Notas prévias e convenções utilizadas

- Os slides de apoio às aulas teóricas baseiam-se na matéria da sebenta Texto de Apoio de Álgebra Linear, e vários dos seus esquemas e/ou figuras provêm da sebenta ou são versões modificadas de esquemas e figuras da sebenta.
- A matéria exposta nestes slides deve ser complementada com a leitura dessa sebenta.
- Vamos usualmente escrever a vermelho as definições, a azul o texto a destacar e a magenta os exercícios e desafios para os alunos.

Um pequeno exemplo para motivar :)

Um famoso conjunto de dados foi obtido por Ronald Fisher medindo o comprimento e a largura das sépalas e pétalas (em cm) de 50 lírios de cada uma das 3 espécies distintas, *Setosa*, *Versicolor* e *Virgínica*.

Autor Diego Mariano https://pt.wikipedia.org/wiki/Conjunto_de_dados_flor_Iris#/media/Ficheiro:Flores_de_Íris.png

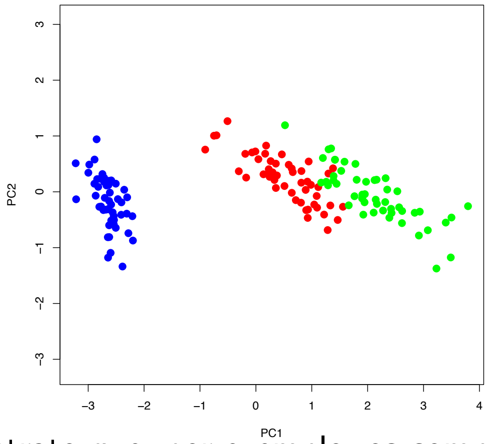
A tabela a seguir apresenta os valores obtidos para alguns dos lírios.

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
⋮	⋮	⋮	⋮	⋮	⋮
150	5.9	3.0	5.1	1.8	virginica

O melhor retrato do conjunto de dados dos lírios

O conjunto de dados dos 150 lírios origina uma núvem de 150 pontos num espaço a 4 dimensões que não conseguimos visualizar, em que o vetor de coordenadas de cada ponto contém o comprimento e a largura das sépalas e pétalas de cada lírio (vetor com 4 componentes).

Usando métodos de Álgebra Linear podemos projetar esta núvem de pontos num plano de modo a obter-se o melhor retrato possível (num certo sentido):



Pode-se observar no retrato que, por exemplo, os comprimentos e as larguras das sépalas e pétalas diferenciam claramente os lírios da espécie *Setosa* dos lírios das restantes 2 espécies *Versicolor* e *Virgínica*. . .

O conjunto $\mathbb{R}^n$

- Recordemos que $\mathbb{R}$ denota o conjunto dos números reais.
- O conjunto dos vetores do plano é o conjunto dos vetores com 2 componentes reais que se denota por $\mathbb{R}^2$, ou seja,

$$\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$$

  Por exemplo, $(1, -\pi) \in \mathbb{R}^2$
- Analogamente, o conjunto dos vetores do espaço é o conjunto dos vetores com 3 componentes reais, denotado $\mathbb{R}^3$, isto é,

$$\mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}.$$

  Por exemplo, $(1, -\pi, 0) \in \mathbb{R}^3$
- Vamos trabalhar com vetores com um número arbitrário de componentes reais: dado um inteiro $n \geq 2$, denotamos o conjunto dos vetores com n componentes reais por $\mathbb{R}^n$, ou seja,

$$\mathbb{R}^n = \Big\{(x_1, x_2, \ldots, x_n) : x_1, x_2, \ldots, x_n \in \mathbb{R}\Big\}$$

  x_i: componente do vetor x que se encontra na posição i

  Por exemplo, se $x = (1, -\pi, 0, 2, 3, -4) \in \mathbb{R}^6$, $x_4 = 2$

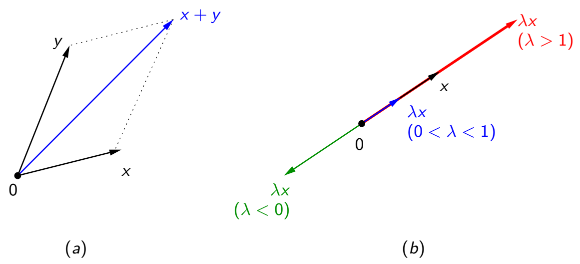
Operações sobre vetores do plano

Recordemos as operações algébricas bem conhecidas sobre vetores do plano ($\mathbb{R}^2$). Se $x = (x_1, x_2)$ e $y = (y_1, y_2)$ são vetores de $\mathbb{R}^2$ e $\lambda \in \mathbb{R}$:

- Adição de vetores:
$$x + y = (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2).$$
- Produto de um vetor por um escalar:
$$\lambda x = \lambda(x_1, x_2) = (\lambda x_1, \lambda x_2).$$
- Produto escalar (ou interno) de vetores:
$$x \cdot y = (x_1, x_2) \cdot (y_1, y_2) = x_1 y_1 + x_2 y_2.$$

Por exemplo, se $x = (3, 1)$, $y = (2, 5)$ e $\lambda = 2$, obtém-se
$$x + y = (5, 6), \quad 2(3, 2) = (6, 4), \quad (3, 1) \cdot (2, 5) = 11.$$

Interpretação geométrica das operações sobre vetores



Recordemos que o produto escalar está relacionado com o cosseno do ângulo θ formado pelo 2 vetores pela relação bem conhecida,

$$x \cdot y = \cos(\theta) \, \|x\| \, \|y\|,$$

onde $\|x\|$ e $\|y\|$ representam os comprimentos do vetores x e y.

A extensão das operações algébricas anteriores para vetores com um número arbitrário de componentes faz-se de modo óbvio.

Operações sobre vetores de $\mathbb{R}^n$

Definição

- Adição de vetores:

$$x + y = (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),$$

isto é, somam-se as componentes homólogas dos vetores

- Produto de um vetor por um escalar:

$$\lambda x = \lambda(x_1, \ldots, \lambda x_n) = (\lambda x_1, \ldots, \lambda x_n),$$

isto é, multiplicam-se todas as componentes do vetor pelo escalar

- Produto escalar (ou interno) de vetores:

$$x \cdot y = (x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Dar exemplos em $\mathbb{R}^4$ para as 3 operações anteriores.

Propriedades das operações sobre vetores

Adição de vetores e o produto de vetores por escalares verificam várias propriedades que decorrem imediatamente das propriedades dos números reais (falaremos mais adiante nas propriedades do produto escalar).

Propriedades das operações algébricas

Sejam x, y, z vetores de $\mathbb{R}^n$, $\vec{0} = (0, \cdots, 0) \in \mathbb{R}^n$ e $\lambda, \mu \in \mathbb{R}$. Tem-se,

1. $x + y = y + x$ (comutativa)
2. $(x + y) + z = x + (y + z)$ (associativa)
3. $x + \vec{0} = x$ (existência de el. neutro)
4. $x + (-x) = \vec{0}$ (existência de el. simétrico)
5. $\lambda(x + y) = \lambda x + \lambda y$ (distributiva...)
6. $(\lambda + \mu)x = \lambda x + \mu x$ (distributiva...)
7. $(\lambda\mu)x = \lambda(\mu x)$ (compatibilidade dos produtos...)
8. $1\,x = x$ (el. identidade da multiplicação por escalar)

Conceito de matriz

Os números reais serão também designados por *escalares* por oposição a vetores

Definição de matriz

Sejam m, n inteiros positivos. Chama-se matriz do tipo $m \times n$ *a uma coleção $A = [a_{ij}]$ de mn números reais dispostos em m linhas e n colunas,*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}_{m\times n}.$$

a_{ij}: *elemento da matriz que se encontra na linha i e coluna j da matriz*

O índice i percorre as linhas da matriz e designa-se por índice de linha *O índice j percorre as colunas da matriz e designa-se por* índice de coluna

As matrizes constituem uma extensão dos vetores adequada ao estudo dos sistemas lineares

Exemplos

- $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}_{4\times 3} = \begin{bmatrix} 10 & 2 & 3 \\ 5 & 0 & 0 \\ 1 & 3 & -1 \\ 5 & 1 & -2 \end{bmatrix}_{4\times 3}$

  O elemento de A que se encontra na linha 4 e coluna 1 é $a_{41} = 5$

- $A = [a_{ij}]_{2\times 3}$ definida por $a_{1j} = 10$ e $a_{2j} = \pi$, para todo o j, é

$$A = \begin{bmatrix} 10 & 10 & 10 \\ \pi & \pi & \pi \end{bmatrix}_{2\times 3}$$

- $A = [a_{ij}]_{3\times 3}$ definida por $a_{ij} = i + j$, para $i, j = 1, 2, 3$, é

$$A = \begin{bmatrix} 1+1 & 1+2 & 1+3 \\ 2+1 & 2+2 & 2+3 \\ 3+1 & 3+2 & 3+3 \end{bmatrix}_{3\times 3} = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix}_{3\times 3}$$

Matriz-linha e matriz-coluna ou vetor

- Se $m = 1$, $A_{1\times n} = \begin{bmatrix} a_{11} & \dots & a_{1n} \end{bmatrix}$ designa-se por *matriz-linha*.
  Por exemplo, $A = \begin{bmatrix} 1 & 3 & -2 \end{bmatrix}$ matriz-linha do tipo 1×3.
- Se $n = 1$, $A_{m\times 1} = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}$ designa-se por *matriz-coluna* ou *vetor*.
  Por exemplo, $(2, 3, -1) = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} \in \mathbb{R}^3$
  - Em geral, $x \in \mathbb{R}^m$ pode ser representado como m-uplo de números reais ou como matriz-coluna do tipo $m \times 1$:

$$x = (x_1, \cdots, x_m) = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}.$$

Matriz definida por vetores e matriz quadrada

- Se $v_1, v_2, \dots, v_n \in \mathbb{R}^m$, $A = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$ denota a matriz do tipo $m \times n$ cujas colunas são os n vetores $v_1, v_2, \dots, v_n$.
  Por exemplo, se $v_1 = (1, 0, 2)$, $v_2 = (-1, 1, 1)$ e $v_2 = (1, 10, 0)$,

$$A = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 10 \\ 2 & 1 & 0 \end{bmatrix}_{3\times 3}.$$

- Se uma matriz A é do tipo $n \times n$, A diz-se *quadrada de ordem n*.
  Por exemplo, a matriz $\begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$ anterior é quadrada de ordem 3.
  - Chama-se *diagonal principal* de uma matriz quadrada $A = [a_{ij}]$ ao conjunto dos elementos a_{ii}, $i = 1, \dots n$:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Matriz triangular e matriz diagonal

$A = [a_{ij}]$ matriz quadrada de ordem n

- A diz-se *triangular superior* se $a_{ij} = 0$ para $i > j$, ou seja, se todos os elementos abaixo da diagonal principal são nulos.
  Por exemplo, $\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ é triangular superior de ordem 3.
- A definição de *triangular inferior* é análoga e fica como exercício.
- A diz-se *diagonal* se $a_{ij} = 0$ para todo $i \neq j$, isto é, se todos os elementos fora da diagonal principal de A forem nulos, e pode ser representada por $A = \mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn})$.
  Por exemplo, $\mathrm{diag}(2, -1, 3) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ é uma matriz diagonal de ordem 3.

Matriz escalar e matriz identidade

- Uma matriz diagonal A de ordem n diz-se *escalar* se todas as entradas da diagonal principal forem iguais entre si, isto é, se para algum $\lambda \in \mathbb{R}$,

$$A = \mathrm{diag}(\lambda, \lambda, \ldots, \lambda) = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix}_{n\times n}$$

- Se $\lambda = 1$, A designa-se por *matriz identidade de ordem n* e denota-se por I_n (ou simplesmente por I). A matriz identidade representa o elemento neutro da multiplicação de matrizes como veremos mais adiante

  Por exemplo, a matriz identidade de ordem 3 é a matriz

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Igualdade entre matrizes e matriz transposta

- $A = [a_{ij}]$ e $B = [b_{ij}]$ do mesmo tipo dizem-se *iguais* se os elementos homólogos forem iguais, isto é, se $a_{ij} = b_{ij}$, $\forall i, j$

  Por exemplo, $\begin{bmatrix} 5 & x \\ y & 6 \end{bmatrix} = \begin{bmatrix} z & 3 \\ 2 & w \end{bmatrix} \Leftrightarrow \begin{cases} x = 3 \\ y = 2 \\ z = 5 \\ w = 6 \end{cases}$

- A *transposta* de $A = \left[a_{ij}\right]$ do tipo $m \times n$ é a matriz $A^T = [a_{ji}]_{n\times m}$ do tipo $n \times m$, cujas colunas são as linhas de A pela mesma ordem.

  Por exemplo, se $A = \begin{bmatrix} 2 & 3 \\ 4 & 1 \\ 5 & 6 \end{bmatrix}$, então $A^T = \begin{bmatrix} 2 & 4 & 5 \\ 3 & 1 & 6 \end{bmatrix}$.

  Tem-se, obviamente, $(A^T)^T = A$.

- $A = [a_{ij}]$ quadrada diz-se *simétrica*, se $A^T = A$, isto é, $a_{ij} = a_{ji}$, $\forall i, j$.

  Por exemplo, $A = \begin{bmatrix} 2 & -2 & 1 \\ -2 & 1 & 5 \\ 1 & 5 & 10 \end{bmatrix}$ é simétrica.

Operações algébricas sobre matrizes: adição de matrizes

As operações algébricas sobre matrizes estendem as operações da adição de vetores, do produto de um vetor por um escalar e do produto escalar de vetores definidas anteriormente.

Se $A = \left[a_{ij}\right]_{m\times n}$ e $B = \left[b_{ij}\right]_{m\times n}$ são matrizes <u>do mesmo tipo</u> define-se a *soma de A com B*, por

$$A + B = \left[a_{ij} + b_{ij}\right]_{m\times n}.$$

Por outras palavras, os elementos de $A + B$ obtêm-se somando os elementos homólogos de A e de B.

Por exemplo, se $A = \begin{bmatrix} 2 & -1 & 0 \\ 4 & 5 & 3 \end{bmatrix}$ e $B = \begin{bmatrix} 0 & 2 & 1 \\ -3 & 1 & 4 \end{bmatrix}$, tem-se

$$A + B = \begin{bmatrix} 2+0 & -1+2 & 0+1 \\ 4-3 & 5+1 & 3+4 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 6 & 7 \end{bmatrix}.$$

Produto de uma matriz por um escalar

Se $\lambda \in \mathbb{R}$ e $A = [a_{ij}]_{m\times n}$ define-se o *produto de A pelo escalar* λ, por

$$\lambda A = [\lambda a_{ij}]_{m\times n}$$

Por outras palavras, λA obtém-se multiplicando cada elemento de A por λ

Se $\lambda = -1$, λA denota-se simplesmente por $-A$

Por exemplo, se $\lambda = 3$ e $A = \begin{bmatrix} 20 & -1 & 13 \\ 18 & -2 & 81 \end{bmatrix}$, então

$$\lambda A = 3\begin{bmatrix} 20 & -1 & 13 \\ 18 & -2 & 81 \end{bmatrix} = \begin{bmatrix} 3\cdot 20 & 3\cdot(-1) & 3\cdot 13 \\ 3\cdot 18 & 3\cdot(-2) & 3\cdot 81 \end{bmatrix} = \begin{bmatrix} 60 & -3 & 39 \\ 54 & -6 & 243 \end{bmatrix}$$

Propriedades da adição de matrizes e do produto de escalares por matrizes

Sejam A, B e C matrizes do tipo $m \times n$ e $\lambda, \mu \in \mathbb{R}$. Tem-se,

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + [0]_{m\times n} = A$ ($[0]_{m\times n}$ matriz cujos elementos são todos nulos)
4. $A + (-A) = [0]_{m\times n}$
5. $\lambda(A + B) = \lambda A + \lambda B$
6. $(\lambda + \mu)A = \lambda A + \mu A$
7. $\lambda(\mu A) = (\lambda\mu)A$
8. $1.A = A$
9. $(A + B)^T = A^T + B^T$
10. $(\lambda A)^T = \lambda\, A^T$
11. $(A^T)^T = A$

Propriedades das operações algébricas sobre matrizes

A matriz nula $[0]_{m\times n}$ é portanto o elemento neutro da adição de matrizes.

As propriedades (1)-(8) decorrem das propriedades da adição e do produto de números reais e são análogas às propriedades da adição e do produto por escalar para vetores.

As restantes três propriedades são evidentes.

Exercício na aula

- Sejam $A = \begin{bmatrix} 2 & 1 \\ 4 & -4 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 2 \\ 3 & 0 \end{bmatrix}$ e I_2 a matriz identidade de ordem 2. Simplifique expressão $((A^T + B)^T + 4I_2)^T$ indicando as propriedades do slide anterior que utilizar e calcule o seu valor.

TPC

Mostre que se A é uma matriz quadrada então $A + A^T$ é simétrica.

Produto de matrizes

- Duas matrizes A e B dizem-se *encadeadas*, se

número de colunas de A = número de linhas de B.

Por exemplo, $A = \begin{bmatrix} 2 & 3 \\ -1 & 4 \\ 0 & 5 \end{bmatrix}_{3\times 2}$ e $B = \begin{bmatrix} 1 & -1 & 0 & 5 \\ 2 & 1 & -1 & 3 \end{bmatrix}_{2\times 4}$ são encadeadas pois o número de colunas de A é igual ao número de linhas de B. Mas B e A não são encadeadas !

Definição do produto de matrizes

Se $A = [a_{ij}]_{m\times n}$ e $B = [b_{jk}]_{n\times p}$ são encadeadas, define-se o *produto* de A por B, denotado AB, como sendo a matriz $C = [c_{ik}]_{m\times p}$ tal que

$$\begin{aligned} c_{ik} &= (\text{linha } i \text{ de } A)\cdot(\text{coluna } k \text{ de } B) \\ &= (a_{i1}, a_{i2}, \ldots, a_{in})\cdot(b_{1k}, b_{2k}, \ldots, b_{nk}) \\ &= a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk}. \end{aligned}$$

Produto de matrizes

Por exemplo, $A = \begin{bmatrix} 2 & 4 & -1 \\ 4 & 5 & 0 \\ 1 & 8 & 5 \end{bmatrix}_{3\times 3}$ e $B = \begin{bmatrix} 1 & -1 \\ 2 & 4 \\ 0 & 5 \end{bmatrix}_{3\times 2}$ são encadeadas, tendo-se

$$\begin{aligned} AB &= \left[\begin{array}{ccc} 2 & 4 & -1 \\ \hline 4 & 5 & 0 \\ \hline 1 & 8 & 5 \end{array}\right] \left[\begin{array}{c|c} 1 & -1 \\ 2 & 4 \\ 0 & 5 \end{array}\right] \\ &= \begin{bmatrix} 2+8+0 & -2+16-5 \\ 4+10+0 & -4+20+0 \\ 1+16+0 & -1+32+25 \end{bmatrix} \\ &= \begin{bmatrix} 10 & 9 \\ 14 & 16 \\ 17 & 56 \end{bmatrix}_{3\times 2} \end{aligned}$$

Por exemplo, o elemento de AB que se encontra na linha 3 e coluna 1 é o produto escalar da terceira linha de A pela primeira coluna de B, isto é, $(1, 8, 5) \cdot (1, 2, 0) = 17$.

Produto escalar via produto de matrizes...

- O produto de matrizes estende o conceito de produto escalar de vetores: se $x, y \in \mathbb{R}^n$ então

$$\boxed{x^T y = x \cdot y}$$

Por exemplo, se $x = (-1, 1, 3) = \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}$ e $y = (1, 0, 1) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$,

$$x^T y = [-1 \ \ 1 \ \ 3]_{1\times 3} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}_{3\times 1} = [(-1,1,3)\cdot(1,0,1)]_{1\times 1} = [2]_{1\times 1} = 2(^1) = x\cdot y$$

- Note-se que $xy^T = \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}_{3\times 1} [1 \ \ 0 \ \ 1]_{1\times 3} \begin{bmatrix} -1 & 0 & -1 \\ 1 & 0 & 1 \\ 3 & 0 & 3 \end{bmatrix}_{3\times 1}$

[1] As matrizes 1×1 identificam-se com a seu único elemento.

Potência de uma matriz quadrada

Potência inteira não negativa

Dada uma matriz quadrada A de ordem n definem-se as *potências inteiras não negativas de A* por,

$$A^0 = I_n \qquad \text{e} \qquad A^k = \underbrace{A \cdots A}_{\text{k vezes}} \quad (k \in \mathbb{N}).$$

TPC

Calcular A^3 com $A = \begin{bmatrix} -1 & 1 \\ 3 & 2 \end{bmatrix}_{2\times 2}$.

Propriedades

Propriedades do produto de matrizes

Sejam A, B, C matrizes, I a matriz identidade de ordem conveniente, $[0]$ a matriz nula de tipo conveniente e $\lambda \in \mathbb{R}$ e k um inteiro não negativo. Sempre que as operações estejam definidas, tem-se:

1. $(AB)C = A(BC)$ (associativa)
2. $A(B + C) = AB + AC$ (distributiva)
3. $(A + B)C = AC + BC$ (distributiva)
4. $AI = IA = A$ (el. neutro da mult.)
5. $A[0] = [0]A = 0$ (el. absorvente da mult.)
6. $\lambda(AB) = (\lambda A)B = A(\lambda B)$ (compatibilidade dos produtos)
7. $(AB)^T = B^T A^T$ (!)
8. $(A^k)^T = (A^T)^k$

“Não propriedades” do produto de matrizes

Ao contrário do que sucede com a adição, algumas propriedades do produto de números reais **não se generalizam** para o produto de matrizes.

Exercício na aula

Calcular os produtos AB e BA com

$$A = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}_{2\times 2}, \qquad B = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}_{2\times 2}$$

O que observa ?

“Não propriedades” do produto de matrizes

- O produto de matrizes não é comutativo, ou seja, em geral,
$$AB \neq BA.$$
- A lei do anulamento do produto também não é válida, ou seja, em geral,
$$AB = [\,0\,] \quad \not\Rightarrow \quad (A = [\,0\,] \text{ ou } B = [\,0\,]).$$
- A lei do corte também não é válida ou seja, em geral, dadas matrizes A, B e C, com $A \neq [\,0\,]$([2]),
$$AB = AC \quad \not\Rightarrow \quad B = C.$$

([0] denota uma matriz nula de ordem conveniente)

TPC

Dar exemplos de 3 matrizes quadradas de ordem 2, A, B e C, para as quais a lei do corte falhe

[2] Para a lei do corte ser válida para matrizes devemos substituir a condição $A \neq [\,0\,]$ por A invertível, conceito que daremos mais adiante.

Uma consequência inesperada...

Se A e B são matrizes quadradas da mesma ordem não permutáveis, isto é, $AB \neq BA$, obtém-se aplicando as propriedades distributivas do produto de matrizes,

- $(A+B)(A-B) = A^2 - AB + BA - B^2 \neq A^2 - B^2$.
- $(A+B)^2 = (A+B)(A+B) = A^2 + AB + BA + B^2 \neq A^2 + 2AB + B^2$.
- $(A-B)^2 = (A-B)(A-B) = A^2 - AB - BA + B^2 \neq A^2 - 2AB + B^2$.

A não comutatividade do produto de matrizes teve como consequência que não são válidos para o produto de matrizes quadradas os análogos dos casos notáveis da multiplicação de números reais!

Observação

Deve-se ter uma particular atenção ao simplificar expressões que envolvam produtos de matrizes!

Simplificação de expressões...

Exercício na aula

Considere

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2\times 2}, \qquad B = \begin{bmatrix} 0 \\ -1 \end{bmatrix}_{2\times 1}, \qquad C = \begin{bmatrix} 2 & 1 \end{bmatrix}_{1\times 2}.$$

Desenvolva e calcule $((BC)^T + A)^2$.

Ainda o produto de matrizes...

Observação

Se $A_{m\times n}$ e $B_{n\times p} = \begin{bmatrix} v_1 & v_2 & \cdots & v_p \end{bmatrix}$ então

$$AB = A\begin{bmatrix} v_1 & v_2 & \cdots & v_p \end{bmatrix} = \begin{bmatrix} Av_1 & Av_2 & \cdots & Av_p \end{bmatrix},$$

ou seja, a k-ésima coluna do produto AB é o produto de A pela k-ésima coluna de B.

Por exemplo, se $A = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 3 & 2 \end{bmatrix} = [v_1 \ \ v_2]$ com $v_1 = (-1, 0, 3)$ e $v_2 = (1, 1, 2)$, tem-se

$$AB = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 5 \\ -6 & 2 \end{bmatrix} = \begin{bmatrix} Av_1 & Av_2 \end{bmatrix}$$

De facto,

$$Av_1 = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ -6 \end{bmatrix} \quad \text{e} \quad Av_2 = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

Sistema de equações lineares

Sistema linear

Um sistema linear a m equações e n variáveis $x_1, \ldots, x_n$ é um sistema de equações da forma,

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ & & & & \vdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{cases},$$

- $a_{ij} \in \mathbb{R}$: *coeficiente da variável x_j na i-ésima equação*.
- $b_i \in \mathbb{R}$: *termo constante* ou *membro direito* da i-ésima equação.

- *Solução* de um sistema linear é uma solução comum a todas as equações desse sistema.

Exemplo de um sistema linear a 3 equações e 3 variáveis

$$\begin{cases} 2\,x_1 + \,x_2 + 2\,x_3 = 3 \\ 4\,x_1 + 6\,x_2 + 6\,x_3 = 4 \\ -2\,x_1 - \,x_3 = -3 \end{cases}$$

Com a notação do slide anterior tem-se, por exemplo,

- $a_{11} = 2$: coeficiente da variável x_1 na primeira equação
- $a_{23} = 6$: coeficiente da variável x_3 na segunda equação
- $b_2 = 4$: termo constante ou membro direito da segunda equação
- $b_3 = -3$: termo constante ou membro direito da terceira equação

Conjunto de soluções e classificação de um sistema linear

Resolver um sistema linear é determinar o seu conjunto de soluções (CS).
Um sistema linear é *classificado* como:

- **impossível** (IMP) se não possuir soluções
- **possível** se possuir pelo menos uma solução, sendo:
  - **determinado** (PD), se possuir uma única solução
  - **indeterminado** (PI), se possuir uma infinidade de soluções

Por exemplo, o sistema linear a 2 equações e 2 variáveis,

$$\begin{cases} 2x_1 - x_2 & = & 3 \\ x_1 + 2x_2 & = & 4 \end{cases}$$

é PD com $CS = \{(2, 1)\}$ (verifique!).

TPC

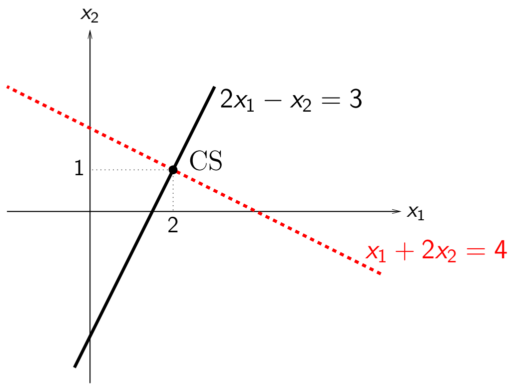
Adaptando o sistema linear anterior dê exemplos de sistemas lineares a 2 equações e 2 variáveis que sejam PI e IMP, indicando em cada caso o respectivo CS.

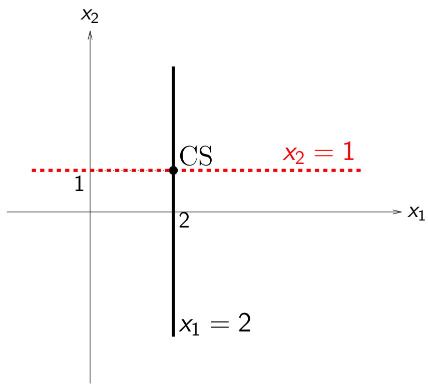
Sistemas equivalentes

- Dois sistemas lineares a m equações e n variáveis dizem-se *equivalentes* se possuem o mesmo conjunto de soluções (CS).

São equivalentes os seguintes sistemas a 2 equações e 2 variáveis:

$$\begin{cases} 2x_1 - x_2 &= 3 \\ x_1 + 2x_2 &= 4 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x_1 &= 2 \\ x_2 &= 1 \end{cases} \text{(sistema reduzido)}$$





- As equações de quaisquer duas retas concorrentes no ponto $(2, 1)$ definem um sistema linear equivalente aos sistemas anteriores.

Matriz ampliada de um sistema a m equações e n variáveis

Consideremos o sistema linear a m equações e n variáveis,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

- $A_{m\times n} = [a_{ij}]$ chama-se *matriz dos coeficientes* do sistema linear,
- $b = (b_1, \ldots, b_m)$ chama-se o *vetor dos termos constantes* ou *membros direitos* do sistema,
- $x = (x_1, \ldots, x_n)$ chama-se *vetor das incógnitas* ou *variáveis* do sistema e finalmente,
- $[A|b] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right]$, chama-se *matriz ampliada do sistema* e contém toda a sua informação relevante

Matriz em escada e matriz reduzida

- Uma matriz diz-se em *escada* se o primeiro elemento não nulo de cada linha, que se designa por *pivot*, estiver à direita do primeiro elemento não nulo da linha anterior e todas as linhas nulas, caso existam, aparecerem no fim
- Uma matriz diz-se *reduzida* se
  - estiver em escada,
  - todos os pivots forem iguais a 1,
  - em cada coluna com pivot apenas o pivot é não nulo.

Exemplos de matrizes em escada e reduzida com os pivots a vermelho,

$$\begin{bmatrix} 0 & 1 & 2 & 1 & 4 \\ 0 & 0 & 3 & -6 & 0 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -6 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Um sistema linear diz-se em *escada/reduzido* se a respectiva matriz dos coeficientes estiver em escada/reduzida.

Operações elementares sobre as linhas de uma matriz

(I) "Apagador" - Adicionar a uma linha i uma linha $j \neq i$ multiplicada por um escalar λ ($L_i + \lambda L_j$).

(II) Multiplicar uma linha i por um escalar $\lambda \neq 0$ (λL_i).

(III) Permutar uma linha i com uma linha j ($L_i \leftrightarrow L_j$).

A notação entre parênteses é uma simplificação da notação usada no Texto de Apoio!

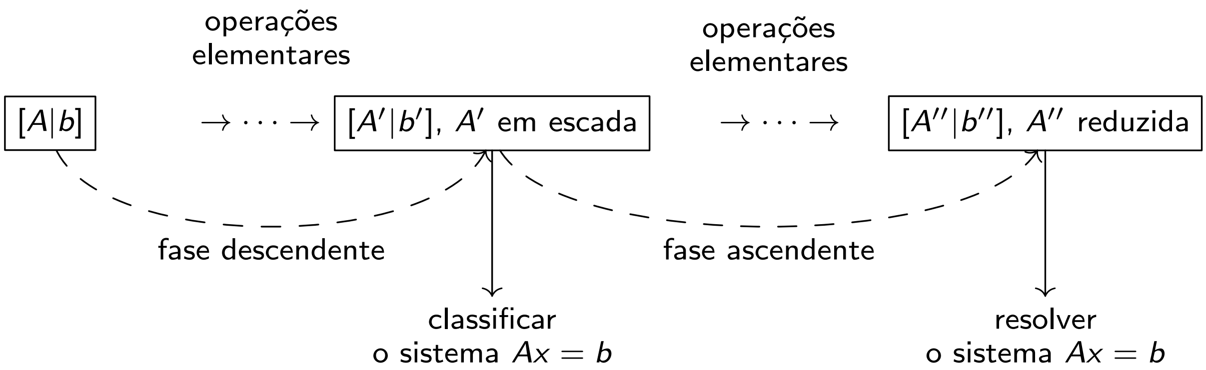
Teorema

As operações elementares (I), (II) e (III) transformam a matriz ampliada de um sistema linear na matriz ampliada de um sistema linear equivalente, ou seja, com o mesmo CS.

Definem-se de modo análogo operações elementares sobre as equações de um sistema linear.

Método de eliminação de Gauss para redução de sistemas

O método de eliminação de Gauss desenvolve-se em duas fases (descendente e ascendente), aplicando operações elementares sobre as linhas da matriz ampliada $[A|b]$ de um sistema linear, de acordo com o seguinte esquema:



Exemplificação do método de Gauss

Exemplo na aula

Vejamos como se processa o método de eliminação de Gauss no sistema

$$\begin{cases} x_1 - x_2 + 2x_3 = -1 \\ 3x_1 + x_2 + x_3 = 6 \\ 2x_1 + 6x_2 = 10 \end{cases}$$

que representa a intersecção de 3 planos em $\mathbb{R}^3$.

TPC

Resolver os seguintes sistemas lineares:

$$(a) \begin{cases} 2x_1 + x_2 + 2x_3 = 3 \\ 4x_1 + 6x_2 + 6x_3 = 4 \\ -2x_1 - x_3 = -3 \end{cases} \qquad (b) \left[\begin{array}{ccc|c} 1 & 3 & 3 & 0 \\ 3 & 4 & 1 & -2 \\ 1 & -2 & -5 & 0 \end{array}\right]$$

Redução do sistema: fase descendente

Aplicando a fase descendente do método de eliminação de Gauss à matriz ampliada $[A|b]$ do sistema do exemplo do slide anterior obtém-se:

$$[A|b] \quad = \quad \left[\begin{array}{rrr|r} 1 & -1 & 2 & -1 \\ 3 & 1 & 1 & 6 \\ 2 & 6 & 0 & 10 \end{array}\right] \xrightarrow[L_3-2L_1]{L_2-3L_1} \left[\begin{array}{rrr|r} 1 & -1 & 2 & -1 \\ 0 & 4 & -5 & 9 \\ 0 & 8 & -4 & 12 \end{array}\right]$$

$$\xrightarrow{L_3-2L_2} \left[\begin{array}{rrr|r} 1 & -1 & 2 & -1 \\ 0 & 4 & -5 & 9 \\ 0 & 0 & 6 & -6 \end{array}\right] = [A'|b'].$$

Matriz dos coeficientes A' em **escada** $\Rightarrow$ podemos classificar o sistema:

- Todas as linhas de $[A'|b']$ correspondem a equações possíveis, isto é, não são do tipo $0\ 0\ 0 \,|\, *$ com $* \neq 0$(3). Logo o sistema é possível.
- Todas as colunas de A' têm *pivot* e portanto não há variáveis livres(4). Logo o sistema é determinado e o CS é um ponto em $\mathbb{R}^3$.

[3] Que correspondem à equação impossível $0x_1 + 0x_2 + 0x_3 = *$.

[4] Cada variável fica determinada por uma equação.

Conclusão da redução do sistema: fase ascendente

$$[A'|b'] \quad = \quad \left[\begin{array}{rrr|r} 1 & -1 & 2 & -1 \\ 0 & 4 & -5 & 9 \\ 0 & 0 & 6 & -6 \end{array}\right] \xrightarrow{\frac{1}{6}L_3} \left[\begin{array}{rrr|r} 1 & -1 & 2 & -1 \\ 0 & 4 & -5 & 9 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

$$\xrightarrow[L_2+5L_3]{L_1-2L_3} \left[\begin{array}{rrr|r} 1 & -1 & 0 & 1 \\ 0 & 4 & 0 & 4 \\ 0 & 0 & 1 & -1 \end{array}\right] \xrightarrow{\frac{1}{4}L_2} \left[\begin{array}{rrr|r} 1 & -1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

$$\xrightarrow{L_1+L_2} \left[\begin{array}{rrr|r} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right] = [A''|b''].$$

Matriz dos coeficientes A'' está **reduzida** e $[A''|b'']$ corresponde à matriz ampliada do sistema reduzido,

$$\begin{cases} x_1 & & & = & 2 \\ & x_2 & & = & 1 \\ & & x_3 & = & -1 \end{cases}$$

Logo, $CS = \{(2, 1, -1)\}$.

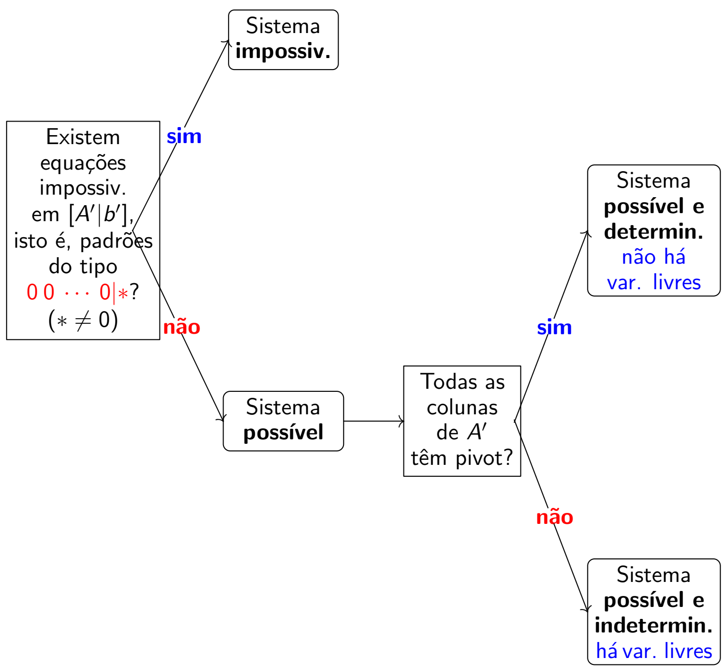
Algoritmo de eliminação de Gauss: fase descendente

- **Input:** Matriz ampliada $[A|b]$ de um sistema linear
- **Objectivo:** Redução do sistema linear
- **Fase descendente:**
  - Aplicando operações elementares do tipo III trocar, se necessário, linhas em $[A|b]$ de modo a que o pivot da primeira linha se encontre na coluna não nula mais à esquerda da matriz dos coeficientes
  - Usando operações elementares do tipo I ("Apagador") e o pivot da primeira linha, eliminar os restantes elementos da coluna abaixo desse pivot
  - Repetir os procedimentos anteriores relativamente à submatriz que se obtém ignorando a primeira linha e assim sucessivamente enquanto existirem linhas não nulas na matriz dos coeficientes dessa submatriz

No final da fase descendente obtém-se uma matriz $[A'|b']$ com A' em escada e podemos classificar o sistema.

A matriz $[A'|b']$ **não é única**, i.e, depende das operações efetuadas.

Classificação do sistema em escada



Variáveis *pivot* e variáveis livres

Observação

- As variáveis associadas às colunas sem *pivot* na matriz em escada designam-se por variáveis livres e podem tomar qualquer valor em $\mathbb{R}$.
- As variáveis associadas às colunas com *pivot* na matriz em escada designam-se por variáveis *pivot* ou variáveis determinadas e são escritas em função das variáveis livres.

Algoritmo de eliminação de Gauss: fase ascendente

- **Fase ascendente:** (apenas se aplica aos sistemas possíveis)
  - Usando operações elementares do tipo II e I tornar o pivot que se encontra mais à direita na matriz A' igual a 1 e usar esse pivot para eliminar os elementos da coluna acima desse pivot
  - Repetir os procedimentos do passo anterior relativamente à coluna com pivot imediatamente anterior e assim sucessivamente enquanto existirem colunas com pivot (percorrendo a matriz da direita para a esquerda)

No final da fase ascendente obtém-se uma matriz $[A''|b'']$ com A'' reduzida, donde resulta imediatamente o CS do sistema, escrevendo as variáveis pivot em função das variáveis livres. Observemos que:

- A matriz $[A''|b'']$ é **única**, isto é, não depende da sequência de operações elementares efectuada
- Dois sistemas com m equações lineares e n variáveis são **equivalentes** se e só se aplicando o método de Gauss às respetivas matrizes ampliadas obtemos **a mesma matriz reduzida**.

Método de eliminação de Gauss

Exercícios na aula

Aplicando o método de Gauss reduza os seguintes sistemas lineares:

$$(a) \quad \begin{cases} 2x_1 + x_2 + 2x_3 + x_4 = 3 \\ 4x_1 + 6x_2 + 6x_3 = 4 \\ -2x_1 - x_3 - x_4 = -3 \end{cases}$$

$$(b) \quad \left[\begin{array}{rrr|r} 1 & 3 & 3 & 0 \\ 3 & 4 & 1 & -2 \\ 1 & -2 & -5 & 0 \end{array}\right]$$

Redução do sistema do exemplo (a): fase descendente

Aplicando a fase descendente do método de eliminação de Gauss à matriz ampliada $[A|b]$ do 1º sistema do slide anterior obtém-se:

$$[A|b] = \left[\begin{array}{rrrr|r} 2 & 1 & 2 & 1 & 3 \\ 4 & 6 & 6 & 0 & 4 \\ -2 & 0 & -1 & -1 & -3 \end{array}\right] \xrightarrow[L_3 + L_1]{L_2 - 2L_1} \left[\begin{array}{rrrr|r} 2 & 1 & 2 & 1 & 3 \\ 0 & 4 & 2 & -2 & -2 \\ 0 & 1 & 1 & -0 & 0 \end{array}\right]$$

$$\xrightarrow{L_2 \leftrightarrow L_3} \left[\begin{array}{rrrr|r} 2 & 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 4 & 2 & -2 & -2 \end{array}\right] \xrightarrow{L_3 - 4L_2} \left[\begin{array}{rrrr|r} \underline{2} & 1 & 2 & 1 & 3 \\ 0 & \underline{1} & 1 & 0 & 0 \\ 0 & 0 & \underline{-2} & -2 & -2 \end{array}\right] = [A'|b'].$$

A matriz dos coeficientes A' está em **escada**, tendo-se:

- Não há linhas do tipo $0\ 0\ 0\ 0 \,|\, *$ com $* \neq 0$, isto é, não há equações impossíveis.
- A 4ª coluna de A' não tem pivot logo x_4 é variável livre, isto é, pode tomar qualquer valor.

Logo o sistema é **possível indeterminado** (PI).

Redução do sistema do exemplo (a): fase ascendente

Aplicando a fase ascendente do método de eliminação de Gauss à matriz em escada $[A'|b']$ do slide anterior obtém-se:

$$[A'|b'] = \left[\begin{array}{cccc|c} 2 & 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & -2 & -2 & -2 \end{array}\right] \xrightarrow{-\frac{1}{2}L_3} \left[\begin{array}{cccc|c} 2 & 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right]$$

$$\xrightarrow[L_2 - L_3]{L_1 - 2L_3} \left[\begin{array}{cccc|c} 2 & 1 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right] \xrightarrow{L_1 - L_2} \left[\begin{array}{cccc|c} 2 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{\frac{1}{2}L_1} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right] = [A''|b''] \dashrightarrow \begin{cases} x_1 & & & & = & 1 \\ & x_2 & & -x_4 & = & -1 \\ & & x_3 & +x_4 & = & 1 \end{cases}$$

A matriz dos coeficientes A'' está em **reduzida**. Passando nas duas últimas equações a variável livre x_4 para o membro direito, podemos escrever as variáveis *pivot* (a azul) à custa da variável livre x_4, obtendo-se

$$CS = \{(x_1, x_2, x_3, x_4) \,:\, x_1 = 1,\ x_2 = -1 + x_4,\ x_3 = 1 - x_4,\ x_4 \in \mathbb{R}\},$$

que possui uma infinidade de soluções.

TPC: dar exemplos de soluções do sistema linear anterior

Redução do exemplo (b): fase descendente

Aplicando a fase descendente do método de eliminação de Gauss à matriz ampliada $[A|b]$ do 2º sistema do slide 45 obtém-se:

$$[A|b] = \left[\begin{array}{ccc|c} 1 & 3 & 3 & 0 \\ 3 & 4 & 1 & -2 \\ 1 & -2 & -5 & 0 \end{array}\right] \xrightarrow[L_3 - L_1]{L_2 - 3L_1} \left[\begin{array}{ccc|c} 1 & 3 & 3 & 0 \\ 0 & -5 & -8 & -2 \\ 0 & -5 & -8 & 0 \end{array}\right]$$

$$\xrightarrow{L_3 - L_2} \left[\begin{array}{ccc|c} 1 & 3 & 3 & 0 \\ 0 & -5 & -8 & -2 \\ 0 & 0 & 0 & 2 \end{array}\right] = [A'|b'].$$

- A última linha da matriz $[A'|b']$ é do tipo $0\ 0\ 0\,|\,*$ com $* \neq 0$ e corresponde portanto a uma equação impossível.

Logo o sistema é **impossível** (IMP) e $CS = \emptyset$.

A equação matricial $Ax = b$

- Consideremos a equação matricial $Ax = b$ com

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 1 \\ 2 & 6 & 0 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} -1 \\ 6 \\ 10 \end{bmatrix}.$$

- Tem-se,

$$\begin{aligned} Ax = b \quad & \Leftrightarrow \quad \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 1 \\ 2 & 6 & 0 \end{bmatrix}_{3\times 3} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{3\times 1} = \begin{bmatrix} -1 \\ 6 \\ 10 \end{bmatrix}_{3\times 1} \\ & \Leftrightarrow \quad \begin{bmatrix} x_1 - x_2 + 2x_3 \\ 3x_1 + x_2 + x_3 \\ 2x_1 + 6x_2 \end{bmatrix}_{3\times 1} = \begin{bmatrix} -1 \\ 6 \\ 10 \end{bmatrix}_{3\times 1} \\ & \Leftrightarrow \quad \begin{cases} x_1 - x_2 + 2x_3 = -1 \\ 3x_1 + x_2 + x_3 = 6 \\ 2x_1 + 6x_2 = 10 \end{cases}. \end{aligned}$$

- Obteve-se uma relação importante - a equação matricial $Ax = b$ é equivalente ao sistema linear cuja matriz ampliada é $[A|b]$

Sistemas lineares na forma matricial Ax=b

Efetuando o mesmo tipo de cálculos pode-se mostrar facilmente que se tem, em geral, a equivalência

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \quad \Leftrightarrow \quad Ax = b,$$

com $A = [a_{ij}]$, $x = (x_1, \ldots, x_n)$ e $b = (b_1, \ldots, b_m)$, isto é, entre o sistema linear com matriz ampliada $[A|b]$ e a equação matricial $Ax = b$, o que permite traduzir os sistemas lineares para a linguagem das matrizes.

Observações

- Por abuso de linguagem, iremos ainda designar por sistema linear tanto a equação matricial $Ax = b$ como a respetiva a matriz ampliada $[A|b]$.
- Uma solução do sistema linear com matriz ampliada $[A|b]$ é uma solução de $Ax = b$, isto é, um vetor $u \in \mathbb{R}^n$ tal que $Au = b$.

TPC: traduzindo o sistema do slide 49 para a equação matricial equivalente, mostre que $(2, 1 - 1)$ é solução desse sistema.

Interpretação geométrica de sistemas lineares - exercício

Exercício na aula

Considere o sistema a 2 equações e 3 variáveis, com parâmetros $c, d \in \mathbb{R}$,

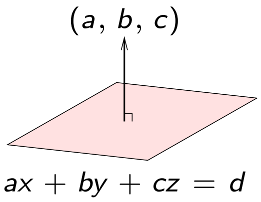
$$\begin{cases} x + 2y + 3z & = & 6 \\ 2x + 4y + cz & = & d \end{cases}$$

Discuta e interprete geometricamente o sistema anterior para todos os valores dos parâmetros $c, d \in \mathbb{R}$.

Resolução: comecemos por recordar que cada equação linear do tipo

$$ax + by + cz = d,$$

onde $a, b, c, d \in \mathbb{R}$ com a, b, c não todos nulos define um plano no espaço ($\mathbb{R}^3$) com vetor normal (a, b, c).



Logo o sistema anterior representa a intersecção de 2 planos no espaço.

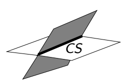
Resolução do exerecício (cont.)

Aplicando a fase descendente do método de eliminação de Gauss à matriz ampliada $[A|b]$ do sistema linear do slide 51 obtém-se:

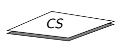
$$[A\,|\,b] = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & 4 & c & d \end{array}\right] \xrightarrow{L_2 - 2L_1} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 0 & c-6 & d-12 \end{array}\right] = [A'\,|\,b'].$$

Discussão e interpretação geométrica do sistema:

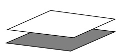
- Para $c \neq 6$ e d arbitrário o sistema é PI com 1 variável livre (y). Neste caso os **vetores normais aos planos não são múltiplos entre si** e o sistema representa 2 planos concorrentes numa reta. Logo $CS =$ reta.



- Para $c = 6$ os **vetores normais são múltiplos entre si** e temos 2 casos:
  - Se $d = 12$ o sistema é PI com 2 variáveis livres (y e z). Neste caso as duas equações são equivalentes e o sistema representa 2 planos coincidentes. Logo $CS =$ plano.



  - Se $d \neq 12$ o sistema é IMP. Neste caso o sistema representa 2 planos paralelos, não coincidentes. Logo $CS = \emptyset$.



Interpretação geométrica de sistemas de equações lineares

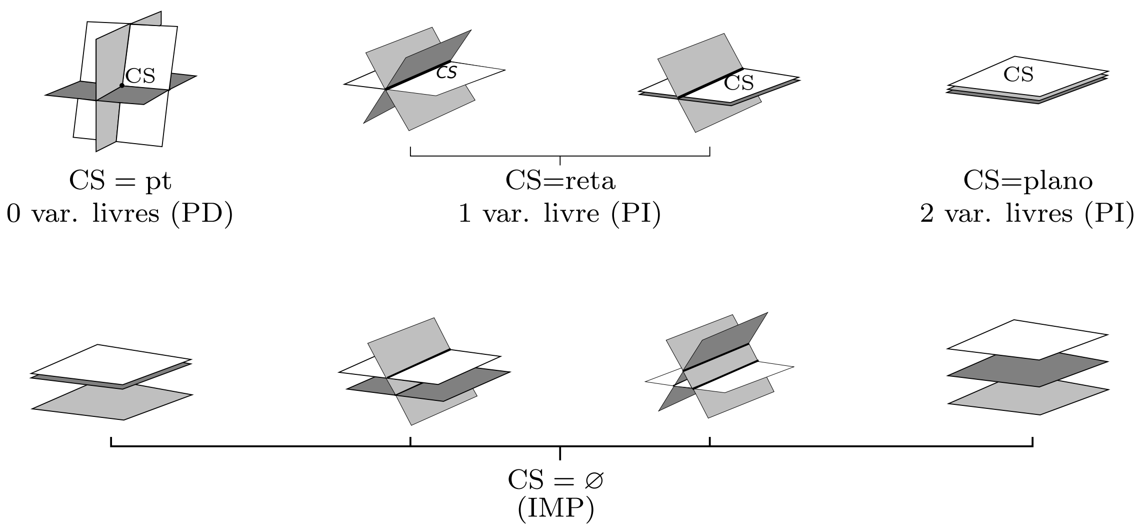
Em geral, tem-se o seguinte:

- Um sistema linear a m equações e n variáveis representa a **intersecção** de:
  - m retas em $\mathbb{R}^2$ (plano), se $n = 2$ (por exemplo, slide 33),
  - m planos em $\mathbb{R}^3$ (espaço), se $n = 3$ (por exemplo, slide 51),
  - m hiperplanos em $\mathbb{R}^n$, se $n \geq 4$ (por exemplo, (a) do slide 45).
- O número de variáveis livres([5]) de um sistema linear (possível) determina o tipo de CS que esse sistema possui. Por exemplo:
  - Se o número de variáveis livres for zero, o CS é um **ponto**
  - Se o número de variáveis livres for um, o CS é uma **reta**
  - Se o número de variáveis livres for dois, o CS é um **plano**
- Iremos principalmente interpretar sistemas lineares com 2 e 3 variáveis, ou seja, cujos CS estão contidos no plano e no espaço.

[5] Também designado por *grau de indeterminação do sistema.*

Geometria dos sistemas lineares a 3 equações e 3 variáveis

- Um sistema a 3 equações e 3 variáveis representa a interseção de 3 planos no espaço ($\mathbb{R}^3$) e geometricamente temos 8 casos:



TPC (Desafio)

Dar exemplos de sistemas com 3 equações e 3 variáveis para cada um dos 8 casos anteriores

Inversa de uma matriz

Definição de inversa

Uma matriz quadrada A de ordem n diz-se *invertível* ou *não singular* se existir uma matriz quadrada B da mesma ordem tal que

$$AB = I_n \qquad \text{e} \qquad BA = I_n,$$

onde I_n denota a matriz identidade de ordem n. Caso contrário, A diz-se *singular*. A matriz B, quando existe, designa-se por *inversa* de A

Por exemplo, $A = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$ é invertível com inversa $B = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$.

De facto, tem-se

$$\begin{aligned} AB &= \left[\begin{array}{cc} 2 & 1 \\ \hline 3 & 2 \end{array}\right] \left[\begin{array}{c|c} 2 & -1 \\ -3 & 2 \end{array}\right] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\ BA &= \left[\begin{array}{cc} 2 & -1 \\ \hline -3 & 2 \end{array}\right] \left[\begin{array}{c|c} 2 & 1 \\ 3 & 2 \end{array}\right] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Unicidade da inversa

Proposição

A inversa de uma matriz A, quando existe, é única e denota-se por A^{-1}.

Por exemplo, considerando as matrizes A e B do slide anterior, tem-se que B é a única inversa de A e podemos escrever, $A^{-1} = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$.

Demonstração da proposição

Consideremos inversas arbitrárias B e C de A. Queremos mostrar que têm que ser a mesma. Como B é inversa de A tem-se, em particular,

$$BA = I.$$

Multiplicando ambos os membros desta igualdade à direita por C tem-se

$$BAC = C,$$

donde resulta imediatamente $B = C$ uma vez que $AC = I$. $\square$

Ainda sobre a inversa...

Observação

Se A e B são matrizes quadradas de ordem n pode-se mostrar que

$$AB = I_n \quad \Leftrightarrow \quad BA = I_n,$$

Logo para provar que uma matriz quadrada A é invertível basta mostrar existe uma matriz quadrada B da mesma ordem que verifique uma das relações,

$$AB = I_n \quad \text{ou} \quad BA = I_n,$$

concluindo-se nessa altura que A e B são inversas uma da outra.

Exercícios na aula

- Mostre que se B é a inversa de A^2 então AB é inversa de A
- Determine, caso exista, a inversa da matriz $\text{diag}(2, 3)$.

TPC

Mostre que se $A_{n\times n}$ verifica $A^3 - 3A - I_n = 0$ então A é invertível e indique a respetiva inversa.

Propriedades da inversa

Proposição

Sejam A, B matrizes invertíveis da mesma ordem. Então:

1. A^{-1} é invertível e tem-se $(A^{-1})^{-1} = A$.
2. λA é invertível para todo o $\lambda \neq 0$ e tem-se $(\lambda A)^{-1} = \lambda^{-1}A^{-1}$.
3. AB é invertível e tem-se $(AB)^{-1} = B^{-1}A^{-1}$, ou seja, a inversa do produto é o produto das inversas, pela ordem inversa([6]).
4. A^k é invertível para todo o $k \in \mathbb{N}$ e tem-se $(A^k)^{-1} = (A^{-1})^k$.
5. A^T é invertível e tem-se $(A^T)^{-1} = (A^{-1})^T$.

(A demonstração das propriedades será feita nas aulas práticas.)

Potências negativas de matrizes invertíveis

Se A é uma matriz invertível, define-se

$$A^{-k} = (A^{-1})^k, \qquad k \in \mathbb{N}.$$

[6] Mais geralmente, se $A_1, A_2, \ldots, A_k$ são invertíveis da mesma ordem, então $A_1A_2\cdots A_k$ é também invertível e tem-se $(A_1A_2\cdots A_k)^{-1} = A_k^{-1}\cdots A_2^{-1}A_1^{-1}$.

Cálculo da inversa de uma matriz - exemplo

- Consideremos $A = \begin{bmatrix} 1 & 4 \\ 1 & 3 \end{bmatrix}$. Para determinarmos a inversa de A, caso exista, vamos resolver a equação matricial

$$AX = I_2,$$

onde X é uma matriz de incógnitas quadrada de ordem 2. Nessa altura sabemos que a solução desta equação, quando existe, corresponde à inversa de A. E em particular, tem que ser única.

- Escrevendo, $X = [\,x \quad y\,] = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix}$ e $I_2 = [\,e_1 \quad e_2\,] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, tem-se

$$\begin{aligned} AX = I_2 \quad &\Leftrightarrow \quad A\,[\,x \quad y\,] = [\,e_1 \quad e_2\,] \\ &\Leftrightarrow \quad [\,Ax \quad Ay\,] = [\,e_1 \quad e_2\,] \quad (\text{ver o slide 29}) \\ &\Leftrightarrow \quad \begin{cases} Ax = e_1 \dashrightarrow [A|e_1] = \left[\begin{array}{cc|c} 1 & 4 & 1 \\ 1 & 3 & 0 \end{array}\right] \\ \\ Ay = e_2 \dashrightarrow [A|e_2] = \left[\begin{array}{cc|c} 1 & 4 & 0 \\ 1 & 3 & 1 \end{array}\right] \end{cases} \end{aligned}$$

Redução simultânea de sistemas lineares e inversa

- Obtivemos 2 sistemas com a mesma matriz de coeficientes A.
- Podemos reduzir simultaneamente ambos os sistemas $Ax = e_1$ e $Ay = e_2$, ampliando A com e_1, e_2, isto é, com a matriz identidade I_2.

  Aplicando a fase descendente tem-se,

$$[\,A\,|\,I_2\,] = \left[\begin{array}{cc|cc} 1 & 4 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right] \xrightarrow{L_2 - L_1} \left[\begin{array}{cc|cc} 1 & 4 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{array}\right] \dashrightarrow \begin{cases} \left[\begin{array}{cc|c} 1 & 4 & 1 \\ 0 & -1 & -1 \end{array}\right] \\ \left[\begin{array}{cc|c} 1 & 4 & 0 \\ 0 & -1 & 1 \end{array}\right] \end{cases}$$

  Como os sistemas $Ax = e_1$ e $Ay = e_2$ são ambos PD, $AX = I_2$ é possível com solução única $X = A^{-1}$. Em particular, A é invertível.

- Aplicando a fase ascendente tem-se,

$$\left[\begin{array}{cc|cc} 1 & 4 & 1 & 0 \\ 0 & -1 & 1 & -1 \end{array}\right] \xrightarrow{-L_2} \left[\begin{array}{cc|cc} 1 & 4 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right] \xrightarrow{L_1 - 4L_2} \left[\begin{array}{cc|cc} 1 & 0 & -3 & 4 \\ 0 & 1 & 1 & -1 \end{array}\right].$$

  Logo as soluções de $Ax = e_1$ e $Ay = e_2$ são $x = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ e $y = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$, e obtém-se

$$A^{-1} = X = [\,x \quad y\,] = \begin{bmatrix} -3 & 4 \\ 1 & -1 \end{bmatrix}.$$

Inversa e redução simultânea de sistemas (caso geral)

- Pode-se mostrar analogamente que resolver a equação matricial $AX = I_n$ com A matriz arbitrária de ordem n, equivale a resolver n sistemas lineares com matrizes ampliadas

$$[\,A\,|\,e_1\,], \quad [\,A\,|\,e_2\,], \quad \cdots, \quad [\,A\,|\,e_n\,], \tag{1}$$

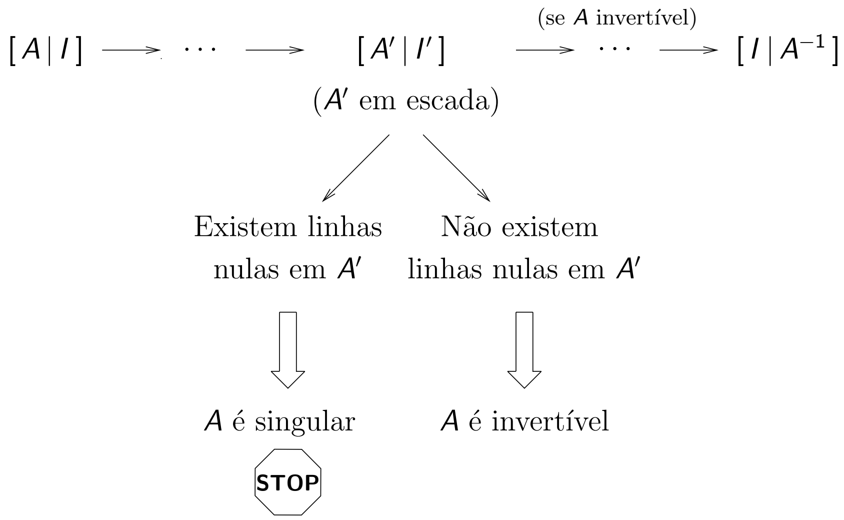
onde $e_1, e_2, \ldots, e_n$ são as colunas da matriz identidade I_n.

- Como os n sistemas têm a mesma matriz de coeficientes A podem ser resolvidos simultaneamente aplicando o método de Gauss à matriz ampliada $[\,A\,|\,e_1\ e_2\ \cdots\ e_n\,] = [\,A\,|\,I_n\,]$.
- Pela **unicidade da inversa**, um dos dois casos tem de ocorrer:
  - os n sistemas (1) são todos PD - nessa altura A é invertível e as n colunas de A^{-1} são as soluções dos n sistemas (1);
  - pelo menos um dos n sistemas (1) é impossível - nessa altura A é singular.
- A redução simultânea de sistemas pode também ser aplicada para resolver equações matriciais mais gerais, do tipo $AX = B$, aplicando o método de Gauss à matriz ampliada $[\,A\,|\,B\,]$...

Algoritmo da inversa

Das considerações dos slide anterior deduz-se o seguinte algoritmo.

- `Input:` Matriz quadrada A
- `Objectivo:` Decidir sobre a invertibilidade de A e calcular A^{-1}



Algoritmo da inversa

Exercício na aula

Aplicando o algoritmo da inversa, decida sobre a invertibilidade de

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 1 & 0 & -1 \end{bmatrix},$$

e determine a sua inversa (caso exista).

Qual é a solução do sistema linear $Ax = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$?

(**Sugestão**: ver o slide 61.)

Resolução aplicando o algoritmo da inversa

$$[A\,|\,I] = \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{L_3 - L_1} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right] = [A'\,|\,I'].$$

A matriz dos coeficientes A' está em escada e não tem linhas nulas. Logo A é invertível. Aplicando a fase ascendente do método de eliminação de Gauss vem,

$$[\,A'\,|\,I'\,] = \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right] \xrightarrow{L_1 + 2L_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 2 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{\frac{1}{2}L_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 2 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right] = [\,I\,|\,A^{-1}\,].$$

Podemos confirmar que A^{-1} está bem calculada verificando a relação $AA^{-1} = I$:

$$\left[\begin{array}{ccc} 1 & 0 & -2 \\ \hline 0 & 2 & 0 \\ \hline 1 & 0 & -1 \end{array}\right] \left[\begin{array}{c|c|c} -1 & 0 & 2 \\ 0 & \frac{1}{2} & 0 \\ -1 & 0 & 1 \end{array}\right] = \begin{bmatrix} 1 & 0 & o \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \checkmark$$

Recorde que vetor $(1, 0, 0)$ é a 1ª coluna da matriz identidade, que denotámos por e_1, e portanto a solução de $Ax = e_1$ é a primeira coluna de A^{-1}, isto é, $(-1, 0, -1)$.

Caso particular: inversa de matrizes diagonais

Efetuando trocas de linhas conclui-se facilmente que qualquer matriz em escada A' obtida a partir de uma matriz diagonal,

$$A = \operatorname{diag}(a_1, a_2 \ldots, a_n) = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix},$$

não tem linhas nulas se e só se $a_1, a_2, \ldots, a_n \neq 0$. Nessa altura, podemos obter imediatamente a inversa de A.

Proposição

A matriz $\operatorname{diag}(a_1, a_2, \ldots, a_n)$ é invertível se e só se $a_1, a_2, \ldots, a_n \neq 0$, tendo-se nessa caso

$$\operatorname{diag}^{-1}(a_1, a_2, \ldots, a_n) = \operatorname{diag}(a_1^{-1}, a_2^{-1}, \ldots, a_n^{-1}).$$

Por exemplo,

$$\operatorname{diag}^{-1}(2, 5, \pi) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & \pi \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{5} & 0 \\ 0 & 0 & \frac{1}{\pi} \end{bmatrix} = \operatorname{diag}(\tfrac{1}{2}, \tfrac{1}{5}, \tfrac{1}{\pi}).$$

Interlúdio: característica de uma matriz

Definição de característica

A característica de uma matriz A, denotada $\operatorname{car}(A)$, é o número de pivots de qualquer matriz em escada que seja obtida a partir de A, por aplicação de operações elementares do método de eliminação de Gauss nas linhas de A.

- A característica está bem definida uma vez que coincide com o número de pivots da matriz reduzida, que é única, e a fase ascendente do método de Gauss não altera o número de pivots.
- $\operatorname{car}(A)$ corresponde também ao número de linhas não nulas de qualquer matriz em escada obtida a partir de A.
- Uma vez que não pode haver mais que um pivot em cada linha e em cada coluna de uma matriz em escada, a característica de uma matriz $A_{m\times n}$ não pode ultrapassar o número de linhas m nem o número de colunas n de A, isto é,
$$\operatorname{car}(A) \leq \min\{m, n\}.$$
- Pode-se provar que para qualquer matriz $A_{m\times n}$ se tem a relação:
$$\operatorname{car}(A^T) = \operatorname{car}(A).$$

Exemplo

- Se, por exemplo,

$$A_{4\times5} \rightarrow \overset{\text{oper. elementares}}{\cdots} \rightarrow A' = \begin{bmatrix} 0 & 1 & 2 & 1 & 4 \\ 0 & 0 & 3 & -6 & 0 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{4\times5},$$

tem-se $\text{car}(A) = 3 \leq \min\{4, 5\}$.

Questão

Considerando a mesma matriz A do exemplo acima, qual seria o número de linhas nulas de uma matriz em escada que fosse obtida a partir de A^T por aplicação de operações elementares?

Inversa e característica

Uma vez que uma matriz **quadrada** em **escada** só possui linhas nulas se existirem colunas sem pivot, obtém-se imediatamente o seguinte critério de invertibilidade.

Teorema (critério de invertibilidade)

Uma matriz quadrada A de ordem n é invertível se e só se $\text{car}(A) = n$.

Exercício na aula

Para que valores de $\alpha \in \mathbb{R}$, a matriz $A = \begin{bmatrix} -1 & 0 & \alpha \\ 0 & \alpha & 4 \\ 1 & 2\alpha & 6 \end{bmatrix}$ é invertível ?

Resolução: Aplicando a fase descendente do método de Gauss tem-se:

$$A = \begin{bmatrix} -1 & 0 & \alpha \\ 0 & \alpha & 4 \\ 1 & 2\alpha & 6 \end{bmatrix} \xrightarrow{L_3 + L_1} \begin{bmatrix} -1 & 0 & \alpha \\ 0 & \alpha & 4 \\ 0 & 2\alpha & 6+\alpha \end{bmatrix} \xrightarrow{\frac{1}{2}L_2} \begin{bmatrix} -1 & 0 & \alpha \\ 0 & \alpha & 4 \\ 0 & 0 & \alpha - 2 \end{bmatrix}.$$

Tem-se que $A_{3\times3}$ é ínvertível se e só se $\text{car}(A) = 3$ se e só se $\alpha \neq 0, 2$.

Aplicação da matriz inversa aos sistemas lineares

- A equação linear $ax = b$ em que $a, b \in \mathbb{R}$ com $a \neq 0$, admite a solução única $x = \frac{b}{a}$, que se pode escrever na forma $x = a^{-1}b$.
- A noção de inversa de uma matriz permite obter a solução de um sistema do tipo $Ax = b$ com A invertível, de uma forma análoga.

Proposição

Se A é uma matriz invertível então o sistema linear $Ax = b$ é PD com solução única $x = A^{-1}b$ qualquer que seja $b \in \mathbb{R}^n$.

De facto, se A é invertível, existe A^{-1} e podemos escrever, multiplicando à esquerda ambos os membros da equação matricial $Ax = b$ por A^{-1},

$$\begin{aligned} Ax = b \quad &\Rightarrow \quad A^{-1}(Ax) = A^{-1}b \\ &\Leftrightarrow \quad (A^{-1}A)x = A^{-1}b \\ &\Leftrightarrow \quad I\,x = A^{-1}b \\ &\Leftrightarrow \quad x = A^{-1}b. \end{aligned}$$

Hors-d'oeuvre: transformações lineares

Uma matriz $A_{m\times n}$ define de forma natural uma transformação $T = T_A$,

$$\begin{aligned} T : \mathbb{R}^n &\to \mathbb{R}^m \\ x &\mapsto A\,x \end{aligned}$$

Exemplo

Se $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & -1 \end{bmatrix}_{2\times 3}$, obtém-se a transformação $T : \mathbb{R}^3 \to \mathbb{R}^2$, definida por

$$T(x) = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ x_1 - x_3 \end{bmatrix},$$

ou seja, em termos de coordenadas vem dada por,

$$T(x_1, x_2, x_3) = (x_1 + 2x_2 + 3x_3\,,\ x_1 - x_3).$$

Vejamos alguns casos bem conhecidos de transformações geométricas no plano e no espaço que podem ser definidas por matrizes([7]).

[7] Nem todas as transformações geométricas no plano e no espaço podem ser definidas a partir de matrizes como acima.

Transformações geométricas no plano definidas por matrizes $A_{2\times 2}$

Exemplos de transformações geométricas do plano definidas por matrizes $A_{2\times 2}$ (algumas dessas transformações estão ilustradas na parte superior do slide 74):

- Se $A = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} = \alpha I_2$ com $\alpha > 0$ obtém-se a homotetia de razão α,
$$H_\alpha(x) = T_{\alpha I_2}(x) = (\alpha I_2)\, x = \alpha x,$$
isto é, em coordenadas,
$$H_\alpha(x_1, x_2) = (\alpha\, x_1, \alpha\, x_2),$$
que corresponde a uma dilatação [contração] se $\alpha > 1$ [$\alpha < 1$].

- Se $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (matriz de permutação), obtém-se
$$S(x) = T_A(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix},$$
isto é,
$$S(x_1, x_2) = (x_2, x_1),$$
que corresponde à reflexão relativamente à bissectriz dos quadrantes ímpares.

Transformações geom. no plano definidas por matrizes $A_{2\times 2}$ (cont.)

- Se $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ obtém-se
$$R(x) = T_A(x) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix},$$
isto é,
$$R(x_1, x_2) = (-x_2, x_1),$$
que corresponde a uma rotação de $\frac{\pi}{2}$ radianos (no sentido anti-horário).

- Mais geralmente, se $A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, obtém-se a rotação de θ radianos (no sentido anti-horário)
$$R_\theta(x) = T_{A_\theta}(x),$$
que corresponde à rotação de $\frac{\pi}{2}$ radianos (no sentido anti-horário) e se deixa como exercício os alunos descreverem em termos de coordenadas. Note que $R = R_{\frac{\pi}{2}}$.

- Se $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ obtém-se a projeção no eixo dos xx,
$$P_1(x_1, x_2) = x_1.$$

Transformações geométricas no espaço definidas por matrizes $A_{3\times 3}$

Algumas transformações geométricas do espaço definidas por matrizes $A_{3\times 3}$ (algumas dessas transformações estão ilustradas na parte inferior do slide 74):

- Se $A = \alpha I_3$ com $\alpha > 0$ obtém-se a homotetia no plano de razão α:

  $$H_\alpha(x) = T_{\alpha I_3}(x) = (\alpha I_3)x = \alpha x,$$

  isto é, em coordenadas,

  $$H_\alpha(x_1, x_2, x_3) = (\alpha\, x_1, \alpha\, x_2, \alpha\, x_3),$$

  que corresponde uma dilatação [contração] se $\alpha > 1$ [$\alpha > 1$].

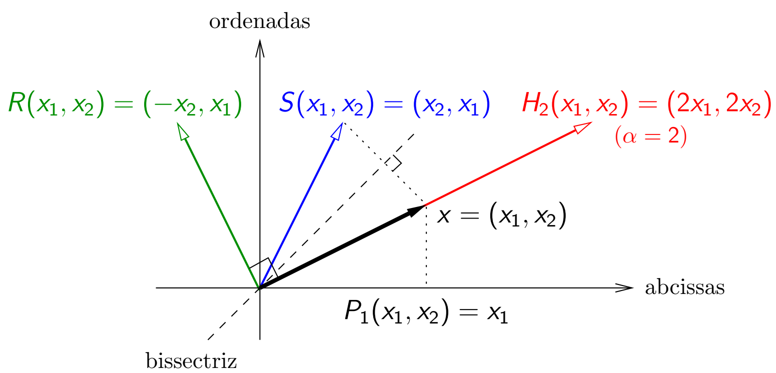
- Se $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$, obtém-se a simetria em relação ao plano xOy,

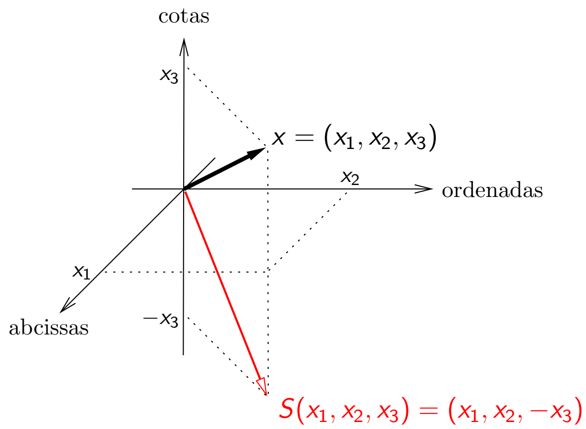
  $$S(x) = Ax = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ -x_3 \end{bmatrix}$$

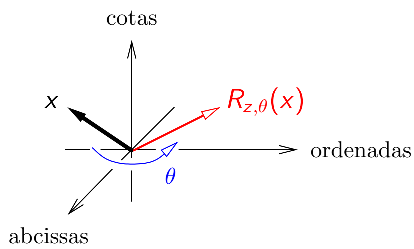
  isto é, em coordenadas, $S_z(x_1, x_2, x_3) = (x_1, x_2, -x_3)$. (não foi dado na aula).

- Se $A = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$, obtém-se a rotação de θ radianos em torno do eixo dos zz (no sentido direto), $R_{z,\theta} = Ax$, que se deixa como exercício para os alunos descreverem em termos de coordenadas.

Ilustração de algumas das transformações geométricas do plano e do espaço







Definição de transformação linear e propriedades

Definição de transformação linear

Uma transformação $T : \mathbb{R}^n \to \mathbb{R}^m$ diz-se linear se verificar as seguintes propriedades para todo o $x, y \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$:

- $T(x+y) = T(x) + T(y)$ (aditividade)
- $T(\alpha\, x) = \alpha\, T(x)$ (homogeneidade)

Algumas consequências

Se $T : \mathbb{R}^n \to \mathbb{R}^m$ é linear então

- $T(\vec{0}) = \vec{0}$.
- Para todo o $x, y \in \mathbb{R}^n$ e $\alpha, \beta \in \mathbb{R}$ tem-se
$$T(\alpha x + \beta y) = T(\alpha x) + T(\beta y) = \alpha\, T(x) + \beta\, T(y).$$
- Mais geralmente, para todo o $u_1, \ldots, u_k \in \mathbb{R}^n$ e $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, tem-se,
$$T(\alpha_1 u_1 + \cdots + \alpha_k u_k) = \alpha_1 T(u_1) + \cdots + \alpha_k T(u_k).$$

Transformações lineares

Teorema (equivalência entre transf. matricial e linear)

- Toda a transformação matricial $T_A : \mathbb{R}^n \to \mathbb{R}^m$ definida por uma matriz $A_{m\times n}$ é linear.
- Reciprocamente, se $T : \mathbb{R}^n \to \mathbb{R}^m$ é uma transformação linear então T é definida por uma matriz, mais precisamente, $T = T_A$, com
$$A = \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_n) \end{bmatrix},$$
onde $e_1, e_2, \ldots, e_n$ são as n colunas da matriz identidade.
A matriz A designa-se por matriz *standard* da transformação linear T.

A demonstração do primeiro ponto é imediata. De facto, se $x, y \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, tem-se

$$T_A(x+y) = A(x+y) = Ax + Ay = T_A(x) + T_A(y),$$
$$T_A(\alpha\, x) = A(\alpha\, x) = \alpha\, Ax = \alpha\, T_A(x).$$

Relativamente ao 2º ponto vamos apenas mostrar como se pode obter a matriz da transformação linear num exemplo.

Matriz de uma transformação linear - exemplo

Consideremos a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^2$ definida por

$$T(x_1, x_2, x_3) = (x_1 + x_2 + x_3\ ,\ x_1 - x_3).$$

Vejamos como podemos obter a matriz desta transformação linear. Podemos escrever,

$$\begin{aligned} T(x_1, x_2, x_3) &= T(x_1(1,0,0) + x_2(0,1,0) + x_3(0,0,1)) \\ &= x_1 \underbrace{T(1,0,0)}_{(1,1)} + x_2 \underbrace{T(0,1,0)}_{(1,0)} + x_3 \underbrace{T(0,0,1)}_{(1,-1)} \\ &= x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \end{aligned}$$

Logo $T(x) = Ax$, com $A = [\,T(e_1) \quad T(e_2) \quad T(e_3)\,] = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}_{2\times 3}$.

$$\begin{aligned} \text{De facto,} \quad Ax &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= \begin{bmatrix} x_1 + x_2 + x_3 \\ x_1 + 0x_2 - x_3 \end{bmatrix} = (x_1 + x_2 + x_3\ ,\ x_1 - x_3) = T(x_1, x_2, x_3). \end{aligned}$$

Composição de transformações lineares

Dadas matrizes encadeadas $A_{m\times n}$ e $B_{n\times p}$ tem-se,

$$T_A(T_B(x)) = A(Bx) = (AB)x = T_{AB}(x),$$

para todo o $x \in \mathbb{R}^p$, ou seja, a composição das transformações lineares definidas pelas matrizes encadeadas $A_{m\times n}$ e $B_{n\times p}$,

$$\mathbb{R}^p \xrightarrow{T_B} \mathbb{R}^n \xrightarrow{T_A} \mathbb{R}^m,$$

é a transformação definida pelo produto $(AB)_{m\times p}$,

$$\mathbb{R}^p \xrightarrow{T_{AB}} \mathbb{R}^m,$$

o que permite interpretar o produto de matrizes em termos de composição de transformações lineares.

Composição de transformações lineares - exemplos

Exemplo

Mantendo as notações das transformações geométricas do plano,

$$R(x) = T_A(x), \quad \text{com} \quad A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

e

$$S(x) = T_B(x), \quad \text{com} \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

tem-se

$$AB = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Portanto a composição de R com S vem dada por,

$$(R \circ S)(x) = T_A(T_B(x)) = T_{AB}(x) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (-x_1, x_2),$$

que corresponde à reflexão no plano relativamente ao eixo dos yy.

De facto, em coordenadas, $R(x_1, x_2) = (-x_2, x_1)$ e $S(x_1, x_2) = (x_2, x_1)$. Logo,

$$R(S(x_1, x_2)) = R(x_2, x_1) = (-x_1, x_2).$$

Inversa de uma transformação linear

Uma transformação linear

$$T_A : \mathbb{R}^n \to \mathbb{R}^n,$$

diz-se invertível se existir uma transformação

$$S : \mathbb{R}^n \to \mathbb{R}^n,$$

tal que

$$T_A \circ S = S \circ T_A = \mathrm{id}_{\mathbb{R}^n}.$$

Nessa altura prova-se que S é também linear e tem-se, escrevendo $S = T_B$,

$$T_A \circ T_B = T_B \circ T_A = \mathrm{id}_{\mathbb{R}^n} = T_{I_n} \Leftrightarrow AB = BA = I_n,$$

ou seja, A é invertível com inversa B. A transformação $S = T_B$ designa-se inversa de A e denota-se T_A^{-1}. Obteve-se então o resultado.

Proposição

Uma transformação linear $T_A : \mathbb{R}^n \to \mathbb{R}^n$ é invertível se e só se A é invertível e nessa altura a sua inversa é $T_A^{-1} = T_{A^{-1}}$.

Inversa de uma transformação linear - exemplos

Exemplos

- A inversa da rotação $R_{z,\theta}$ em torno do eixo dos zz de θ radianos é a rotação $R_{z,-\theta}$ (verifique).
- Considerando $A = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$, obtém-se

$$T_A(x_1, x_2) = (2x_1 + x_2, 3x_1 + 2x_2) \qquad (\textit{verifique}).$$

Como A é invertível com inversa $A^{-1} = \begin{bmatrix} 2 & -1 \\ -3 & 1 \end{bmatrix}$, tem-se

$$T_A^{-1}(x) = T_{A^{-1}}(x) = \begin{bmatrix} 2 & -1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_2 \\ -3x_1 + 2x_2 \end{bmatrix},$$

isto é, $T_A^{-1}(x_1, x_2) = (2x_1 - x_2, -3x_1 + 2x_2)$.

De facto, $T_A(T_A^{-1}(x_1, x_2)) = T_A^{-1}(T_A(x_1, x_2)) = (x_1, x_2)$ (confirme).

O espaço vetorial $\mathbb{R}^n$

No slide 8 mencionámos as seguintes propriedades das operações, adição de vetores de $\mathbb{R}^n$ e produto de um vetor de $\mathbb{R}^n$ por um escalar, que decorrem imediatamente de propriedades análogas dos números reais.

Propriedades das operações algébricas com vetores

Sejam x, y, z vetores de $\mathbb{R}^n$, $\vec{0} = (0, \cdots, 0) \in \mathbb{R}^n$ e $\lambda, \mu \in \mathbb{R}$. Tem-se,

1. $x + y = y + x$ (comutativa)
2. $(x + y) + z = x + (y + z)$ (associativa)
3. $x + \vec{0} = x$ (existência de el. neutro)
4. $x + (-x) = \vec{0}$ (existência de el. simétrico)
5. $\lambda(x + y) = \lambda x + \lambda y$ (distributiva...)
6. $(\lambda + \mu)x = \lambda x + \mu x$ (distributiva...)
7. $(\lambda\mu)x = \lambda(\mu x)$ (compatibilidade dos produtos)
8. $1\,x = x$ (el. identidade da multiplicação por escalar)

Estas 8 propriedades podem ser resumidas dizendo que $\mathbb{R}^n$ munido da adição de vetores e do produto de vetores por escalares é um *espaço vetorial*...

Subespaço vetorial de $\mathbb{R}^n$

Vamos estudar os subconjuntos não vazios $V \subset \mathbb{R}^n$ para os quais se podem adicionar vetores de V e multiplicar vetores de V por escalares sem sair de V, isto é, de modo a ainda se obterem vetores de V.

Definição de subespaço vetorial

Um subconjunto $V \subset \mathbb{R}^n$ diz-se um *subespaço vetorial* de $\mathbb{R}^n$ se

- $V \neq \varnothing$
- V é fechado para a adição, isto é, para todo o $u, v \in V$ tem-se $u + v \in V$
- V é fechado para o produto por escalar, isto é, para todo o $u \in V$ e $\alpha \in \mathbb{R}$ tem-se $\alpha u \in V$

Observação

É imediato verificar que se V é um subespaço vetorial de $\mathbb{R}^n$ ainda são válidas as propriedades $(1), \ldots, (8)$ relativamente aos vetores de V, isto é, que V munido das operações adição de vetores e multiplicação de vetores por escalares, herda a estrutura de espaço vetorial de $\mathbb{R}^n$.

Conceito de subespaço vetorial

Exercício na aula

Quais dos seguintes conjuntos V definem subespaços vetoriais de $\mathbb{R}^2$?

- $V = \{(x, y) : xy = 0\}$ (eixos coordenados de $\mathbb{R}^2$)
  Não (não é fechado para a adição - se considerarmos, por exemplo, $u = (1, 0)$ e $v = (0, 1)$ tem-se $u, v \in V$ mas $u + v = (1, 1) \notin V$)
- $V = \{(x, y) : x, y \geq 0\}$ (1^{o} quadrante de $\mathbb{R}^2$)
  Não (não é fechado para o produto por escalar - se considerarmos, por exemplo, $u = (1, 0) \in V$ e $\alpha = -2 \in \mathbb{R}$ tem-se $\alpha u = -2u = (-2, 0) \notin V$)
- $V = \{(x, y) : y = 0\}$ (eixo dos xx).
  Sim (veremos que o conjunto de soluções de um sistema linear em que os termos indpendentes são nulos - sistema linear homogéneo - é sempre um subespaço vetorial)

Subespaço mininal e subespaço maximal

As 3 condições da definição de subespaço vetorial de $\mathbb{R}^n$ do slide 83 são trivialmente verificadas quando $V = \{\vec{0}\}$ e $V = \mathbb{R}^n$, obtendo-se os seguintes 2 subespaços vetoriais especiais de $\mathbb{R}^n$:

- $\{\vec{0}\}$ subespaço vetorial minimal (ou trivial).
- $\mathbb{R}^n$ subespaço vetorial maximal.

Condição **necessária** para ser subespaço vetorial...

Vamos agora estabelecer uma condição necessária (mas não suficiente) para um subconjunto de $\mathbb{R}^n$ definir um subespaço vetorial.

Teorema

Se V é subespaço vetorial de $\mathbb{R}^m$ então $\vec{0} \in V$.

Demonstração: Se V é subespaço vetorial de $\mathbb{R}^m$, tem-se:

- $V \neq \emptyset$, logo existe um vetor $v \in V$.
- V é fechado para o produto por escalar, logo $\lambda v \in V, \forall \lambda \in \mathbb{R}$.
- Considerando, em particular, $\lambda = 0$, conclui-se que $0v = \vec{0} \in V$ como se pretendia. $\square$

Exemplo

$V = \{(x, y) : x^2 + y^2 = 1\}$ não é subespaço vetorial de $\mathbb{R}^2$, pois $(0, 0) \notin V$ $(0^2 + 0^2 \neq 1)$. O que representa geometricamente V?

Espaço nulo de uma matriz

O seguinte conceito introduz o primeiro dos subespaços vetoriais fundamentais associados a matrizes que vamos considerar.

Definição de espaço nulo de uma matriz

Seja A uma matriz do tipo $m \times n$. Chama-se *espaço nulo de A* e denota-se por $\mathcal{N}(A)$, ao conjunto de soluções do sistema linear $Ax = \vec{0}$, isto é,

$$\mathcal{N}(A) = \left\{x \in \mathbb{R}^n \ : \ Ax = \vec{0}\right\} \subset \mathbb{R}^n.$$

- O espaço nulo de A corresponde ao CS do sistema linear $Ax = \vec{0}$, dito homogéneo, em que o vetor dos termos constantes é o vetor nulo.
- Um sistema homogéneo nunca é impossível uma vez que possui sempre a solução trivial $x = \vec{0}$ (pois $A\vec{0} = \vec{0}$). Em particular $\mathcal{N}(A) \neq \emptyset$.
- Para calcularmos $\mathcal{N}(A)$ temos que resolver o sistema homogéneo $Ax = \vec{0}$, isto é, temos que reduzir a matriz ampliada $\left[\, A \,|\, \vec{0} \,\right]$ ([8]).

[8] O vetor dos termos constantes pode ser omitido, uma vez que é sempre nulo ao longo do método de Gauss.

Espaço nulo - exercícios na aula

Exercícios na aula

Determine os espaços nulos das seguintes matrizes:

1. $A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & -1 & 1 \end{bmatrix}_{2\times 3}$.
2. $A = \begin{bmatrix} 1 & 4 \\ 2 & -1 \end{bmatrix}$

Reduzindo a matriz ampliada $[\, A \,|\, \vec{0}\,]$ do 1º sistema obtém-se,

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 2 & -1 & 1 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right] \rightsquigarrow \left\{ \begin{array}{rl} x_1 & = 0 \\ x_2 - x_3 & = 0. \end{array}\right.$$

Logo,

$$\begin{aligned} \mathcal{N}(A) &= \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = 0,\ x_2 = x_3,\ x_3 \in \mathbb{R}\} \\ &= \{(0, x_3, x_3) : x_3 \in \mathbb{R}\} = \{x_3(0,1,1) : x_3 \in \mathbb{R}\} \end{aligned}$$

- Geometricamente $\mathcal{N}(A)$ define uma reta de $\mathbb{R}^3$ (porque o sistema $Ax = \vec{0}$ possui uma variável livre), que passa na origem (porque o sistema é homogéneo) com vetor diretor $v = (0, 1, 1)$.

Cálculo do espaço nulo do 2º exercício

Consideremos agora a matriz $A = \begin{bmatrix} 1 & 4 \\ 2 & -1 \end{bmatrix}_{2\times 2}$.

- Aplicando a fase descendente do método de Gauss obtém-se,

  $$[A \mid \vec{0}] = \left[\begin{array}{cc|c} 1 & 4 & 0 \\ 2 & -1 & 0 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 4 & 0 \\ 0 & -9 & 0 \end{array}\right] = [A' \mid \vec{0}].$$

  Uma vez que todas as colunas de A' têm pivot o sistema $Ax = \vec{0}$ é determinado e portanto possui apenas a solução trivial $x_1 = x_2 = 0$, isto é, $CS = \{(0,0)\}$.([9])
- Logo $\mathcal{N}(A) = \{\vec{0}\}$, isto é, $\mathcal{N}(A)$ é o subespaço minimal de $\mathbb{R}^2$.

Critério para $\mathcal{N}(A) = \{\vec{0}\}$ (subespaço minimal)

$\mathcal{N}(A) = \{\vec{0}\} \Leftrightarrow Ax = \vec{0}$ é determinado $\Leftrightarrow$ car(A)=nº colunas de A.

[9] Confirme que aplicando a fase ascendente à matriz ampliada $[A' \mid \vec{0}]$ se obtém a matriz $[I_2 \mid \vec{0}]$, com I_2 a matriz identidade de ordem 2, e portanto que a solução (única) do sistema é $x_1 = x_2 = 0$.

O espaço nulo é um subespaço vetorial...

Teorema

Se A é uma matriz do tipo $m \times n$ então $\mathcal{N}(A)$ é um subespaço vetorial de $\mathbb{R}^n$.

Demonstração

Temos que verificar as 3 condições da definição de subespaço vetorial do slide 83:

- $\mathcal{N}(A) \neq \emptyset$ como vimos no slide 87.
- $\mathcal{N}(A)$ fechado para a adição: se $u, v \in \mathcal{N}(A)$ então u e v são soluções de $Ax = \vec{0}$, isto é, $Au = Av = \vec{0}$ e portanto $A(u+v) = Au + Av = \vec{0} + \vec{0} = \vec{0}$, o que significa que $u+v$ é também solução de $Ax = \vec{0}$. Logo $u + v \in \mathcal{N}(A)$.
- $\mathcal{N}(A)$ fechado para o produto por escalar fica para os alunos mostrarem... $\square$

Subespaços vetoriais definidos por CS

O CS de qualquer sistema linear homogéneo com n variáveis define um subespaço vetorial de $\mathbb{R}^n$, pois corresponde ao espaço nulo da matriz dos coeficientes desse sistema (que possui n colunas).

- Por exemplo, o seguinte CS de um sistema homogéneo,

  $V = \{(x_1, x_2, x_3, x_4) : x_1 + 2x_2 + 3x_3 + 4x_4 = 0,\ -x_1 + 3x_2 + x_4 = 0\}$,

  é um subespaço vetorial de $\mathbb{R}^4$, pois $V = \mathcal{N}\left(\begin{bmatrix} 1 & 2 & 3 & 4 \\ -1 & 3 & 0 & 1 \end{bmatrix}\right)$.

O CS de um sistema linear não homogéneo nunca define um subespaço vetorial uma vez que não contém o vetor nulo (origem).

- Por exemplo, o plano de $\mathbb{R}^3$ definido pela equação não homogénea $x_1 + x_2 + 2x_3 = 1$,

  $$V = \{(x_1, x_2, x_3) : x_1 + x_2 + 2x_3 = 1\},$$

  não define um subespaço vetorial porque não passa na origem.

Possíveis subespaços vetoriais de $\mathbb{R}^n$

Mas afinal quais são os conjuntos que definem subespaços vetoriais de $\mathbb{R}^n$?

- Subespaços vetoriais do plano ($\mathbb{R}^2$):

  $$\{\vec{0}\}, \quad \text{retas que passam na origem}, \quad \mathbb{R}^2.$$

- Subespaços vetoriais do espaço ($\mathbb{R}^3$):

  $$\{\vec{0}\}, \quad \text{retas e planos que passam na origem}, \quad \mathbb{R}^3.$$

- Subespaços vetoriais de $\mathbb{R}^n$, com $n \geq 4$:

  $$\{\vec{0}\}, \quad \text{retas, \ldots e hiperplanos que passam na origem}, \quad \mathbb{R}^n.$$

  (um hiperplano é um conjunto definido por uma equação linear do tipo $a_1x_1 + \cdots + a_nx_n = b$, com os coeficientes $a_1, \ldots, a_n$, não todos nulos.)

Combinação linear de vetores

Definição de combinação linear

Um vetor $b \in \mathbb{R}^m$ é **combinação linear** (CL) de $v_1, \ldots, v_n \in \mathbb{R}^m$ se existirem escalares $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tais que

$$b = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

Os escalares $\alpha_1, \ldots, \alpha_n$ chamam-se *coeficientes* da combinação linear.

Por outras palavras, b é CL de $v_1, \ldots, v_n$ se puder ser obtido como soma de múltiplos desses vetores.

Exemplos de combinações lineares de vetores

- $(-2, -4, -2) = -2(1, 2, 1)$.
- $\begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- O vetor nulo $\vec{0} \in \mathbb{R}^n$ é CL de qualquer conjunto de m vetores $v_1, \ldots, v_m \in \mathbb{R}^n$:

$$\vec{0} = 0\, v_1 + 0\, v_2 + \cdots + 0\, v_m$$

- E cada um dos vetores v_i é CL dos vetores $v_1, v_2, \ldots, v_m$:

$$\begin{aligned} v_1 &= 1\, v_1 + 0\, v_2 + \cdots + 0\, v_m, \\ v_2 &= 0\, v_1 + 1\, v_2 + \cdots + 0\, v_m, \\ &\vdots \\ v_m &= 0\, v_1 + 0\, v_2 + \cdots + 1\, v_m. \end{aligned}$$

Determinação da combinação linear de vetores - exemplo

- Será que $b = (2, 5, 1)$ é CL de $v_1 = (2, 2, 1)$ e $v_2 = (2, 3, 1)$?
- Por outras palavras, será que existem escalares $\alpha_1, \alpha_2 \in \mathbb{R}$ tais que

$$b = \alpha_1 v_1 + \alpha_2 v_2 \ ?$$

Ora,

$$\begin{aligned} b = \alpha_1 v_1 + \alpha_2 v_2 \quad &\Leftrightarrow \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix} = \alpha_1 \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \\ &\Leftrightarrow \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 2\alpha_1 + 2\alpha_2 \\ 2\alpha_1 + 3\alpha_2 \\ \alpha_1 + \alpha_2 \end{bmatrix} \\ &\Leftrightarrow \begin{cases} 2\alpha_1 + 2\alpha_2 = 2 \\ 2\alpha_1 + 3\alpha_2 = 5 \\ \alpha_1 + \alpha_2 = 1 \end{cases} \end{aligned}$$

Logo (α_1, α_2) é solução do sistema cuja matriz ampliada é $[\, v_1 \ \ v_2 \mid b \,]$!

Determinação da combinação linear do exemplo (concl.)

- Aplicando a fase descendente do método de Gauss a $[\, v_1 \ \ v_2 \mid b \,]$, obtém-se

$$[\, v_1 \ \ v_2 \mid b \,] = \left[\begin{array}{cc|c} 2 & 2 & 2 \\ 2 & 3 & 5 \\ 1 & 1 & 1 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right].$$

- Como o sistema é possível podemos escrever b como CL de v_1 e v_2.
- Para determinarmos os coeficientes α_1 e α_2 da CL aplicamos a fase ascendente:

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right] \to \begin{cases} \alpha_1 = -2 \\ \alpha_2 = 3 \end{cases}$$

- Assim, $b = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix} = -2 \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = -2v_1 + 3v_2$

Observação

Vimos que o vetor $u = (-2, 3)$ dos coeficientes da CL é solução do sistema com matriz ampliada $[\, v_1 \ v_2 \mid b \,]$, ou seja, da equação matricial $Ax = b$ com $A = [\, v_1 \ v_2 \,]$. Isto significa que $b = Au$ e portanto multiplicar uma matriz A por um vetor u corresponde a efetuar a CL das colunas de A com coeficientes dados pelas componentes do vetor u.

Determinação de combinações lineares

Exemplo

Considerando $c = (0, 0, 1)$ e novamente os vetores $v_1 = (2, 2, 1)$ e $v_2 = (2, 3, 1)$ do exemplo do slide 95, tem-se que o sistema $Ax = c$ com $A = [\, v_1 \quad v_2 \,]$ é impossível (verifique). Logo c não é CL de v_1 e v_2.

Escrever b como CL de vetores $v_1, \ldots, v_n$ - resumo

Consideremos $b, v_1, \ldots, v_n \in \mathbb{R}^m$ e seja $A = [\, v_1 \quad v_2 \quad \cdots \quad v_n \,]$. Tem-se:

- Se $Ax = b$ for impossível então b não é CL de $v_1, \ldots, v_n$.
- Se $Ax = b$ for possível então b é CL de $v_1, \ldots, v_n$, tendo-se:
  - Se $Ax = b$ é PD então b escreve-se como CL de $v_1, \ldots, v_n$ de uma única forma.
  - Se $Ax = b$ for PI então b escreve-se como CL de $v_1, \ldots, v_n$ de infinitas maneiras distintas.

  Cada solução $u = (\alpha_1, \ldots, \alpha_n)$ de $Ax = b$ dá origem a uma CL $b = \alpha_1 v_1 + \cdots + \alpha_n v_n$, que podemos escrever como $b = Au$.

Espaço gerado e espaço das colunas

Espaço gerado por vetores e espaço das colunas de uma matriz

Sejam $v_1, \ldots, v_n \in \mathbb{R}^m$ e $A = [\, v_1 \; v_2 \; \cdots \; v_n \,]$.

- Chama-se *espaço gerado* por $v_1, \ldots, v_n$, e denota-se por $\langle v_1, \ldots, v_n \rangle$, ao subconjunto dos vetores de $\mathbb{R}^m$ que são CL de $v_1, \ldots, v_n$, isto é,

$$\begin{aligned} \langle v_1, \ldots, v_n \rangle &= \{ b \in \mathbb{R}^m \;:\; b \text{ é CL de } v_1, \ldots, v_n \} \\ &= \Big\{ b \in \mathbb{R}^m \colon Ax = b \text{ é possível} \Big\} \end{aligned}$$

- Chama-se *espaço das colunas* de A, e denota-se por $\mathcal{C}(A)$, ao espaço gerado pelos vetores que constituem as n colunas de A, isto é,

$$\mathcal{C}(A) = \langle v_1, \ldots, v_n \rangle = \Big\{ b \in \mathbb{R}^m \colon Ax = b \text{ é possível} \Big\},$$

  que define o segundo subespaço vetorial fundamental associado a uma matriz.

Observação

$b \in \langle v_1, \ldots, v_n \rangle = \mathcal{C}(A) \quad \Leftrightarrow \quad b$ é CL de $v_1, \ldots, v_n \quad \Leftrightarrow \quad Ax = b$ é possível.

Exemplo do slide 95 revisitado

Consideremos novamente os vetores $v_1 = (2, 2, 1)$ e $v_2 = (2, 3, 1)$ do slide 95 e seja $A = [\ v_1 \quad v_2\ ]$. Tem-se:

- $\langle v_1, v_2 \rangle = \mathcal{C}(A) = \{b \in \mathbb{R}^3 \ : \ Ax = b \text{ é possível}\}$ e obtém-se, aplicando o método de eliminação de Gauss à matriz $[A|b]$,

$$[A|b] = \left[\begin{array}{cc|c} 2 & 2 & b_1 \\ 2 & 3 & b_2 \\ 1 & 1 & b_3 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 1 & b_3 \\ 0 & 1 & b_2 - 2b_3 \\ 0 & 0 & b_1 - 2b_3 \end{array}\right] = [A'|b']$$

- Logo o sistema $Ax = b$ é possível sse $b_1 - 2b_3 = 0$ e portanto,

$$\langle v_1, v_2 \rangle = \mathcal{C}(A) = \{(b_1, b_2, b_3) \ : \ b_1 - 2b_3 = 0\},$$

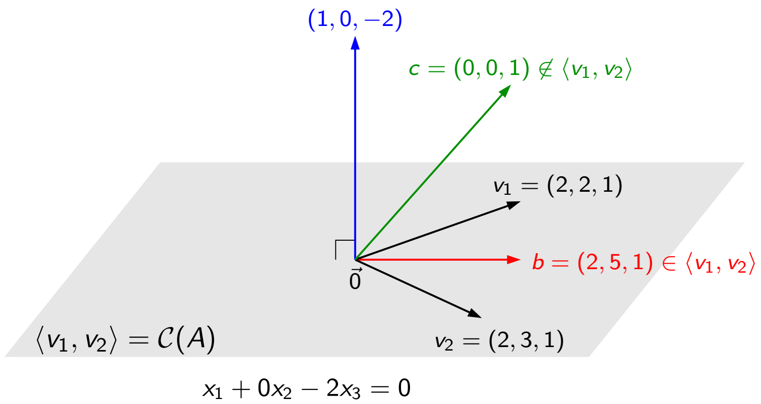
que define o plano de $\mathbb{R}^3$ de equação cartesiana $x_1 + 0x_2 - 2x_3 = 0$, que passa na origem e tem vetor normal $(1, 0, -2)$.

Exemplo do slide 95 revisitado - interpretação geométrica

Consideremos novamente os vetores $b = (2, 5, 1)$ do slide 95 e $c = (0, 0, 1)$ do slide 97.

- Vimos que $b = (2, 5, 1) = -2v_1 + 3v_2$ (ver o slide 95) é CL de v_1 e v_2 e portanto pertence ao espaço gerado $\langle v_1, v_2 \rangle = \mathcal{C}(A)$, o que se comprova pois b **satisfaz** a equação $x_1 + 0x_2 - 2x_3 = 0$ que define $\langle v_1, v_2 \rangle$.

- Vimos que $c = (0, 0, 1)$ do slide 97 não é CL de v_1 e v_2 e portanto não pertence ao espaço gerado $\langle v_1, v_2 \rangle = \mathcal{C}(A)$, o que se comprova uma vez que c **não satisfaz** a equação $x_1 + 0x_2 - 2x_3 = 0$ que define $\langle v_1, v_2 \rangle$.



Algoritmo para determinar o espaço gerado/espaço das colunas

O mesmo tipo de procedimento que foi aplicado no exemplo do slide 99 pode ser utilizado para determinar o espaço gerado/espaço das colunas $\langle v_1, \ldots, v_n \rangle = \mathcal{C}(A)$, com $A = [\, v_1 \ \cdots \ v_n \,]$ arbitrária.

Algoritmo

Aplica-se a fase descendente do método de Gauss a $[A|b]$, com $b = (b_1, \ldots, b_m)$ vetor genérico. Seja $[A'|b']$ matriz em escada obtida a partir de $[A|b]$. Tem-se:

- Se A' **não possui** linhas nulas então $v_1, \ldots, v_n$ **geram** $\mathbb{R}^m$, isto é,

$$\langle v_1, \ldots, v_n \rangle = \mathcal{C}(A) = \mathbb{R}^m.$$

- Se A' **possui** linhas nulas então $v_1, \ldots, v_n$ **não geram** $\mathbb{R}^m$ e obtém-se um sistema de equações definidoras para $\langle v_1, \ldots, v_n \rangle = \mathcal{C}(A)$,

$$\{(b_1, \ldots, b_m) : b'_{i_1} = 0,\ b'_{i_2} = 0,\ \cdots, b'_{i_k} = 0\},$$

onde $b'_{i_1}, \ldots, b'_{i_k}$ são as componentes do vetor b' associadas às linhas nulas da matriz em escada A'.

Espaço gerado/espaço das colunas é o subespaço maximal

Exercício na aula

Aplicando o algoritmo do slide 101 calcule

$$\langle v_1, v_2, v_3, v_4 \rangle = \langle (1,1,0), (1,0,1), (0,1,1), (1,1,1) \rangle.$$

Resolução

Seja $A = [v_1\ v_2\ v_3\ v_4]$. Tem-se

$$\langle v_1, v_2, v_3, v_4 \rangle = \mathcal{C}(A) = \{b = (b_1, b_2, b_3) : Ax = b \ \text{ é possível}\}.$$

Aplicando a fase descendente do método de eliminação de Gauss a $[A|b]$ obtém-se,

$$[A|b] = \left[\begin{array}{cccc|c} 1 & 1 & 0 & 1 & b_! \\ 1 & 0 & 1 & 1 & b_2 \\ 0 & 1 & 1 & 1 & b_3 \end{array}\right] \xrightarrow{L_2 - L_1} \left[\begin{array}{cccc|c} 1 & 1 & 0 & 1 & b_1 \\ 0 & -1 & 1 & 1 & b_2 - b_1 \\ 0 & 1 & 1 & 1 & b_3 \end{array}\right]$$

$$\xrightarrow{L_3 + L_2} \left[\begin{array}{cccc|c} 1 & 1 & 0 & 1 & b_1 \\ 0 & -1 & 1 & 1 & b_2 - b_1 \\ 0 & 0 & 2 & 1 & b_3 + b_2 - b1 \end{array}\right] = [A'|b'].$$

Como a matriz em escada A' não tem linhas nulas (1º caso do algoritmo do slide 101), não há restrições a impor ao vetor $b = (b_1, b_2, b_3)$ para o sistema ser possível.

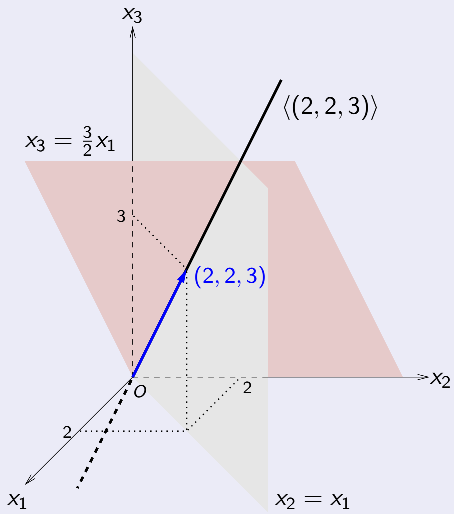
Logo $\langle v_1, v_2, v_3, v_4 \rangle = \mathcal{C}(A) = \mathbb{R}^3$, isto é, v_1, v_3, v_3, v_4 geram o subespaço maximal.

Espaço gerado/espaço das colunas é definido por equações

Exercício na aula

Considere o subespaço $\langle(2,2,3)\rangle = \{b : b = \alpha(2,2,3) \text{ com } \alpha \in \mathbb{R}\}$, que define a reta de $\mathbb{R}^3$ que passa na origem e tem vetor diretor $v = (2,2,3)$. Aplicando o algoritmo do slide 101 mostre que esta reta corresponde à intersecção dos planos abaixo:

$$\begin{cases} -x_1 + x_2 & = 0 \\ -\frac{3}{2}x_1 \quad + x_3 & = 0 \end{cases}$$



Resolução do exercício

Denotando por A a matriz com a única coluna $v = (2,2,3)$, $A = [\, v\,]$, tem-se

$$\begin{aligned} \langle(2,2,3)\rangle = \mathcal{C}(A) &= \{b = (b_1, b_2, b_3) \ : \ Ax = b \text{ é possível}\} \\ &= \{b = (b_1, b_2, b_3) \ : \ [A|b] \text{ é possível}\}. \end{aligned}$$

- Aplicando a fase descendente à matriz $[\,A|b\,]$ vem,

$$[A|b] = \left[\begin{array}{c|c} 2 & b_1 \\ 2 & b_2 \\ 3 & b_3 \end{array}\right] \xrightarrow[L_3 - \frac{3}{2}L_1]{L_2 - L_1} \left[\begin{array}{c|c} 2 & b_1 \\ 0 & b_2 - b_1 \\ 0 & b_3 - \frac{3}{2}b_1 \end{array}\right] = [A'|b'].$$

- Tem-se que o sistema $Ax = b$ é possível se e só se as componentes do vetor b' associadas às linhas nulas da matriz em escada A' forem nulas, isto é, $b'_2 = b_2 - b_1 = 0$ e $b'_3 = b_3 - \frac{3}{2}b_1 = 0$ (2º caso do algoritmo do slide 101). Logo,

$$\langle(2,2,3)\rangle = \mathcal{C}(A) = \left\{(b_1, b_2, b_3) \ : \ b_2 - b_1 = 0, \ b_3 - \frac{3}{2}b_1 = 0\right\},$$

que corrresponde à interseção dos planos $x_2 - x_1 = 0$ e $x_3 - \frac{3}{2}x_1 = 0$ como se pretendia mostrar.

Independência linear

Definição de independência linear

Sejam $v_1, \ldots, v_n \in \mathbb{R}^m$.

- $\{v_1, \ldots, v_n\}$ diz-se linearmente independente (l.i.) se

$$\forall \alpha_1, \ldots, \alpha_n \in \mathbb{R}\colon \alpha_1 v_1 + \cdots + \alpha_n v_n = \vec{0} \Rightarrow \alpha_1 = \cdots = \alpha_n = 0,$$

  isto é, se a combinação linear com **todos** os coeficientes nulos,

$$0 v_1 + 0 v_2 + \cdots + 0 v_n = \vec{0},$$

  for a **única** forma de escrever o vetor nulo como CL de $v_1, \ldots, v_n$.

- Caso contrário $\{v_1, \ldots, v_n\}$ diz-se linearmente dependente (l.d.).

  Por outras palavras, $\{v_1, \ldots, v_n\}$ é linearmente dependente se existem coeficientes **não todos nulos** $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tais que,

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \vec{0}.$$

(In)dependência linear - exemplos

- $\{(1,3,-1)\}$ é l.i. pois,

$$\alpha\,(1,3,-1) = (0,0,0) \;\Rightarrow\; (\alpha, 3\,\alpha\,, -\alpha) = (0,0,0) \;\Rightarrow\; \alpha = 0.$$

- $\{\vec{0}\}$ é l.d. pois $\alpha\,\vec{0} = \vec{0} \not\Rightarrow \alpha = 0$ (por exemplo, $2\,\vec{0} = \vec{0}$).

- $\{(1,3,-1),(2,6,-2)\}$ é l.d. uma vez que conseguimos obter o vetor nulo como CL dos vetores $(1,3,-1)$ e $(2,6,-2)$, com os coeficientes **não todos nulos**, por exemplo, como,

$$-2\,(1,3,-1) + 1\,(2,6,-2) = (0,0,0).$$

  Neste caso os vetores são múltiplos entre si, isto é, são colineares.

- $\{(1,3,-1),(0,1,5)\}$ é l.i. pois,

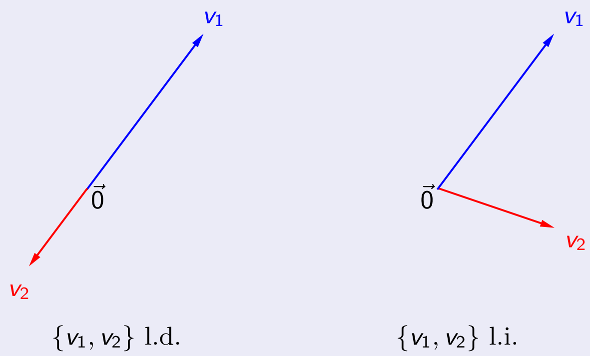
$$\alpha_1 \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} \alpha_1 = 0 \\ 3\,\alpha_1 + \alpha_2 = 0 \\ -\alpha_1 + \alpha_2 = 0 \end{cases} \Rightarrow \begin{cases} \alpha_1 = 0 \\ \alpha_2 = 0 \end{cases}$$

  Neste caso os vetores não são múltiplos entre si, isto é, são não colineares.

Caso dos conjuntos de cardinalidade ≤ 2

Independência linear de conjuntos com um e com dois vetores

- $\{\vec{v}\}$ é linearmente independente $\Leftrightarrow$ $v \neq \vec{0}$.
- $\{v_1, v_2\}$ é linearmente independente $\Leftrightarrow$ v_1 e v_2 são não colineares.



Decidir sobre a independência linear de conjuntos com mais que dois vetores não é, em geral, imediato. Por essa razão vamos começar por dar uma caracterização alternativa de conjunto linearmente independente.

Caracterização alternativa de (in)dependência linear

Teorema

Consideremos $v_1, \ldots, v_n \in \mathbb{R}^m$ e sejam $A = [\, v_1 \;\, v_2 \;\, \cdots \;\, v_n \,]$ e A' matriz em escada obtida a partir de A aplicando operações elementares. Tem-se que:

- $\{v_1, \ldots, v_n\}$ é linearmente **inde**pendente sse $Ax = \vec{0}$ for **de**terminado, isto é, <u>todas as colunas de A' tiverem pivot</u> ($\Leftrightarrow$ car$(A) = n$).
- $\{v_1, \ldots, v_n\}$ é linearmente **de**pendente sse $Ax = \vec{0}$ for **inde**terminado, isto é, <u>existirem as colunas sem pivot em A'</u> ($\Leftrightarrow$ car$(A) < n$).

  Neste caso tem-se $\mathcal{N}(A) \neq \{\vec{0}\}$ e cada solução $(\alpha_1, \ldots, \alpha_n) \in \mathcal{N}(A)$ origina a uma CL distinta,
  $$\alpha_1 v_1 + \cdots + \alpha_n v_n = \vec{0}.$$

O teorema decorre imediatamente da definição de independência linear e dos resultados do slide 97 com $b = \vec{0}$, que mostram que $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ verificam

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = \vec{0},$$

se e só se $(\alpha_1, \ldots, \alpha_n)$ é solução de $Ax = \vec{0}$, isto é, $(\alpha_1, \ldots, \alpha_n) \in \mathcal{N}(A)$.

(In)dependência linear via método de Gauss

A caracterização anterior permite usar o método de eliminação de Gauss para decidir a independência linear de conjuntos com n vetores em que n arbitrário.

Exercício na aula

Considere os vetores $v_1 = (1, \alpha, 1)$, $v_2 = (0, 1, -1)$ e $v_3 = (\alpha, 3, 3)$, com $\alpha \in \mathbb{R}$. Decida sobre a independência linear de $\{v_1, v_2, v_3\}$ em função de α.

Resolução

Consideremos a matriz $A = [v_1 \; v_2 \; v_3]$. Aplicando a fase descendente do método de eliminação de Gauss à matriz A vem,

$$A = \begin{bmatrix} 1 & 0 & \alpha \\ \alpha & 1 & 3 \\ 1 & -1 & 3 \end{bmatrix} \xrightarrow[L_3 - L_2]{L_2 - \alpha L_1} \begin{bmatrix} 1 & 0 & \alpha \\ 0 & 1 & 3-\alpha^2 \\ 0 & -1 & 3-\alpha \end{bmatrix}$$

$$\xrightarrow{L_3 + L_2} \begin{bmatrix} 1 & 0 & \alpha \\ 0 & 1 & 3-\alpha^2 \\ 0 & 0 & 6-\alpha-\alpha^2 \end{bmatrix} = A'.$$

Tem-se $\{v_1, v_2, v_3\}$ l.i. $\Leftrightarrow$ todas as colunas de A' têm pivot $\Leftrightarrow 6 - \alpha - \alpha^2 \neq 0 \Leftrightarrow$ $\alpha \neq \frac{1 \pm \sqrt{1-4(-1)6}}{2(-1)}$ (fórmula resolvente) $\Leftrightarrow \alpha \neq -3, 2$.
Daqui resulta ainda que $\{v_1, v_2, v_3\}$ l.d. $\Leftrightarrow \alpha = -3$ ou $\alpha = 2$.

Cardinalidade máxima de um conjunto l.i.

Da relação $\text{car}(A_{m\times n}) \leq \min\{m, n\}$ e da caracterização de independência linear de um conjunto $\{v_1, \ldots, v_n\}$ dada no slide 108,

$$\{v_1, \ldots, v_n\} \text{ l.i.} \quad \Leftrightarrow \quad \text{car}(A) = n,$$

onde $A = [v_1 \; v_2 \cdots v_n]$, deduz-se que se $\{v_1, \ldots, v_n\}$ é l.i. então $n \leq m$. Temos portanto o seguinte resultado.

Proposição

Um conjunto linearmente independente de vetores de $\mathbb{R}^m$ possui no máximo m vetores

- Por exemplo, o conjunto $\{v_1, v_2, v_3, v_4, v_5\}$, em que $v_1 = (1, 2, 1, 1)$, $v_2 = (0, 1, -1, 1)$, $v_3 = (1, 3, 1, 4)$, $v_4 = (0, 0, 1, 1)$ e $v_5 = (0, 0, 0, 1)$, é linearmente dependente pois é formado por 5 de vetores de $\mathbb{R}^4$.

Base para espaço nulo de uma matriz - exercício 2

Exercício na aula

Indicar uma base e a dimensão do espaço nulo da matriz

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -1 & -1 & -1 \end{bmatrix}.$$

Resolução: aplicando a fase descendente à matriz $[A|\vec{0}]$ obtém-se,

$$[A|\vec{0}] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ -1 & -1 & -1 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & -2 & 0 \end{array}\right] = [A'|\vec{0}].$$

Neste caso não há colunas sem pivot em A', isto é, não há variáveis livres. Logo $Ax = \vec{0}$ é determinado e portanto $\mathcal{N}(A) = \{\vec{0}\}$.

Logo $\{\}$ é a base de $\mathcal{N}(A)$, tendo-se $\dim \mathcal{N}(A) = 0$.

Dependência linear e combinações lineares

Vimos que um conjunto formado por 2 vetores era l.d. se e só se um dos vetores era múltiplo do outro. Vejamos o que se passa com conjuntos com 3 vetores.

Consideremos $u, v, w \in \mathbb{R}^m$.

- Se w é CL de u e v, isto é, $w = \alpha u + \beta v$ com $\alpha, \beta \in \mathbb{R}$, então $\{u, v, w\}$ é l.d. pois podemos escrever o vetor nulo como CL de u, v, w com coeficientes **não todos nulos**. De facto,

$$\alpha\, u + \beta\, v - 1\, w = \vec{0}.$$

- Reciprocamente se $\{u, v, w\}$ é l.d. então por definição existem α, β, γ, **não todos nulos** tais que

$$\alpha u + \beta v + \gamma w = \vec{0}.$$

Se $\alpha \neq 0$ podemos escrever $u = -\frac{\beta}{\alpha} v - \frac{\gamma}{\alpha} w$ e portanto u é CL de v e w. Analogamente se mostra que se $\beta \neq 0$ [$\gamma \neq 0$] então v [w] é CL dos restantes 2 vetores (que fica como exercício para os alunos).

Logo $\{u, v, w\}$ é l.d. se e só se um dos vetores é CL dos restantes 2 vetores.

Aplicando o mesmo tipo de ideias a conjuntos com um número arbitrário de vetores obtemos a caracterização de dependência linear do próximo slide.

Noção mais intuitiva do conceito de dependência linear

Teorema (Caracterização da dependência linear em termos de CL)

Um conjunto com dois ou mais vetores é linearmente dependente se e só se pelo menos um dos vetores do conjunto for combinação linear dos restantes vetores do conjunto.

Consequências

- Um conjunto de vetores que contenha o vetor nulo é l.d.
- Um conjunto de vetores que contenha um conjunto l.d. de vetores é l.d. Em particular, se o conjunto contiver vetores colineares é l.d.
- Reciprocamente, um conjunto de vetores não vazio que esteja contido um conjunto linearmente independente ainda é linearmente independente.

Exemplos

- $\{v_1, v_2, v_3\} = \{(1, 2, 3, 4), (1, 1, 1, 1), (2, 3, 4, 5)\}$ é l.d. pois $v_3 = v_1 + v_2$ e logo $\{v_1, v_2, v_3, v_4\}$ é tb l.d. para todo o $v_4 \in \mathbb{R}^4$, uma vez que $v_3 = v_1 + v_2 + 0v_4$.
- $\{(1, 1, 1, 1), (1, 1, 1, 0), (1, 1, 0, 0), (1, 0, 0, 0)\}$ é l.i. (verifique). Logo o subconjunto $\{(1, 1, 1, 1), (1, 1, 1, 0), (1, 1, 0, 0)\}$ é também l.i.

TPC: verifique que no 1º caso $(1, 1, -1, 0) \in \mathcal{N}(A)$ com $A = [v_1\ v_2\ v_3\ v_4]$. Porquê?

Base de um subespaço vetorial

Vamos dar agora aquele que é, possivelmente, o **conceito mais central em Álgebra Linear**.

Intuitivamente uma base de um subespaço vetorial V é um subconjunto de vetores de V tal que: (i) não contém vetores "redundantes" no sentido em que nenhum dos vetores da base se pode obter como CL dos restantes vetores da base; (ii) todo o vetor do subespaço V é CL linear dos vetores da base.

Mais precisamente, temos a seguinte definição.

Definição de base

Sejam V subespaço vetorial de $\mathbb{R}^m$ e $\underline{v_1, \ldots, v_k \in V}$. Diz-se que $\{v_1, \ldots, v_k\}$ é uma **base** de V se verificar as condições:

(i) $\{v_1, \ldots, v_k\}$ é linearmente independente.

(ii) $\langle v_1, \ldots, v_k \rangle = V$, isto é, $v_1, \ldots, v_k$ geram V.

Exemplo fundamental: base canónica de $\mathbb{R}^m$

Vejamos que $\{e_1, e_2, e_3\}$ é base de $\mathbb{R}^3$, onde $e_1 = (1,0,0)$, $e_2 = (0,1,0)$ e $e_3 = (0,0,1)$ são as colunas da matriz identidade de ordem 3:

(i) $\{e_1, e_2, e_3\}$ é linearmente independente.
De facto, a matriz $A = [\, e_1 \;\; e_2 \;\; e_3 \,]$ é a matriz identidade de ordem 3, I_3, que já está em escada e possui todas as colunas com pivot.

(ii) $\langle e_1, e_2, e_3 \rangle = \mathbb{R}^3$.
De facto, tem-se para qualquer $b = (b_1, b_2, b_3) \in \mathbb{R}^3$,

$$\begin{aligned} (b_1, b_2, b_3) &= b_1(1,0,0) + b_2(0,1,0) + b_3(0,0,1) \\ &= b_1 e_1 + b_2 e_2 + b_3 e_3. \end{aligned}$$

Logo qualquer $b \in \mathbb{R}^3$ é CL de e_1, e_2 e e_3 e portanto $\langle e_1, e_2, e_3 \rangle = \mathbb{R}^3$.

A base anterior generaliza-se para $\mathbb{R}^m$ com m arbitrário.

Base canónica de $\mathbb{R}^m$

O conjunto $\{e_1, e_2, \ldots, e_m\}$ formado pelas m colunas da matriz identidade I_m,

$$e_1 = (1,0,\ldots,0), \;\; e_2 = (0,1,\ldots,0), \;\; \ldots, e_m = (0,0,\ldots,1),$$

constitui uma base de $\mathbb{R}^m$ que se designa por base canónica (b.c.).

Dimensão de um subespaço vetorial

Teorema-definição

Todas as bases de um mesmo subespaço vetorial V possuem o mesmo número de vetores a que chamamos dimensão de V e denotamos por $\dim V$.

Dimensão do subespaço minimal

- Convenciona-se que $\{\}$ é a base do subespaço mininal $\{\vec{0}\}$. Uma vez que esta base não possui vetores tem-se,

$$\boxed{\dim \{\vec{0}\} = 0}$$

Dimensão do subespaço maximal

- Vimos que $\mathbb{R}^m$ admitia uma base especial, dita base canónica, constituída pelas m colunas da matriz identidade de ordem m. Daqui resulta que,

$$\boxed{\dim \mathbb{R}^m = m}$$

- Logo qualquer outra base para $\mathbb{R}^m$ também possui m vetores.

Caracterização das bases do subespaço maximal

Sejam $v_1, \ldots, v_m \in \mathbb{R}^m$ e consideremos $A = [v_1 \ v_2 \ \cdots \ v_m]$ que é uma matriz quadrada de ordem m. Seja A' matriz em escada obtida a partir de A aplicando operações elementares. Uma vez que A' é também uma matriz **quadrada** obtêm-se, aplicando os resultados dos slides 108 e 101, as seguintes equivalências:

$$\begin{aligned} \{v_1, \ldots, v_m\} \text{ lin. indep.} &\Leftrightarrow \text{ todas as colunas de } A' \text{ têm pivot} \\ &\Leftrightarrow \text{ não há linhas nulas em } A' \\ &\Leftrightarrow \langle v_1, \ldots, v_m \rangle = \mathcal{C}(A) = \mathbb{R}^m. \end{aligned}$$

Logo por definição de base provámos o seguinte resultado.

Teorema (Critério para definir base do subespaço maximal $\mathbb{R}^m$)

Sejam $v_1, \ldots, v_m \in \mathbb{R}^m$. As seguintes afirmações são equivalentes:

- $\{v_1, \ldots, v_m\}$ é base de $\mathbb{R}^m$.
- $\{v_1, \ldots, v_m\}$ é linearmente independente.
- $\langle v_1, \ldots, v_m \rangle = \mathbb{R}^m$.

Bases de subespaços maximais - exemplo

Pelo teorema do slide anterior e pelo resultado do slide 110 tem-se o seguinte resultado.

Corolário

As bases de $\mathbb{R}^m$ são os conjuntos linearmente independentes com m vetores, ou seja, os conjuntos lin. indep. de cardinalidade máxima([10]).

Exemplo na aula

Quais dos seguintes conjuntos são lin. indep. / geram $\mathbb{R}^3$ / base de $\mathbb{R}^3$?

1. $\{(1,0,0),(2,5,0)\}$. (Sim / Não / Não)
2. $\{(1,0,0),(2,5,0),(3,5,0)\}$. (Não / Não / Não)
3. $\{(1,0,0),(2,5,0),(3,5,9)\}$. (Sim / Sim / Sim)
4. $\{(1,2,3),(4,5,6),(7,8,9),(10,11,12)\}$. (Não / Sim / Não)

[10] E são também os conjuntos de geradores de $\mathbb{R}^m$ de cardinalidade mínima.

Construção de bases para subespaços vetoriais

Um subespaço vetorial V pode ser definido de duas formas distintas:

- *Como CS de um sistema de equações lineares homogéneas* / *espaço nulo de uma matriz*. Por exemplo,
  - $V = \{(x_1, x_2, x_3) : x_1 - x_2 + 2x_3 = 0,\ 3x_1 + x_2 + x_3 = 0, 2x_1 + 6x_2 = 0\}$,

    ou seja, $\quad V = \mathcal{N}(A), \quad \text{com } A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 1 \\ 2 & 6 & 0 \end{bmatrix}.$

- *Gerado por um conjunto de vetores* / *espaço das colunas de uma matriz*. Por exemplo,
  - $V = \langle (1,1,0,1), (1,2,-1,0), (1,0,1,2), (2,1,1,3) \rangle$,

    ou seja, $\quad V = \mathcal{C}(A), \quad \text{com } A = \begin{bmatrix} 1 & 1 & 1 & 2 \\ 1 & 2 & 0 & 1 \\ 0 & -1 & 1 & 1 \\ 1 & 0 & 2 & 3 \end{bmatrix}.$

Base para o espaço nulo de uma matriz - exercício

Exercício na aula

Indicar uma base e a dimensão do espaço nulo da matriz

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

TPC

Determinar uma base do hiperplano de $\mathbb{R}^4$,

$$V = \{(x_1, x_2, x_3, x_4) : x_1 - 2x_2 + x_3 - 5x_4 = 0\}.$$

Resolução do exercício do slide anterior

Reduzindo a matriz $[\,A\,|\,\vec{0}\,]$ obtém-se (verifique),

$$[\,A\,|\,\vec{0}\,] \;=\; \left[\begin{array}{rrrr|r} 1 & 0 & 2 & 4 & 0 \\ 1 & 2 & 0 & 2 & 0 \\ -1 & -1 & -1 & -3 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{rrrr|r} 1 & 0 & 2 & 4 & 0 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\rightsquigarrow \left\{\begin{array}{rcl} x_1 \qquad 2x_3 + 4x_4 & = & 0 \\ x_2 - x_3 - \;\; x_4 & = & 0 \\ 0 & = & 0 \end{array}\right. .$$

Tem-se que $\mathcal{N}(A) \neq \{\vec{0}\}$ uma vez que existem variáveis livres (x_3 e x_4), obtendo-se,

$$\begin{aligned} \mathcal{N}(A) &= \{(x_1, x_2, x_3, x_4) : x_1 = -2\,x_3 - 4\,x_4,\; x_2 = x_3 + x_4,\; x_3,\; x_4 \in \mathbb{R}\} \\ &= \{(-2\,x_3 - 4\,x_4, x_3 + x_4, x_3, x_4) :\; x_3,\; x_4 \in \mathbb{R}\} \\ &= \{(-2\,x_3, x_3, x_3, 0) + (-4\,x_4, x_4, 0, x_4) :\; x_3,\; x_4 \in \mathbb{R}\} \\ &= \underbrace{\{x_3(-2,1,1,0) + x_4(-4,1,0,1) :\; x_3,\; x_4 \in \mathbb{R}\}}_{\text{todas as somas de múltiplos de } (-2,1,1,0) \text{ e } (-4,1,0,1)} \\ &= \langle(-2,1,1,0), (-4,1,0,1)\rangle = \langle v_1, v_2\rangle \end{aligned}$$

Logo $\mathcal{N}(A) = \langle v_1, v_2\rangle$. Como v_1 e v_2 não são múltiplos entre si, $\{v_1, v_2\}$ é lin. indep.
Logo por definição $\{v_1, v_2\}$ é base de $\mathcal{N}(A)$ e $\dim \mathcal{N}(A) =$ nº de vetores da base $= 2$.

Observações

- O processo descrito no slide anterior para determinar uma base de $\mathcal{N}(A)$, pode ser generalizado para uma matriz A arbitrária (desde que o sistema $Ax = \vec{0}$ possua variáveis livres) e conduz sempre a bases de $\mathcal{N}(A)$, não sendo necessário provar que o conjunto é linearmente independente.

- O primeiro vetor da base do slide anterior, $v_1 = (-2, 1, 1, 0)$, corresponde à solução do sistema $Ax = \vec{0}$ considerando a variável livre $x_3 = 1$ e a variável livre $x_4 = 0$:

$$(-2\,x_3 - 4\,x_4, x_3 + x_4, x_3, x_4) \xrightarrow{\substack{x_3 = 1 \\ x_4 = 0}} (-2, 1, 1, 0) = v_1.$$

Analogamente, o segundo vetor da base, $v_2 = (-4, 1, 0, 1)$, corrresponde à solução do sistema $Ax = \vec{0}$ com $x_3 = 0$ e $x_4 = 1$:

$$(-2\,x_3 - 4\,x_4, x_3 + x_4, x_3, x_4) \xrightarrow{\substack{x_3 = 0 \\ x_4 = 1}} (-4, 1, 0, 1) = v_2.$$

Base para o espaço nulo de uma matriz - algoritmo

Algoritmo

`Input:` Matriz A do tipo $m \times n$.

`Objectivo:` Base para $\mathcal{N}(A)$.

- Resolver o sistema $Ax = \vec{0}$ aplicando o método de Gauss a $[\,A\,|\,\vec{0}\,]$.
  Seja k o número de variáveis livres do sistema.
- Se $k = 0$, isto é, se não há variáveis livres então $\mathcal{N}(A) = \{\vec{0}\}$ e $\{\}$ é a base de $\mathcal{N}(A)$, tendo-se $\dim \mathcal{N}(A) = 0$.
- Se $k > 0$, associamos alternadamente a cada variável livre a solução do sistema em que essa variável livre toma o valor 1 (ou qualquer valor não nulo) e as restantes variáveis livres o valor zero.

  O conjunto das k soluções de $Ax = \vec{0}$ obtidas deste modo constitui uma base para $\mathcal{N}(A)$.

Em particular,

$$\boxed{\dim \mathcal{N}(A) = \text{nº de variáveis livres} = n - \text{car}(A)}$$

Base para espaço nulo de uma matriz - exercício 2

Exercício na aula

Indicar uma base e a dimensão do espaço nulo da matriz

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -1 & -1 & -1 \end{bmatrix}.$$

Resolução: aplicando a fase descendente à matriz $[A|\vec{0}]$ obtém-se,

$$[A|\vec{0}] = \left[\begin{array}{rrr|r} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ -1 & -1 & -1 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{rrr|r} 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & -2 & 0 \end{array}\right] = [A'|\vec{0}].$$

Neste caso não há colunas sem pivot em A', isto é, não há variáveis livres. Logo $Ax = \vec{0}$ é determinado e portanto $\mathcal{N}(A) = \{\vec{0}\}$.
Logo $\{\}$ é a base de $\mathcal{N}(A)$, tendo-se $\dim \mathcal{N}(A) = 0$.

Base para o espaço de colunas - exercício

Vamos agora ver como se podem determinar bases para o espaço de colunas de uma matriz.

Exercício na aula

Indicar uma base e a dimensão para o espaço nulo das colunas da matriz

$$A = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 1 & 2 & 0 & 2 \\ -1 & -2 & -1 & -3 \end{bmatrix} = [\, v_1 \;\; v_2 \;\; v_3 \;\; v_4 \,].$$

Resolução: vamos começar por determinar $\mathcal{C}(A)$. Aplicando a fase descendente do método de Gauss a $[\, A \mid b \,] = [\, v_1 \; v_2 \; v_3 \; v_4 \mid b \,]$ vem,

$$\left[\begin{array}{cccc|c} 1 & 2 & 2 & 4 & b_1 \\ 1 & 2 & 0 & 2 & b_2 \\ -1 & -2 & -1 & -3 & b_3 \end{array}\right] \to \cdots \to \left[\begin{array}{cccc|c} 1 & 0 & 2 & 4 & b_1 \\ 0 & 0 & 1 & 1 & b_1 + b_3 \\ 0 & 0 & 0 & 0 & b_1 + b_2 + 2b_3 \end{array}\right].$$

Logo para o sistema $Ax = b$ ser possível, $b_1 + b_2 + 2b_3 = 0$ e portanto

$$\mathcal{C}(A) = \langle v_1, v_2, v_3, v_4 \rangle = \{ b = (b_1, b_2, b_3) \, : \, b_1 + b_2 + 2b_3 = 0 \}.$$

Base para o espaço de colunas - exercício (concl.)

Observação

- A sequência efetuada de operações elementares do método de Gauss apenas depende das colunas que estão associadas às colunas com pivot na matriz em escada.
- As colunas sem pivot em A' não têm influência na discussão do sistema em escada $[\, A' \mid b' \,]$.

De facto, aplicando a fase descendente apenas aos vetores que estão associados às colunas com pivot em A', vem (confirme),

$$[\, v_1 \;\; v_3 \mid b \,] = \left[\begin{array}{cc|c} 1 & 2 & b_1 \\ 1 & 0 & b_2 \\ -1 & -1 & b_3 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 2 & b_1 \\ 0 & 1 & b_1 + b_3 \\ 0 & 0 & b_1 + b_2 + 2b_3 \end{array}\right],$$

e portanto,

$$\langle v_1, v_3 \rangle = \{ b = (b_1, b_2, b_3) \, : \, b_1 + b_2 + 2b_3 = 0 \} = \mathcal{C}(A) = \langle v_1, v_2, v_3, v_4 \rangle.$$

Logo os vetores associados às colunas sem pivot em A', v_2 e v_4, são redundantes. Como $= \langle v_1, v_3 \rangle = \mathcal{C}(A)$ e $\{v_1, v_3\}$ é l.i. (porque estão associados às colunas com pivot em A'), $\{v_1, v_3\}$ é uma base de $\mathcal{C}(A)$ (contida no conjunto inicial de geradores). Em particular, $\dim \mathcal{C}(A) =$ nº pivots em $A' = 2$.

Observações

- Mais geralmente, pode-se mostrar que dados $v_1, \ldots, v_n \in \mathbb{R}^m$ e $A = [v_1 \; v_2 \cdots \; v_n] \to A'$ (escada), com $\text{car}(A) = k$, se tem,

$$\mathcal{C}(A) = \langle v_1, v_2, \ldots, v_n \rangle = \langle v_{i_1}, v_{i_2}, \ldots, v_{i_k} \rangle,$$

  onde $v_{i_1}, v_{i_2}, \ldots, v_{i_k}$ são as k colunas de A associadas às colunas com pivot em A', isto é, que as colunas de A que estão associadas às colunas sem pivot em A' **não são necessárias para gerarem** $\mathcal{C}(A)$ (são vetores redundantes).

- Por outro lado, $\{v_{i_1}, v_{i_2}, \ldots, v_{i_k}\}$ é linearmente independente, pois é constituído por vetores que estão associados a colunas com pivot na matriz em escada A'.

- Das considerações anteriores e da definição de base do slide 114 deduz-se o algoritmo do próximo slide.

Base para o espaço das colunas/espaço gerado - algoritmo

Algoritmo

`Input:` $A = [\, v_1 \quad \cdots \quad v_n \,]$ com $v_1 \ldots, v_n \in \mathbb{R}^m$.

`Objectivo:` Base para $\mathcal{C}(A) = \langle v_1, \ldots, v_n \rangle$.

- Aplicar a fase descendente do método de Gauss à matriz A:
$$A \to \cdots \to A' \text{ com } A' \text{ escada.}$$
- O subconjunto das colunas de A que correspondem às colunas com pivot em A' constitui uma base de $\mathcal{C}(A) = \langle v_1, \ldots, v_n \rangle$, contida no conjunto inicial de geradores $v_1, \ldots, v_n$.

Em particular, tem-se

$$\boxed{\dim \langle v_1, \ldots, v_n \rangle = \dim \mathcal{C}(A) = \text{número de pivots em } A' = \text{car(A)}.}$$

Obs: a característica de uma matriz A é muitas vezes definida como $\dim \mathcal{C}(A)$.

Exercício do slide 125 revisitado

- Aplicando o algoritmo do slide anterior à matriz $A = [\, v_1 \;\; v_2 \;\; v_3 \;\; v_4 \,]$ do exercício do slide 125 (não é necessário ampliar com o vetor genérico b) vem,

$$A = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 1 & 2 & 0 & 2 \\ -1 & -2 & -1 & -3 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & 2 & 4 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = A',$$

  donde se deduz que $\{v_1, v_3\} = \{(1,1,-1),(2,0,-1)\}$ é uma base de $\mathcal{C}(A) = \langle v_1, v_2, v_3, v_4 \rangle$, pois são as colunas de A que estão associadas às colunas com pivot em A', tendo-se $\dim \mathcal{C}(A) = \mathrm{car}(A) = 2$. Note-se que a base anterior está contida no conjunto inicial de geradores v_1, v_2, v_3, v_4.

- Alternativamente, viu-se que $\mathcal{C}(A) = \{(b_1, b_2, b_3) : b_1 + b_2 + 2\, b_3 = 0\}$ e portanto $\mathcal{C}(A)$ é o CS da equação definidora $b_1 + b_2 + 2\, b_3 = 0$.

  Logo pode-se escrever $\mathcal{C}(A) = \mathcal{N}(B)$, onde $B = [1 \;\; 1 \;\; 2]$ é matriz do sistema homogéneo cuja a única equação é $b_1 + b_2 + 2\, b_3 = 0$.

  Aplicando o algoritmo da base para o espaço nulo do slide 123 à matriz B deduz-se a nova base de $\mathcal{C}(A)$, $\{(-1,1,0),(-2,0,1)\}$ (verifique), que já não está contida no conjunto inicial de geradores v_1, v_2, v_3, v_4.

Relação entre as dimensões de $\mathcal{N}(A)$ e de $\mathcal{C}(A)$

- Seja A matriz do tipo $m \times n$ e A' matriz em escada obtida a partir de A. Pelos resultados do slides 123 e 128 tem-se:
  - $\dim \mathcal{N}(A) = n - \mathrm{car}(A)$ (nº de colunas sem pivot em A').
  - $\dim \mathcal{C}(A) = \mathrm{car}(A)$ (nº de colunas com pivot em A').

- Daqui resulta imediatamente a seguinte resultado que estabelece uma relação importante entre as dimensões dos dois subespaços fundamentais associados à matriz A.

Teorema

Se A é uma matriz do tipo $m \times n$ tem-se

$$\boxed{\dim \mathcal{N}(A) + \dim \mathcal{C}(A) = \text{número de colunas de } A = n.}$$

Subespaço vetorial e dimensão

- O conhecimento da dimensão de um subespaço vetorial permite conhecer o tipo de conjunto que esse subespaço vetorial define
- Para os subespaços vetoriais do plano ($\mathbb{R}^2$) e do espaço ($\mathbb{R}^3$), tem-se o seguinte.

	dimensão do subespaço	tipo de subespaço vetorial
$\mathbb{R}^2$	0	$\{\vec{0}\}$
	1	reta que passa na origem
	2	$\mathbb{R}^2$
$\mathbb{R}^3$	0	$\{\vec{0}\}$
	1	reta que passa na origem
	2	plano que passa na origem
	3	$\mathbb{R}^3$

Têm-se ainda as seguintes caracterizações dos subespaços minimal e maximal de $\mathbb{R}^m$ com m arbitrário, em função das suas dimensões:

- $V = \{\vec{0}\} \;\Leftrightarrow\; \dim V = 0$.
- $V = \mathbb{R}^m \;\Leftrightarrow\; \dim V = m$.

Vetor pertence ao espaço nulo / espaço das colunas de uma matriz

Recordatória

Sejam $A_{m\times n}$, $u \in \mathbb{R}^n$ e $b \in \mathbb{R}^m$. Então:

- $u \in \mathcal{N}(A) \quad \Leftrightarrow \quad Au = \vec{0}$.
- $b \in \mathcal{C}(A) \quad \Leftrightarrow \quad [\,A\,|\,b\,]$ é possível.

Exemplo

Seja $A = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & -1 & -1 \\ 2 & 1 & 1 & 3 \end{bmatrix}_{3\times 4}$.

- Vejamos que $u = (-2, 1, 0, 1) \in \mathcal{N}(A)$. De facto,

$$Au = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & -1 & -1 \\ 2 & 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \vec{0}.$$

- Vejamos que $b = (1, -1, 5) \in \mathcal{C}(A)$. De facto,

$$[A|b] = \left[\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 1 \\ 0 & 1 & -1 & -1 & -1 \\ 2 & 1 & 1 & 3 & 5 \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 1 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] = [A'|b']$$

é possível.

Critério para definir base de um subespaço vetorial

- Vimos anteriormente que as bases de $\mathbb{R}^m$ (cuja dimensão é m) são os conjuntos linearmente independententes formados por m vetores de $\mathbb{R}^m$ (conjuntos l.i. de vetores de $\mathbb{R}^m$ de cardinalidade **máxima**)
- Temos uma caracterização análoga para qualquer subespaço vetorial V cuja dimensão se conhece!

Teorema

As bases de um subespaço vetorial V de dimensão $k > 0$ são os conjuntos linearmente independentes formados por k vetores de V.

(Isto é, conjuntos l.i. de vetores de V de cardinalidade **máxima**).
Nos exercícios podemos aplicar o teorema anterior com a seguinte formulação.

Teorema (Critério para definir base de V)

Sejam V subespaço vetorial de $\mathbb{R}^m$ e $v_1, \ldots, v_k \in \mathbb{R}^m$. Tem-se que $\{v_1, \ldots, v_k\}$ é **base de** V se e só se verificar as seguintes 3 condições:

- $v_1, \ldots, v_k \in V$.
- $\{v_1, \ldots, v_k\}$ é linearmente independente.
- $\dim V = k$ (nº de vetores do conjunto).

Critério para definir base de um subespaço vetorial - exercício

Exercício na aula

Considere $v_1 = (-2, 1, 0, 1)$, $v_2 = (-1, 0, -1, 1)$ e a matriz do exemplo do slide 132,

$$A = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & -1 & -1 \\ 2 & 1 & 1 & 3 \end{bmatrix}_{3\times 4}.$$

Mostre que $\{v_1, v_2\}$ é base de $\mathcal{N}(A)$.

Pelo critério do slide 133 basta verificar as seguintes condições:

- $v_1, v_2 \in \mathcal{N}(A)$. De facto, tem-se $Av_1 = \vec{0}$ e $Av_2 = \vec{0}$ (confirme).
- $\{v_1, v_2\}$ é linearmente independente. De facto, v_1 e v_2 são não colineares.
- $\dim \mathcal{N}(A) = 2$ (nº de vetores do conjunto). De facto, a matriz em escada A' obtida a partir de A tem 2 colunas sem pivot (confirme).

Logo $\{v_1, v_2\}$ é base de $\mathcal{N}(A)$.

Componentes de um vetor numa base de um subespaço

Teorema (Representação única na base de um subespaço)

Seja V um subespaço vetorial de $\mathbb{R}^m$ e $\mathfrak{B} = \{v_1, \ldots, v_k\}$ uma base de V. Para todo o $b \in V$ existem escalares únicos, $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tais que

$$b = \alpha_1 v_1 + \cdots + \alpha_k v_k. \qquad (2)$$

Os escalares $\alpha_1, \ldots, \alpha_k$ designam-se por componentes de b na base $\mathfrak{B}$.

Observação

O vetor $u = (\alpha_1, \ldots, \alpha_k)$ das componentes de b que verificam a relação (2) relativamente à base $\mathfrak{B}$ de V (assumindo esta base ordenada), é a solução única do sistema **PD** $Ax = b$, com $A = [\, v_1 \; v_2 \; \cdots \; v_k \,]$, isto é, verifica $Au = b$, e pode ser obtido reduzindo a matriz $[\, v_1 \; v_2 \; \cdots \; v_k \,|\, b \,]$.

De facto, por definição de base (slide 114) e pelo critério do slide 108,
ii) $b \in V = \langle v_1, \ldots, v_k \rangle = \mathcal{C}(A)$ e portanto o sistema $Ax = b$ é **possível**;
i) $\{v_1, \ldots, v_k\}$ é l.i. logo $\text{car}(A) = k$ e portanto $Ax = b$ é **determinado**.

Componentes de um vetor numa base de um subespaço - exemplos

Exemplos

- O vetor das componentes de $b = (1, 4, 2)$ relativamente à base canónica de $\mathbb{R}^3$, $\{e_1, e_2, e_3\}$, é o próprio vetor $(1, 4, 2)$ pois,

$$(1, 4, 2) = 1\,(1, 0, 0) + 4\,(0, 1, 0) + 2\,(0, 0, 1) = 1\,e_1 + 4\,e_2 + 2\,e_3.$$

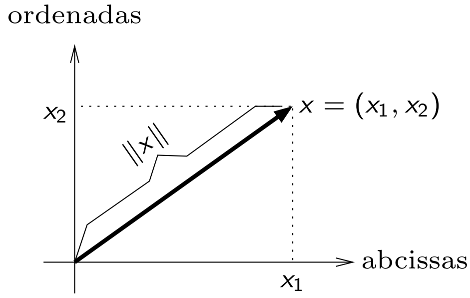
- O vetor das componentes de $b = (1, 4, 2)$ relativamente à base de $\mathbb{R}^3$, $\{v_1, v_2, v_3\} = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$, é $(-3, 2, 2)$ pois,

$$(1, 4, 2) = -3\,(1, 0, 0) + 2\,(1, 1, 0) + 2\,(1, 1, 1) = -3\,v_1 + 2\,v_2 + 2\,v_3.$$

  O vetor $(-3, 2, 2)$ corresponde à solução (única) de $Ax = b$ com $A = [\, v_1 \; v_2 \; v_3 \,]$, e é calculado reduzindo a matriz $[\, v_1 \; v_2 \; v_3 \,|\, b \,]$.

TPC: determine as componentes do vetor genérico $b = (b_1, b_2, b_3)$ na base canónica (ver o slide 115) e na base $\{v_1, v_2, v_3\}$ anterior.

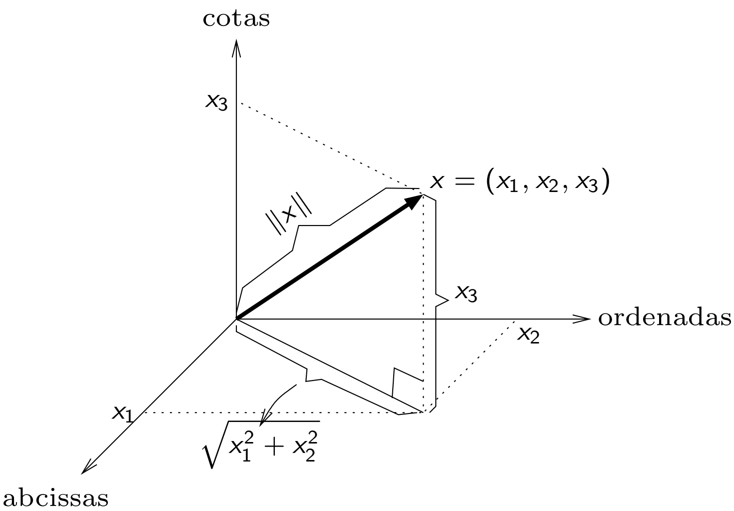
Recordatória: norma (comprimento) de um vetor do plano



$\|x\|$ representa a norma ou comprimento do vetor x, ou seja, a distância do vetor à origem. Pelo teorema de Pitágoras obtém-se,

$$\|x\| = \|(x_1, x_2)\| = \sqrt{x_1^2 + x_2^2} = \sqrt{(x_1, x_2) \cdot (x_1, x_2)} = \sqrt{x \cdot x} = \sqrt{x^T x}$$

Norma de um vetor do espaço



Analogamente, tem-se pelo teorema de Pitágoras,

$$\|x\| = \|(x_1, x_2, x_3)\| = \sqrt{\left(\sqrt{x_1^2 + x_2^2}\right)^2 + x_3^2} = \sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{x \cdot x}$$

Norma de um vetor de $\mathbb{R}^n$

Definição de norma (caso geral)

Dado $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ define-se a sua norma (comprimento) por,

$$\boxed{\|x\| = \sqrt{x^T x} = \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}}$$

Ex.: $\|(4,2,-1,2)\| = \sqrt{(4,2,-1,2)\cdot(4,2,-1,2)} = \sqrt{4^2+2^2+(-1)^2+2^2} = 5.$

Propriedades da norma

Para qualquer $x \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$ tem-se

1. $\|x\| \geq 0$.
2. $\|x\| = 0$ se e só se $x = \vec{0}$.
3. $\|\lambda\, x\| = |\lambda|\, \|x\|$.
4. $\|x\|^2 = x \cdot x = x^T x$.

Dem (do ponto 3): $\|\lambda x\| = \sqrt{(\lambda x)^T \lambda x} = \sqrt{\lambda\, x^T \lambda\, x} = \sqrt{\lambda^2}\sqrt{x^T x} = |\lambda| \|x\|$.

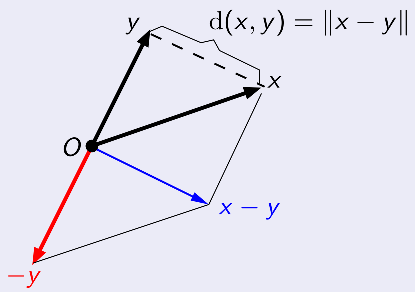
Distância (euclideana) entre vetores de $\mathbb{R}^n$

A partir da norma define-se a *distância (euclideana)* entre vetores de $\mathbb{R}^n$.

Definição de distância euclideana

Dados $x, y \in \mathbb{R}^n$, $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$, define-se a distância entre x e y por,

$$\boxed{d(x,y) = \|x - y\| = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + \cdots + (x_n-y_n)^2}.}$$



Por exemplo,

$$\begin{aligned} d((1,3,2,1),(1,4,3,-1)) &= \|(1,3,2,1) - (1,4,3,-1)\| \\ &= \|(0,-1,-1,2)\| = \sqrt{6}. \end{aligned}$$

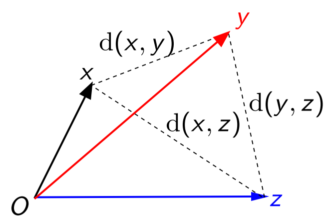
Propriedades da distância

Propriedades da distância

Dados $x, y, z \in \mathbb{R}^n$, tem-se

1. $\mathrm{d}(x, y) \geq 0$.
2. $\mathrm{d}(x, y) = 0 \iff x = y$.
3. $\mathrm{d}(x, y) = \mathrm{d}(y, x)$.
4. $\mathrm{d}(x, z) \leq \mathrm{d}(x, y) + \mathrm{d}(y, z)$.

A propriedade 4. designa-se por desigualdade triangular e significa que o comprimento do lado de qualquer triângulo é inferior ou igual à soma dos comprimentos dos outros dois lados.



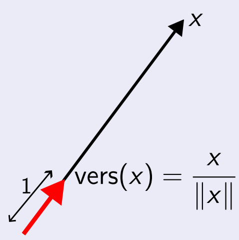
A 3 primeiras propriedades significam que $d(x, y)$ define uma dissemelhança.
Uma distância é uma dissemelhança que verifica ainda a propriedade triangular.

Vetor unitário e versor

Definições de vetor unitário e versor

- $x \in \mathbb{R}^n$ diz-se *unitário* se $\|x\| = 1$, isto é, se $x \cdot x = \|x\|^2 = 1$
- A cada vetor $x \in \mathbb{R}^n$, $x \neq \vec{0}$, associamos o **único** vetor unitário com a mesma direção e sentido que x, designado versor de x,

$$\mathrm{vers}(x) = \frac{x}{\|x\|}.$$



Por exemplo, $\mathrm{vers}(3, 4) = \frac{(3,4)}{\|(3,4)\|} = \frac{(3,4)}{5} = \left(\frac{3}{5}, \frac{4}{5}\right)$.

Note-se que $\|\mathrm{vers}(3, 4)\| = \|\left(\frac{3}{5}, \frac{4}{5}\right)\| = \sqrt{\frac{9}{25} + \frac{16}{25}} = 1$.

Ortogonalidade entre vetores

Vetores ortogonais

Dois vetores $u, v \in \mathbb{R}^m$ dizem-se *ortogonais* ($u \perp v$) se $u \cdot v = 0$, ou equivalentemente, usando a notação matricial, $u^T v = 0$.

Por exemplo, os vetores $u = (-4, 1, 2, 1) = \begin{bmatrix} -4 \\ 1 \\ 2 \\ 1 \end{bmatrix}$ e

$v = (1, 1, 1, 1) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ são ortogonais pois

$$u \cdot v = u^T v = \begin{bmatrix} -4 & 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = -4 + 1 + 2 + 1 = 0.$$

Ortogonalidade entre um vetor e um subespaço vetorial

Consideremos o plano de $\mathbb{R}^3$ que passa na origem,

$$V = \{(x_1, x_2, x_3) : \ x_1 + 2x_2 + 3x_3 = 0\}.$$

Podemos escrever,

$$V = \{(x_1, x_2, x_3) : (x_1, x_2, x_3) \cdot (1, 2, 3) = 0\},$$

o que significa que o vetor normal ao plano V, $(1, 2, 3)$, é perpendicular a todos os vetores de V. Diz-se então que $(1, 2, 3)$ é ortogonal ao subespaço V, que se denota por $(1, 2, 3) \perp V$.

Mais geralmente tem-se a seguinte definição.

Vetor ortogonal a um subespaço vetorial

Sejam $u \in \mathbb{R}^m$ e V subespaço vetorial de $\mathbb{R}^m$. Diz-se que é *u ortogonal a V* e denota-se $u \perp V$ se u for ortogonal a todos os vetores de V isto é, se $u^T v = 0$ para qualquer $v \in V$.

Vetor ortogonal a um subespaço dado por geradores

O seguinte resultado mostra que para verificarmos que um vetor é ortogonal a um subespaço vetorial basta mostrar que é ortogonal a um conjunto de geradores desse subespaço.

Teorema (resultado-chave)

Sejam $u \in \mathbb{R}^m$ e $V = \langle v_1, \ldots, v_n \rangle$ subespaço vetorial de $\mathbb{R}^m$.
Tem-se,

$$u \perp V \quad \Leftrightarrow \quad \begin{cases} u \perp v_1, \\ \vdots \\ u \perp v_n. \end{cases}$$

Exercício na aula

Exercício

Sejam $v_1 = (1, 2, -1)$, $v_2 = (2, 0, 2)$, $V = \langle v_1, v_2 \rangle$ e $b = (1, -1, -1)$.
Prove que $b \perp V$.

Resolução: tem-se:

- $b \perp v_1$ pois $v_1 \cdot b = v_1^T b = \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} = 0.$
- $b \perp v_2$ pois $v_2 \cdot b = v_2^T b = \begin{bmatrix} 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} = 0.$

Como $b \perp v_1$ e $b \perp v_2$, conclui-se pelo teorema do slide 145 que $b \perp V = \langle v_1, v_2 \rangle$.

Complemento ortogonal de um subespaço vetorial

Definição de complemento ortogonal

Seja V subespaço vetorial de $\mathbb{R}^m$. Chama-se complemento ortogonal de V e denota-se por $V^\perp$, ao conjunto de todos os vetores de $\mathbb{R}^m$ que são ortogonais a V, isto é,

$$V^\perp = \{x \in \mathbb{R}^m \ : \ x \perp V\}.$$

- Note-se que por definição de complemento ortogonal, tem-se

$$\boxed{u \perp V \Leftrightarrow u \in V^\perp.}$$

- Vejamos como determinar o complemento ortogonal num exemplo, que nos irá sugerir também um método geral para calcular complementos ortogonais de subespaços vetoriais arbitrários.

Exemplo

Consideremos novamente os vetores $v_1 = (1, 2, -1)$, $v_2 = (2, 0, 2)$ e seja $A = [\, v_1 \ \ v_2 \,]$. Vamos determinar $\mathcal{C}(A)^\perp = \{x \in \mathbb{R}^3 : x \perp \mathcal{C}(A)\}$.

Dado $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ tem-se:

$$x \perp \mathcal{C}(A) = \langle v_1, v_2 \rangle \quad \Leftrightarrow \quad \begin{cases} v_1 \perp x \\ v_2 \perp x \end{cases} \Leftrightarrow \begin{cases} v_1^T x = 0 \\ v_2^T x = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \\ \begin{bmatrix} 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \end{cases}$$

$$\Leftrightarrow \begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow A^T x = \vec{0}.$$

Logo, $\mathcal{C}(A)^\perp = \{x \in \mathbb{R}^3 \ : \ A^T x = \vec{0}\} = \mathcal{N}(A^T) = \cdots = \langle(-1, 1, 1)\rangle$ (verifique).

Uma relação fundamental

A relação estabelecida no slide 148 pode ser generalizada para o espaço das colunas de uma matriz arbitrária A. Mais precisamente tem-se o seguinte.

Complemento ortogonal do espaço de colunas / espaço gerado

Sejam $v_1, \ldots, v_n \in \mathbb{R}^m$ e $A = [\, v_1 \quad \cdots \quad v_n \,]$. Tem-se:

$$\boxed{\mathcal{C}(A)^\perp = \langle v_1, \ldots, v_n \rangle^\perp = \mathcal{N}(A^T).}$$

- Em geral, se V é subespaço vetorial $\mathbb{R}^m$ de dimensão $k > 0$ e $\{v_1, \ldots, v_k\}$ é uma base de V, podemos escrever $V = \mathcal{C}(A)$, onde $A_{m\times k} = [v_1 \ \cdots \ v_k]$ é a matriz da base.
- Logo $V^\perp = \mathcal{C}(A)^\perp = \mathcal{N}(A^T)$ é também um subespaço vetorial de $\mathbb{R}^m$ e tem-se atendendo a que o número de colunas de A^T é m,

$$\dim V^\perp = \dim \mathcal{N}(A^T) = m - \operatorname{car}(A^T) = m - \operatorname{car}(A) = m - \dim V.$$

Estas e outras propriedades são enunciadas no próximo slide.

Propriedades do complemento ortogonal

Teorema

Seja V um subespaço vetorial de $\mathbb{R}^m$. Então $V^\perp$ é também subespaço vetorial de $\mathbb{R}^m$ e tem-se:

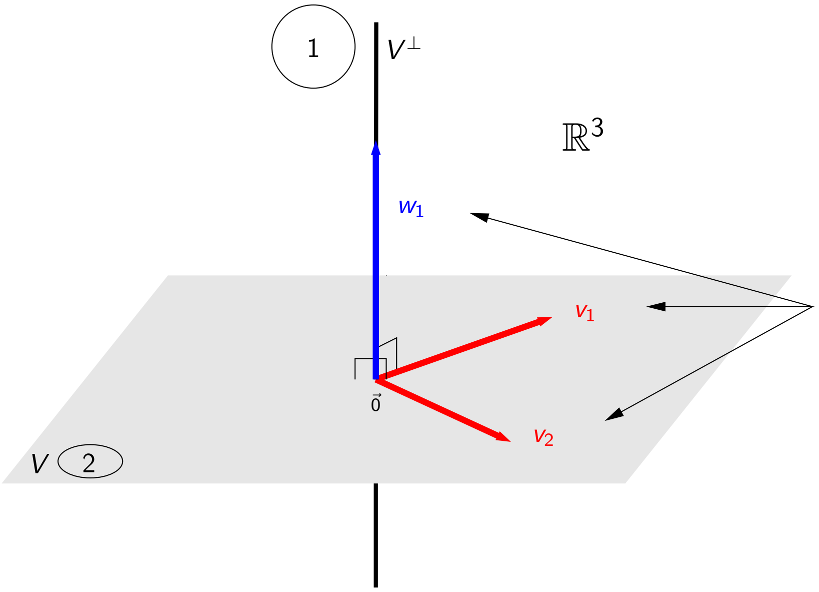
- $V \cap V^\perp = \{\vec{0}\}$.
- $(V^\perp)^\perp = V$.
- $\dim V + \dim V^\perp = \dim \mathbb{R}^m = m$.
- Se $\{v_1, \ldots, v_n\}$ é base de V e $\{w_1, \ldots, w_{m-n}\}$ é base de $V^\perp$, então

$$\{\underbrace{v_1, \ldots, v_n}_{\text{base de } V}, \underbrace{w_1, \ldots, w_{m-n}}_{\text{base de } V^\perp}\},$$

é **base de** $\mathbb{R}^m$.

A relação $V \cap V^\perp = \{\vec{0}\}$ significa que o vetor nulo é o **único** vetor que é ortogonal a si próprio. A relação $(V^\perp)^\perp = V$ significa que se $W = V^\perp$ então $W^\perp = V$.

Ilustração das propriedades do complemento ortogonal quando V é um plano de $\mathbb{R}^3$ ($m = 3$ e $n = \dim V = 2$)



Neste exemplo:

base de $\mathbb{R}^3$: $\{v_1, v_2, w_1\}$

$\dim \mathbb{R}^3 = 3$

$\dim V = 2$ (V plano)

$\dim V^\perp = 3 - 2 = 1$ ($V^\perp$ reta)

Subespaços vetoriais e respetivos complementos ortogonais

Quadros-resumo do complemento ortogonal de subespaços vetoriais de $\mathbb{R}^2$ e $\mathbb{R}^3$

$V \subset \mathbb{R}^2$	$V^\perp \subset \mathbb{R}^2$	$\dim V + \dim V^\perp$
$\{\vec{0}\}$	$\mathbb{R}^2$	0+2
reta que passa na origem	reta perpendicular que passa na origem	1+1
$\mathbb{R}^2$	$\{\vec{0}\}$	2+0

$V \subset \mathbb{R}^3$	$V^\perp \subset \mathbb{R}^3$	$\dim V + \dim V^\perp$
$\{\vec{0}\}$	$\mathbb{R}^3$	0+3
reta que passa na origem	plano perpendicular que passa na origem	1+2
plano que passa na origem	reta perpendicular que passa na origem	2+1
$\mathbb{R}^3$	$\{\vec{0}\}$	3+0

Para qualquer $m \geq 2$ têm-se ainda as relações:

- $(\mathbb{R}^m)^\perp = \{\vec{0}\}$, isto é, subespaço maximal$^\perp$ = subespaço minimal.
- $\{\vec{0}\}^\perp = \mathbb{R}^m$, isto é, subespaço minimal$^\perp$ = subespaço maximal.

Complemento ortogonal de $\mathcal{N}(A)$ / CS do sistema homogéneo

- Se A é uma matriz do tipo $m \times n$ tem-se pelo slide 149 aplicado a A^T que $\mathcal{C}(A^T)^\perp = \mathcal{N}((A^T)^T) = \mathcal{N}(A)$.
- Aplicando a 2ª propriedade do complemento ortogonal do slide 150 à relação $\mathcal{N}(A) = \mathcal{C}(A^T)^\perp$ do ponto anterior tem-se,

$$\mathcal{N}(A)^\perp = (\mathcal{C}(A^T)^\perp)^\perp = \mathcal{C}(A^T).$$

Obtivemos a 2ª relação fundamental sobre o complemento ortogonal de um subespaço vetorial associado a uma matriz.

Complemento ortogonal do espaço nulo

$$\boxed{\mathcal{N}(A)^\perp = \mathcal{C}(A^T).}$$

Exercício

Exercício na aula

Considere $V = \{(x_1, x_2, x_3) : x_1 + 2x_2 + 3x_3 = 0\}$ que define um plano de $\mathbb{R}^3$ com vetor normal $(1, 2, 3)$. Determine $V^\perp$.

Resolução: tem-se $V = \mathcal{N}([\,1\ 2\ 3\,])$ e portanto pela relação do slide anterior vem,

$$V^\perp = \mathcal{N}([\,1\ 2\ 3\,])^\perp = \mathcal{C}\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right) = \langle (1, 2, 3) \rangle,$$

isto é, o complemento ortogonal do plano de $\mathbb{R}^3$ que passa na origem com vetor **normal** $(1, 2, 3)$ é a reta perpendicular ao plano que passa na origem com vetor **diretor** $(1, 2, 3)$.

Só para recordar :)

Pelos resultados dos slides 149 e 153 podemos escrever o seguinte.

Mnemónica

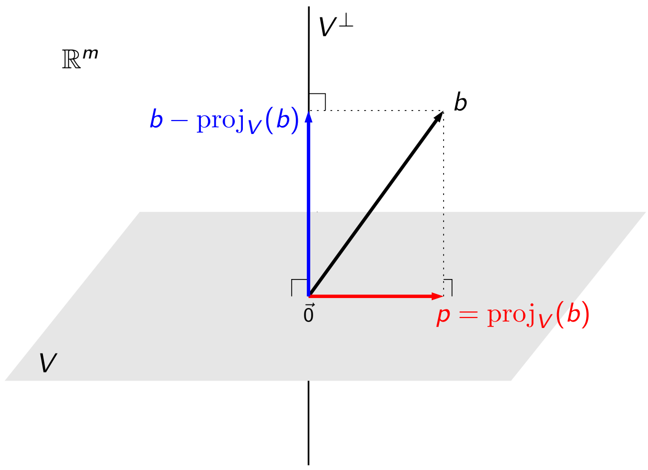
O complemento ortogonal dum subespaço vetorial associado a uma matriz, "troca" o espaço de colunas com o espaço nulo e transpõe essa matriz:

$$\begin{aligned} \mathcal{C}(A)^{\perp} &= \mathcal{N}(A^T) \\ \mathcal{N}(A)^{\perp} &= \mathcal{C}(A^T) \end{aligned}$$

Conceito de projeção ortogonal

Teorema-definição

- Seja V um subespaço vetorial de $\mathbb{R}^m$. Para todo o $b \in \mathbb{R}^m$ existe um e um só $p \in V$ tal que $b - p \in V^{\perp}$, isto é, tal que $b - p \perp V$.
- O vetor p é designado por projeção ortogonal de b sobre V e denota-se por $\text{proj}_V(b)$.



Conceito de projeção ortogonal

A definição anterior significa que o vetor $p = \text{proj}_V(b)$ é caracterizado por duas propriedades:

- $p \in V \rightarrow$ projecta b sobre o subespaço vetorial V.
- $(b - p) \perp V \rightarrow$ a direção da projeção é perpendicular a V.

Exercício na aula

Sejam $b = (-1, 1, 3)$, $v_1 = (1, 1, 2)$, $v_2 = (-1, 1, 0)$ e $V = \langle v_1, v_2 \rangle$.
Mostre que $\text{proj}_V(b) = (0, 2, 2)$.

Resolução: por definição é necessário mostrar que $p = (0, 2, 2)$ verifica as seguintes 2 condições:

- $p \in V$.
- $(b - p) \perp V$, isto é, $(b - p) \in V^{\perp}$.

Resolução do exercício na aula (cont.)

Tem-se:

- $p = (0, 2, 2) \in V = \langle v_1, v_2 \rangle = \mathcal{C}(A)$, onde $A = [\, v_1 \quad v_2 \,]$, se e só se $Ax = p$ for possível. Ora,

$$[\, A \mid p \,] = \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 1 & 1 & 2 \\ 2 & 0 & 2 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{array}\right].$$

  Como o sistema $Ax = p$ é possível, $p = (0, 2, 2) \in V$. ✓

- $b - p = (-1, 1, 3) - (0, 2, 2) = (-1, -1, 1) \in V^{\perp} = \mathcal{C}(A)^{\perp} = \mathcal{N}(A^T)$ se e só se $A^T(b - p) = \vec{0}$. De facto,

$$A^T(b - p) = \begin{bmatrix} 1 & 1 & 2 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \vec{0}.$$

  Logo $(b - p) \in V^{\perp}$ ✓ (alternativamente pode-se mostrar que $b - p$ é ortogonal aos geradores de V, i.e., $(b - p) \cdot v_1 = (b - p) \cdot v_2 = 0$).

Uma vez que as duas condições são verificadas, $p = \text{proj}_V(b)$.

Casos triviais: projeção sobre os subespaços maximal e minimal

A projeção ortogonal sobre o subespaço maximal $\mathbb{R}^m$ ou sobre o subespaço minimal $\{\vec{0}\}$ decorre imediatamente por definição:

- $\text{proj}_{\mathbb{R}^m}(b) = b$ para todo o $b \in \mathbb{R}^m$.

  De facto,
  - $p = b \in \mathbb{R}^m$.
  - $b - p = b - b = \vec{0} \in (\mathbb{R}^m)^\perp = \{\vec{0}\}$.
- $\text{proj}_{\{\vec{0}\}}(b) = \vec{0}$ para todo o $b \in \mathbb{R}^m$.

  De facto,
  - $p = \vec{0} \in \{\vec{0}\}$.
  - $b - \vec{0} = b \in \{\vec{0}\}^\perp = \mathbb{R}^m$.

Observação

Dado V subespaço vetorial de $\mathbb{R}^m$ e $b \in \mathbb{R}^m$, tem-se em geral,

$$\boxed{\text{proj}_V(b) = b \quad \Leftrightarrow \quad b \in V}$$

como decorre facilmente por definição de projeção ortogonal (exercício).

Caso não trivial mais simples: projeção sobre uma reta

E a projeção ortogonal sobre outros subespaços vetoriais?

- O caso não trivial mais simples corresponde a calcular a projeção ortogonal sobre um subespaço vetorial de dimensão um, isto é, sobre uma reta que passa na origem.

Fórmula da projeção ortogonal sobre uma reta

Seja $V = \langle v \rangle$ com $v \in \mathbb{R}^m$ e $v \neq \vec{0}$. Para qualquer $b \in \mathbb{R}^m$ tem-se

$$\boxed{\text{proj}_V(b) = \text{proj}_{\langle v \rangle}(b) = \frac{v^T b}{v^T v} v = \frac{v \cdot b}{v \cdot v} v}$$

A demonstração deste resultado será feita no próximo slide.

Projeção ortogonal sobre uma reta - demonstração

Demonstração: Por definição de projeção ortogonal,

- $p \in V = \langle v \rangle = \mathcal{C}(v)$. Logo existe $\alpha \in \mathbb{R}$ tal que $p = \alpha v$.
- $(b - p) \in V^{\perp} = \mathcal{N}(v^T)$. Logo $v^T(b - p) = 0$.

Têm-se as equivalências,

$$\begin{aligned} v^T(b-p) = 0 \quad &\Leftrightarrow \quad v^T(b - \alpha v) = 0 \Leftrightarrow v^T b - \alpha v^T v = 0 \\ &\Leftrightarrow \quad \alpha v^T v = v^T b \Leftrightarrow \alpha = \frac{v^T b}{v^T v} = \frac{v \cdot b}{v \cdot v}.(^{11}) \end{aligned}$$

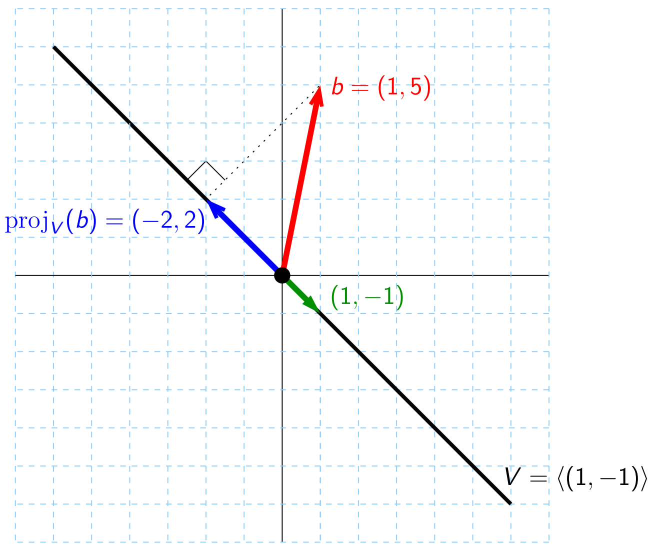
Logo $p = \alpha v = \dfrac{v^T b}{v^T v} v = \dfrac{v \cdot b}{v \cdot v} v.$ $\square$

[11] Note-se que $v \cdot v = \|v\|^2 \neq 0$ pois $v \neq 0$.

Projeção ortogonal sobre uma reta - exemplo

Sejam $b = (1, 5)$ e $V = \{(x_1, x_2) \ : \ x_1 = -x_2\}$ a bissectriz dos quadrantes pares. Então V é uma reta que passa na origem com vetor director $(1, -1)$ (por exemplo), tendo-se $V = \langle (1, -1) \rangle$. Logo,

$$\text{proj}_V(b) = \text{proj}_{\langle (1,-1) \rangle}(b) = \frac{(1,-1) \cdot (1,5)}{(1,-1) \cdot (1,-1)}(1,-1) = (-2, 2).$$



Projeção ortogonal sobre um vetor

Fórmula da projeção ortogonal sobre um vetor

Seja $v \in \mathbb{R}^m$ e $v \neq \vec{0}$. Dado $b \in \mathbb{R}^m$ define-se a projeção ortogonal de b sobre o vetor v, denotada $\text{proj}_v(b)$, como sendo a projeção de b sobre a reta definida por v, isto é,

$$\text{proj}_v(b) = \text{proj}_{\langle v \rangle}(b) = \frac{v \cdot b}{v \cdot v} v$$

Voltando ao exemplo do slide anterior, tem-se

$$\begin{aligned} \text{proj}_{(-4,4)}(1,5) &= \text{proj}_{\langle(-4,4)\rangle}(1,5) = \frac{(-4,4)\cdot(1,5)}{(-4,4)\cdot(-4,4)}(-4,4) \\ &= \frac{1}{2}(-4,4) = (-2,2). \end{aligned}$$

Uma decomposição fundamental

Observação

Se $p = \text{proj}_V(b)$ tem-se por definição:

- $p \in V$
- $b - p \in V^\perp$

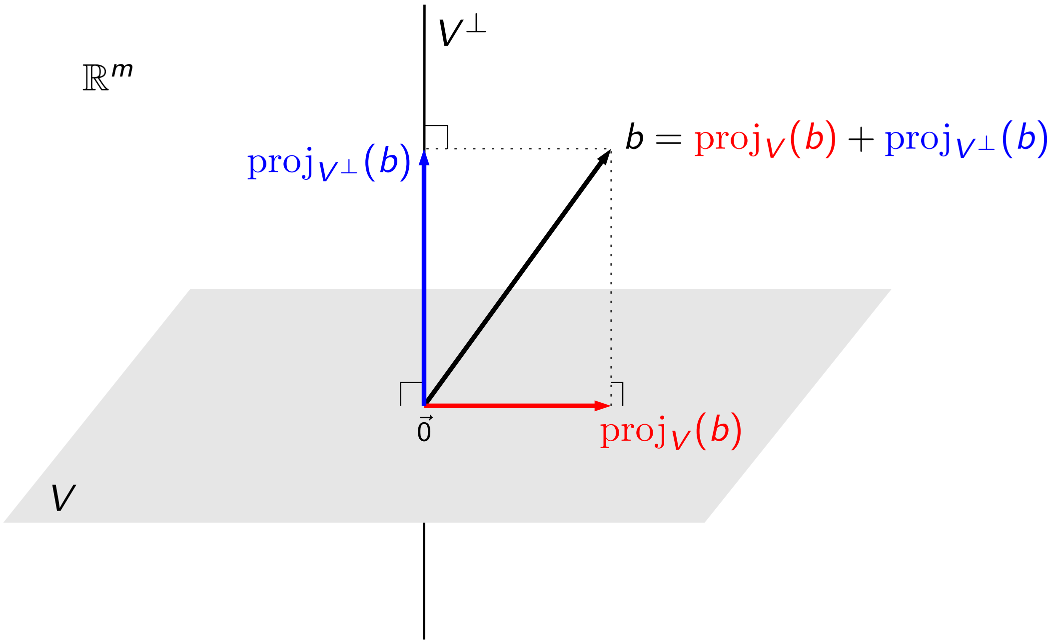
Logo $b - p = \text{proj}_{V^\perp}(b)$. De facto,

- $q = b - p \in V^\perp$
- $b - q = b - (b - p) = p \in V = (V^\perp)^\perp$

Como $b = p + (b - p)$ obtivemos a seguinte decomposição (única) de b segundo V e $V^\perp$:

$$\boxed{b = \text{proj}_V(b) + \text{proj}_{V^\perp}(b)}$$

$b = \text{proj}_V(b) + \text{proj}_{V^\perp}(b)$



Uma aplicação da decomposição do slide 165

Se V for um subespaço vetorial de $\mathbb{R}^m$ de dimensão $m-1$, isto é, se $V^\perp$ tiver dimensão um, então $V^\perp = \langle w \rangle$ com $w \neq \vec{0}$, e pode-se aplicar a fórmula da projeção ortogonal sobre uma reta dada no slide 160 ao complemento ortogonal $V^\perp$ o que, juntamente com a decomposição do slide 165, permite obter a projeção ortogonal de um vetor $b \in \mathbb{R}^m$ sobre V:

$$\boxed{\text{proj}_V(b) = b - \text{proj}_{V^\perp}(b) = b - \frac{w \cdot b}{w \cdot w} w.}$$

Exercício na aula

Considere $V = \{(x_1, x_2, x_3) : x_1 + 2x_2 + 3x_3 = 0\}$ e $b = (1, 1, 1)$.
Calcule $\text{proj}_V(b)$.

TPC: calcule $\text{proj}_V(b)$ onde $V = \langle (1, 1, 2), (-1, 1, 0) \rangle$ e $b = (-1, 1, 3)$ e compare o resultado obtido com o exercício do slide 157.

Resolução do exercício na aula

Escrevendo $V = \mathcal{N}(A)$ com $A = [\,1 \;\; 2 \;\; 3\,]$ conclui-se que $\dim V = 2$ (nº de variáveis livres) e $\dim V^\perp = \dim \mathbb{R}^3 - \dim V = 1$, obtendo-se pela fórmula do slide 155 (ver também o slide 154),

$$V^\perp = \mathcal{N}(A)^\perp = \mathcal{C}(A^T) = \mathcal{C}\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right) = \langle (1,2,3) \rangle.$$

Logo $V^\perp$ define uma reta com vetor diretor $(1,2,3)$ e tem-se pelo resultado do slide 166,

$$\begin{aligned}
\mathrm{proj}_V(b) &= b - \mathrm{proj}_{V^\perp}(b) \\
&= (1,1,1) - \mathrm{proj}_{\langle (1,2,3) \rangle}((1,1,1)) \\
&= (1,1,1) - \frac{(1,2,3)\cdot(1,1,1)}{(1,2,3)\cdot(1,2,3)}(1,2,3) \\
&= (1,1,1) - \frac{6}{14}(1,2,3) = \frac{1}{7}(4,1,-2).
\end{aligned}$$

Distância de um vetor a um subespaço vetorial

Definição de distância de um vetor a um subespaço vetorial

Dados $b \in \mathbb{R}^m$ e V subespaço vetorial de $\mathbb{R}^m$ define-se distância de b a V, que se denota-se por $d(b, V)$, como sendo a distância entre b e o vetor p de V que se encontra mais próximo de b([12]), isto é,

$$d(b,V) = d(b,p) = \min_{v \in V} d(b,v).$$

Intuitivamente o vetor de V mais próximo de b é o vetor de V que se encontra na reta que passa em b e tem direção perpendicular a V. Mais precisamente, tem-se o resultado do próximo slide.

[12] que se pode mostrar que existe sempre e é único!

Vetor mais próximo de um subespaço vetorial é a projeção ortogonal

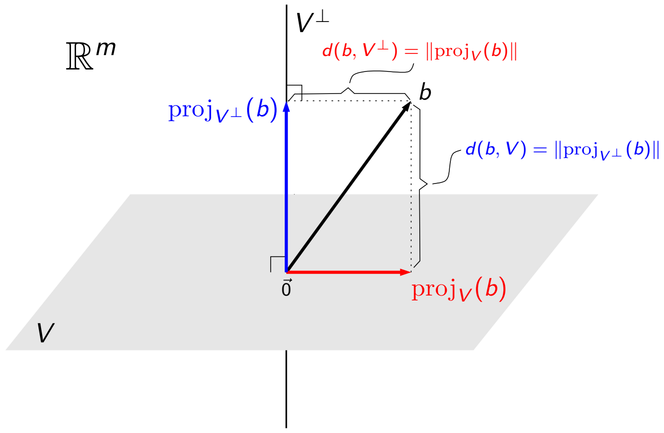
Teorema

Sejam $b \in \mathbb{R}^m$ e V subespaço vetorial de $\mathbb{R}^m$. O vetor de V que se encontra mais próximo de b é $p = \mathrm{proj}_V(b)$ e tem-se,

$$\boxed{d(b, V) = d(b, \mathrm{proj}_V(b)) = \|b - \mathrm{proj}_V(b)\| = \|\mathrm{proj}_{V^\perp}(b)\|.}$$

Pelo teorema anterior aplicado a $V^\perp$, obtém-se ainda a relação,

- $\boxed{d(b, V^\perp) = d(b, \mathrm{proj}_{V^\perp}(b)) = \|b - \mathrm{proj}_{V^\perp}(b)\| = \|\mathrm{proj}_V(b)\|.}$

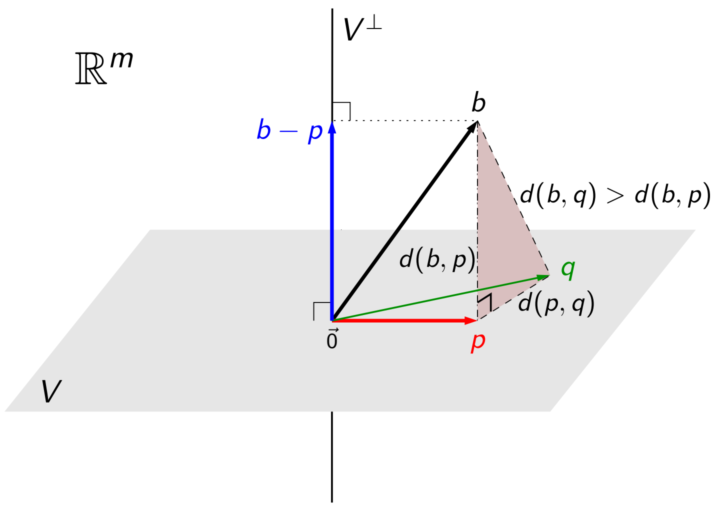


Demonstração do teorema do slide 169

Podemos supor que $b \notin V$. Sejam $p = \mathrm{proj}_V(b)$ e $q \in V$, $q \neq p$. Tem-se que $d(b, q)$ corresponde ao comprimento da hipotenusa do triângulo rectângulo assinalado na figura, triângulo esse com catetos de comprimentos $d(b, p) > 0$ e $d(q, p) > 0$. Pelo análogo do teorema de Pitágoras para vetores de $\mathbb{R}^m$ tem-se,

$$d(b, q)^2 = d(b, p)^2 + d(q, p)^2 > d(b, p)^2.$$

Logo $d(b, q) > d(b, p)$, para todo o $q \in V$, $q \neq p$. $\square$



Equações normais

Sejam V subespaço vetorial de $\mathbb{R}^m$, $\{v_1, \ldots, v_n\}$ base de V e $A = [\, v_1 \ \cdots \ v_n \,]$ matriz da base. Sejam $b \in \mathbb{R}^m$ e $p = \text{proj}_V(b)$. Por definição tem-se:

(i) $p \in V = \mathcal{C}(A)$. Logo $Ax = p$ é possível e podemos escrever $p = A\bar{x}$ com $\bar{x}$ solução do sistema $Ax = p$.

(ii) $(b - p) \in V^\perp = \mathcal{C}(A)^\perp = \mathcal{N}(A^T)$. Logo, $A^T(b - p) = \vec{0}$.

Por (i) e (ii) tem-se,

$$\begin{aligned} A^T(b-p) = \vec{0} \quad &\Leftrightarrow \quad A^T(b - A\,\bar{x}) = \vec{0} \Leftrightarrow A^T b - A^T A\,\bar{x} = \vec{0} \\ &\Leftrightarrow \quad A^T A\,\bar{x} = A^T b. \end{aligned}$$

Logo $\bar{x}$ é também solução do sistema,

$$\boxed{A^T A\, x = A^T b,}$$

que se designa por sistema das equações normais. Uma vez que A é a matriz de uma base com n vetores, $\text{car}(A) = n$, e pode-se mostrar que a matriz $A^T A$ é invertível (de ordem n). Das considerações anteriores conclui-se que

$$\text{proj}_V(b) = p = A\bar{x}$$

com $\bar{x}$ solução única do sistema de equações normais ($\bar{x} = (A^T A)^{-1} A^T b$), donde se deduz método das equações normais do próximo slide.

Método das equações normais

Algoritmo

`Input:` V subespaço vetorial de $\mathbb{R}^m$ e $b \in \mathbb{R}^m$.
`Objectivo:` Calcular $\text{proj}_V(b)$.

1. Determinar uma base para V, $\{v_1, \ldots, v_n\}$.
2. Determinar a solução (única) $\bar{x}$ do sistema das equações normais,

$$A^T A\, x = A^T b,$$

   onde $A = [\, v_1 \ \cdots \ v_n \,]$ é a matriz da base de V.
3. $\text{proj}_V(b) = A\,\bar{x}$.

Exercício na aula

Determine a projeção ortogonal de $b = (1, 0, 4)$ sobre subespaço $V = \big\langle (1, 0, 1), (1, 1, 2) \big\rangle$ utilizando o método das equações normais. Indique ainda as distâncias de b a V e a $V^\perp$.

Exercício na aula (resolução)

Aplicando o algoritmo do slide anterior tem-se:

1. $\{v_1, v_2\} = \{(1,0,1),(1,1,2)\}$ é base de $V = \langle v_1, v_2\rangle$ (justifique).
2. Seja $A = [\, v_1 \;\; v_2 \,]$ a matriz da base. Tem-se:
   - $A^TA = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix}$.
   - $A^Tb = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \\ 9 \end{bmatrix}$.

   Reduzindo o sistema das equações normais $A^TAx = A^Tb$ obtém-se

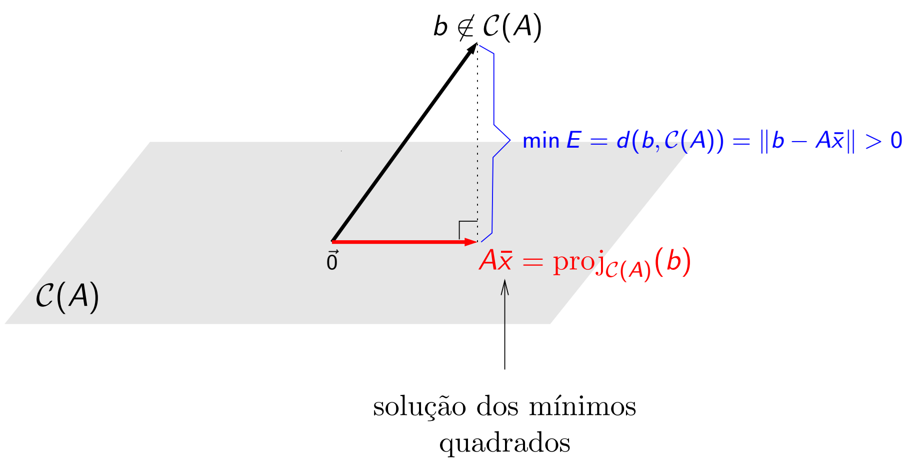
   $$[\, A^TA \mid A^Tb \,] = \left[\begin{array}{cc|c} 2 & 3 & 5 \\ 3 & 6 & 9 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right],$$

   cuja a única solução é $\bar{x} = (1,1)$.
3. $\mathrm{proj}_V(b) = A\bar{x} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$.

$d(b,V) = \|\mathrm{proj}_{V^\perp}(b)\| = \|b - \mathrm{proj}_V(b)\| = \sqrt{3}, \quad d(b, V^\perp) = \|\mathrm{proj}_V(b)\| = \sqrt{14}.$

Solução no sentido dos mínimos quadrados de $Ax = b$

- A solução $\bar{x}$ do sistema das equações normais $A^TAx = A^Tb$ verifica $\mathrm{proj}_V(b) = A\bar{x}$ com $V = \mathcal{C}(A)$ e portanto é solução de $Ax = \mathrm{proj}_{\mathcal{C}(A)}(b)$.
- Quando $b \notin \mathcal{C}(A)$ o sistema $Ax = b$ é impossível e $\bar{x}$ é o vetor que melhor se aproxima de ser uma solução de $Ax = b$ no sentido em que é o vetor que minimiza a diferença (erro) $E = \|Ax - b\|$, e designa-se por solução de $Ax = b$ no sentido dos mínimos quadrados.
- Quando $b \in \mathcal{C}(A)$, a solução $\bar{x}$ no sentido dos mínimos quadrados é também uma solução de $Ax = b$ no sentido usual, isto é, verifica, $A\bar{x} = b$.



A projeção ortogonal como uma transformação linear

Nas condições do slide 171 tem-se, multiplicando à esquerda por $(A^T A)^{-1}$ ambos os membros do sistema das equações normais,

$$\begin{aligned} A^T A\bar{x} = A^T b \quad &\Leftrightarrow \quad (A^T A)^{-1} A^T A\bar{x} = (A^T A)^{-1} A^T b \\ &\Leftrightarrow \quad \bar{x} = (A^T A)^{-1} A^T b, \end{aligned}$$

e portanto,

$$\mathrm{proj}_V(b) = A\bar{x} = A(A^T A)^{-1} A^T b.$$

Logo a projeção ortogonal sobre V pode ser vista como a transformação linear associada à matriz $P = A(A^T A)^{-1} A^T$,

$$\begin{aligned} \mathbb{R}^m \quad &\to \quad \mathbb{R}^m \\ b \quad &\mapsto \quad \mathrm{proj}_V(b) = Pb, \end{aligned}$$

o que motiva a definição do próximo slide.

Matriz de projeção

Definição da matriz de projeção sobre um subespaço vetorial

Sejam V subespaço vetorial de $\mathbb{R}^m$, $\{v_1, \ldots, v_n\}$ uma base para V e $A = [\, v_1 \ \cdots \ v_n \,]$ a matriz dessa base. A matriz de projeção sobre V é a matriz quadrada de ordem m,

$$\boxed{P = A(A^T A)^{-1} A^T,}$$

tendo-se, para todo o $b \in \mathbb{R}^m$,

$$\boxed{\mathrm{proj}_V(b) = P\, b.}$$

Observação

Nas condições da definição anterior, tem-se ainda para todo o $b \in \mathbb{R}^m$,

$$\mathrm{proj}_{V^\perp}(b) = b - \mathrm{proj}_V(b) = b - Pb = (I_m - P)b,$$

onde I_m denota a matriz identidade de ordem m, donde se conclui que $(I_m - P)$ é a matriz de projeção sobre $V^\perp$.

Propriedades

Propriedades da matriz de projeção

A matriz de projeção P sobre um subespaço V de $\mathbb{R}^m$ não depende da escolha da base de V e verifica as seguintes propriedades:

- $P^T = P$ (simétrica).
- $P^2 = P$ (idempotente).

(exercício 33.6 da sebenta de exercícios)

A demonstração das seguintes propriedades decorre do facto da projeção ortogonal definir uma transformação linear associada a uma matriz e fica como exercício para os alunos.

Linearidade da projeção ortogonal

Para todo o $u, v \in \mathbb{R}^m$ e $\alpha \in \mathbb{R}$ tem-se:

- $\mathrm{proj}_V(u + v) = \mathrm{proj}_V(u) + \mathrm{proj}_V(v)$.
- $\mathrm{proj}_V(\alpha\, u) = \alpha\, \mathrm{proj}_V(u)$.

Matriz de projeção ortogonal sobre uma reta (dimensão 1)

Observação

Se $V = \langle v \rangle$ é a reta com vetor diretor $v \neq \vec{0}$, a matriz da base de V reduz-se a v e obtém-se a expressão muito elegante para a matriz de projeção sobre V,

$$\boxed{P = v(v^T v)^{-1} v^T = \frac{v\, v^T}{v^T v}.}$$

Exemplo

A matriz de projeção sobre a reta $V = \langle(-1, -1, 1)\rangle$, que passa na origem com vetor diretor $(-1, -1, 1)$, é a matriz

$$P = \frac{v\, v^T}{v^T v} = \frac{\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & 1 \end{bmatrix}}{\begin{bmatrix} -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}} = \frac{1}{3} \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}.$$

TPC: qual a projeção ortogonal do vetor genérico $b = (b_1, b_2, b_3)$ sobre V?

Matriz de projeção ortogonal sobre um subespaço de dimensão 2

Se V é subespaço vetorial de $\mathbb{R}^m$ de dimensão 2 e $A_{m\times 2}$ é a matriz de uma base de V, a matriz de projeção sobre V,

$$P = A(A^TA)^{-1}A^T,$$

envolve a inversa da matriz quadrada de ordem 2, A^TA, que pode ser facilmente obtida usando a seguinte mnemónica.

Mnemónica para calcular a inversa de uma matriz 2×2

"Switch diagonally, negate the wings and divide by a cross" (*):

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

(*) https://www.dam.brown.edu/people/mchb/la/matrix_algebra.pdf
Nota: a matriz é invertível $\Leftrightarrow ad - bc \neq 0$, que se designa por determinante da matriz.

Exercício na aula

Determinar as matrizes de projeção sobre $V = \langle (1,0,1),(1,1,2)\rangle$ e $V^\perp$.

Resolução do exercício na aula

Uma base de V é $\{v_1, v_2\} = \{(1,0,1),(1,1,2)\}$ (justifique) e tem-se, designando por $A = [\, v_1 \;\; v_2 \,]$ a matriz desta base,

$$A^TA = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix}.$$

Usando a mnemónica anterior para calcular $(A^TA)^{-1}$ e determinando a matriz de projeção P sobre V obtém-se,

$$\begin{aligned} P &= A(A^TA)^{-1}A^T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix}^{-1}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix}\frac{1}{2\times 6 - 3\times 3}\begin{bmatrix} 6 & -3 \\ -3 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 3 & -3 & 0 \\ -1 & 2 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}. \end{aligned}$$

Resolução do exercício na aula (concl.)

A matriz de projeção sobre $V^\perp$ é $I_3 - P$, em que I_3 é a matriz identidade de ordem 3, e portanto vem dada por

$$\begin{aligned} I_3 - P &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}. \end{aligned}$$

TPC: calcule $V^\perp$ e confirme esta matriz com o resultado obtido no exercício do slide 178.

Aplicação ao famoso conjunto de dados dos lírios de Fisher

Consideremos o conjunto $X_{150\times 4}$ dos dados dos 150 lírios de 3 espécies, *Setosa*, *Versicolor* e *Virginica* do slide 2 que origina uma núvem de 150 pontos em $\mathbb{R}^4$ e denotemos por $y_i \in \mathbb{R}^4$ o vetor dos comprimentos e diâmetros das pétalas e sépalas do lírio i e por $Y = [\, y_1 \; y_2 \; \cdots \; y_{150} \,]_{4\times 150} = X^T$ a matriz destes vetores.

Usando certos resultados de Álgebra Linear determinamos a base $\{u, v\}$, com

$$\begin{aligned} u &= (0.36138659, -0.08452251, 0.85667061, 0.35828920), \\ v &= (0.65658877, 0.73016143, -0.17337266, -0.07548102), \end{aligned}$$

do subespaço vetorial V de dimensão 2, que permite obter o melhor retrato bidimensional (projeção ortogonal sobre V) dos líríos, no sentido em que melhor preserva a informação (variabilidade) dos dados originais.

Os vetores u e v são unitários e ortogonais entre si e obtém-se $A^T A = I_2$, onde $A = [\, u \; v \,]$ denota a matriz da base $\{u, v\}$ de V. Logo a matriz de projeção $P = A(A^T A)^{-1} A^T$sobre V vem dada simplesmente por $P = AA^T$ e tem-se,

$$P = \begin{bmatrix} 0.56170908 & 0.44887050 & 0.1957547 & 0.07992092 \\ 0.44887050 & 0.54027978 & -0.1989980 & -0.08539683 \\ 0.19575473 & -0.19899799 & 0.7639426 & 0.32002217 \\ 0.07992092 & -0.08539683 & 0.3200222 & 0.13406853 \end{bmatrix}_{4\times 4}.$$

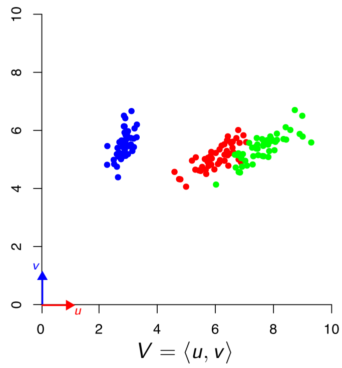
Aplicação ao famoso iris dataset

A projeção da núvem dos 150 lírios no subespaço V é então dada pela matriz,

$$Y' = PY = [y_1'\, y_2' \cdots y_{150}'] = \begin{bmatrix} 4.7258039 & 4.3890268 & \cdots & 5.802902 \\ 3.8845422 & 3.5246282 & \cdots & 3.100571 \\ 1.4353802 & 1.4957283 & \cdots & 5.030106 \\ 0.5835525 & 0.6102668 & \cdots & 2.088779 \end{bmatrix}_{150\times 4},$$

onde cada vetor $y_i' = Py_i$ representa a projeção do i-ésimo lírio no subespaço vetorial V. Note que $\text{car}(Y') = 2$ uma vez que $y_i' \in V = \langle u, v\rangle$ para todo o i.

Para cada $i = 1, \ldots, 150$, existem escalares $\alpha_i, \beta_i \in \mathbb{R}$ tais que $y_i' = \alpha_i u + \beta_i v$. Projetando no referencial ortonormado definido por u e v, cada lírio y_i no ponto de coordenadas (α_i, β_i), obtém-se o retrato abaixo da núvem dos 150 lírios.



Conjunto ortogonal de vetores

Definição de conjunto ortogonal de vetores

Sejam $v_1, \ldots, v_k \in \mathbb{R}^m$. Diz-se que $\{v_1, \ldots, v_k\}$ é um conjunto ortogonal se os vetores forem 2 a 2 perpendiculares entre si, isto é, se $v_i \cdot v_j = 0$, $\forall i, j = 1, \ldots, k$, $i \neq j$.

Exemplos

- A base canónica de $\mathbb{R}^m$ é um conjunto ortogonal de vetores
- $\{v_1, v_2, v_3\} = \{(0,1,1), (1,2,-2), (4,-1,1)\}$ é um conjunto ortogonal de vetores de $\mathbb{R}^3$. De facto,

$$\begin{aligned} v_1 \cdot v_2 &= (0,1,1)\cdot(1,2,-2) = 0, \text{ isto é, } v_1 \perp v_2, \\ v_1 \cdot v_3 &= (0,1,1)\cdot(4,-1,1) = 0, \text{ isto é,} v_1 \perp v_3, \\ v_2 \cdot v_3 &= (1,2,-2)\cdot(4,-1,1) = 0, \text{ isto é, } v_2 \perp v_3. \end{aligned}$$

Ortogonalidade e independência linear

Consideremos novamente o conjunto ortogonal de vetores $\{v_1, v_2, v_3\} = \{(0,1,1),(1,2,-2),(4,-1,1)\}$. Tem-se:

- $v_3 \perp v_1, v_2 \;\Rightarrow\; v_3 \in \langle v_1, v_2\rangle^{\perp} \Rightarrow\; v_3 \notin \langle v_1, v_2\rangle$ $(v_3 \neq \vec{0})$.
  Logo v_3 não é CL de v_1 e v_2.
- $v_2 \perp v_1, v_3 \;\Rightarrow\; v_2 \in \langle v_1, v_3\rangle^{\perp} \Rightarrow\; v_2 \notin \langle v_1, v_3\rangle$ $(v_2 \neq \vec{0})$.
  Logo v_2 não é CL de v_1 e v_3.
- $v_1 \perp v_2, v_3 \;\Rightarrow\; v_1 \in \langle v_2, v_3\rangle^{\perp} \Rightarrow\; v_1 \notin \langle v_2, v_3\rangle$ $(v_1 \neq \vec{0})$.
  Logo v_1 não é CL de v_2 e v_3.

Como nenhum dos vetores é CL dos restantes vetores do conjunto concluímos que $\{v_1, v_2, v_3\}$ é linearmente independente (ver o slide 113).

Em geral, tem-se o seguinte resultado.

Teorema

Todo o conjunto ortogonal de vetores <u>não nulos</u> de $\mathbb{R}^m$ é linearmente independente.

Base ortogonal = base + conjunto ortogonal

Definição de base ortogonal

Uma base ortogonal de um subespaço vetorial V é uma base de V que é simultaneamente um conjunto ortogonal.

Por exemplo, o conjunto ortogonal considerado anteriormente,

$$\{v_1, v_2, v_3\} = \{(0,1,1),(1,2,-2),(4,-1,1)\},$$

define uma base ortogonal de $\mathbb{R}^3$ porque é um conjunto linearmente independente formado por 3 vetores de $\mathbb{R}^3$.

Do teorema do slide 185 e da definição de base deduz-se o seguinte.

Teorema

Seja $V = \langle v_1, \ldots, v_k\rangle$ tal que $\{v_1, \ldots, v_k\}$ é um conjunto ortogonal de vetores <u>não nulos</u> de $\mathbb{R}^m$. Então $\{v_1, \ldots, v_k\}$ é uma base ortogonal de V.

A projeção sobre um subespaço vetorial munido de uma base ortogonal é direta e estende a fórmula da projeção sobre uma reta.

Projeção sobre um espaço munido de uma base ortogonal

Teorema

Seja $\{v_1, \ldots, v_k\}$ uma base ortogonal de um subespaço vetorial V de $\mathbb{R}^m$.
Para todo o $b \in \mathbb{R}^m$ tem-se,

$$\boxed{\text{proj}_V(b) = \text{proj}_{v_1}(b) + \cdots + \text{proj}_{v_k}(b) = \frac{b \cdot v_1}{v_1 \cdot v_1} v_1 + \cdots + \frac{b \cdot v_k}{v_k \cdot v_k} v_k.}$$

Muito importante: o resultado é falso se a base não for ortogonal (!!!)

Exercício na aula

Calcular $\text{proj}_V(b)$ em que $V = \langle (0, 1, 1), (1, 2, -2) \rangle$ e $b = (-1, 0, 4)$.

Resolução: Sejam $v_1 = (0, 1, 1)$ e $v_2 = (1, 2, -2)$. Como $v_1 \cdot v_2 = 0$ com v_1 e v_2 não nulos, $\{v_1, v_2\}$ é uma base **ortogonal** de V (ver o slide 186) e obtém-se,

$$\begin{aligned}
\text{proj}_V(b) &= \text{proj}_{v_1}(b) + \text{proj}_{v_2}(b) = \frac{b \cdot v_1}{v_1 \cdot v_1} v_1 + \frac{b \cdot v_2}{v_2 \cdot v_2} v_2 \\
&= \frac{(-1, 0, 4) \cdot (0, 1, 1)}{(0, 1, 1) \cdot (0, 1, 1)} (0, 1, 1) + \frac{(-1, 0, 4) \cdot (1, 2, -2)}{(1, 2, -2) \cdot (1, 2, -2)} (1, 2, -2) \\
&= \tfrac{4}{2}(0, 1, 1) + \tfrac{-9}{9}(1, 2, -2) = (-1, 0, 4).
\end{aligned}$$

Como obter bases ortogonais de um subespaço vetorial?

O seguinte algoritmo permite construir base ortogonais para subespaços vetoriais a partir de bases não ortogonais desses mesmos subespaços vetoriais.

Algoritmo - Método de ortogonalização de Gram-Schmidt

`Input:` Base "original" $\{u_1, \ldots, u_n\}$ de um subespaço vetorial V.
`Objectivo:` Determinar uma base ortogonal de V, $\{v_1, \ldots, v_n\}$

- $v_1 = u_1$.
- $v_2 = u_2 - \text{proj}_{v_1}(u_2) = u_2 - \frac{v_1 \cdot u_2}{v_1 \cdot v_1} v_1$.
- $v_3 = u_3 - \text{proj}_{v_1}(u_3) - \text{proj}_{v_2}(u_3) = u_3 - \frac{v_1 \cdot u_3}{v_1 \cdot v_1} v_1 - \frac{v_2 \cdot u_3}{v_2 \cdot v_2} v_2$.

  $\vdots$
- $v_n = u_n - \text{proj}_{v_1}(u_n) - \text{proj}_{v_2}(u_n) - \cdots - \text{proj}_{v_{n-1}}(u_n)$
  $= u_n - \frac{v_1 \cdot u_n}{v_1 \cdot v_1} v_1 - \frac{v_2 \cdot u_n}{v_2 \cdot v_2} v_2 - \cdots - \frac{v_{n-1} \cdot u_n}{v_{n-1} \cdot v_{n-1}} v_{n-1}$.

Note-se que no caso em que a base original $\{u_1, \ldots, u_n\}$ já é ortogonal, o método de Gram-Schmidt devolve a própria base!

Método de Gram-Schmidt

Observações

- Podemos multiplicar cada vetor v_i da base ortogonal por um escalar não nulo que ainda obtemos uma base ortogonal de V.
- Em particular, tomando os versores dos vetores $v_1, \ldots, v_n$ da base ortogonal obtém-se a base ortogonal em que todos os vetores são unitários, dita base ortonormada (o.n.) de V, $\left\{ \frac{v_1}{\|v_1\|}, \ldots, \frac{v_n}{\|v_n\|} \right\}$.

Exercício na aula

- A partir da base não ortogonal de $\mathbb{R}^3$,

$$\{u_1, u_2, u_3\} = \{(1, -1, 1), (1, 0, 1), (1, 1, 2)\},$$

  obtenha uma **base ortogonal** de $\mathbb{R}^3$ usando Gram-Schmidt.
- Transforme a base ortogonal anterior numa base ortonormada de $\mathbb{R}^3$.

Exercício na aula (resolução)

Aplicando o método de Gram-Schmidt à base $\{u_1, u_2, u_3\}$ obtém-se:

- $v_1 = u_1 = (1, -1, 1)$
- $v_2 = u_2 - \text{proj}_{v_1}(u_2) = u_2 - \frac{u_2 \cdot v_1}{v_1 \cdot v_1} v_1 = (1, 0, 1) - \frac{(1,0,1)\cdot(1,-1,1)}{(1,-1,1)\cdot(1,-1,1)}(1, -1, 1) = (1, 0, 1) - \frac{2}{3}(1, -1, 1) = \frac{1}{3}(1, 2, 1) \rightsquigarrow (1, 2, 1)$
- $v_3 = u_3 - \text{proj}_{v_1}(u_3) - \text{proj}_{v_2}(u_3)) = u_3 - \frac{u_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{u_3 \cdot v_2}{v_2 \cdot v_2} v_2 = (1, 1, 2) - \frac{(1,1,2)\cdot(1,-1,1)}{(1,-1,1)\cdot(1,-1,1)}(1, -1, 1) - \frac{(1,1,2)\cdot(1,2,1)}{(1,2,1)\cdot(1,2,1)}(1, 2, 1) = (1, 1, 2) - \frac{2}{3}(1, -1, 1) - \frac{5}{6}(1, 2, 1) = \frac{1}{2}(-1, 0, 1) \rightsquigarrow (-1, 0, 1)$

Obteve-se a base ortogonal de $\mathbb{R}^3$,

$$\{v_1, v_2, v_3\} = \{(1, -1, 1), (1, 2, 1), (-1, 0, 1)\}.$$

Normalizando esta base ortogonal, dividindo cada vetor pela sua norma, obtém-se a base ortonormada de $\mathbb{R}^3$,

$$\{\text{vers}(v_1), \text{vers}(v_2), \text{vers}(v_3)\} = \left\{ \frac{(1, -1, 1)}{\sqrt{3}}, \frac{(1, 2, 1)}{\sqrt{6}}, \frac{(-1, 0, 1)}{\sqrt{2}} \right\}.$$

Método de ortogonalização de Gram-Schmidt

Observação

- Se $\{v_1, \ldots, v_n\}$ for uma base ortogonal de um subespaço vetorial $V \subset \mathbb{R}^m$ e $\{w_1, \ldots, w_{m-n}\}$ uma base ortogonal de $V^\perp$ então $\{v_1, \ldots, v_n, w_1, \ldots, w_{m-n}\}$ é uma base ortogonal de $\mathbb{R}^m$.

Note-se que se tem $v_i \perp w_j$, para todo o i e j, uma vez que por definição $V^\perp$ é constituído pelos vetores que são ortogonais a todos os vetores de V (e vice-versa).

Exercício na aula

- Calcule a projeção ortogonal de $b = (1, 1, 3)$ sobre $V = \langle (1, -1, 1), (1, 0, 1) \rangle$ ortogonalizando uma base de V.
- Estenda a base ortogonal de V determinada na alínea anterior de modo a obter uma base ortogonal de $\mathbb{R}^3$ usando a observação acima.

Exercício na aula (resolução)

- Uma base (não ortogonal) de V é $\{u_1, u_2\} = \{(1, -1, 1), (1, 0, 1)\}$ (justifique!).
- Aplicando o método de Gram-Schmidt à base anterior (ver o slide 154 - os vetores são os mesmos) obtém-se:

$$\begin{aligned} v_1 &= u_1 = (1, -1, 1) \\ v_2 &= u_2 - \operatorname{proj}_{v_1}(u_2) = u_2 - \frac{u_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \frac{1}{3}(1, 2, 1) \rightsquigarrow (1, 2, 1) \end{aligned}$$

Uma base ortogonal de V é portanto $\{v_1, v_2\} = \{(1, -1, 1), (1, 2, 1)\}$, tendo-se (ver o teorema do slide 187),

$$\begin{aligned} \operatorname{proj}_V(b) &= \operatorname{proj}_{v_1}(b) + \operatorname{proj}_{v_2}(b) = \frac{v_1 \cdot b}{v_1 \cdot v_1} v_1 + \frac{v_2 \cdot b}{v_2 \cdot v_2} v_2 \\ &= \frac{3}{3}(1, -1, 1) + \frac{6}{6}(1, 2, 1) = (2, 1, 2). \end{aligned}$$

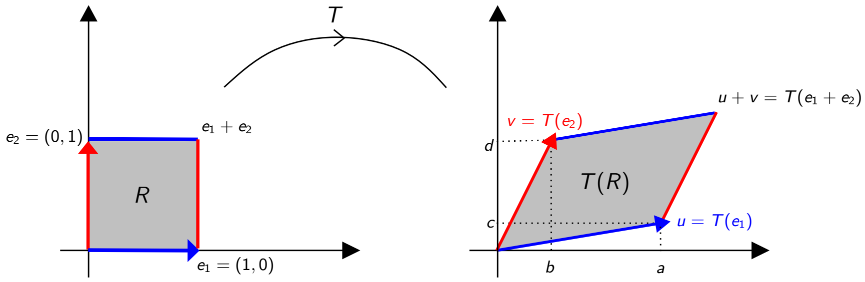
- Calculando $V^\perp = \mathcal{C}(A)^\perp = \mathcal{N}(A^T)$ com $A = [\, u_1 \;\; u_2 \,]$ (também se pode considerar $A = [\, v_1 \;\; v_2 \,]$), obtém-se $V^\perp = \langle (-1, 0, 1) \rangle$.
Tem-se portanto a base ortogonal([13]) $\{w\} = \{(-1, 0, 1)\}$ de $V^\perp$.
- Reunindo a base ortogonal de V com a base ortogonal de $V^\perp$ obtém-se a base ortogonal de $\mathbb{R}^3$ que estende a base ortogonal de V,

$$\{v_1, v_2, w\} = \{(1, -1, 1), (1, 2, 1), (-1, 0, 1)\}.$$

[13] Uma base de um subespaço vetorial de dimensão um é sempre ortogonal.

Motivação do determinante: caso das matrizes 2×2

- Consideremos uma transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$, definida por uma matriz $A_{2\times 2} = [\, u \;\, v \,] = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, isto é, $T(x) = Ax$.
- Tem-se $T(e_1) = u = \begin{bmatrix} a \\ c \end{bmatrix}$, $T(e_2) = v = \begin{bmatrix} b \\ d \end{bmatrix}$ e pode-se mostrar que T envia o quadrado unitário R no paralelogramo $T(R)$ definido por u e v,



- O determinante da matriz A é como veremos,

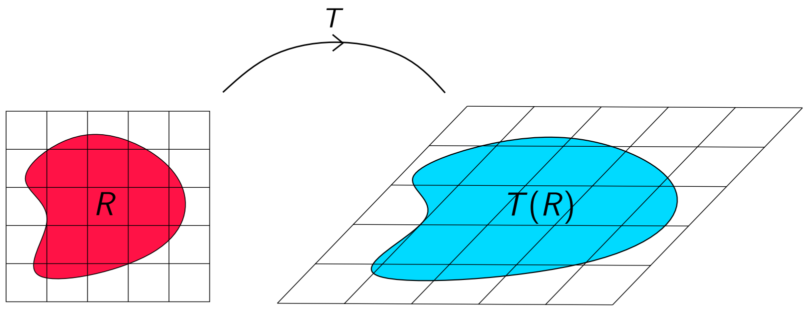
$$\det(A) = ad - bc,$$

e verifica a propriedade (ver o exercício 36 da sebenta),

$$\boxed{\text{área de } T(R) = |\det(A)| \;\times\; \text{área de } R = |\det(A)|.}$$

Extensão a outras de regiões do plano

- A propriedade anterior estende-se a outras regiões R do plano com "boas propriedades":



- Aproximando sucessivamente a área de uma região R usando uma grelha cada vez mais fina e aplicando a fórmula do slide 193 às áreas das imagens dos quadrados dessa grelha, mostra-se que se tem

$$\boxed{\text{área de } T(R) = |\det(A)| \;\times\; \text{área de } R.}$$

Desafio

A partir da área do círculo de raio 1 deduza a fórmula da área da elipse de semi-eixos a e b, aplicando uma transformação linear conveniente.

Motivação do determinante: caso das matrizes 3×3

- Considerando agora a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^3$, definida por uma matriz $A_{3\times 3} = [\,u \ v \ w\,]$, tem-se que o valor absoluto do determinante de A verifica a relação análoga,

$$\boxed{\text{volume de } T(R) = |\det(A)| \ \times \ \text{volume de } R = |\det(A)|,}$$

  onde R é o cubo unitário definido pelos vetores da base canónica de $\mathbb{R}^3$ (cujo volume é 1) e $T(R)$ o paralelipípedo definido pelos vetores u, v e w, obtido como imagem por meio de T de R.

- Esta relação pode ser estendida a regiões do espaço R mais gerais com "boas propriedades", tendo-se ainda

$$\boxed{\text{volume}(T(R)) = |\det(A)| \ \times \ \text{volume}(R).}$$

- A expressão para o determinante de uma matriz quadrada A de ordem 3 é mais complicada e será dada mais adiante no slide 198.

Determinante de matrizes de ordem ≤ 2

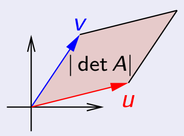
- Se $n = 1$, define-se,

$$\det [a] = a.$$

- Se $n = 2$, define-se,

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

Interpretação geométrica do determinante de matrizes 2×2



Recordemos que o valor absoluto do determinante de

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = [\,u \quad v\,], \text{ com } u = (a, c) \text{ e } v = (b, d),$$

corresponde à área do paralelogramo definido por u e v. Esta área é não nula se e só se u e v são não colineares!

Determinante de matrizes de ordem 2 - exemplo

Por exemplo, o determinante da matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ é,

$$\det(A) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = 1 \times 4 - 2 \times 3 = -2 \neq 0$$

e em particular, o paralelogramo definido pelos vetores $u = (1, 3)$ e $v = (2, 4)$ tem área 2.

Determinante de matrizes 3×3: regra de Sarrus

$$\det(A) = \det \left[\begin{matrix} a & b & c \\ d & e & f \\ g & h & i \end{matrix}\right] \begin{matrix} a & b \\ d & e \\ g & h \end{matrix}$$

$$= aei + bfg + cdh - (ceg + afh + bdi)$$

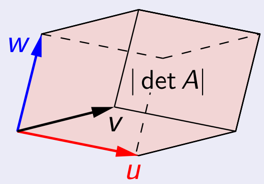
Por exemplo,

$$\det A = \det \begin{bmatrix} 1 & 0 & -2 \\ 2 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} \begin{matrix} 1 & 0 \\ 2 & 1 \\ -1 & 3 \end{matrix} = 1 \times 1 \times 1 + 0 \times 1 \times (-1) + (-2) \times 2 \times 3$$

$$-((-2) \times 1 \times (-1) + 1 \times 1 \times 3 + 0 \times 2 \times 1) = 1 + 0 - 12 - (2 + 3 + 0) = -16 \neq 0$$

Determinantes 3x3

Interpretação geométrica do determinante de matrizes 3×3



Recordemos que o valor absoluto do determinante da matriz $A = [\, u \;\; v \;\; w \,]_{3\times 3}$ com $u, v, w \in \mathbb{R}^3$ corresponde ao volume do paralelipípedo definido por u, v e w, tendo-se que este volume é não nulo se e só se o paralelipípedo for não degenerado, ou seja, u, v, w forem não complanares.

- Por exemplo, o paralelipípedo definido pelas 3 colunas da matriz A do slide anterior, $u = (1, 2, -1)$, $\mathbf{v} = (\mathbf{0}, \mathbf{1}, \mathbf{3})$ e $w = (-2, 1, 1)$, tem volume 16.
- A regra de Sarrus só se aplica a matrizes 3x3!
- E no caso geral de matrizes $n \times n$, com n arbitrário?

Menores e co-factores

Definições de menor complementar e co-factor

Sejam A uma matriz quadrada de ordem n e $1 \leq i, j \leq n$.

- Chama-se *menor complementar da entrada* (i, j), denotado por A_{ij}, ao determinante da submatriz que se obtém eliminando a linha i e a coluna j de A.
- Chama-se *complemento algébrico ou co-factor da entrada* (i, j) a

$$\Delta_{ij} = (-1)^{i+j} A_{ij}.$$

Por exemplo, o menor complementar da entrada $(1, 2)$ de $A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 2 & 3 \\ -1 & 2 & 1 \end{bmatrix}$, é o determinante da submatriz que se obtém eliminando a linha 1 e coluna 2 de A, isto é,

$$A_{12} = \det \begin{bmatrix} 0 & 3 \\ -1 & 1 \end{bmatrix} = 0 \times 1 - 3 \times (-1) = 3,$$

e o co-factor da entrada $(1, 2)$ é $\Delta_{12} = (-1)^{1+2} A_{12} = (-1) \times 3 = -3$.

Regra de Laplace

Teorema (Regra de Laplace)

Seja $A = [a_{ij}]$ uma matriz quadrada de ordem $n \geq 2$. Então

- Para qualquer $i = 1, \ldots, n$, tem-se

$$\det A = a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in}.$$

(Expansão do determ. em co-factores ao longo da linha i)

- Para qualquer $j = 1, \ldots, n$, tem-se

$$\det A = a_{1j}\Delta_{1j} + a_{2j}\Delta_{2j} + \cdots + a_{nj}\Delta_{nj}.$$

(Expansão do determ. em co-factores ao longo da coluna j)

- A regra de Laplace reduz o cálculo do determinante de uma matriz $n \times n$ ao cálculo de n determinantes de matrizes $(n-1) \times (n-1)$.
- Devem escolher-se linhas ou colunas com o maior número possível de zeros.
- O resultado não depende da escolha da linha ou da coluna.

Regra de Laplace: exemplo 3 × 3

Consideremos a matriz $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 2 & 3 \\ -1 & 2 & 1 \end{bmatrix}$.

- Expandindo o determinante ao longo da 2ª linha tem-se,

$$\begin{aligned} \det A &= a_{21}\Delta_{21} + a_{22}\Delta_{22} + a_{23}\Delta_{23} = 0(-1)^{2+1}\det\begin{bmatrix} -2 & 3 \\ 2 & 1 \end{bmatrix} + \\ & \quad 2(-1)^{2+2}\det\begin{bmatrix} 1 & 3 \\ -1 & 1 \end{bmatrix} + 3(-1)^{2+3}\det\begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} \\ &= 0 + 2 \times (1+3) - 3 \times 0 = 8. \end{aligned}$$

- Expandindo o determinante ao longo da 1ª coluna tem-se,

$$\begin{aligned} \det A &= a_{11}\Delta_{11} + a_{21}\Delta_{21} + a_{31}\Delta_{31} = 1(-1)^{1+1}\det\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} + \\ & \quad 0(-1)^{2+1}\det\begin{bmatrix} -2 & 3 \\ 2 & 1 \end{bmatrix} + (-1)(-1)^{3+1}\det\begin{bmatrix} -2 & 3 \\ 2 & 3 \end{bmatrix} \\ &= 1 \times (2-6) + 0 + (-1) \times (-6-6) = 8. \end{aligned}$$

Regra de Laplace: exemplo 4×4

Exercício na aula

Calcular o determinante de

$$A = \begin{bmatrix} -1 & 2 & -1 & 0 \\ 0 & 1 & -1 & 1 \\ 2 & 0 & 1 & 0 \\ 0 & 1 & 2 & -1 \end{bmatrix}$$

Resolução: aplicando a regra de Laplace ao longo da terceira linha que possui 2 zeros, obtém-se:

$$\begin{aligned} \det A &= a_{31}\Delta_{31} + a_{32}\Delta_{32} + a_{33}\Delta_{33} + a_{34}\Delta_{34} \\ &= 2\,(-1)^{3+1} \begin{vmatrix} 2 & -1 & 0 \\ 1 & -1 & 1 \\ 1 & 2 & -1 \end{vmatrix} + 0 + 1\,(-1)^{3+3} \begin{vmatrix} -1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{vmatrix} + 0 \\ &= 2(-4) + 0 + 1 \cdot 2 + 0 = -6. \end{aligned}$$

TPC: confirme os valores dos 2 determinantes 3×3 do cálculo anterior, usando a regra de Sarrus no primeiro e a regra de Laplace no segundo.

Consequências da regra de Laplace

Tem-se (ver também o exercício 38 da sebenta):

- Se A possui uma linha ou uma coluna de zeros então $\det A = 0$.
- Se A possui linhas ou colunas múltiplas entre si então $\det A = 0$.
- Se A é uma matriz triangular superior (ou inferior) então $\det A =$ produto dos elementos da diagonal principal:

$$\det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} = a_{11}a_{22}\cdots a_{nn}.$$

Em particular,

- $\det\left(\mathrm{diag}(a_1, \ldots, a_n)\right) = \det \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix} = a_1 a_2 \cdots a_n.$
- $\det(\alpha I_n) = \det\left(\mathrm{diag}(\alpha, \ldots, \alpha)\right) = \alpha^n.$
- $\det I_n = 1.$

Propriedades do determinante

Proposição

Sejam A e B matrizes quadradas de ordem n e $\alpha \in \mathbb{R}$. Tem-se:

- $\det(AB) = \det A \det B$.
- $\det(A^T) = \det A$.
- $\det(\alpha A) = \alpha^n \det A$ (atenção!)

Em geral, $\det(A + B) \neq \det A + \det B$ (atenção!)

- O 3º ponto é consequência do 1º ponto e dos resultados do slide anterior:
$$\det(\alpha A) = \det(\alpha(I_n\, A)) = \det((\alpha I_n)\, A) = \det(\alpha I_n)\, \det A = \alpha^n \det A.$$
- Para termos uma ideia da justificação do 1º ponto vamos admitir que A e B têm ordem 2. Se T_A e T_B são as transformações lineares definidas por essas matrizes, R é o quadrado unitário definido pela base canónica de $\mathbb{R}^2$ e $S = T_B(R)$ a imagem desse quadrado por T_B, tem-se (ver os slides 193 e 78),
$$\begin{aligned} |\det(AB)| &= \text{área de } T_{AB}(R) = \text{área de } (T_A \circ T_B)(R) \\ &= \text{área de } T_A(T_B(R)) = \text{área de } T_A(S) \\ &= |\det A| \times \text{ área de } S = |\det A|\, |\det B|. \end{aligned}$$

Determinante e inversa

- Vimos antes que o valor absoluto do determinante de uma matriz 2×2 [3×3] correspondia à área do paralelogramo [volume do paralelipípedo] definido pelas colunas dessa matriz.
- Logo o determinante dessa matriz é não nulo se e só se as suas colunas forem não colineares [não complanares], isto é, definirem um conjunto linearmente independente de vetores, ou seja, A for invertível.
- Tem-se um resultado análogo para matrizes quadradas de ordem arbitrária, como veremos no próximo slide.

Invertibilidade de uma matriz e determinante

Teorema

Seja $A = [\, v_1 \ v_2 \ \cdots \ v_n \,]$, com $v_i \in \mathbb{R}^n$, uma matriz quadrada de ordem n. As seguintes afirmações são equivalentes:

- A é invertível.
- $\text{car}(A) = n$.
- $\{v_1, \ldots, v_n\}$ é linearmente independente.
- $\{v_1, \ldots, v_n\}$ é base de $\mathbb{R}^n$.
- $\det(A) \neq 0$.

Nas condições anteriores tem-se ainda,

$$\boxed{\det(A^{-1}) = \frac{1}{\det A}.}$$

Para deduzir $\det(A^{-1})$ basta notar que pelo primeiro ponto do slide 205,

$$1 = \det I_n = \det(A\,A^{-1}) = \det A \,\det(A^{-1}).$$

Valores e vetores próprios - motivação

Consideremos a matriz $A = \begin{bmatrix} 1/2 & 3/2 \\ 0 & 2 \end{bmatrix}$.

- Tem-se:

$$\begin{aligned} A \begin{bmatrix} x_1 \\ 0 \end{bmatrix} &= \begin{bmatrix} 1/2 & 3/2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2}x_1 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} x_1 \\ 0 \end{bmatrix}, \\ A \begin{bmatrix} x_1 \\ x_1 \end{bmatrix} &= \begin{bmatrix} 1/2 & 3/2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}x_1 + \frac{3}{2}x_1 \\ 2x_1 \end{bmatrix} = 2 \begin{bmatrix} x_1 \\ x_1 \end{bmatrix}. \end{aligned}$$

- Logo a transformação linear definida por A, $T(x) = Ax$, induz uma contração de razão $\lambda_1 = \frac{1}{2}$ na direção do eixo do x_1 e uma dilatação de razão $\lambda_2 = 2$ na direção da bisssectriz dos quadrantes ímpares $x_2 = x_1$.
- Logo 1º eixo coordenado fica **invariante** por ação da transformação linear definida por A e qualquer vetor diretor u deste eixo (por exemplo $u = (1, 0)$) verifica a relação $Au = \frac{1}{2}u$, dizendo-se nessa altura que u é um **vetor próprio** de A associado ao **valor próprio** $\lambda_1 = \frac{1}{2}$.
- Analogamente, a bissectriz dos quadrantes ímpares é **invariante** por ação da transformação definida por A e qualquer vetor diretor v desta bissectriz (por exemplo, $v = (1, 1)$), verifica a relação $Av = 2v$, dizendo-se nessa altura que v é um **vetor próprio** de A associado ao **valor próprio** $\lambda_2 = 2$.

Conceitos de vetor e valor próprio

Definições de vetor próprio e valor próprio

Sejam A matriz quadrada de ordem n, $v \in \mathbb{R}^n$ com $v \neq \vec{0}$.
Diz-se que v é um vetor próprio de A se existir $\lambda \in \mathbb{R}$ tal que

$$\boxed{Av = \lambda\, v}$$

λ designa-se por valor próprio associado ao vetor próprio v.

Exemplo

Considerando $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$, tem-se que $v = (1,1,1)$ é vetor próprio de A associado ao valor próprio $\lambda = 2$ uma vez que,

$$A\,v = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 2\, v$$

Polinómio característico de uma matriz

Observação

Se A é uma matriz quadrada de ordem n e λ uma variável então a expressão $\det(A - \lambda I)$ define um polinómio de grau n na variável λ, que se designa por polinómio característico de A e se denota $p_A(\lambda)$.

A importância do polinómio característico fica evidente no próximo resultado.

Teorema

Tem-se que $\alpha \in \mathbb{R}$ é valor próprio de uma matriz quadrada A se e só se $p_A(\alpha) = 0$, isto é, α for raíz do polinómio $p_A(\lambda)$.

Demonstração: $\alpha \in \mathbb{R}$ é valor próprio de uma matriz A $\Leftrightarrow$ existe um vetor próprio $v \neq \vec{0}$ tal que $Av = \alpha v \Leftrightarrow Av - \alpha v = \vec{0} \Leftrightarrow$ $(A - \alpha I)v = \vec{0} \Leftrightarrow (A - \alpha I)x = \vec{0}$ admite uma solução $v \neq \vec{0} \Leftrightarrow$ $(A - \alpha I)$ não invertível $\Leftrightarrow p_A(\alpha) = \det(A - \alpha I) = 0$. $\square$

Exemplo

Exemplo

Consideremos $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ e vejamos que $\lambda = 2$ é valor próprio de A como se concluiu no slide 209. De facto, tem-se

$$\begin{aligned} p_A(2) &= \det(A - 2I) = \det\left| \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \right| \\ &= \begin{vmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{vmatrix} = 0, \end{aligned}$$

uma vez que a matriz possui uma linha de zeros.

Multiplicidade algébrica de um valor próprio

Observação

Pelo teorema do slide 210 os valores próprios de uma matriz A de ordem n são as raízes (reais e complexas) de $p_A(\lambda)$, que podem ser repetidas.

Definição de multiplicidade algébrica de um valor próprio

Chama-se multiplicidade algébrica de um valor próprio λ, denotada m.a.(λ), ao número de vezes que λ aparece repetido como raíz na factorização de $p_A(\lambda)$.

Exemplo

Consideremos a matriz $A = \begin{bmatrix} 1 & 1 \\ -4 & 5 \end{bmatrix}$. Tem-se,

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 \\ -4 & 5-\lambda \end{vmatrix} \\ &= (1-\lambda)(5-\lambda) + 4 = \lambda^2 - 6\lambda + 9 = (\lambda - 3)^2. \end{aligned}$$

Logo $p_A(\lambda)$ admite apenas raíz dupla $\lambda = 3$ e portanto A tem apenas o valor próprio $\lambda = 3$ com multiplicidade algébrica 2 (m.a.(3)=2).

Subespaço próprio e multiplicidade geométrica

Definição de subespaço próprio e multiplicidade geométrica

Sejam A matriz quadrada de ordem n e $\lambda \in \mathbb{R}$ valor próprio de A. Chama-se subespaço próprio de A associado a λ ao subespaço vetorial,

$$E(\lambda) = \mathcal{N}(A - \lambda I).$$

A dimensão de $E(\lambda)$ designa-se por multiplicidade geométrica de λ e denota-se por $\text{m.g.}(\lambda)$.

Teorema

Os vetores próprios de A associados a um dado valor próprio λ de A são os vetores não nulos do subespaço próprio $E(\lambda)$.

Demonstração: de facto,

$$\begin{aligned} v \in \mathbb{R}^n \text{ é vetor próprio de } A \text{ associado a } \lambda \quad &\Leftrightarrow \quad Av = \lambda v, \ v \neq \vec{0} \\ &\Leftrightarrow \quad (A - \lambda I)v = \vec{0}, \ v \neq \vec{0} \\ &\Leftrightarrow \quad v \in \mathcal{N}(A - \lambda I) \setminus \{\vec{0}\}. \quad \square \end{aligned}$$

Exemplo do slide 209 revisitado

Vamos aplicar os conceitos dos slides anteriores à matriz $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

- Tem-se,

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I) = \det\left(\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & & \\ & \lambda & \\ & & \lambda \end{bmatrix} \right) \\ &= \det \begin{bmatrix} 1-\lambda & 1 & 0 \\ 0 & 2-\lambda & 0 \\ 0 & 1 & 1-\lambda \end{bmatrix} \\ &\quad \underset{\text{Laplace na } 1^{\underline{a}} \text{ coluna}}{\uparrow} \\ &= (1-\lambda)(-1)^{1+1} \det \begin{bmatrix} 2-\lambda & 0 \\ 1 & 1-\lambda \end{bmatrix} + 0 + 0 \\ &= (1-\lambda)(2-\lambda)(1-\lambda) = (1-\lambda)^2(2-\lambda). \end{aligned}$$

- Logo $p_A(\lambda)$ admite a raíz **dupla** $\lambda = 1$ uma vez que aparece repetida 2 vezes na factorização do polinómio e a raíz **simples** $\lambda = 2$.
- Portanto A admite valores próprios distintos, $\lambda = 1, 2$, com $\text{m.a.}(1) = 2$ e $\text{m.a.}(2) = 1$.

Exemplo do slide 209 revisitado (cont.)

Relativamente ao subespaço próprio $E(1) = \mathcal{N}(A - I)$ tem-se:

- Aplicando o método de Gauss vem,

$$\left[\, A - I \,|\, \vec{0} \,\right] = \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \text{ e portanto,}$$

$$\begin{aligned} \mathcal{N}(A - I) &= \{(x_1, x_2, x_3) : x_2 = 0, x_1, x_3 \in \mathbb{R}\} \\ &= \{(x_1, 0, x_3) : x_1, x_3 \in \mathbb{R}\} = \big\langle (1,0,0), (0,0,1) \big\rangle. \end{aligned}$$

- Logo $E(1) = \langle (1,0,0), (0,0,1) \rangle$. Uma base para $E(1)$ é portanto $\{(1,0,0), (0,0,1)\}$, tendo-se $\text{m.g.}(1) = \dim E(1) = 2$.
- Geometricamente $E(1)$ define o plano de $\mathbb{R}^3$ que passa na origem com vetores diretores $(1,0,0)$ e $(0,0,1)$.
- Os vetores próprios de A associados ao valor próprio $\lambda = 1$ são os vetores <u>não nulos</u> de $E(1)$. Por exemplo, tomando $x_1 = 1$ e $x_3 = -2$ obtém-se o vetor próprio $(1,0,2)$ de A associado a $\lambda = 1$.

Exemplo do slide 209 revisitado (concl.)

Relativamente ao subespaço próprio $E(2) = \mathcal{N}(A - 2I)$ tem-se:

- $A - 2I = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} - \begin{bmatrix} 2 & & \\ & 2 & \\ & & 2 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}.$
- Aplicando o método de Gauss ao sistema $\left[\, A - 2I \,|\, \vec{0} \,\right]$ obtém-se,

$$\left[\begin{array}{ccc|c} -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

- Logo, $E(2) = \mathcal{N}(A - 2I) = \{(x_3, x_3, x_3) : x_3 \in \mathbb{R}\} = \big\langle (1,1,1) \big\rangle$, e uma base para $E(2)$ é $\{(1,1,1)\}$, tendo-se $\text{m.g.}(2) = \dim E(2) = 1$.
- Geometricamente $E(2)$ é a reta que passa na origem com vetor diretor $(1,1,1)$.
- Os vetores próprios de A associados ao valor próprio $\lambda = 2$ são os vetores <u>não nulos</u> de $E(2)$, isto é, os vetores da forma (a, a, a) com $a \neq 0$.

A informação dita **espectral** sobre a matriz A pode ser organizada numa tabela:

λ	m.a.(λ)	m.g.(λ)	base de $E(\lambda)$
1	2	2	$\{(1,0,0), (0,0,1)\}$
2	1	1	$\{(1,1,1)\}$

Só para relembrar :)

Resumo

- Reconhecer / verificar que $v \neq \vec{0}$ é vetor próprio de A
  $\longrightarrow$ Mostrar que $Av = \lambda v$ para algum $\lambda \in \mathbb{R}$.
  λ é o valor próprio associado a v.

- Reconhecer / verificar que $\alpha \in \mathbb{R}$ é valor próprio de A
  $\longrightarrow$ Mostrar que $p_A(\alpha) = det(A - \alpha I) = 0$.

- Determinar os valores próprios de A
  $\longrightarrow$ Determinar as raízes (reais e complexas) de $p_A(\lambda) = \det(A - \lambda I)$.
  A multiplicidade algébrica de cada valor próprio λ, m.a.(λ), é o número de vezes que λ aparece repetido como raíz na factorização do polinómio característico $p_A(\lambda)$.

- Determinar os vetores próprios de A associados ao valor próprio λ
  $\longrightarrow$ Determinar os vetores <u>não nulos</u> de $E(\lambda) = \mathcal{N}(A - \lambda I)$.
  A multiplicidade geométrica de λ é m.g.$(\lambda) = \dim E(\lambda)$.

Propriedades dos valores próprios

Teorema

Seja A uma matriz quadrada de ordem n. Então:

- Para todo o valor próprio λ de A tem-se
  $$\boxed{1 \leq \text{m.g.}(\lambda) \leq \text{m.a.}(\lambda)}$$
- A matriz A possui n valores próprios (reais e/ou complexos) contando com repetições, ou seja, a soma das multiplicidades algébricas dos valores próprios **distintos** de A é igual à ordem da matriz A.
- A soma dos valores próprios de A, contando com repetições (m.a.), é igual ao traço de A, $\text{tr}(A)$, que se define como a soma dos elementos da diagonal principal de A.
- O produto dos valores próprios de A, contando com repetições (m.a.), é igual ao $\det A$.
- $\lambda = 0$ é valor próprio de A se e só se A é não invertível.

Propriedades dos valores próprios - exemplo

Consideremos a matriz A do exemplo do slide 209 e a respectiva informação espectral do slide 216,

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

λ	m.a.(λ)	m.g.(λ)	base de $E(\lambda)$
1	2	2	$\{(1,0,0),(0,0,1)\}$
2	1	1	$\{(1,1,1)\}$

Constata-se que:

- $2 = \text{m.g.}(1) \leq \text{m.a.}(1) = 2$
  $1 = \text{m.g.}(2) \leq \text{m.a.}(2) = 1$.
- $\text{m.a.}(1) + \text{m.a.}(2) = 2 + 1 = 3 = n$ (ordem da matriz A).
- A soma dos valores próprios de A contando com repetições (m.a.), $1 + 1 + 2 = 4$, coincide com $tr(A) = 1 + 2 + 1 = 4$ (soma das entradas da diagonal principal).
- O produto dos valores próprios de A, contando com repetições (m.a.), $1 \times 1 \times 2$ coincide com $\det(A) = 2$ (verifique).
- Como $\lambda = 0$ não é valor próprio a matriz A é invertível.

Diagonalização de matrizes

Definição de matriz diagonalizável

Uma matriz A de ordem n diz-se diagonalizável se existir uma matriz invertível P e uma matriz diagonal D tal que

$$P^{-1} A P = D.$$

A matriz P designa-se por matriz de diagonalização para A.

Exemplo

Consideremos $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ e $P = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Tem-se (verifique),

$$P^{-1} A P = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = D.$$

Logo A é diagonalizável com matriz de diagonalização P.
Como obter uma matriz P de diagonalização (caso exista)?

Diagonalização e base própria

Seja A uma matriz quadrada de ordem n e $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$ uma matriz diagonal. Dada uma matriz invertível $P = [\, v_1 \quad \cdots \quad v_n \,]$, isto é, uma matriz de uma base $\{v_1, \ldots, v_n\}$ de $\mathbb{R}^n$, têm-se as equivalências

$$P^{-1}AP = D \quad \Leftrightarrow \quad P\,P^{-1}AP = PD \quad \Leftrightarrow \quad AP = PD.$$

Como se tem,

$$\begin{aligned} AP &= A[\, v_1 \;\; v_2 \;\; \cdots \;\; v_n \,] = [\, Av_1 \;\; Av_2 \;\; \cdots \;\; Av_n \,], \quad \text{e} \\ PD &= [\, v_1 \;\; v_2 \;\; \cdots \;\; v_n \,] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\, \lambda_1 v_1 \;\; \lambda_2 v_2 \;\; \cdots \;\; \lambda_n v_n \,], \end{aligned}$$

conclui-se que $AP = PD$ se só se $Av_i = \lambda_i v_i$, para todo o $i = 1, \ldots, n$.

Logo A é diagonalizável com matriz de diagonalização $P = [\, v_1 \quad \cdots \quad v_n \,]$ tal que

$$P^{-1}AP = D = \text{diag}(\lambda_1, \ldots, \lambda_n),$$

se e só se $\{v_1, \ldots, v_n\}$ for uma base de $\mathbb{R}^n$ formada por vetores próprios de A associados aos valores próprios $\lambda_1, \ldots, \lambda_n$.

Propriedades dos vetores próprios

Observação

Sejam A matriz quadrada de ordem n e $\lambda'_1, \ldots, \lambda'_k$ os valores próprios distintos de A.

- Pode-se mostrar que existe uma base de $\mathbb{R}^n$ formada por vetores próprios de A se e só se

  $$\text{m.g.}(\lambda'_1) + \cdots + \text{m.g.}(\lambda'_k) = n.$$

  Esta base é obtida reunindo bases dos subespaços próprios $E(\lambda'_1), \ldots, E(\lambda'_k)$.

- Uma vez que soma das mult. alg. dos valores próprios distintos de A é igual à ordem da matriz A e que a multiplicidade geométrica de qualquer valor próprio é sempre inferior ou igual à sua multiplicidade algébrica, a condição anterior é equivalente à condição

  $$\text{m.g.}(\lambda'_i) = \text{m.a.}(\lambda'_i), \quad i = 1, \ldots, k.$$

Critérios de diagonalização

Teorema

Seja A uma matriz quadrada de ordem n. As seguintes afirmações são equivalentes:

(i) A é diagonalizável.

(ii) Existe uma base de $\mathbb{R}^n$ formada por vetores próprios de A.

(iii) A soma das multiplicidades geométricas dos valores próprios distintos de A é n.

(iv) $\text{m.g.}(\lambda) = \text{m.a.}(\lambda)$ para qualquer valor próprio λ de A.

Nas condições equivalentes anteriores a matriz $P = [\, v_1 \;\; \cdots \;\; v_n \,]$, onde $\{v_1, \ldots, v_n\}$ é uma base de $\mathbb{R}^n$ formada por vetores próprios de A, obtida reunindo bases de todos os subespaços próprios de A, é uma matriz de diagonalização para A, tendo-se

$$P^{-1}AP = \text{diag}(\lambda_1, \ldots, \lambda_n),$$

com λ_i, $i = 1, \ldots, n$, valor próprio de A associado ao vetor próprio v_i.

Exemplo de uma matriz diagonalizável

Consideremos novamente a matriz A do exemplo do slide 209 e a respectiva informação espectral do slide 216,

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

λ	$m.a.(\lambda)$	$m.g.(\lambda)$	base de $E(\lambda)$
1	2	2	$\{(1,0,0),(0,0,1)\}$
2	1	1	$\{(1,1,1)\}$

- Uma vez que $\text{m.g.}(1) = \text{m.a.}(1) = 2$ e a $\text{m.g.}(2) = \text{m.a.}(2) = 1$, a matriz A é **diagonalizável** e o conjunto

$$\{u_1, u_2, u_3\} = \{(1,0,0,),(0,0,1),(1,1,1)\},$$

obtido reunindo a base $\{(1,0,0),(0,0,1)\}$ de $E(1)$ com a base $\{(1,1,1)\}$ de $E(2)$ é uma **base de $\mathbb{R}^3$ formada por vetores próprios de** A.

- Logo a matriz desta base própria, $P = [\, u_1 \;\; u_2 \;\; u_3 \,] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ é uma matriz de diagonalização para A, tendo-se (verifique),

$$P^{-1}AP = \begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Exemplo de uma matriz não diagonalizável

- Consideremos agora a matriz $D = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 5 \end{bmatrix}$ do exercício 44, cujos valores próprios distintos são $\lambda = 1$ e $\lambda = 6$, tendo-se m.a.$(1) = 2 >$ m.g.$(1) = 1$ e m.a.$(6) =$ m.g.$(6) = 1$ (ver a solução do exercício 44.)
- Como m.g.$(1) \neq$ m.a.(1) não existe uma base de $\mathbb{R}^3$ formada por vetores próprios de D e portanto D não é diagonalizável.
- Neste caso a cardinalidade (número de vetores) máxima de um conjunto linearmente independente formado por vetores próprios de D é m.g.$(1) +$ m.g.$(6) = 2 < 3 = \dim \mathbb{R}^3$.

Introdução à Programação Linear (PL)

Problema 1

Uma exploração agrícola dispõe de 80 ha de terreno para produzir tomate e trigo. Para além do terreno, os recursos susceptíveis de limitar a produção das duas culturas são a água e a mão de obra: sabe-se que **cada hectare de tomate** necessita de 8000 m^3 de água e de 40 h de mão de obra e que **cada hectare de trigo** requer apenas 20 h de mão de obra. A exploração agrícola dispõe de 320000 m^3 de água e 2000 horas de mão de obra. As receitas, por cada hectare de tomate e trigo cultivados são, respetivamente, 300 € e 200 €. Pretende-se determinar a **área a destinar a cada cultura** por forma a maximizar a receita total.

Dados do problema:

	Utilização de recursos		
	Água	Mão de obra	Receita (max)
Tomate	8000 m^3/ha	40 h/ha	300 €/ha
Trigo		20 h/ha	200 €/ha
Disponibilidades	$\leq$ 320000 m^3	$\leq$ 2000 h	
Terreno	$\leq$ 80 ha		

Construção do modelo matemático

Variáveis de decisão

Vamos considerar duas variáveis, x e y, que representam as áreas (em hectares) a destinar ao cultivo do tomate e do trigo, respetivamente.

Função objetivo

A função objetivo (f.o.) traduz a relação entre o valor da receita total (em €) e as receitas obtidas pelo cultivo de x hectares de tomate e y hectares de trigo:

$$z = 300\,x + 200\,y.$$

Restrições funcionais

As restrições funcionais traduzem as limitações dos recursos disponíveis:

- A área total de **terreno** cultivado não pode exceder 80 ha $\longrightarrow$ $x + y \leq 80$.
- O consumo de **água** não pode exceder 320000 m^3 $\longrightarrow$ $8000\,x \leq 320000$.
- A **mão de obra** utilizada não pode exceder 2000 h $\longrightarrow$ $40\,x + 20\,y \leq 2000$.

Restrições de sinal

Pela sua natureza as variáveis não podem tomar valores negativos $\longrightarrow$ $x, y \geq 0$.

Formulação do Problema 1 em PL

O Problema 1 pode então ser formulado em PL como,

$$\begin{array}{lrcll}
\max & z = 300\,x + 200\,y & & & \\
\text{s.a} & x + y & \leq & 80 & (T) \\
 & 8000\,x & \leq & 320000 & (A) \\
 & 40\,x + 20\,y & \leq & 2000 & (MO) \\
 & x, y & \geq & 0 &
\end{array}$$

em que

- x = área (em ha) destinada à cultura de tomate,
- y = área (em ha) destinada à cultura de trigo.

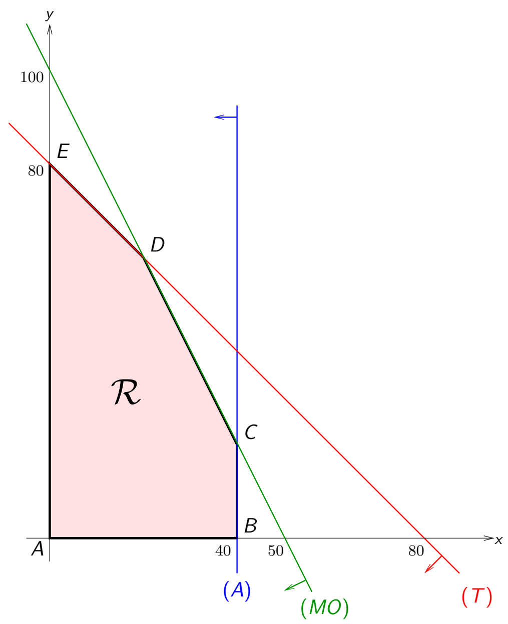
Repare-se que apesar da cultura de trigo não necessitar de água e requerer menos horas de mão de obra que a cultura de tomate, também gera menos receita, pelo que não é óbvia qual a área a destinar a cada uma das culturas de modo a maximizar a receita.

Solução e região admissível de um problema em PL

- A região admissível de um problema em PL é o conjunto das suas soluções admissíveis, isto é, o conjunto das soluções que satisfazem todas as restrições funcionais e de sinal.
- Dividindo a segunda restrição da formulação do slide anterior do Problema 1 do slide anterior por 8000 e a terceira por 20 obtêm-se restrições lineares mais simples, passando a região admissível $\mathcal{R}$ do Problema 1 a ser definida por:

$$\begin{aligned} x + y &\leq 80 \\ x \quad &\leq 40 \\ 2x + y &\leq 100 \\ x, y &\geq 0 \end{aligned}$$

Região admissível do problema em PL



- A inequação linear $x + y \leq 80$ define o semi-plano (assinalado por meio de $\rightarrow$) que contém a origem (porque $0 + 0 \leq 80$) e cuja fronteira é a reta de suporte (a vermelho) de equação $x + y = 80$. Se $y = 0$ nesta equação, obtém-se $x = 80$ e se $x = 0$ então $y = 80$, concluindo-se que a reta de suporte intersecta os eixos coordenados nos pontos $(80, 0)$ e $(0, 80)$.
- A inequação $x \leq 40$ define o semi-plano (assinalado por meio de $\rightarrow$) com fronteira dada pela reta vertical de suporte $x = 40$ (a azul).
- A inequação $2x + y \leq 100$ define o semi-plano (assinalado por meio de $\rightarrow$), que contém a origem e cuja fronteira é a reta de suporte (a verde) de equação $2x + y = 100$, que intersecta os eixos coordenados em $(50, 0)$ e $(0, 100)$.

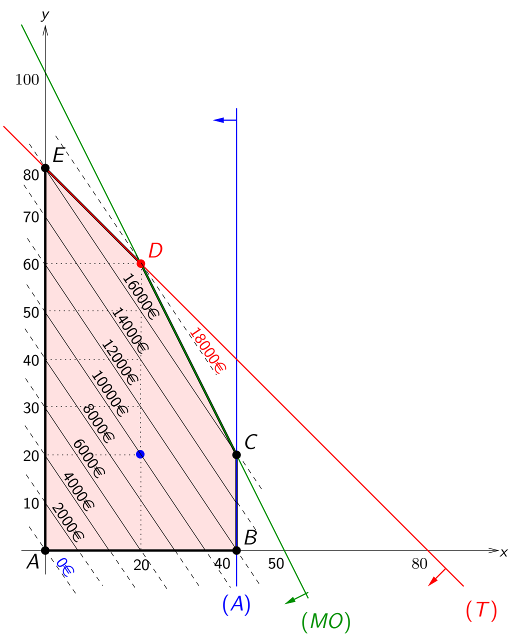
A região admissível $\mathcal{R}$ obtém-se intersectando os 3 semi-planos descritos acima com o primeiro quadrante definido pelas restrições de sinal $x, y \geq 0$, e define o polígono $[ABCDE]$.

Conjuntos de nível da função objetivo

- Dado $k \in \mathbb{R}$ define-se o conjunto de nível k da função objectivo (f.o.) $z = 300\,x + 200\,y$, como
$$C_k = \{(x,y)\ :\ 300\,x + 200\,y = k\},$$
que representa o conjunto dos pontos do plano em que a f.o. toma o valor k. Os conjuntos C_k, $k \in \mathbb{R}$, definem retas paralelas entre si, uma vez que são todas perpendiculares ao mesmo vetor normal $(300, 200)$.
- O conjunto das soluções que geram uma dada receita k€ é a parte do conjunto de nível C_k contida na região admissível $\mathcal{R}$, ou seja,
$$\{(x,y) \in \mathcal{R}\ :\ 300\,x + 200\,y = k\},$$
que pode obviamente ser vazia.
- Por exemplo, cultivar 20 hectares de tomate e 20 hectares de trigo corresponde à solução admissível $(20, 20) \in \mathcal{R}$ e gera uma receita de $k = 300x20 + 200x20 = 10000$€.
- O conjunto das soluções admissíveis que geram a mesma receita que a solução $(20, 20)$ é o conjunto
$$\{(x,y) \in \mathcal{R}\ :\ 300\,x + 200\,y = 10000\},$$
que corresponde à parte da reta de nível C_{10000} contida em $\mathcal{R}$.

Resolução gráfica do problema de PL

Representam-se na figura abaixo a solução admissível $(20, 20) \in \mathcal{R}$ e as retas de nível da f.o. para diferentes valores de receita k€, com a parte fora da região admissível a tracejado.



Torna-se evidente pela figura que o valor máximo da receita é atingido no vértice D, cujas coordenadas podem ser obtidas intersectando as retas de suporte $x + y = 80$ e $2x + y = 100$, isto é, como solução do sistema,
$$\begin{cases} x + y = 80 \\ 2x + y = 100 \end{cases} \Leftrightarrow \begin{cases} x = 20 \\ y = 60 \end{cases}$$

O vértice $D = (20, 60)$ designa-se por solução ótima do problema de PL e corresponde a cultivar 20 hectares de tomate e 60 ha de trigo, originando uma receita máxima de 18000€.

A solução ótima $D = (20, 60)$ utiliza a totalidade da mão de obra e do terreno disponíveis, uma vez que está na interseção das retas de suporte das correspondentes restrições funcionais, e portanto estas 2 restrições são satisfeitas com igualdade ($x + y = 20 + 60 = 80$ e $2x + y = 40 + 60 = 100$), dizendo-se que estão saturadas.

Vértices do polígono de admissibilidade e solução ótima

- As coordenadas de cada vértice da região admissível $\mathcal{R}$ do Problema 1 obtêm-se intersectando as retas de suporte que contêm o vértice. Por exemplo, as coordenadas do vértice C obtêm-se como

$$C: \quad \begin{cases} x = 40 \\ 2x + y = 100 \end{cases} \Leftrightarrow \begin{cases} x = 40 \\ y = 20 \end{cases}$$

- Veremos a seguir que uma solução ótima do problema de PL ocorre num vértice do polígono de admissibilidade $\mathcal{R}$, uma vez que a região admissível $\mathcal{R}$ é não vazia e limitada.
- Calculando o valor da função objectivo em cada vértice de $\mathcal{R}$ constata-se que o valor mais elevado é obtido no vértice D, concluindo-se novamente que uma solução ótima ocorre no vértice D:

vértice (x, y)	$z = 300x + 200y$
$A = (0, 0)$	0€
$B = (40, 0)$	12000€
$C = (40, 20)$	16000€
$D = (20, 60)$	18000€ (máx.)
$E = (0, 80)$	16000€

Formulação de um problema de PL: caso geral

- Num problema de PL, pretende-se determinar o(s) valor(es) de um conjunto de variáveis de decisão $x_1, \ldots, x_k$ que otimizam (maximizam ou minimizam), uma função linear designada por função objetivo (f.o.), e que satisfazem um conjunto de restrições funcionais (restrições lineares) (1),. . . , (m) e de sinal (m+1):

$$\begin{array}{llr}
\text{max / min} & z = c_1\, x_1 + c_2\, x_2 + \cdots + c_k\, x_k & \text{(f.o.)} \\
\quad s.a & a_{11}\, x_1 + a_{12}\, x_2 + \cdots + a_{1k}\, x_k \geq, \leq \text{ ou } = b_1 & (1) \\
 & a_{21}\, x_1 + a_{22}\, x_2 + \cdots + a_{2k}\, x_k \geq, \leq \text{ ou } = b_2 & (2) \\
 & \vdots & \\
 & a_{m1}\, x_1 + a_{m2}\, x_2 + \cdots + a_{mk}\, x_k \geq, \leq \text{ ou } = b_m & (m) \\
 & x_1 \geq 0, \leq 0 \text{ ou livre}, \ldots, x_k \geq 0, \leq 0 \text{ ou livre} & (m+1)
\end{array}$$

- c_j, a_{ij} e b_i, com $i = 1, \ldots, m$ e $j = 1, \ldots, k$, são os parâmetros do problema.
- O conjunto de pontos que satisfazem as restrições funcionais (1),. . . , (m) e as restrições de sinal (m+1) designa-se por região admissível do problema, denotada $\mathcal{R}$ e define um poliedro de $\mathbb{R}^k$ chamado poliedro de admissibilidade.
- Cada ponto da região admissível $\mathcal{R}$ designa-se por solução admissível.
- Uma solução admissível que otimize (maximize ou minimize) a f.o. designa-se por solução ótima.
- A cada restrição linear do tipo $a_{i1}\, x_1 + a_{i2}\, x_2 + \ldots + a_{ik}\, x_k \leq (\geq)\, b_i$ associamos a equação linear $a_{i1}\, x_1 + a_{i2}\, x_2 + \cdots + a_{ik}\, x_k = b_i$ que se designa por hiperplano de suporte da região admissível $\mathcal{R}$ se intersetar a fronteira de $\mathcal{R}$.

Vértices da região admissível e soluções ótimas

Teorema fundamental

Consideremos o problema $\mathcal{P}$ de programação linear,

$$\begin{array}{ll} \max \ (\text{ou } \min) & z = c_1 x_1 + c_2 x_2 + \cdots + c_k x_k \\ \text{s. a} & (x_1, x_2, \ldots, x_k) \in \mathcal{R}, \end{array}$$

com $c_1, \ldots, c_k \in \mathbb{R}$ e $\mathcal{R}$ região admissível definida por restrições funcionais e de sinal como descrito no slide 234. Se $\mathcal{R}$ for **limitada e não vazia** tem-se:

1. Existe um vértice de $\mathcal{R}$ que é solução ótima do problema $\mathcal{P}$.
2. Se q vértices $v_1, \ldots, v_q$ são soluções ótimas do problema $\mathcal{P}$ então qualquer combinação convexa de $v_1, \ldots, v_q$,

$$\lambda_1 v_1 + \cdots + \lambda_q v_q,$$

em que $\lambda_1, \ldots, \lambda_q \geq 0$ com $\lambda_1 + \cdots + \lambda_q = 1$, é ainda solução ótima de $\mathcal{P}$

Observações

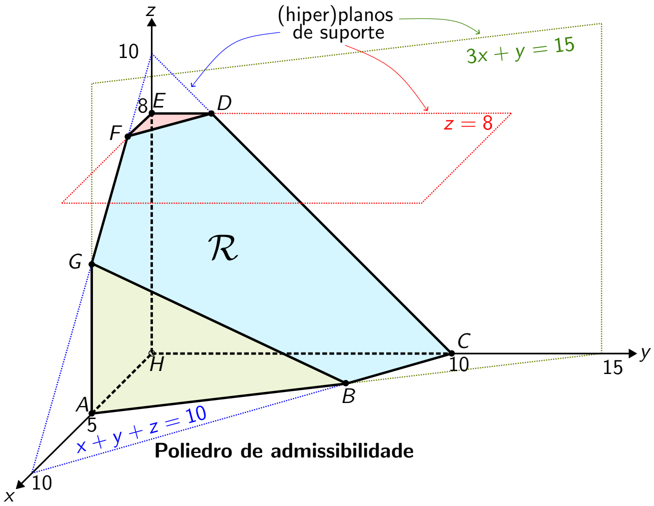
- O 1º ponto do teorema anterior reduz o problema de determinar uma solução ótima de um problema de PL com região admissível limitada e não vazia $\mathcal{R}$, ao problema de determinar os vértices de $\mathcal{R}$ (que são em número finito) e identificar o(s) vértice(s) onde a função objectivo atinge o maior ou menor valor, consoante o problema seja de maximização ou minimização.
- As combinações convexas de 2 vértices v_1 e v_2 são os pontos do segmento de reta que une v_1 a v_2. Logo pelo 2º ponto do teorema anterior se v_1 e v_2 são soluções ótimas de $\mathcal{P}$, qualquer ponto desse segmento de reta é ainda uma solução ótima de $\mathcal{P}$, que se designa por solução ótima alternativa.
- De modo análogo se conclui que se v_1, v_2 e v_3 são soluções ótimas não colineares de $\mathcal{P}$ então qualquer o ponto do triângulo de vértices v_1, v_2 e v_3 é uma solução ótima alternativa de $\mathcal{P}$.
- ...

Aplicação a um problema em PL com 3 variáveis

Consideremos o problema de PL,

$$\begin{array}{lrl} \max & 2x + y + z & \\ \text{s. a} & x + y + z & \leq 10 \\ & 3x + y & \leq 15 \\ & z & \leq 8 \\ & x, y, z & \geq 0 \end{array}$$

A região admissível é o poliedro $\mathcal{R}$ representado na seguinte figura.



Poliedro de admissibilidade

Valores da f.o. nos vértices do poliedro de admissibilidade

Como $\mathcal{R}$ é não vazia e limitada para obter uma solução ótima basta determinar o(s) vértice(s) de $\mathcal{R}$ onde a função objetivo atinge o valor máximo. Ora tem-se,

Vértice de $\mathcal{R}$ (x, y, z)	Valor da f.o. $2x + y + z$
$A = (5, 0, 0)$	10
$B = (2.5, 7.5, 0)$	12.5
$C = (0, 10, 0)$	10
$D = (0, 2, 8)$	10
$E = (0, 0, 8)$	8
$F = (2, 0, 8)$	12
$G = (5, 0, 5)$	15 (máx.)
$H = (0, 0, 0)$	0

Logo uma solução ótima do problema ocorre no vértice $(5, 0, 5)$ com valor ótimo 15.

Problema de PL com soluções ótimas alternativas

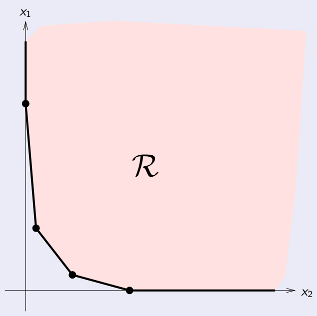
- Se no problema de PL do slide 237 alterarmos a f.o. para $0x + 0y + z$ (cota) e recalcularmos os valores da f.o. nos vértices do poliedro de admissibilidade $\mathcal{R}$, concluímos que o máximo (de valor 8) ocorre nos vértices D, E e F (verifique).

- Logo pelas observações do slide 236, tem-se que qualquer ponto do triângulo $[DEF]$, que é uma combinação convexa dos vértices D, E e F, é uma solução ótima alternativa do problema de PL do slide 237 para nova f.o. $0x + 0y + z$, o que é geometricamente evidente, uma vez que todos os pontos desse triângulo estão à cota máxima 8.

Regiões admissíveis não limitadas

Observação

Se a região admissível de um problema de PL for não limitada (como na região da figura abaixo) pode não existir um vértice onde ocorra uma solução ótima.



Por exemplo, no exercício 59 da sebenta de exercícios existe um vértice onde ocorre o mínimo da f.o., mas não existe um vértice onde ocorra o máximo (uma vez que este é $+\infty$).

Resolução de problemas de PL no caso geral?

- Problemas de PL usados em contextos mais realistas podem chegar a envolver milhões de variáveis de decisão!
- Nos casos em que a dimensão do problema é relativamente pequena podemos usar o suplemento Solver que vem incluído no programa de folha de cálculo Excel, que permite especificar até 200 variáveis (células de valor ajustável). Veremos nas aulas práticas como implementar e resolver problemas em PL usando o suplemento Solver (consultar também os ficheiros disponibilizados no separador **Material de Apoio** no Fénix).
- Embora não seja, em geral, viável representar graficamente regiões admissíveis de problemas de PL com 3 ou mais variáveis de decisão, podemos determinar ou reconhecer os vértices dessas regiões de forma indireta, como veremos a seguir. Mas para isso temos primeiro que converter a formulação do problema para uma certa forma canónica...

Problema de PL na forma *standard*

Definição de problema de PL na forma *standard*

Um problema de PL diz-se na forma *standard* se as todas as restrições funcionais forem equações lineares e as variáveis de decisão tomarem valores não negativos.

- No que se segue iremos sempre considerar problemas de PL em que as variáveis de decisão $x_1, \ldots, x_k$ tomam valores não negativos, ou seja, vamos assumir desde o início que as restrições de sinal são todas do tipo $x_1, \ldots, x_k \geq 0$.
- Acrescentando novas variáveis auxiliares, ditas variáveis de folga, podemos converter um problema em PL para a forma *standard*, como explicado no próximo slide.

Conversão de um problema de PL para a forma *standard*

- Substituímos cada restrição funcional do tipo,

$$a_1x_1 + a_2x_2 + \cdots + a_kx_k \leq b,$$

pela restrição,

$$a_1x_1 \;+\; a_2x_2 \;+\; \cdots \;+\; a_kx_n + f \;\; = b,$$

onde $f = b - a_1x_1 \;+\; a_2x_2 \;+\; \cdots \;+\; a_kx_n$ é uma nova variável de folga e acrescentamos a restrição de sinal dessa variável, $f \geq 0$.

- Substituímos cada restrição funcional do tipo,

$$a_1'x_1 \;+\; a_2'x_2 \;+\; \cdots \;+\; a_{ik}'x_k \geq b',$$

pela restrição,

$$a_1'x_1 \;+\; a_2'x_2 \;+\; \cdots \;+\; a_k'x_n - f' \;\; = \;\; b,'$$

onde $f' = a_1'x_1 \;+\; a_2'x_2 \;+\; \cdots \;+\; a_k'x_n - b'$ é uma nova variável de folga e acrescentamos a restrição de sinal dessa variável, $f' \geq 0$.

- As restrições funcionais do tipo,

$$a_1''x_1 \;+\; a_2''x_2 \;+\; \cdots \;+\; a_k''x_k = b'',$$

as restrições de sinal das variáveis de decisão e a f.o. ficam **inalteradas**.

Exemplo: conversão do Problema 1 para a forma *standard*

Problema original

$$\begin{array}{lllll}
\max & z = 300\,x + 200\,y & & & \\
\text{s.a} & \quad x + \quad y & & \leq & 80 \\
 & 8000\,x & & \leq & 320000 \\
 & \; 40\,x + 20y & & \leq & 2000 \\
 & \quad x,\; y & \geq 0 & &
\end{array}$$

$- - \rightarrow$

Problema na forma *standard*

$$\begin{array}{ll}
\max & z = 300\,x + 200\,y \\
\text{s.a} & x + \quad y + f_1 \qquad\qquad = 80 \\
 & 8000\,x \qquad\qquad + f_2 \quad = 320000 \\
 & 40\,x + 20y \qquad\qquad + f_3 = 2000 \\
 & x,\; y, f_1, f_2, f_3 \geq 0
\end{array}$$

em que,

- $f_1 = 80 - (x + y)$ representa a área de terreno disponível (em ha) que não foi utilizada.
- $f_2 = 320000 - 8000x$ representa a água disponível (em m^3) e não utilizada.
- $f_3 = 2000 - (40x + 20y)$ representa as horas de mão-de-obra disponíveis e não utilizadas.

Problema 2

Problema 2

Uma empresa produz três tipos de fertilizantes, A, B e C. Cada tonelada de fertilizante A, B e C gera 50, 40 e 60 unidades de resíduos tóxicos e origina um lucro de 10, 5 e 10 euros, respetivamente. A empresa tem capacidade para produzir 15 mil toneladas de fertilizantes por mês. Compromissos já assumidos obrigam a empresa a entregar mensalmente 5 mil toneladas de fertilizante A a um cliente. Pretende-se determinar o plano de produção mensal que gera a menor quantidade possível de resíduos tóxicos de modo a obter-se um lucro mensal de pelo menos 100 mil euros e uma produção mensal nunca inferior a 80% da capacidade de produção da empresa.

Dados do problema:

	Resíduos	Lucro	
Fertlizante A	50 unid./t	10 €/t	$\geq$ 5000 t/mês
Fertlizante B	40 unid./t	5 €/t	
Fertlizante C	60 unid./t	10 €/t	
	min	$\geq$ 100000 €	
Capacidade mensal	$\leq$ 15000 t		
Produção mensal	$\geq$ .80 $\times$ 15000 t		

Formulação do Problema 2 em PL

O Problema 2 pode então ser formulado em PL como,

$$
\begin{array}{lllll}
\min & z = 50x_A + 40X_B + 60x_C & & & \\
\text{s.a} & x_A + x_B + x_C & \leq & 15000 & (1) \\
& x_A & \geq & 5000 & (2) \\
& 10x_A + 5x_B + 10x_C & \geq & 100000 & (3) \\
& x_A + x_B + x_C & \geq & 12000 & (4) \\
& x_A,\ x_B,\ x_C \geq 0 & & &
\end{array}
$$

em que

- x_A, x_B e x_C representam, respetivamente, as quantidades, em toneladas, de fertilizante dos tipos A, B e C a produzir mensalmente.

Vamos converter esta formulação para a forma *standard* aplicando as regras do slide 243.

Conversão do Problema 2 para a forma *standard*

$$\begin{array}{llr}
\min & z = 50x_A + 40X_B + 60x_C & \\
\text{s.a} & x_A + x_B + x_C + f_1 = 15000 & (1) \\
& x_A - f_2 = 5000 & (2) \\
& 10\,x_A + 5\,x_B + 10\,x_C - f_3 = 100000 & (3) \\
& x_A + x_B + x_C - f_4 = 12000 & (4) \\
& x_A,\ x_B,\ x_C,\ f_1,\ f_2,\ f_3,\ f_4 \geq 0 &
\end{array}$$

em que as varíveis de folga, f_1, f_2, f_3 e f_4 têm o seguinte significado:

- $f_1 = 15000 - (x_A + x_B + x_C)$ que representa a capacidade de produção mensal de fertilizantes (em toneladas) não utilizada.
- $f_2 = x_A - 5000$ que representa a quantidade de fertilizante A produzida (em toneladas) para além do compromisso assumido.
- $f_3 = 10x_a + 5x_B + 10x_C - 100000$ que representa o lucro obtido (em €) acima do lucro mínimo pretendido de 100000€.
- $f_4 = x_A + x_B + x_C - 12000$ que representa a produção mensal de fertilizantes (em toneladas) produzida acima de 80% da capacidade mensal de produção.

Região admissível $\mathcal{F}$ do problema na forma *standard*

Dado um problema de PL na forma *standard*, com k variáveis de decisão $x_1, \ldots, x_k$ e s variáveis de folga $f_1, \ldots, f_s$, vamos denotar por $\mathcal{F}$ a respetiva região admissível, que podemos descrever matricialmente como,

$$\mathcal{F} = \left\{\bar{x} = (x_1, \ldots, x_k, f_1, \ldots, f_s) \in \mathbb{R}^{k+s} \ :\ A\,\bar{x} = b,\ \bar{x} \geq 0\right\},$$

isto é, $[\,A\,|\,b\,]$ é a matriz ampliada do sistema que define as restrições funcionais na forma *standard*, com A matriz do tipo $m \times n$, onde m é número de restrições funcionais do problema e $n = k + s$ o número de variáveis, contando com variáveis de folga([14]), $b \in \mathbb{R}^m$ e $\bar{x} \geq 0$ significa $x_1, \ldots, x_k, f_1, \ldots, f_s \geq 0$.

- Por exemplo, a região admissível na forma *standard* do Problema 2 do slide anterior pode ser escrita usando notação matricial como,

$$\mathcal{F} = \left\{\bar{x} = (x_A, x_B, x_C, f_1, f_2, f_3, f_4) \ :\ A\bar{x} = b, \quad \bar{x} \geq 0\right\}, \quad \text{em que}$$

$$\begin{array}{c}
\begin{array}{ccccccc} x_A & x_B & x_C & f_1 & f_2 & f_3 & f_4 \end{array} \\
[\,A\,|\,b\,] = \left[\begin{array}{ccccccc|c}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 15000 \\
1 & 0 & 0 & 0 & -1 & 0 & 0 & 5000 \\
10 & 5 & 10 & 0 & 0 & -1 & 0 & 100000 \\
1 & 1 & 1 & 0 & 0 & 0 & -1 & 12000
\end{array}\right].
\end{array}$$

[14] Iremos sempre assumir que a car(A) é igual ao número de restrições funcionais m, condição que é verificada para todos os exercícios da sebenta.

Vértice de $\mathcal{R} \leftrightarrow$ s.b.a. de $\mathcal{F}$

Tem-se o seguinte resultado fundamental que permite reconhecer se um dado vetor é vértice da região admissível de um problema de PL.

Teorema

Com as notações do slide anterior tem-se que um vetor $x = (x_1, \ldots, x_k)$ é vértice de $\mathcal{R}$ se e só se o vetor $\bar{x} = (x_1, \ldots, x_k, f_1, \ldots, f_s)$, obtido acrescentando os valores das variáveis de folga, verifica as 4 condições:

1. Todas as componentes de $\bar{x}$ são **não negativas**.
2. O número de componentes nulas de $\bar{x}$ é **superior ou igual** a $n - m$, onde $n = k + s$ é o número de variáveis (contando com variáveis de folga) e m o número de restrições funcionais do problema.
3. $\bar{x}$ verifica as restrições funcionais na forma standard, isto é, $A\bar{x} = b$.
4. As colunas de A associadas às componentes não nulas de $\bar{x}$ formam um conjunto **linearmente independente** de vetores.

Nas condições acima, diz-se que $\bar{x} = (x_1, \ldots, x_k, f_1, \ldots, f_s)$ é uma solução básica admissível (s.b.a.) de $\mathcal{F}$ (associada ao vértice x).

Como reconhecer se uma solução admissível é vértice ?

- Consideremos a solução do Problema 2 que consiste em produzir 5000 toneladas de fertilizante A, nenhuma de fertilizante B e 10000 de fertilizante C.
- A solução anterior,
$$x = (x_A, x_B, x_C) = (5000, 0, 10000),$$
verifica todas as restrições funcionais,
$$\begin{aligned} x_A + x_B + x_C &= 15000 \leq 15000, && (1)\\ x_A &= 5000 \geq 5000, && (2)\\ 10x_A + 5x_B + 10x_C &= 150000 \geq 100000, && (3)\\ x_A + x_B + x_c &= 15000 \geq 12000, && (4) \end{aligned}$$
e de sinal $x_A, x_B, x_C \geq 0$ e é portanto uma solução admissível do problema.
- Será que $x = (x_A, x_B, x_C) = (5000, 0, 10000)$ corresponde a um vértice da região admissível do problema?

Averiguar se $x = (5000, 0, 10000)$ é vértice...

Vamos averiguar se a solução $x = (x_A, x_B, x_C) = (5000, 0, 10000)$ do Problema 2, que consiste em produzir 5000 toneladas de fertilizante A, nenhuma de fertilizante B e 10000 de fertilizante C, é um vértice da região admissível $\mathcal{R}$ usando o teorema do slide 249.

Calculando os valores das variáveis de folga (ver o slide 247) obtém-se,

$$
\begin{aligned}
f_1 &= 15000 - (x_A + x_B + x_C) = 15000 - 15000 = 0,\\
f_2 &= x_A - 5000 = 5000 - 5000 = 0,\\
f_3 &= 10x_A + 5x_B + 10x_C - 100000 = 150000 - 100000 = 50000,\\
f_4 &= x_A + x_B + x_C - 12000 = 15000 - 12000 = 3000.
\end{aligned}
$$

Tem-se então que $x = (x_A, x_B, x_C) = (5000, 0, 10000)$ é vértice de $\mathcal{R}$ se e só se a solução ampliada com as respetivas folgas,

$$\bar{x} = (x_A, x_B, x_C, f_1, f_2, f_3, f_4) = (5000, 0, 10000, 0, 0, 50000, 3000),$$

é uma s.b.a. da região admissível $\mathcal{F}$ do problema na forma *standard*, ou seja, verifica as 4 condições do slide seguinte:

Exemplo (cont.)

1. Todas as componentes de $\bar{x}$ são não negativas, o que se verifica. ✓
2. O número de componentes nulas de $\bar{x}$ é superior ou igual ao número de variáveis (contando com variáveis de folga) menos o número de restrições funcionais: $3 \geq 7 - 4$. ✓
3. $\bar{x}$ verifica as restrições funcionais do problema na forma standard, isto é, verifica $A\,\bar{x} = b$, onde $[\,A\,|\,b\,]$ é a matriz do slide 248, que define a região admissível $\mathcal{F}$. De facto,

$$A\,\bar{x} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 10 & 5 & 10 & 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 5000 \\ 0 \\ 10000 \\ 0 \\ 0 \\ 50000 \\ 3000 \end{bmatrix} = \begin{bmatrix} 15000 \\ 5000 \\ 100000 \\ 12000 \end{bmatrix} = b. \quad \checkmark$$

4. O conjunto das colunas de A associadas às componentes não nulas de $\bar{x} = (5000, 0, 10000, 0, 0, 50000, 3000)$, isto é, o conjunto das colunas de A assinaladas a vermelho,

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 10 & 5 & 10 & 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & -1 \end{bmatrix},$$

é **linearmente independente**.

Exemplo (concl.)

De facto, aplicando o método de eliminação de Gauss à matriz formada por esse conjunto de colunas obtém-se,

$$B = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 10 & 10 & -1 & 0 \\ 1 & 1 & 0 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} = B'.$$

Como todas as colunas da matriz em escada B' têm pivot, o conjunto formado pelas 1ª, 3ª, 6ª e 7ª colunas de A é linearmente independente.✓

(Pode-se provar que o conjunto formado pelas 4 colunas é linearmente independente mostrando alternativamente que $\det B \neq 0$, o que fica como exercício para os alunos.)

Uma vez que as 4 condições são verificadas concluímos que $\bar{x}$ é uma s.b.a. de $\mathcal{F}$, o que significa que x é de facto um vértice de $\mathcal{R}$.

UFF !!

Definição equivalente de solução básica admissível

Definição de solução básica admissível

Consideremos um sistema linear $A\bar{x} = b$ com $A_{m\times n}$ tal que $\text{car}(A) = m < n$, e seja B um subconjunto linearmente independente de m colunas de A. Vamos designar as variáveis associadas às colunas de B, x_β, $\beta \in B$, por variáveis básicas e as restantes $n - m$ variáveis por variáveis não básicas ou livres.([15])

- Chama-se solução básica associada a B denotada por x_B, à solução de $A\bar{x} = b$ que é obtida resolvendo o sistema $[B|b]$ em ordem às m variáveis básicas x_β, $\beta \in B$ (sistema PD), e acrescentando as restantes $n - m$ variáveis não básicas com o valor zero.
- Se todas as componentes da solução básica x_B forem não negativas, x_B diz-se uma solução básica admissível e denota-se por s.b.a.
  Caso contrário x_B diz-se não admissível e denota-se por s.b.n.a.

Observação

Como existem $\binom{n}{m}$ formas distintas de escolher m colunas de um conjunto de n, o número de soluções básicas de $A\bar{x} = b$ não pode ultrapassar $\binom{n}{m}$.

[15] Por abuso de linguagem, denotamos ainda por B o subconjunto dos índices das colunas de B.

“Toy example”

Exercício na aula

Determinar as soluções básicas admissíveis e não admissíveis do sistema $A\bar{x} = b$ com $[A \mid b] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 1 & -2 & -5 & -2 \end{array}\right]$.

Resolução: A matriz A tem $n = 3$ colunas e $m = 2$ linhas, tendo-se $\text{car}(A) = m = 2$ e portanto existem $\binom{3}{2} = 3$ maneiras distintas de escolher 2 colunas (var. básicas) em 3:

- Considerando B o conjunto **linearmente independente** formado pela 1ª e 2ª colunas de A, tem-se $[B|b] = \left[\begin{array}{cc|c} 1 & 1 & 4 \\ 1 & -2 & -2 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 2 \end{array}\right]$. Logo $x_1 = x_2 = 2$. Fazendo $x_3 = 0$ obtém-se a s.b.a. $\bar{x}_{1,2} = (2, 2, 0)$.
- Considerando B o conjunto **linearmente independente** formado pela 1ª e 3ª colunas de A, tem-se $[B|b] = \left[\begin{array}{cc|c} 1 & 1 & 4 \\ 1 & -5 & -2 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 1 \end{array}\right]$. Logo $x_1 = 3$ e $x_3 = 1$. Fazendo $x_2 = 0$ obtém-se a s.b.a. $\bar{x}_{1,3} = (3, 0, 1)$.
- Considerando B o conjunto **linearmente independente** formado pela 2ª e 3ª colunas de A, tem-se $[B|b] = \left[\begin{array}{cc|c} 1 & 1 & 4 \\ -2 & -5 & -2 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & 6 \\ 0 & 1 & -2 \end{array}\right]$. Logo $x_2 = 6$ e $x_3 = -2$. Fazendo $x_1 = 0$ obtém-se a s.b.n.a. $\bar{x}_{2,3} = (0, 6, -2)$ (uma vez que possui uma componente negativa).

Problema 1 revisitado

Aplicando as mesmas ideias do slide anterior ao sistema $A\bar{x} = b$ onde

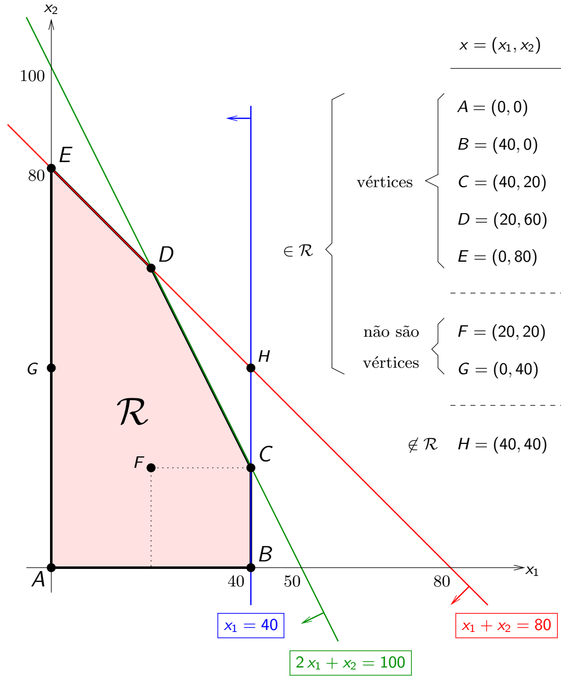
$$[\,A \mid b\,] = \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & 0 & 80 \\ 2 & 1 & 0 & 1 & 0 & 100 \\ 1 & 0 & 0 & 0 & 1 & 40 \end{array}\right],$$

é a matriz ampliada que define a região admissível do Problema 1 na forma standard,

$$\begin{array}{ll} \max & z = 300\,x_1 + 200\,x_2 \\ \text{s.a} & x_1 + x_2 + f_1 = 80 \\ & 2x_1 + x_2 + f_2 = 100 \\ & x_1 + f_3 = 40 \\ & x_1,\ x_2,\ f_1,\ f_2,\ f_3 \geq 0 \end{array}$$

podemos determinar todas as s.b.a. da região admissível $\mathcal{F}$ do problema na forma standard e consequentemente todos os vértices da região admissível $\mathcal{R}$ e as relações de adjacência entre esses vértices, uma vez que 2 vértices são adjacentes se e só forem definidos a partir conjuntos de variáveis básicas que apenas diferem entre si de uma variável básica.

Problema 1 revisitado



		$x = (x_1, x_2)$	$\bar{x} = (x_1, x_2, f_1, f_2, f_3)$		
$\in \mathcal{R}$	vértices	$A = (0, 0)$	$\bar{A} = (0, 0, 80, 100, 40)$	s.b.a.	$\in \mathcal{F}$
		$B = (40, 0)$	$\bar{B} = (40, 0, 40, 20, 0)$		
		$C = (40, 20)$	$\bar{C} = (40, 20, 20, 0, 0)$		
		$D = (20, 60)$	$\bar{D} = (20, 60, 0, 0, 20)$		
		$E = (0, 80)$	$\bar{E} = (0, 80, 0, 20, 40)$		
	não são vértices	$F = (20, 20)$	$\bar{F} = (20, 20, 40, 40, 20)$	não são s.b.a.	
		$G = (0, 40)$	$\bar{G} = (0, 40, 40, 60, 40)$		
$\notin \mathcal{R}$		$H = (40, 40)$	$\bar{H} = (40, 40, 0, -20, 0)$	s.b.n.a	$\notin \mathcal{F}$ (tem uma componente negativa)

Variáveis de folga:

$f_1 = 80 - (x_1 + x_2)$

$f_2 = 100 - (2x_1 + x_2)$

$f_3 = 40 - x_2$

Var. básicas	**1,2,3**	**1,2,4**	**1,2,5**	**1,3,4**	**1,3,5**	**1,4,5**	**2,3,4**	**2,3,5**	**2,4,5**	**3,4,5**
s.b.a de $\mathcal{F}$	$\bar{C}$		$\bar{D}$	$\bar{B}$			-		$\bar{E}$	$\bar{A}$
vértice de $\mathcal{R}$	C		D	B			-		E	A
s.b.n.a.		$\bar{H}$			TPC	(80,0,0,-60,-40)	-	TPC		
"pseudo vértice"		H			TPC	(80,0)	-	TPC		

Problema 1 revisitado: identificação dos vértices A e C

A título de exemplo vamos detalhar como se obtiveram, na tabela do slide anterior, os vértices A e C da região admissível $\mathcal{R}$ do Problema 1:

- Considerando o conjunto linearmente independente B formado pelas 3ª, 4ª e 5ª colunas de A vem $[B|b] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 80 \\ 0 & 1 & 0 & 100 \\ 0 & 0 & 1 & 40 \end{array}\right]$, obtendo-se para as variáveis básicas, $f_1 = 80$, $f_2 = 100$ e $f_3 = 40$. Fazendo as variáveis não básicas $x_1 = x_2 = 0$, tem-se $\bar{x}_{3,4,5} = (0, 0, 80, 100, 40) = \bar{A}$ que define uma s.b.a de $\mathcal{F}$. Esquecendo as folgas obtém-se o vértice $A = (0, 0)$ de $\mathcal{R}$.

- Considerando agora B constituído pelas 1ª, 2ª e 3ª colunas de A, e aplicando o método de Gauss vem
$$[B|b] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 80 \\ 2 & 1 & 0 & 100 \\ 1 & 0 & 0 & 40 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 40 \\ 0 & 1 & 0 & 20 \\ 0 & 0 & 1 & 20 \end{array}\right].$$
Uma vez que o sistema é PD, B é linearmente independente e obtêm-se os valores das variáveis básicas, $x_1 = 40$, $x_2 = 20$ e $f_1 = 20$. Fazendo as variáveis não básicas $f_2 = f_3 = 0$ obtém-se $\bar{x}_{1,2,3} = (40, 20, 20, 0, 0) = \bar{C}$, que define uma s.b.a. de $\mathcal{F}$. Esquecendo os valores das variáveis de folga obtém-se o correspondente vértice $C = (40, 20)$ da região admissível $\mathcal{R}$.

Problema 2 revisitado e vértices adjacentes

- Vimos no slide 251 que $x = (5000, 0, 10000)$ era um vértice da região admissível $\mathcal{R}$ associado à s.b.a. $\bar{x} = (5000, 0, 10000, 0, 0, 50000, 3000)$.
- A solução x está associada ao conjunto de variáveis básicas $1, 3, 6, 7$ (uma vez que as restantes 3 variáveis têm valor zero) e denota-se $x_{1,3,6,7}$.

Vejamos se trocando no conjunto B a coluna 3 pela coluna 2 ainda se obtém um vértice de $\mathcal{R}$. Ora tem-se,

$$[B|b] = \left[\begin{array}{rrrr|c} 1 & 1 & 0 & 0 & 15000 \\ 1 & 0 & 0 & 0 & 5000 \\ 10 & 5 & -1 & 0 & 100000 \\ 1 & 1 & 0 & -1 & 12000 \end{array}\right] \to \cdots \to \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 5000 \\ 0 & 1 & 0 & 0 & 10000 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3000 \end{array}\right].$$

Daqui resulta que $x_{1,2,6,7} = (5000, 10000, 0)$ é vértice de $\mathcal{R}$ associado à s.b.a. degenerada ([16]) $\bar{x} = (5000, 10000, 0, 0, 0, 0, 3000)$.

O vértice $x_{1,2,6,7}$ é adjacente ao vértice $x_{1,3,6,7}$ uma vez que os respectivos conjuntos de variáveis básicas apenas diferem numa variável básica, tendo-se ainda que melhora (isto é, baixa) o valor da f.o. $z = 50x_A + 40x_B + 60x_C$.

[16]Porque o número de componentes nulas é estritamente superior ao número de variáveis menos o número de restrições.

Método do simplex

"O método do simplex foi desenvolvido em 1949 pelo matemático americano George B. Dantzig para resolver problemas de PL. A ideia central do método consiste em mover-se de uma s.b.a. para s.b.a. adjacente enquanto o valor da f.o. for melhorando, até se atingir uma solução ótima, em que as soluções adjacentes já não melhoram esse valor da f.o. [...]"

(Retirado da página 159 do Texto de Apoio)