Algebra Linear

1. Sistema de Equações Lineares

2. Ortonormalização

Sistema de Equações Lineares

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{A \in \mathbb{R}^{m\times n}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{x \in \mathbb{R}^n} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{y \in \mathbb{R}^m}$$

- $a_{ij}, y_i \in \mathbb{R}$: dados do sistema; $x_i \in \mathbb{R}$: incógnitas
- Solução para $Ax = y$ quando $m > n$, $m = n$ ou $m < n$

- Problema: dados A e y

1. $\exists x : \; Ax = y$?
2. Se existe solução, qual o número de soluções LI?

Sistema de Equações Lineares

Definição O espaço Range da matriz A é o espaço gerado pelas colunas de A:

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

Exemplo $y = \begin{bmatrix} 1 & -1 \end{bmatrix}^T$ está no $\mathcal{R}(A)$, com

$$A = \begin{bmatrix} 1 & -1 & -3 \\ 0 & 10 & 0 \end{bmatrix} \ ?$$

Em outras palavras, $y = Ax$ para algum x? Resposta positiva, pois as colunas de A geram todo o espaço 2-dimensional

Rank (ou posto) da matriz $A \in \mathbb{R}^{m \times n}$ é definido como a dimensão do range de A (ou, equivalentemente, é dado pelo número de colunas LI de A, gerando o espaço $\mathbb{R}^m$). O posto é denotado por $\rho(A)$

- posto de colunas de A é a dimensão do $\mathcal{R}(A)$
- posto de linhas de A é a dimensão do $\mathcal{R}(A^*)$
- $\dim\mathcal{R}(A) = \dim\mathcal{R}(A^*) = r = \text{posto}(A)$ ou $\texttt{rank}(A)$

Sistema de Equações Lineares

Definição O espaço nulo de A é o espaço gerado por todos os vetores x satisfazendo $Ax = 0$, i.e.:

$$\mathcal{N}(A) \triangleq \{x \in \mathbb{R}^n \ : \ Ax = 0\} \subseteq \mathbb{R}^n$$

Exemplo $x = \begin{bmatrix} 1 & 10 & 0 \end{bmatrix}^T$ está no espaço $\mathcal{N}(A)$ com $A = \begin{bmatrix} 1 & -1 & -3 \\ 0 & 10 & 0 \end{bmatrix}$?

Em outras palavras, $Ax = 0$? Note que

$$\begin{bmatrix} 1 & -1 & -3 \\ 0 & 10 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 10 \\ 0 \end{bmatrix} = \begin{bmatrix} -9 \\ 100 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

e portanto a resposta é negativa

- A dimensão do espaço nulo é chamada de **nulidade** da matriz A e denotada por $\nu(A)$

Sistema de Equações Lineares

Note que $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ pode ser escrita como $\boldsymbol{y} = \boldsymbol{x_1 a_1} + \boldsymbol{x_2 a_2} + \cdots + \boldsymbol{x_n a_n}$ e, portanto, $\boldsymbol{x_i} \in \mathbb{R}$, $\boldsymbol{i = 1, \ldots, n}$ são as ponderações das colunas de $\boldsymbol{A}$

Propriedade

$\leadsto$ Para $\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{m \times n}}$:

$$\boldsymbol{\nu(A) = n - \rho(A)}$$

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

Sistema de Equações Lineares

$\rightsquigarrow$ Se $\nu(A) = 0$ então $\mathcal{N}(A) = \{0\}$ e as seguintes afirmações são equivalentes:

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

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

Sistema de Equações Lineares

Exemplo: Considere a matriz $A \in \mathbb{R}^{3\times 5}$ dada por

$$A = \begin{bmatrix} 0 & 1 & 1 & 2 & -1 \\ 1 & 2 & 3 & 4 & -1 \\ 2 & 0 & 2 & 0 & 2 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \end{bmatrix}$$

$$Ax = x_1 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

$$a_3 = a_1 + a_2 \;\; ; \;\; a_4 = 2a_2 \;\; ; \;\; a_5 = a_3 - a_4 = a_1 + a_2 - 2a_2 = a_1 - a_2$$

$$Ax = (x_1 + x_3 + x_5)a_1 + (x_2 + x_3 + 2x_4 - x_5)a_2$$

Como a_1 e a_2 são LI $\quad \Rightarrow \quad \rho(A) = 2$

Sistema de Equações Lineares

$$Ax = 0 \quad \Longleftrightarrow \quad \begin{cases} x_1 + x_3 + x_5 = 0 \\ x_2 + x_3 + 2x_4 - x_5 = 0 \end{cases}$$

Número de Equações $= \rho(A) = 2$

Número de Incógnitas $= 5$

Número de Graus de Liberdade $= 3$

Possíveis soluções LI:

$$v_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \; ; \; v_2 = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} \; , \; v_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$\{v_1, v_2, v_3\}$ formam uma base de $\mathcal{N}(A)$. Nulidade: $\nu(A) = 3$

Sistema de Equações Lineares

Teorema

1. Dada uma matriz $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ e um vetor $\boldsymbol{y} \in \mathbb{R}^{m \times 1}$, existe uma solução $\boldsymbol{x} \in \mathbb{R}^{n \times 1}$ da equação $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ se e somente se $\boldsymbol{y} \in \mathcal{R}(\boldsymbol{A})$ ou, de forma equivalente:

$$\rho(\boldsymbol{A}) = \rho\left(\left[\begin{array}{cc} \boldsymbol{A} & \boldsymbol{y} \end{array}\right]\right)$$

2. Dada uma matriz $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, uma solução $\boldsymbol{x} \in \mathbb{R}^{n}$ de $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ existe para todo $\boldsymbol{y} \in \mathbb{R}^{m}$ se e somente se $\rho(\boldsymbol{A}) = m$ (rank completo de linhas)

Sistema de Equações Lineares

Teorema (Parametrização de todas as soluções) Dada uma matriz $A \in \mathbb{R}^{m\times n}$ e um vetor $y \in \mathbb{R}^m$, seja x_p uma solução $x \in \mathbb{R}^n$ da equação $y = Ax$. Seja $k = n - \rho(A) = \nu(A)$ a nulidade de A.

⇝ Se $\rho(A) = n$ (rank completo de colunas) então a solução x_p é única

⇝ Se $k > 0$, então para todo $\alpha_i \in \mathbb{R}$, $i = 1, \ldots, k$ o vetor

$$x = x_p + \alpha_1 n_1 + \alpha_2 n_2 + \cdots + \alpha_k n_k = x_p + \sum_{i=1}^{k} \alpha_i n_i$$

é uma solução de $Ax = y$, sendo $\{n_1, \ldots, n_k\}$ uma base de $\mathcal{N}(A)$

Note que da última condição: $Ax = Ax_p + A\sum_{i=1}^{k} \alpha_i n_i = Ax_p + 0 = y$

Sistema de Equações Lineares

Exemplo Considere a descrição:

$$Ax = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 2 & 3 & 4 \\ 2 & 0 & 2 & 0 \end{bmatrix} x = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \end{bmatrix} x = \begin{bmatrix} -4 \\ -8 \\ 0 \end{bmatrix} = y$$

Note que como $y \in \mathcal{R}(A)$ então $\exists x$ solução para $Ax = y$ (item 1 do Teorema na pag. 10). Resolvendo para A e y, obtém-se uma solução particular:

$$x_p = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -2 \end{bmatrix}. \quad \text{Outra solução?} \quad x_p = \begin{bmatrix} 0 \\ -4 \\ 0 \\ 0 \end{bmatrix}$$

Matlab:

```
A=[0 1 1 2;1 2 3 4;2 0 2 0]; y=[-4;-8;0]; xp=linsolve(A,y)
```

Sistema de Equações Lineares

Note que a matriz $\boldsymbol{A} \in \mathbb{R}^{3\times 4}$ tem posto $\rho(\boldsymbol{A}) = 2$. E.g., $\boldsymbol{a_1}$ e $\boldsymbol{a_2}$ são LI. A 3a. coluna é combinação de $\boldsymbol{a_3} = \boldsymbol{a_1} + \boldsymbol{a_2}$ e a 4a. coluna é $\boldsymbol{a_4} = 2\boldsymbol{a_2}$. Desta forma:

$$\boldsymbol{A}\boldsymbol{x} = (x_1 + x_3)\boldsymbol{a_1} + (x_2 + x_3 + 2x_4)\boldsymbol{a_2}$$

Note que a nulidade de $\boldsymbol{A}$ é $\nu(\boldsymbol{A}) = n - \rho(\boldsymbol{A}) = 4 - 2 = 2$ e, então, tem-se dois vetores que geram o espaço nulo calculados de:

$$\boldsymbol{A}\boldsymbol{x} = 0 \iff \begin{cases} x_1 + x_3 = 0 \\ x_2 + x_3 + 2x_4 = 0 \end{cases}$$

Possíveis soluções LI que formam uma base para $\mathcal{N}(\boldsymbol{A})$ (há diversas soluções!):

$$n_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}; n_2 = \begin{bmatrix} 0 \\ 2 \\ 0 \\ -1 \end{bmatrix}$$. Usando `null(A,'r')`: $n_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}; \; n_2 = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 0 \end{bmatrix}$

Sistema de Equações Lineares

Aglutinando estas informações, via Teorema na pág. 11, tem-se uma solução geral para $\boldsymbol{Ax} = \boldsymbol{y}$ da forma:

$$\boldsymbol{x} = \boldsymbol{x_p} + \boldsymbol{\alpha_1 n_1} + \boldsymbol{\alpha_2 n_2} = \boldsymbol{x_p} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{-2} \end{bmatrix} + \boldsymbol{\alpha_1} \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{-1} \\ \mathbf{0} \end{bmatrix} + \boldsymbol{\alpha_2} \begin{bmatrix} \mathbf{0} \\ \mathbf{2} \\ \mathbf{0} \\ \mathbf{-1} \end{bmatrix}$$

para quaisquer valores de $\boldsymbol{\alpha_1} \in \mathbb{R}$ e $\boldsymbol{\alpha_2} \in \mathbb{R}$

Notas gerais

⇝ Seja $A \in \mathbb{R}^{m \times n}$. Se $C \in \mathbb{R}^{n \times n}$ e $D \in \mathbb{R}^{m \times m}$ são não singulares, então:

$$\rho(AC) = \rho(A) \quad \text{e} \quad \rho(DA) = \rho(A)$$

Isto é, o posto de uma matriz não se altera ao ser pré ou pós-multiplicada por uma matriz não singular

⇝ Considere que $A \in \mathbb{C}^{m \times n}$ e A^* seja sua conjugada transposta. Então:

- se $\rho(A) = n \iff \rho(A^*A) = n \;;\; \det(A^*A) \neq 0$
- se $\rho(A) = m \iff \rho(AA^*) = m \;;\; \det(AA^*) \neq 0$

⇝ Observe que $\rho(A) = n$ implica:

- $n \leq m$
- $A\alpha = 0 \quad \Rightarrow \quad \alpha = 0$

Notas gerais

$\rightsquigarrow$ Para uma matriz quadrada $A \in \mathbb{R}^{n\times n}$, A é invertível ou não-singular se $\det(A) \neq 0$ ou, de forma equivalente:

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

Sistema de Equações Lineares

⇝ Considere o sistema $y = Ax$ com $A \in \mathbb{R}^{m\times n}$ e $m > n$, i.e., há mais equações do que incógnitas (pode não ter solução)...

Aproximação – define-se o erro: $e = Ax - y$ e busca-se x^* que minimiza $\|e\|$.

A solução x^* gera o ponto no $\mathcal{R}(A)$ que está mais próximo de y, ou seja, Ax_a é a projeção de y no $\mathcal{R}(A)$. Escreva na forma de um problema de otimização:

$$f(x) = \min_{x} \ \|e\|^2 \ = \ \min_{x} \ e'e \ = \ \min_{x} \ (Ax - y)'(Ax - y)$$

Condição de Otimalidade – derivando em relação a x: $2x'A'A - 2y'A = 0$

▶ Assumindo que A tem posto completo $\rho(A) = n$ (e portanto $A'A$ é invertível), então uma solução aproximada é:

$$\boxed{x^* = (A'A)^{-1}A'y}$$

Sistema de Equações Lineares

- x^* é a solução exata de $y = Ax$ se $y \in \mathcal{R}(A)$

$\rightsquigarrow$ $A^\dagger \triangleq (A'A)^{-1}A'$ é chamada de pseudo-inversa de A (inversa à esquerda)

$\rightsquigarrow$ A projeção de y no $\mathcal{R}(A)$ é linear e é dada por

$$Ax^* = A(A'A)^{-1}A'y$$

e $A(A'A)^{-1}A'$ é chamada matriz de projeção

$\rightsquigarrow$ Note que o erro:

$$e = Ax^* - y = \left[A(A'A)^{-1}A' - \mathbf{I}\right]y$$

é ortogonal ao $\mathcal{R}(A)$, i.e.,

$$\langle e, Ax \rangle = y' \left[A(A'A)^{-1}A' - \mathbf{I}\right]' Ax = 0$$

Estimador Mínimos Quadrados

$\rightsquigarrow$ Vários problemas em engenharia envolvendo reconstrução de informação, inversão ou estimação podem ser colocados na forma $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\Delta}$, sendo $\boldsymbol{x}$ o vetor a ser estimado ou reconstruído; $\boldsymbol{y}$ o vetor de informações ou medidas e $\boldsymbol{\Delta}$ o vetor de ruídos, erros de medida ou incerteza não modelada

Problema – Obter a estimativa $\hat{\boldsymbol{x}}$ que minimiza a diferença entre os valores medidos $\boldsymbol{y}$ e o que deveria ser obtido se $\boldsymbol{\Delta} = \boldsymbol{0}$

Solução: $\hat{\boldsymbol{x}} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{y}$

Sistemas de Equações Lineares

$\rightsquigarrow$ Para o sistema $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ com $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ e $\boldsymbol{n} > \boldsymbol{m}$, considere o problema de otimização que minimiza a norma-2 de $\boldsymbol{x}$, i.e.:

$$\begin{cases} \min\limits_{\boldsymbol{x}} & \|\boldsymbol{x}\|^2 \quad (= \boldsymbol{x}'\boldsymbol{x}) \\ \text{sujeito a} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y} \end{cases}$$

Como resolvê-lo? Introduza o multiplicador de Lagrange $\boldsymbol{\lambda}$ e o Lagrangeano:

$$\mathbb{L}\,(\boldsymbol{x}, \boldsymbol{\lambda}) = \boldsymbol{x}'\boldsymbol{x} + \boldsymbol{\lambda}'(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y})$$

As condições de otimalidade são obtidas derivando o Lagrangeano:

$$\frac{\partial \mathbb{L}}{\partial \boldsymbol{x}} = 2\boldsymbol{x}' + \boldsymbol{\lambda}'\boldsymbol{A} = \boldsymbol{0} \quad \text{e} \quad \frac{\partial \mathbb{L}}{\partial \boldsymbol{\lambda}} = (\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}) = \boldsymbol{0}$$

Veja que $\boldsymbol{x} = -\boldsymbol{A}'\boldsymbol{\lambda}/2$. Como $(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y})' = \boldsymbol{0} \Rightarrow \ \boldsymbol{\lambda} = -2(\boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{y}$, então:

$$\boxed{\boldsymbol{x}^* = \boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{y} \qquad (\exists\ \boldsymbol{x}^* \text{ se } \boldsymbol{\rho}(\boldsymbol{A}) = \boldsymbol{m} \text{ implicando } \boldsymbol{\rho}(\boldsymbol{A}\boldsymbol{A}') = \boldsymbol{m})}$$

Ortonormalização

$\rightsquigarrow$ Um vetor $\boldsymbol{x}$ é dito estar normalizado se sua norma Euclidiana é igual a $\mathbf{1}$, i.e., $\boldsymbol{x}'\boldsymbol{x} = \langle \boldsymbol{x}, \boldsymbol{x} \rangle = \mathbf{1}$

Dois vetores $\boldsymbol{x}_1$ e $\boldsymbol{x}_2$ são ortogonais se $\boldsymbol{x}_1'\boldsymbol{x}_2 = \boldsymbol{x}_2'\boldsymbol{x}_1 = \mathbf{0}$

Um conjunto de vetores $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ é ortonormal se

$$\boldsymbol{x}_i'\boldsymbol{x}_j = \begin{cases} \mathbf{0} & \text{se } i \neq j \\ \mathbf{1} & \text{se } i = j \end{cases}$$

Ortonormalização

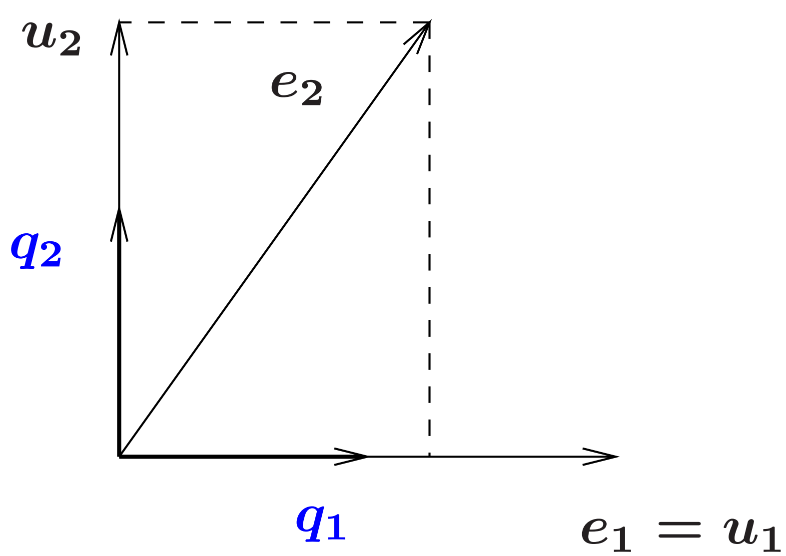
$\rightsquigarrow$ Dado um conjunto de vetores LI $\{e_1, e_2, \ldots, e_n\}$, pode-se obter um conjunto de vetores ortonormais usando o procedimento de ortonormalização de Schmidt:

$$
\begin{aligned}
u_1 &\triangleq e_1 & q_1 &\triangleq u_1/\|u_1\| \\
u_2 &\triangleq e_2 - (q_1' e_2) q_1 & q_2 &\triangleq u_2/\|u_2\| \\
&\vdots & &\vdots \\
u_n &\triangleq x_n - \sum_{k=1}^{n-1} (q_k' e_n) q_k & q_n &\triangleq u_n/\|u_n\|
\end{aligned}
$$

$\looparrowright$ A primeira equação apenas normaliza o vetor e_1, tal que $q_1' q_1 = 1$

Ortonormalização

↬ Na segunda equação, $(q_1'e_2)q_1$ é a projeção do vetor e_2 ao longo de q_1. E para $u_2 = e_2 - (q_1'e_2)q_1$, $q_2 = u_2/\|u_2\|$ é ortogonal ao vetor q_1



Ortonormalização - Propriedades

⇝ Se um conjunto de vetores $\boldsymbol{u_1}$, $\boldsymbol{u_2}$, . . . , $\boldsymbol{u_n}$ é ortonormal então

$$\boldsymbol{U} \triangleq \begin{bmatrix} \boldsymbol{u_1} & \boldsymbol{u_2} & \cdots & \boldsymbol{u_n} \end{bmatrix} \quad \Rightarrow \quad \boldsymbol{U'U = \mathsf{I}_n}$$

Propriedade: Se $\boldsymbol{y = Ux}$, então $\displaystyle\sum_{i=1}^{n} \boldsymbol{y_i^2} = \mathsf{I} \sum_{i=1}^{n} \boldsymbol{x_i^2}$ ou, em outras palavras:

$$\|\boldsymbol{y}\|^2 = \|\boldsymbol{Ux}\|^2 = (\boldsymbol{Ux})'(\boldsymbol{Ux}) = \boldsymbol{x'}\underbrace{\boldsymbol{U'U}}_{=\mathsf{I}}\boldsymbol{x} = \boldsymbol{x'x} = \|\boldsymbol{x}\|^2$$

Neste caso, o mapeamento $\boldsymbol{y = Ux}$ é uma isometria (não altera a norma)

Propriedade: Se $\boldsymbol{y = Ux}$ e $\boldsymbol{\tilde{y} = U\tilde{x}}$ então $\langle \boldsymbol{y, \tilde{y}} \rangle = \langle \boldsymbol{x, \tilde{x}} \rangle$ (a multiplicação por $\boldsymbol{U}$ não altera o produto interno) pois

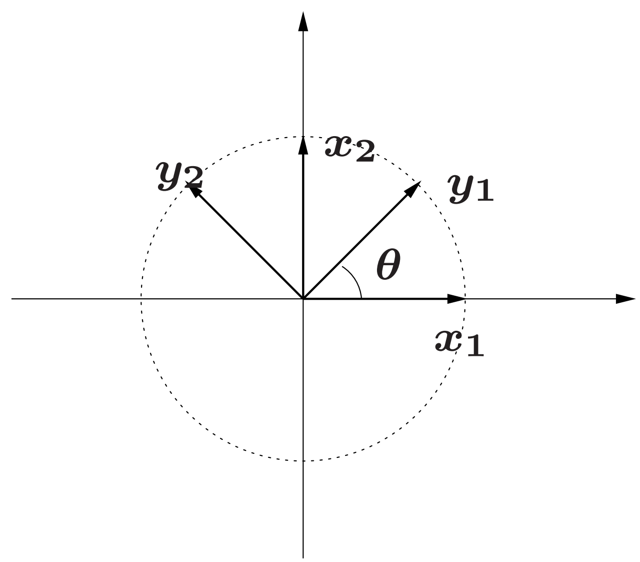
$$\langle \boldsymbol{y, \tilde{y}} \rangle = \langle \boldsymbol{Ux, U\tilde{x}} \rangle = (\boldsymbol{Ux})'(\boldsymbol{U\tilde{x}}) = \boldsymbol{x'}\underbrace{\boldsymbol{U'U}}_{=\mathsf{I}}\boldsymbol{x} = \langle \boldsymbol{x, \tilde{x}} \rangle$$

Ortonormalização

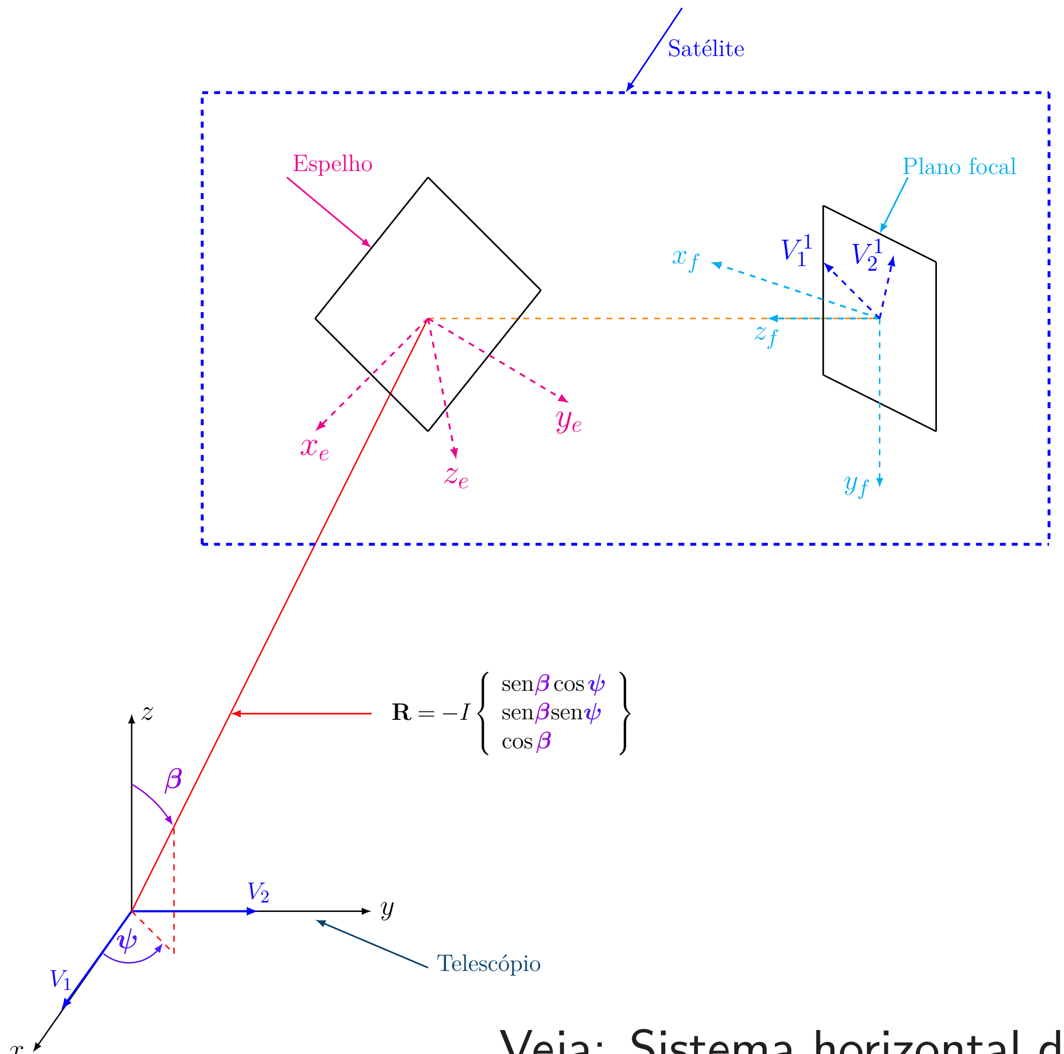
Exemplo: Transformações de rotação, reflexão e translação de vetores via matriz ortogonal

⇝ A transformação que rotaciona um vetor no sentido anti-horário de $\boldsymbol{\theta}$ no $\mathbb{R}^2$ é:

$$\boldsymbol{y = U_\theta x} \quad ; \quad \boldsymbol{U_\theta} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \; ; \; \boldsymbol{U_\theta' U_\theta = I}$$



Rastreamento de satélite – Telescópio terrestre



Quatro sistemas de coordenadas

1) Coordenadas na Terra (x,y,z)

2) Coordenadas do Satélite

3) Espelho (x_e,y_e,z_e)

4) Plano focal (x_f,y_f,z_f)

Posição do Satélite

definida na figura em relação a:

- azimute ψ
- zênite β
- inclinação I

Projeção, Rotação e Reflexão

Veja: Sistema horizontal de coordenadas

Outro Exemplo – Encontrar projeções que minimizem o erro de reconstrução com base ortonormal

▷ A ideia é projetar um conjunto de dados no espaço n-dimensional em outro de dimensão menor, preservando o máximo de informação possível. Obter a melhor aproximação, por exemplo, de dados em 3D no 2D

▷ Para quê?

↬ Comprimir, redução de ruído

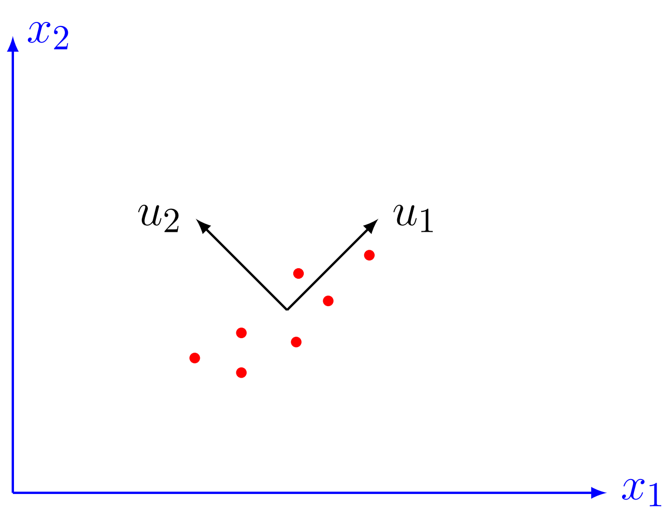
↬ Reduzir a dimensão do espaço de características para fins de classificação

⇝ Pode-se usar uma projeção que minimiza o erro quadrático em reconstruir os dados originais

Encontrar Projeções que minimizem o erro de reconstrução

Por exemplo, suponha que os dados são representados por um conjunto de ϕ vetores com dimensão n cada um deles: $\chi = \left\{ x'_1, x'_2, \ldots, x'_\phi \right\}$

É possível escrever o k-ésimo vetor da forma: $x_k = \sum_{i=1}^{\text{posto}(\chi)} z_i^k u_i$, sendo que os vetores u_i formam um conjunto de vetores ortonormais

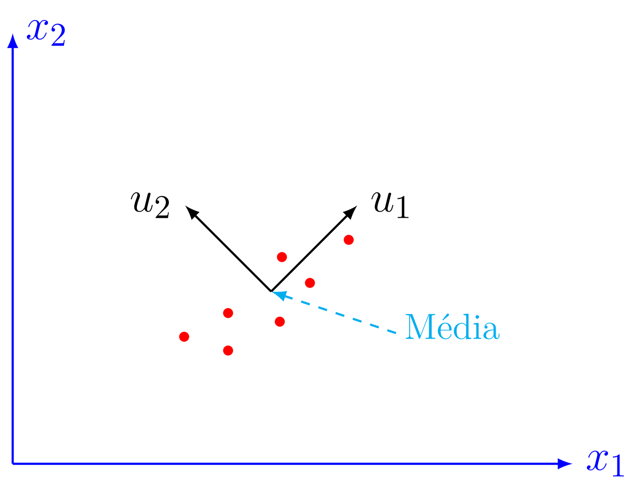


Encontrar Projeções que minimizem o erro de reconstrução

PCA: Dado $\zeta < \phi$ encontre o conjunto de vetores ortonormais $\{u_1, u_2, \ldots, u_\zeta\}$ tal que minimize o erro quadrático:

$$J_\zeta = \sum_{k=1}^{\phi} \|x_k - \hat{x}_k\|^2$$

sendo que $\hat{x}_k = \bar{x} + \sum_{i=1}^{\zeta} z_i^k u_i$ e $\bar{x}$ é a média dada por $\bar{x} = \frac{1}{\phi}\sum_{i=1}^{\phi} x_i$



Ortonormalização

$\rightsquigarrow$ O procedimento de ortonormalização pode ser descrito como uma transformação linear. Considere a matriz $\boldsymbol{X}$ formada por $\boldsymbol{n}$ vetores LI, $\{\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_n}\}$, da forma: $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x_1} & \boldsymbol{x_2} & \cdots & \boldsymbol{x_n} \end{bmatrix}$

Deseja-se determinar uma matriz $\boldsymbol{F}$ tal que

$$\boldsymbol{X}\boldsymbol{F} = \boldsymbol{Q} \triangleq \begin{bmatrix} \boldsymbol{q_1} & \boldsymbol{q_2} & \cdots & \boldsymbol{q_n} \end{bmatrix}$$

A condição para que os vetores $\{\boldsymbol{q_1}, \boldsymbol{q_2}, \ldots, \boldsymbol{q_n}\}$ sejam ortonormais pode ser escrita na forma

$$\boldsymbol{Q}'\boldsymbol{Q} = \mathbf{I} \quad \Longrightarrow \quad \boldsymbol{F}'\boldsymbol{R}\boldsymbol{F} = \mathbf{I} \quad ; \quad \boldsymbol{R} \triangleq \boldsymbol{X}'\boldsymbol{X}$$

Sistema com $\boldsymbol{n(n+1)/2}$ restrições e $\boldsymbol{n^2}$ incógnitas ($\boldsymbol{n(n-1)/2}$ elementos de $\boldsymbol{F}$ são arbitrários). Solução?

Ortonormalização

$\rightsquigarrow$ Uma solução particular pode ser obtida pela **fatorização de Cholesky** da matriz $\boldsymbol{R}$, que produz uma matriz triangular superior $\boldsymbol{L}$ tal que $\boldsymbol{R} = \boldsymbol{L}'\boldsymbol{L}$:

$$\boldsymbol{F}'\boldsymbol{R}\boldsymbol{F} = \mathsf{I} = \boldsymbol{F}'\boldsymbol{L}'\boldsymbol{L}\boldsymbol{F} = (\boldsymbol{L}\boldsymbol{F})'(\boldsymbol{L}\boldsymbol{F})$$

e por inspeção uma solução é dada por $\boldsymbol{F} = \boldsymbol{L}^{-1}$

$\rightsquigarrow$ A **fatorização de Schur** aplicada à matriz $\boldsymbol{R}$ produz uma matriz unitária $\boldsymbol{V}$ e uma matriz diagonal $\boldsymbol{\Omega}$ composta pelos autovalores de $\boldsymbol{R}$ tal que $\boldsymbol{R} = \boldsymbol{V}\boldsymbol{\Omega}\boldsymbol{V}'$, sugerindo como solução

$$\boldsymbol{F}'\boldsymbol{R}\boldsymbol{F} = \mathsf{I} = \boldsymbol{F}'\boldsymbol{V}\boldsymbol{\Omega}\boldsymbol{V}'\boldsymbol{F} = (\boldsymbol{V}\boldsymbol{\Omega}^{\mathbf{0.5}}\boldsymbol{V}'\boldsymbol{F})'\left(\boldsymbol{V}\boldsymbol{\Omega}^{\mathbf{0.5}}\boldsymbol{V}'\boldsymbol{F}\right)$$

$\therefore \quad \boldsymbol{F} = \boldsymbol{V}\boldsymbol{\Omega}^{\mathbf{-0.5}}\boldsymbol{V}'$

Ortonormalização via Otimização

$\rightsquigarrow$ A escolha da transformação $\boldsymbol{F}$ que ortonormaliza $\boldsymbol{X}$ pode ser orientada também por algum critério específico, descrito por problemas de otimização restritos. Por exemplo, considere encontrar $\boldsymbol{Q}$ (que guarda a base ortonormal) e seja a solução para:

$$\begin{cases} \min\limits_{\boldsymbol{Q}} & \mathbf{Tr}\ \{(\boldsymbol{X}-\boldsymbol{Q})'(\boldsymbol{X}-\boldsymbol{Q})\} \\ \text{sujeito a} & \boldsymbol{Q}'\boldsymbol{Q} = \mathbf{I} \end{cases}$$

i.e., $\boldsymbol{Q}$ minimiza a norma ao quadrado de $\boldsymbol{x_i} - \boldsymbol{q_i}$, $\boldsymbol{i = 1, \ldots, n}$ (denotado pelo traço). Esta transformação procura preservar o máximo possível os vetores originais. Note que $\mathbf{Tr}(\boldsymbol{X}-\boldsymbol{Q})$ é uma função linear dos elementos da matriz $\boldsymbol{Q}$

Re-escrevendo o problema de otimização em termos de $\boldsymbol{F}$ tem-se

Ortonormalização via Otimização

Re-escrevendo em termos de F

$$\begin{cases} \min\limits_{F} & \mathsf{Tr}\ \{(X - XF)'(X - XF)\} \\ \text{sujeito a} & F'X'XF = \mathsf{I} \end{cases}$$

$$\Updownarrow$$

$$\begin{cases} \min\limits_{F} & \mathsf{Tr}(X'X + F'X'XF - F'X'X - X'XF) \\ \text{sujeito a} & F'X'XF = \mathsf{I} \end{cases}$$

Substituindo a restrição ($F'X'XF = \mathsf{I}$) na função objetivo tem-se o termo constante ($X'X + \mathsf{I}$). Note que termos constantes na função objetivo não influenciam na solução F, então pode-se escrever:

$$\begin{cases} \min\limits_{F} & \mathsf{Tr}\ (-F'X'X - X'XF) \ = \ \max\ \mathsf{Tr}\ (2X'XF) \\ \text{sujeito a} & F'X'XF = \mathsf{I} \end{cases}$$

... e a solução?

Derivadas de funções matricias

$\rightsquigarrow$ Considere $\boldsymbol{X} \triangleq [\boldsymbol{x_{ij}}]$, $\boldsymbol{X} \in \mathbb{C}^{\boldsymbol{m \times n}}$ e $\boldsymbol{F(X)}$ uma função escalar (real ou complexa) de $\boldsymbol{X}$. Então

$$\frac{\partial}{\partial X} F(X) \triangleq \left[\frac{\partial}{\partial x_{ij}} F(X) \right]$$

$\rightsquigarrow$ Considere $\boldsymbol{A}$ e $\boldsymbol{B}$ matrizes complexas:

$$\frac{\partial}{\partial X} \mathrm{Tr}(AXB) = A'B'$$

$$\frac{\partial}{\partial X} \mathrm{Tr}(AX'B) = BA$$

$$\frac{\partial}{\partial X} \mathrm{Tr}(AXBX) = A'X'B' + B'X'A'$$

$$\frac{\partial}{\partial X} \mathrm{Tr}(AXBX') = A'XB' + AXB$$

Ortonormalização via Otimização

$\rightsquigarrow$ Para obter uma solução, escreva o Lagrangeano $\mathbb{L}\,(\boldsymbol{F},\boldsymbol{\Lambda})$ ($\boldsymbol{\Lambda}$ é a variável dual simétrica associada à restrição original do problema: $\boldsymbol{F}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{F} = \mathbf{I}$):

$$\mathbb{L}\,(\boldsymbol{F},\boldsymbol{\Lambda}) \;=\; \mathbf{Tr}\;[\mathbf{2}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{F} + \boldsymbol{\Lambda}\,(\boldsymbol{F}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{F} - \mathbf{I})]$$

As condições necessárias de otimalidade são dadas por:

$$\frac{\partial\mathbb{L}}{\partial\boldsymbol{F}} = \mathbf{2}(\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{X}'\boldsymbol{X}\boldsymbol{F}\boldsymbol{\Lambda}) = \mathbf{2}\boldsymbol{X}'\boldsymbol{X}(\mathbf{I} + \boldsymbol{F}\boldsymbol{\Lambda}) = \mathbf{0}$$

$$\frac{\partial\mathbb{L}}{\partial\boldsymbol{\Lambda}} = (\boldsymbol{F}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{F} - \mathbf{I}) = \mathbf{0}$$

Ortonormalização via Otimização

$\rightsquigarrow$ Note que X é não singular. Portanto resolvendo em termos de F e de Λ em $2X'X(\mathsf{I} + F\Lambda) = 0$, necessariamente $2X'X \neq 0$ e então $(\mathsf{I} + F\Lambda) = 0$, tal que:

$$F = -\Lambda^{-1} \Longrightarrow FX'X'F = \mathsf{I} \Longrightarrow (-\Lambda^{-1})X'X(-\Lambda^{-1}) = \mathsf{I} \Longrightarrow X'X = \Lambda^2$$

$$\Lambda = \begin{cases} +(X'X)^{0.5} \\ -(X'X)^{0.5} \end{cases} \quad \Longrightarrow \quad F = \begin{cases} -(X'X)^{-0.5} \\ +(X'X)^{-0.5} \end{cases}$$

Veja que o custo inicial é dado por:

$$\min \ \mathsf{Tr}\,(-2X'XF) \ = \ \max \ \mathsf{Tr}\,(2X'XF)$$

e a solução ótima (que minimiza) é obtida para $F = (X'X)^{-0.5}$

$\rightsquigarrow$ Note que a solução $F = -(X'X)^{-0.5}$ fornece a mesma base ortonormalizada Q (apenas trocando o sinal dos vetores) e pode ser interpretada como a transformação F que maximiza a diferença entre as normas (ao quadrado) dos vetores de cada base

Ortonormalização

Exemplo Considere $\boldsymbol{X} \triangleq \begin{bmatrix} \boldsymbol{x_1} & \boldsymbol{x_2} & \boldsymbol{x_3} \end{bmatrix}$, formada pelos vetores LI

$$\boldsymbol{x_1} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}', \ \boldsymbol{x_2} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}', \ \boldsymbol{x_3} = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}'$$

Algoritmo: $\boldsymbol{u_1 = x_1; \ q_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{3}}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.5774 \\ 0.5774 \\ 0.5774 \end{bmatrix}$

$$\boldsymbol{u_2 = x_2 - (q_1' x_2) q_1; \ q_2 = \frac{u_2}{\|u_2\|}} = \begin{bmatrix} -0.7071 & 0 & 0.7071 \end{bmatrix}'$$

$$\boldsymbol{u_3 = x_3 - (q_2' x_3) q_2 - (q_1' x_3) q_1; \ q_3 = \frac{u_3}{\|u_3\|}} = \begin{bmatrix} -0.4082 & 0.8165 & -0.4082 \end{bmatrix}'$$

Ortonormalização

$\rightsquigarrow$ Aplicando Cholesky na matriz $\boldsymbol{R} = \boldsymbol{X}'\boldsymbol{X}$ (MATLAB©):

$$\boldsymbol{L} = \texttt{chol}(\boldsymbol{R}) = \begin{bmatrix} \mathbf{1.7321} & \mathbf{3.4641} & \mathbf{0} \\ \mathbf{0} & \mathbf{1.4142} & \mathbf{-0.7071} \\ \mathbf{0} & \mathbf{0} & \mathbf{1.2247} \end{bmatrix} \Rightarrow \boldsymbol{L}'\boldsymbol{L} = \boldsymbol{R}$$

$$\boldsymbol{Q} = \boldsymbol{X}\boldsymbol{L}^{-1} = \underbrace{\begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{bmatrix}}_{\boldsymbol{U}}, \ \boldsymbol{Q}'\boldsymbol{Q} = \mathbf{I}$$

$\left\{ \begin{array}{c} \text{Uma outra alternativa.} \\ \text{Decomposição Triangular} \\ \text{Ortonormal – QR:} \\ \text{[Q,L]=}\texttt{qr}(\boldsymbol{X}) \end{array} \right.$

Ortonormalização

Veja que de Cholesky:

$$Q = X * \mathtt{inv}(L) = \begin{bmatrix} 0.5774 & -0.7071 & -0.4082 \\ 0.5774 & -0.0000 & 0.8165 \\ 0.5774 & 0.7071 & -0.4082 \end{bmatrix}$$

E usando QR?

$[Q1, L1] = \mathtt{qr}(X) \quad \Rightarrow$ observe o sinal...

$$Q1 = \begin{bmatrix} -0.5774 & 0.7071 & 0.4082 \\ -0.5774 & 0.0000 & -0.8165 \\ -0.5774 & -0.7071 & 0.4082 \end{bmatrix} ;\ L1 = \begin{bmatrix} -1.7321 & -3.4641 & 0 \\ 0 & -1.4142 & 0.7071 \\ 0 & 0 & -1.2247 \end{bmatrix}$$

Ortonormalização

Usando Schur: $[\boldsymbol{V},\boldsymbol{Omega}] = \mathtt{schur}(\boldsymbol{R})$

$$V = \begin{bmatrix} 0.8850 & 0.2376 & -0.4005 \\ -0.4035 & -0.0381 & -0.9142 \\ -0.2325 & 0.9706 & 0.0622 \end{bmatrix} ;\ Omega = \begin{bmatrix} 0.2643 & 0 & 0 \\ 0 & 2.0392 & 0 \\ 0 & 0 & 16.6965 \end{bmatrix}$$

$\rightsquigarrow$ $R^{-0.5} = V\Omega^{-0.5}V' \Rightarrow RR = V * \mathtt{sqrt}(\mathtt{inv}(Omega)) * V'$

$$Q_2 = XR^{-0.5} = X * RR = \begin{bmatrix} 0.9908 & -0.0890 & -0.1021 \\ 0.1348 & 0.5758 & 0.8064 \\ 0.0130 & 0.8127 & -0.5825 \end{bmatrix},\ Q_2'Q_2 = \mathsf{I}$$

Ortonormalização

Ortonormalização via Otimização (pg. 32–36): $\boldsymbol{X'X} = \begin{bmatrix} \mathbf{3} & \mathbf{6} & \mathbf{0} \\ \mathbf{6} & \mathbf{14} & \mathbf{-1} \\ \mathbf{0} & \mathbf{-1} & \mathbf{2} \end{bmatrix}$

$$\boldsymbol{F} = (\boldsymbol{X'X})^{-0.5} = \begin{bmatrix} 1.6020 & -0.6112 & -0.2447 \\ -0.6112 & 0.5222 & 0.1427 \\ -0.2447 & 0.1427 & 0.7658 \end{bmatrix}$$

e

$$\boldsymbol{Q_3} = \boldsymbol{XF} = \begin{bmatrix} 0.9908 & -0.0890 & -0.1021 \\ 0.1348 & 0.5758 & 0.8064 \\ 0.0130 & 0.8127 & -0.5825 \end{bmatrix}, \; Q_3'Q_3 = \mathsf{I}$$

Que é igual a fatorização de Schur Q_2...

Ortonormalização

Exemplo Considere dois vetores LI no plano $\mathbb{R}^2$:

$$\boldsymbol{X} = \left[\begin{array}{c|c} \boldsymbol{x_1} & \boldsymbol{x_2} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{1} & \mathbf{2} \\ \mathbf{3} & \mathbf{0.5} \end{array}\right]$$

Ortonomalização de Schmidt: $\boldsymbol{S} = \left[\begin{array}{c|c} \boldsymbol{s_1} & \boldsymbol{s_2} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{0.32} & \mathbf{0.95} \\ \mathbf{0.95} & \mathbf{-0.32} \end{array}\right]$

Fatorização de Schur: $\boldsymbol{U} = \left[\begin{array}{c|c} \boldsymbol{u_1} & \boldsymbol{u_2} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{0.0995} & \mathbf{0.9950} \\ \mathbf{0.9950} & \mathbf{-0.0995} \end{array}\right]$

Ortonormalização

Usando a proposta de Ortonormalização via Otimização (pg. 32–36), faça

$$\boldsymbol{X'X} = \begin{bmatrix} 1 & 3 \\ 2 & 0.5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 0.5 \end{bmatrix} = \begin{bmatrix} 10 & 3.5 \\ 3.5 & 4.25 \end{bmatrix}$$

$$\boldsymbol{F = (X'X)^{-0.5}} = \begin{bmatrix} 0.3528 & -0.1266 \\ -0.1266 & 0.5608 \end{bmatrix}$$

e

$$\boldsymbol{Q = XF} = \left[\begin{array}{c|c} q_1 & q_2 \end{array}\right] = \left[\begin{array}{c|c} 0.0995 & 0.9950 \\ 0.9950 & -0.0995 \end{array}\right] \quad \text{(E vale a relação:} \boldsymbol{Q'Q} = \mathbf{I}\text{)}$$

Que é igual a fatorização de Schur $\boldsymbol{U}$...

MATLAB

- `trace(A)` – traço de uma matriz quadrada $\boldsymbol{A}$
- `rank(A)` – posto de uma matriz $\boldsymbol{A}$
- `chol(A)` – fatoração de Cholesky de uma matriz $\boldsymbol{A}$
- `qr(A)` – fatoração QR de uma matriz $\boldsymbol{A}$
- `null(A)` – retorna uma base ortonormal para o espaço nulo de $\boldsymbol{A}$
- `orth(A)` – retorna uma base ortonormal para o range de $\boldsymbol{A}$