathbf{A}\mathbf{A}^H\mathbf{U}_1\mathbf{\Sigma}^{-1} = \mathbf{I}_{w\times w}$
  * $\mathbf{U}_2^H\mathbf{A}\mathbf{A}^H\mathbf{U}_2 = \mathbf{0}, (\mathbf{A}^H\mathbf{U}_2)^H\mathbf{A}^H\mathbf{U}_2 = \mathbf{0} \rightarrow \mathbf{A}^H\mathbf{U}_2 = \mathbf{0}$

– Seja $\mathbf{V}_1(m \times w) = \mathbf{A}^H\mathbf{U}_1\mathbf{\Sigma}^{-1} \rightarrow \mathbf{V}_1^H\mathbf{V}_1 = (\mathbf{A}^H\mathbf{U}_1\mathbf{\Sigma}^{-1})^H(\mathbf{A}^H\mathbf{U}_1\mathbf{\Sigma}^{-1}) = \mathbf{\Sigma}^{-1}\mathbf{U}_1^H\mathbf{A}\mathbf{A}^H\mathbf{U}_1\mathbf{\Sigma}^{-1} \rightarrow \mathbf{\Sigma}^{-1}\mathbf{\Sigma}^2\mathbf{\Sigma}^{-1} = \mathbf{I}$.

  Logo $\mathbf{V}_1$ é unitária (colunas são ortonormais).

– Façamos:
  $\mathbf{V}_{m\times m} = [\mathbf{V}_1, \mathbf{V}_2]$, onde $\mathbf{V}_2$ é $m \times (w - m)$ e $\mathbf{V}$ unitária.

  $\mathbf{V}^H\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \end{bmatrix} [\mathbf{V}_1\mathbf{V}_2] = \begin{bmatrix} \mathbf{V}_1^H\mathbf{V}_1 & \mathbf{V}_1^H\mathbf{V}_2 \\ \mathbf{V}_2^H\mathbf{V}_1 & \mathbf{V}_2^H\mathbf{V}_2 \end{bmatrix} = \mathbf{I} \rightarrow \mathbf{V}_1$ e $\mathbf{V}_2$ são unitárias. E $\mathbf{V}_2^H\mathbf{V}_1 = \mathbf{0}$ ou $\mathbf{V}_1^H\mathbf{V}_2 = \mathbf{0}$.

– Portanto, pode-se observar:
  $\mathbf{U}^H\mathbf{A}\mathbf{V} = \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix} \mathbf{A}[\mathbf{V}_1\mathbf{V}_2] = \begin{bmatrix} \mathbf{U}_1^H\mathbf{A}\mathbf{V}_1 & \mathbf{U}_1^H\mathbf{A}\mathbf{V}_2 \\ \mathbf{U}_2^H\mathbf{A}\mathbf{V}_1 & \mathbf{U}_2^H\mathbf{A}\mathbf{V}_2 \end{bmatrix}$
  Como: $\mathbf{V}_1 = \mathbf{A}^H\mathbf{U}_1\mathbf{\Sigma}^{-1} \therefore \mathbf{V}_1\mathbf{\Sigma} = \mathbf{A}^H\mathbf{U}_1 \therefore \mathbf{\Sigma}\mathbf{V}_1^H = \mathbf{U}_1^H\mathbf{A} \therefore \mathbf{\Sigma} = \mathbf{U}_1^H\mathbf{A}\mathbf{V}_1$.

– Como: $\mathbf{A}^H\mathbf{U}_2 = \mathbf{0} \rightarrow \mathbf{U}_2^H\mathbf{A} = \mathbf{0} \rightarrow \mathbf{U}_2^H\mathbf{A}\mathbf{V}_1 = \mathbf{0}$ e $\mathbf{U}_2^H\mathbf{A}\mathbf{V}_2 = \mathbf{0}$
  $(\mathbf{V}_1^H\mathbf{A})\mathbf{V}_2 = \mathbf{\Sigma}\mathbf{V}_1\mathbf{V}_2 = \mathbf{0}$
  Logo: $\mathbf{U}^H\mathbf{A}\mathbf{V} = \begin{bmatrix} \mathbf{\Sigma^2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}_{k\times m}$

– Assim o Teorema SVD vale tanto para $k > m$ como para $k < m$.

– Terminologia e relação à autoanálise:

  * σ_i: valores singulares.
  * $\mathbf{v}_i, \cdots, \mathbf{v}_m$ Vetores singulares à direita de A, (autovetores de $\mathbf{A}^H\mathbf{A}$).
  * $\mathbf{u}_i, \cdots, \mathbf{u}_k$ Vetores singulares à esquerda de A, (autovetores de $\mathbf{A}\mathbf{A}^H$).
  * Número de valores singulares > 0 =rank de $\mathbf{A}$.
  * $\mathbf{U}^H\mathbf{A}\mathbf{V} = \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \rightarrow \mathbf{A}\mathbf{V} = \mathbf{U}\begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \rightarrow \left\{ \begin{array}{c} \mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i, i = 1, \cdots, w \\ \mathbf{A}\mathbf{v}_i = 0, i = w+1, \cdots, k \end{array} \right\}$

– Logo:
$\mathbf{A}\mathbf{A}^H = \mathbf{U}[\mathbf{\Lambda}]\mathbf{V}^H\mathbf{V}[\mathbf{\Lambda}]\mathbf{U}$
$\mathbf{U}^H(\mathbf{A}\mathbf{A}^H)\mathbf{U} = \begin{bmatrix} \mathbf{\Sigma}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \rightarrow \lambda' s$ de $\mathbf{A}\mathbf{A}^H = \lambda' s$ de $\mathbf{A}^H\mathbf{A}$.

– $\mathbf{A}\mathbf{V}\mathbf{V}^H = \mathbf{U}\begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{V}^H \therefore \mathbf{A} = [\mathbf{u}_1, \mathbf{u}_2, \cdots \mathbf{u}_k]_{k\times k}\begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}_{k\times m}\begin{bmatrix} \mathbf{v}_1^H \\ \vdots \\ \mathbf{v}_k^H \end{bmatrix}_{m\times m}$

$\mathbf{A} = [\mathbf{u}_1, \mathbf{u}_2, \cdots \mathbf{u}_k]_{k\times k}\begin{bmatrix} \sigma_1\mathbf{v}_1^H \\ \vdots \\ \sigma_k\mathbf{v}_k^H \end{bmatrix}_{k\times m} \quad \therefore \mathbf{A} = \sum_{i=1}^{w}\sigma_i\mathbf{u}_i\mathbf{v}_i \rightarrow$ Decomposição de $\mathbf{A}$ em w matrizes $\mathbf{u}_i, \mathbf{v}_i^H$.

– Também se tem:
$\mathbf{V}\mathbf{V}^H = I \rightarrow \mathbf{U}^H\mathbf{A}\mathbf{V}\mathbf{V}^H = \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{V}^H \therefore \mathbf{A}^H\mathbf{U} = \mathbf{V}\begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$

– Do que se tem: $\mathbf{A}^H\mathbf{u}_i = \sigma_i\mathbf{v}_i, i = 1, \cdots w$.
$\mathbf{A}^H\mathbf{u}_i = 0, i = w+1, \cdots m$.

– Sendo assim, $\mathbf{A}^H$ é decomposta na forma: $\mathbf{A}^H = \sum_{i=i}^{w}\sigma_i\mathbf{v}_i\mathbf{u}_i^H$.

– Pseudo Inversa

  * $\mathbf{A}$ é $k \times m$ com SVD :$\mathbf{A} = \mathbf{U}\begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{V}^H$
  * Definição de Pseudo Inversa: $\mathbf{A}^\dagger = \mathbf{V}\begin{bmatrix} \mathbf{\Sigma}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{U}^H, \mathbf{\Sigma}^{-1} = \mathrm{diag}(\sigma_1^{-1}, \cdots, \sigma_w^{-1}) \rightarrow$
  $\mathbf{A}^\dagger = \sum_{i=1}^{w}\frac{1}{\sigma_i}\mathbf{v}_i\mathbf{u}_i^H$.

– Caso sobredeterminado $\rightarrow k > m$
Suponha: $w = m \rightarrow (\mathbf{A}^H\mathbf{A})^{-1}\exists$.
Logo: $\mathbf{A}^\dagger = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H; (\mathbf{A}^H\mathbf{A})^{-1} = \mathbf{V}_1\mathbf{\Sigma}^{-2}\mathbf{V}_1^H$ e $\mathbf{A}^H = \mathbf{V}_1\mathbf{\Sigma}\mathbf{U}_1$
$(\mathbf{A}^H\mathbf{A})^{-1} = \mathbf{V}_1\mathbf{\Sigma}^{-1}\mathbf{U}_1^H = \mathbf{V}\begin{bmatrix} \mathbf{\Sigma}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{U}^H = \mathbf{A}^\dagger$

- Caso subdeterminado $\rightarrow m > k$
  Supondo: $w = k$
  $\mathbf{A}^{\dagger} = \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}$, neste caso $(\mathbf{A}\mathbf{A}^H)^{-1} = \mathbf{U}_1\mathbf{\Sigma}^{-2}\mathbf{U}_1^H$ e $\mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1} = (\mathbf{V}_1\mathbf{\Sigma}\mathbf{U}_1^H)(\mathbf{U}_1\mathbf{\Sigma}^{-2}\mathbf{U}_1^H) = \mathbf{V}_1\mathbf{\Sigma}^{-1}\mathbf{U}_1 = \mathbf{A}^{\dagger}$.