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 escrever a vermelho as definições, a azul o texto a destacar e a magenta os exercícios e desafios para os alunos.

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,

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

  x_i: componente do vetor x que se encontra na posição i
  Por exemplo, se $x = (1, -\pi, 0, 2, 3, -4) \in \mathbb{R}^6$, $x_4 = 2$

Cálculo matricial

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 multiplicaçãoo de matrizes

  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 $A = [a_{ij}]$ 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 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:

$$\lambda x = \lambda(x_1, x_2) = (\lambda x_1, \lambda x_2).$$

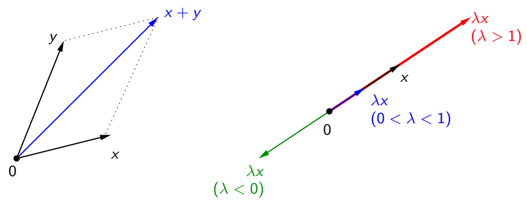
- Produto escalar (ou interno) de vetores:

$$x \cdot y = (x_1, x_2) \cdot (y_1, y_2) = x_1 y_1 + x_2 y_2.$$

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 algébricas...



(a) (b)

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 algébricas 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 operaçẽs algébricas 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)

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.

Se $A = \left[a_{ij}\right]_{m\times n}$ e $B = \left[b_{ij}\right]_{m\times n}$ são matrizes do mesmo tipo 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 (corrigido!)

- 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 $((A^T - 2B)^T + 4I_2)^T$ indicando as propriedades que utilizar e calcule o seu valor.

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

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

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

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

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 $y = (1, 0, 1) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$,

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

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

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,

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

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 - corrigido !

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 ? Qual é o resultado de $(AB)^2$?

“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 A, B e C, com $A \neq [\,0\,]$,
$$AB = AC \quad \not\Rightarrow \quad B = C.$$

(aqui $[\,0\,]$ 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

Uma consequência inesperada...

Se A e B são matrizes quadradas de ordem n 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.

Atenção

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

Exercícios na aula e TPC

- Considere

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

  Simplifique e calcule o valor de $((BC)^2A)^T$.

- Grupos de exercícios 2, 3 e 4.1a) das páginas 2 e 3 da sebenta de exercícios.

- Mostre que se A é uma matriz quadrada de ordem n então as matrizes $A + A^T$ e AA^T são simétricas.

Sistema de equações lineares

Sistema linear

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

Exercício

- O que representa geometricamente cada equação ?
- E o que representa geometricamente o sistema linear ? ([1])

[1] Cada equação representa um plano em $\mathbb{R}^3$ e o sistema representa a interseção de 3 planos em $\mathbb{R}^3$

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!).

Exercício (TPC)

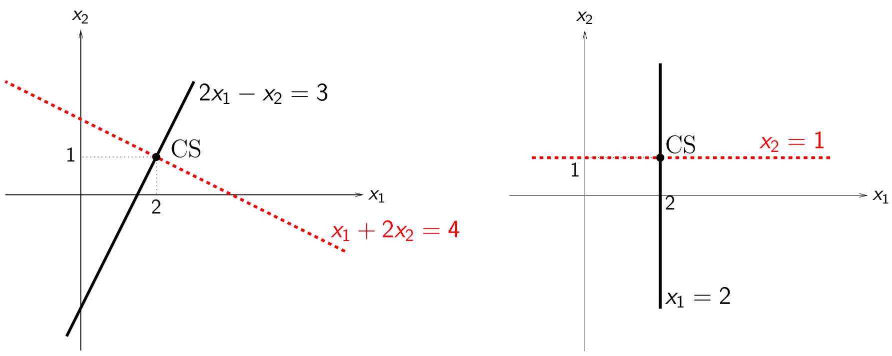
Dar exemplos de sistemas lineares a 2 equações e 2 variáveis que sejam PI e IMP, indicando em cada caso o respectivo CS (Sugestão: adapte o sistema anterior).

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,

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

- $A_{m\times n} = [a_{ij}]$ 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

Exemplo

O sistema linear a 3 equações e 3 variáveis,

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

tem matriz ampliada $\left[\begin{array}{ccc|c} 1 & -1 & 2 & -1 \\ 3 & 1 & 1 & 6 \\ 2 & 6 & 0 & 10 \end{array}\right]$

Para ilustrar o método de eliminação de Gauss vamos resolver este sistema linear aplicando certas operações sobre a sua matriz ampliada. Mas para isso temos que começar por introduzir os conceitos de:

- matriz em escada e matriz reduzida
- operações elementares sobre as linhas de uma matriz

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.
- 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,

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

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 ($L_i + \lambda L_j \to L_i$).

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

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

A notação entre parênteses refere-se à 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.

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 à matriz ampliada de um sistema linear $[A|b]$, de acordo com o seguinte esquema:

$$\begin{tikzcd}
\boxed{[A|b]} \arrow[r, "\substack{\text{operações}\\ \text{elementares}}\ \to \cdots \to"] \arrow[r, dashed, bend right, "\text{fase descendente}"'] & \boxed{[A'|b'],\ A' \text{ em escada}} \arrow[d] \arrow[r, "\substack{\text{operações}\\ \text{elementares}}\ \to \cdots \to"] \arrow[r, dashed, bend right, "\text{fase ascendente}"'] & \boxed{[A''|b''],\ A'' \text{ reduzida}} \arrow[d] \\
 & \substack{\text{classificar}\\ \text{o sistema } Ax = b} & \substack{\text{resolver}\\ \text{o sistema } Ax = b}
\end{tikzcd}$$

Redução de sistemas lineares pelo método de Gauss

Exemplo na aula

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

$$\left\{\begin{array}{rcr} x_1 - x_2 + 2x_3 & = & -1 \\ 3x_1 + x_2 + x_3 & = & 6 \\ 2x_1 + 6x_2 & = & 10 \end{array}\right.$$

TPC

Resolver os seguintes sistemas lineares:

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

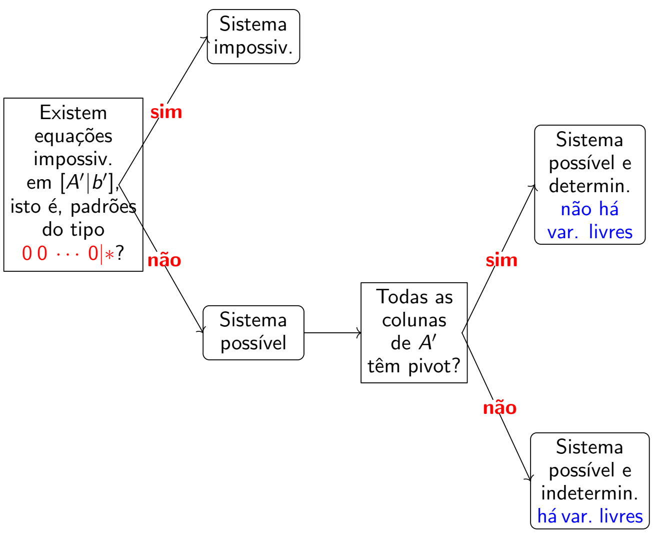
Algoritmo de eliminação de Gauss: fase descendente

- **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.

Discussão do sistema em escada



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 lineares são **equivalentes** se e só se aplicando o método de Gauss às respetivas matrizes ampliadas obtemos **a mesma matriz reduzida**.

Redução de sistemas lineares - exercício na aula

Considere o sistema linear com um parâmetro $a \in \mathbb{R}$,

$$\begin{cases} \quad\ \ -5x_2 - 8x_3 &= -2 \\ x_1 + 3x_2 + 3x_3 &= 0 \\ 3x_1 + 4x_2 \ + x_3 &= a \end{cases}$$

- Aplicando a fase descendente do método de Gauss à matriz ampliada do sistema anterior indique os valores do parâmetro a para os quais o sistema é:
  - IMP
  - PD
  - PI. Nessa altura quantas variáveis livres possui? ([2])
- Aplicando a fase ascendente do método de Gauss à matriz em escada obtida na alínea anterior, determine o CS para o(s) valor(es) de $a \in \mathbb{R}$ para os quais o sistema é possível. Que tipo de CS obtivemos? ([3])

[2] O sistema é IMP para $a \neq -2$ e PI para $a = -2$ com uma variável livre x_3.

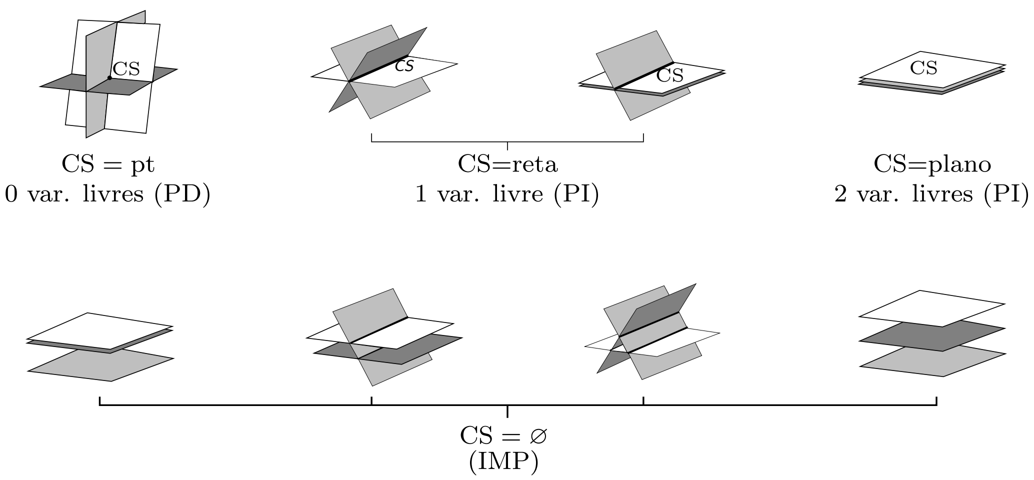
[3] Para $a = -2$, $CS = \{(x_1, x_2, x_3) : x_1 = -\frac{6}{5} + \frac{9}{5}x_3,\ x_2 = \frac{2}{5} - \frac{8}{5}x_3,\ x_3 \in \mathbb{R}\}$, que define uma reta obtida como interseção de 3 planos em $\mathbb{R}^3$

Interpretação geométrica dos sistemas de equações lineares

- Geometricamente um sistema linear a m equações e n variáveis representa a intersecção de:
  - m retas em $\mathbb{R}^2$ (plano), se $n = 2$,
  - m planos em $\mathbb{R}^3$ (espaço), se $n = 3$,
  - m hiperplanos em $\mathbb{R}^n$, se $n \geq 4$.
- O número de variáveis livres 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 interpretar geometricamente principalmente os sistemas lineares com 2 e 3 variáveis, ou seja, os sistemas lineares cujos CS estão contidos em $\mathbb{R}^2$ e em $\mathbb{R}^3$.

Geometria dos sistemas lineares a 3 equações e 3 incógnitas

- Geometricamente existem os 8 casos distintos representados na seguinte figura:



Exercício

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

Uma equivalência fundamental

Se $A = [a_{ij}]$ é uma matriz do tipo $m \times n$, $b = (b_1, \ldots, b_m) \in \mathbb{R}^m$ e $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, tem-se:

$$Ax = b \quad \Leftrightarrow \quad \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}_{m \times n} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}_{m \times 1}$$

$$\Leftrightarrow \quad \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}_{m \times 1} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}_{m \times 1}$$

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

TPC: estabeleça a equivalência acima com $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ e $b = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$.

Sistemas lineares e equações matriciais

Obtivemos portanto a equivalência fundamental,

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

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 $Ax = b$, que permite traduzir os sistemas lineares para a linguagem das matrizes.

Observações

- Uma solução do sistema linear com matriz ampliada $[A|b]$ é portanto uma solução de $Ax = b$, isto é, um vetor $u \in \mathbb{R}^n$ tal que $Au = b$.
- No que se segue iremos frequentemente chamar sistema linear a uma equação matricial do tipo $Ax = b$.

Exemplo - sistemas lineares e equações matriciais

- Considerando novamente o sistema do slide 36, tem-se a equivalência,

$$\left\{\begin{array}{rcl} x_1 - x_2 + 2x_3 & = & -1 \\ 3x_1 + x_2 + x_3 & = & 6 \\ 2x_1 + 6x_2 & = & 10 \end{array}\right. \quad \Leftrightarrow \quad Ax = b$$

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

- Podemos verificar que $u = (2, 1, -1)$ é solução do sistema anterior substituindo $x_1 = 2$, $x_2 = 1$ e $x_3 = -1$ nas equações desse sistema (confirme) ou, em alternativa, verificando que é solução da equação matricial equivalente $Ax = b$, substituindo x por u nessa equação:

$$Au = \begin{bmatrix} 1 & -1 & 2 \\ \hline 3 & 1 & 1 \\ \hline 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 6 \\ 10 \end{bmatrix} = b.$$

- A equação linear $ax = 5$, com $a \neq 0$, admite a solução $x = \frac{5}{a} = a^{-1}5$. Vamos ver que se A for invertível podemos, de forma análoga, exprimir a solução u do sistema $Ax = b$ como $u = A^{-1}b$, onde A^{-1} denota a matriz inversa de A, conceito que vamos definir a seguir.

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, $AB = BA = I_2$ (verifique!).

Teorema

A inversa de uma matriz A é única (e denota-se por A^{-1})

Demonstração: exercício.

Inversa de uma matriz

- Se A e B são matrizes quadradas de ordem n, pode-se mostrar que $AB = I_n \Rightarrow BA = I_n$ (e reciprocamente). Assim:
  - Para mostrar que A é invertível com inversa B basta verificar que $AB = I_n$ ([4]). Nessa altura, B é invertível com inversa A, ou seja, A e B são inversas uma da outra.
  - Para decidir se A é invertível e calcular a sua inversa (caso seja invertível), basta resolver a equação matricial $AX = I_n$ onde X designa uma matriz quadrada de incógnitas de ordem n. A equação matricial $AX = I_n$ é possível se e só se A for invertível e nessa altura a (única) solução dessa equação é $X = A^{-1}$

[4] Ou, em alternativa, verificar que $BA = I_n$

Ainda a inversa...

Exercícios na aula

- Determine (caso exista) a inversa da matriz diagonal

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \mathrm{diag}(2,3),$$

resolvendo a equação matricial $AX = I_2$, onde $X = \begin{bmatrix} x & y \\ z & t \end{bmatrix}_{2\times 2}$ é uma matriz de incógnitas e $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ a matriz identidade de ordem 2.

TPC: mesmo exercício com $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$.

- Mostre que se B é a inversa de A^2 então AB é inversa de A

Algumas propriedades importantes da inversa

Teorema

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

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

Potências negativas de matrizes invertíveis

Se A é uma matriz invertível, define-se

$$A^{-k} = (A^{-1})^k, \qquad k = 1, 2, 3, \ldots$$

[5] 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 $(A_1A_2\cdots A_k)^{-1} = A_k^{-1}\cdots A_2^{-1}A_1^{-1}$

Exercícios que envolvem inversa e as suas propriedades

TPC

- Mostre que se $A_{n\times n}$ verifica $A^3 - 3A - I_n = 0$ com I_n matriz identidade de ordem n, então A é invertível e indique a sua inversa.
- Sejam A e B matrizes invertíveis de ordem n. Determine X em função de A e B (se existir) tal que

$$(A^{-2}XB^T)^{-1} = 3I_n.$$

Soluções: $A^{-1} = A^2 - 3I_n$, $\quad X = \frac{1}{3}A^2(B^{-1})^T$

Inversa e redução simultânea de sistemas - exemplo

- Consideremos $A = \begin{bmatrix} 1 & 4 \\ 1 & 3 \end{bmatrix}$. Queremos determinar a solução X de $AX = I_2$ (caso exista). Nessa altura, sabemos que $A^{-1} = X$
- 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

$$
\begin{aligned}
AX = I_2 \quad &\Leftrightarrow \quad \left[\begin{array}{cc} 1 & 4 \\ \hline 1 & 3 \end{array}\right] \left[\begin{array}{c|c} x_1 & y_1 \\ x_2 & y_2 \end{array}\right] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&\Leftrightarrow \quad \begin{bmatrix} x_1 + 4x_2 & y_1 + 4y_2 \\ x_1 + 3x_2 & y_1 + 3y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&\Leftrightarrow \quad \left\{ \begin{array}{l} \left\{ \begin{array}{lcl} x_1 + 4x_2 & = & 1 \\ x_1 + 3x_2 & = & 0 \end{array} \right. \quad \to \left[\begin{array}{cc|c} 1 & 4 & 1 \\ 1 & 3 & 0 \end{array}\right] = [A|e_1] \\ \\ \left\{ \begin{array}{lcl} y_1 + 4y_2 & = & 0 \\ y_1 + 3y_2 & = & 1 \end{array} \right. \quad \to \left[\begin{array}{cc|c} 1 & 4 & 0 \\ 1 & 3 & 1 \end{array}\right] = [A|e_2] \end{array} \right.
\end{aligned}
$$

Inversa e redução simultânea de sistemas - exemplo (cont.)

- Portanto resolver a equação matricial $AX = I_2$ é equivalente a resolver dois sistemas lineares, $[A|e_1]$ e $[A|e_2]$ com mesma matriz de coeficientes A
- Podemos reduzir simultaneamente ambos os sistemas anteriores ampliando A com os vetores e_1 e e_2, isto é, com a matriz identidade
- Aplicando a fase descendente a este sistema obtém-se,

$$[A|e_1\ e_2] = \left[\begin{array}{cc|cc} 1 & 4 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 4 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{array}\right] = [A'|I']$$

e portanto ambos os sistemas $[A\,|\,e_1]$ e $[A\,|\,e_2]$ são PD. Logo a equação matricial $AX = I_2$ é também PD e portanto A é invertível

- Aplicando a fase ascendente, obtém-se

$$[A'|I'] = \left[\begin{array}{cc|cc} 1 & 4 & 1 & 0 \\ 0 & -1 & 1 & -1 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|cc} 1 & 0 & -3 & 4 \\ 0 & 1 & 1 & -1 \end{array}\right] = [\,I\,|\,x \quad y\,]$$

Logo $x = (-3, 1)$ e $y = (4, -1)$ e tem-se $A^{-1} = X = [\,x \quad y\,] = \begin{bmatrix} -3 & 4 \\ 1 & -1 \end{bmatrix}$

isto é, as colunas de A^{-1} são as soluções dos sistemas com matrizes ampliadas $[A|e_1]$ e $[A|e_2]$.

Inversa e redução simultânea de sistemas (caso geral)

- Analogamente pode-se mostrar que resolver a equação matricial $AX = I_n$ com A matriz arbitrária de ordem n, é equivalente a resolver n sistemas lineares com a mesma matriz de coeficientes A,

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

  onde $e_1, e_2, \ldots, e_n$ são as colunas da matriz identidade I_n
- Nessa altura os n sistemas podem ser resolvidos simultaneamente aplicando o método de Gauss à matriz ampliada $[A \,|\, I_n]$
- Pela unicidade da inversa, ou os n sistemas (1) são todos PD e nessa altura A é invertível ou pelo menos um desses sistemas é impossível e nessa altura A é singular
- As colunas de A^{-1} são as soluções dos n sistemas, $[A|e_i]$, $= 1, \ldots, n$
- 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 slides anteriores deduz-se o seguinte algoritmo

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

$$[\,A\,|\,I\,] \longrightarrow \cdots \longrightarrow [\,A'\,|\,I'\,] \xrightarrow{\text{(se } A \text{ invertível)}} \cdots \longrightarrow [\,I\,|\,A^{-1}\,]$$

(A' em escada)

Existem linhas nulas em A'	Não existem linhas nulas em A'
$\Downarrow$	$\Downarrow$
A é singular	A é invertível
STOP	

Algoritmo da inversa

Exercício na aula

Aplique o algoritmo da inversa para averiguar se a matriz

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

é invertível e determinar a sua inversa (caso exista!)

Qual a solução da equação matricial $Ax = e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$?

(**Sugestão**: ver o slide 53)

E de $Ax = 2e_2 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$?

(**Sugestão**: relacione com a solução de $Ax = e_2$)

Soluções: $A^{-1} = \begin{bmatrix} -1 & 0 & 2 \\ 0 & \frac{1}{2} & 0 \\ -1 & 0 & 1 \end{bmatrix}$, $x = \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}$, $x = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

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.

Teorema

Seja A uma matriz quadrada. As seguintes afirmações são equivalentes:

- A é invertível.
- O sistema linear $Ax = b$ é PD para todo o $b \in \mathbb{R}^n$ (com solução única $x = A^{-1}b$)
- O sistema $Ax = \vec{0}$ tem apenas a solução trivial $x = \vec{0}$ ([6])

TPC: Utilize a inversa para obter novamente as soluções de $Ax = e_1$ e $Ax = 2e_2$, em que A é a matriz do slide 55 e e_1 e e_2 a primeira e segunda colunas da matriz I_3, resp.

[6] Os sistemas lineares $Ax = \vec{0}$ designam-se por sistemas homogéneos e vão ser considerados em detalhe mais adiante.

Critério de invertibilidade

- Para decidir apenas sobre a invertibilidade de uma matriz quadrada A de ordem n, sem pretender calcular a sua inversa, não é necessário ampliar A com a matriz identidade, nem aplicar a fase ascendente do método de Gauss - basta aplicar a fase descendente à matriz A.
- Seja A' uma matriz em escada obtida a partir da matriz quadrada A por aplicação de operações elementares nas linhas de A.

  Tem-se, como vimos no algoritmo da inversa, que:
  - se A' não tem linhas nulas $\Rightarrow$ A é invertível
  - se A' tem linhas nulas $\Rightarrow$ A é singular

Observação

Uma vez que A' é quadrada e está em escada, tem-se que A' não tem linhas nulas se e só se todas as suas colunas tiverem pivot, o que nos vai permitir obter um critério alternativo para decidir sobre a invertibilidade de uma matriz baseado no número de colunas pivot da matriz em escada. Mas para isso temos primeiro que introduzir o conceito de característica de uma matriz...

Interlúdio: característica de uma matriz

Definição de caracterítica

A característica de uma matriz A, denotada $\text{car}(A)$, é o número de pivots de qualquer matriz em escada obtida a partir de A por aplicação do método de eliminação de Gauss

- 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
- $\text{car}(A)$ corresponde também ao número de linhas não nulas de qualquer matriz em escada obtida a partir de A por aplicação do método de eliminação de Gauss
- 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, $\text{car}(A) \leq \min\{m, n\}$.

Exemplo

- Se, por exemplo,

$$A_{4\times 5} \rightarrow^{\text{oper. elementares}} \cdots \rightarrow A' = \begin{bmatrix} 0 & 1 & 2 & 1 & 4 \\ 0 & 0 & 3 & -6 & 0 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{4\times 5},$$

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

Inversa e característica

- Para decidir sobre a invertibilidade de uma matriz é suficiente calcular a sua característica. De facto, tem-se o seguinte critério.

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 ?

Solução: $\alpha \neq 0, 2$

Sistemas homogéneos

Uma classe importante de sistemas lineares é constituída pelos sistemas cujos termos constantes são todos nulos, ou seja, da forma

$$Ax = \vec{0},$$

que se designam por sistemas lineares homogéneos. Tem-se que:

- Os sistemas lineares homogéneos possuem sempre a solução trivial $u = \vec{0}$ pois $A\,u = A\,\vec{0} = \vec{0}$ e portanto **nunca são impossíveis**
- Se u e v são soluções do sistema homogéneo $Ax = \vec{0}$, isto é, se $Au = Av = \vec{0}$ e se $\lambda \in \mathbb{R}$ é um escalar arbitrário, tem-se
  - $A(u + v) = Au + Av = \vec{0} + \vec{0} = \vec{0}$,
  - $A(\lambda u) = \lambda\, Au = \lambda\,\vec{0} = \vec{0}$,

  e portanto $u + v$ e λu são ainda soluções do sistema $Ax = 0$.

As condições anteriores significam, por um lado, que o CS de um sistema homogéno $Ax = 0$ nunca é vazio e por outro lado que é fechado para a adição e para o produto por escalar, o que se traduz dizendo que o CS de um sistema linear homogéneo é um subespaço vetorial de $\mathbb{R}^n$, conceito fundamental que vamos introduzir na próxima aula...

O espaço vetorial $\mathbb{R}^n$

No slide 13 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 sobre os 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 o conjunto $\mathbb{R}^n$ com a adição e a multiplicação por escalares tem uma estrutura de *espaço vetorial*...

Subespaço vetorial de $\mathbb{R}^n$

- Queremos 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.
- Nas condições anteriores pode-se mostrar que são ainda verificadas as propriedades $(1), \ldots, (8)$ relativamente aos vetores de V e portanto que V, com as operações usuais da adição de vetores e da multiplicação de vetores por escalares, herda a estrutura de espaço vetorial que vem de $\mathbb{R}^n$, dizendo-se nessa altura que V é um subespaço vetorial de $\mathbb{R}^n$.

  Mais precisamente, tem-se a seguinte definição.

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 $\alpha u \in V$

Subespaço vetorial

Exercício na aula

Averiguar se os seguintes subconjuntos do plano ($\mathbb{R}^2$) definem subespaços vetoriais:

- $V = \{(x, y) : xy = 0\}$ (eixos coordenados de $\mathbb{R}^2$)
- $V = \{(x, y) : x, y \geq 0\}$ (1^{o} quadrante de $\mathbb{R}^2$)

- Mas afinal que tipo de conjuntos definem os subespaços vetoriais de $\mathbb{R}^n$?
- Antes de respondermos à questão anterior vamos estabelecer uma condição necessária (mas não suficiente) para que um subconjunto de $\mathbb{R}^n$ defina um subespaço vetorial.

Uma condição **necessária** para ser subespaço vetorial...

- Se V é subespaço vetorial de $\mathbb{R}^m$, tem-se em particular que:
  - $V \neq \emptyset$, logo

    existe um vetor $v \in V$

  - V é fechado para o produto por escalar, logo

    $$\lambda v \in V, \text{ para todo o } \lambda \in \mathbb{R}$$

  - Em particular, considerando $\lambda = 0$, obtém-se

    $$\lambda v = 0v = \vec{0} \in V$$

Logo tem-se a seguinte condição necessária para ser subespaço vetorial.

Teorema

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

- Por 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 o conjunto V ?

Subespaços vetoriais de $\mathbb{R}^n$?

Comecemos por observar que as três condições da definição de subespaço vetorial de $\mathbb{R}^n$ (slide 63) são trivialmente verificadas nos seguintes 2 casos:

- $V = \{\vec{0}\}$ que se designa por subespaço vetorial minimal (ou trivial)
- $V = \mathbb{R}^n$ que se designa por subespaço vetorial maximal

Pode-se mostrar que os subconjuntos que definem subespaços vetoriais do plano ($\mathbb{R}^2$), do espaço ($\mathbb{R}^3$) e mais geralmente de $\mathbb{R}^n$ com $n \geq 4$, são dos seguintes tipos:

$\mathbb{R}^2$:	$\{\vec{0}\}$, *retas que passam na origem*, $\mathbb{R}^2$
$\mathbb{R}^3$:	$\{\vec{0}\}$, *retas e planos que passam na origem*, $\mathbb{R}^3$
$\mathbb{R}^n$:	$\{\vec{0}\}$, *retas, planos e "hiperplanos" que passam na origem*, $\mathbb{R}^n$

Espaço nulo de uma matriz

Vamos introduzir o primeiro subespaço vetorial fundamental associado a uma matriz

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 homogéneo $Ax = \vec{0}$, isto é,

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

Antes de provarmos que o espaço nulo define um subespaço vetorial de $\mathbb{R}^n$ vamos calcular esse subespaço nalguns exemplos

Espaço nulo - exemplo

Consideremos a matriz $A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & -1 & 1 \end{bmatrix}_{2\times 3}$. Tem-se:

$$\begin{aligned}
\mathcal{N}(A) &= \left\{x \in \mathbb{R}^3 \ : \ Ax = \vec{0}\right\} \\
&= \left\{x \in \mathbb{R}^3 \ : \ \left[\begin{array}{ccc} 1 & 1 & -1 \\ \hline 2 & -1 & 1 \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\} \\
&= \left\{x \in \mathbb{R}^3 \ : \ \begin{bmatrix} x_1 + x_2 - x_3 \\ 2x_1 - x_2 + x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\} \\
&= \left\{x \in \mathbb{R}^3 \ : \ \begin{cases} x_1 + x_2 - x_3 = 0 \\ 2x_1 - x_2 + x_3 = 0 \end{cases}\right\}
\end{aligned}$$

Logo, para determinar $\mathcal{N}(A)$ tem que se reduzir a matriz ampliada do sistema $Ax = \vec{0}$,

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

Espaço nulo - exemplo (cont.)

Reduzindo a matriz ampliada $[\,A\,|\,\vec{0}\,]$ do slide anterior vem,

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

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) e que tem vetor diretor $v = (0,1,1)$([7])
- Uma vez que $\mathcal{N}(A)$ define uma reta de $\mathbb{R}^3$ que passa na origem, conclui-se pelo slide 66 que $\mathcal{N}(A)$ define um subespaço vetorial de $\mathbb{R}^3$
- Em alternativa, pode-se verificar facilmente que $\mathcal{N}(A)$ verifica as 3 condições da definição de subespaço vetorial do slide 63, o que fica como exercício.

---

[7]Esta reta é obtida como intersecção de 2 planos de $\mathbb{R}^3$ que passam na origem, uma vez que a equação matricial $Ax = \vec{0}$ é equivalente a um sistema linear homogéneo com 2 equações e 3 variáveis (ver o slide anterior)

Cálculo do espaço nulo - exemplo 2

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

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

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

  Uma vez que todas as colunas de A' têm pivot (isto é, $\text{car}(A) = 2 = n$), o sistema $Ax = \vec{0}$ é determinado e portanto possui apenas a solução trivial $x_1 = x_2 = 0$ ([8])
- 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)

Se $A_{m\times n}$ tem-se $\mathcal{N}(A) = \{\vec{0}\} \Leftrightarrow Ax = \vec{0}$ é determinado $\Leftrightarrow \text{car}(A) = n$.

---

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

O espaço nulo é um subespaço vetorial...

Teorema

Seja A uma matriz do tipo $m \times n$. Tem-se que $\mathcal{N}(A)$ define um subespaço vetorial de $\mathbb{R}^n$

Demonstração

Temos que verificar as 3 condições da definição de subespaço vetorial do slide 63:

- $\mathcal{N}(A) \neq \emptyset$, ou seja, o sistema $Ax = \vec{0}$ possui pelo menos uma solução.
- $\mathcal{N}(A)$ é fechado para a adição, ou seja, se u e v são soluções de $Ax = 0$ então $u + v$ ainda é solução de $Ax = 0$.
- $\mathcal{N}(A)$ é fechado para o produto por escalar, ou seja, se u é solução de $Ax = 0$ e $\lambda \in \mathbb{R}$, então λu ainda é solução de $Ax = 0$.

As 3 condições anteriores resultam imediatamente das considerações feitas no slide 61. $\square$

Consequências e exemplos

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.

- 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 contém a origem.

Combinação linear de vetores

Definição de combinação linear

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

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

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

Por outras palavras, $b \in \mathbb{R}^m$ é CL de $v_1, \dots, v_n$ se puder ser obtido como soma de múltiplos desses vetores

- $b = (-2, -4, -2)$ é CL de $v_1 = (1, 2, 1)$ pois $b = -2v_1$ $(\alpha_1 = -2)$
- $b = (3, 1)$ é CL de $v_1 = (1, 1)$ e $v_2 = (1, 0)$, pois
  $b = (3, 1) = 1(1, 1) + 2(1, 0) = v_1 + 2v_2$ $\quad(\alpha_1 = 1$ e $\alpha_2 = 2)$
- $\vec{0} \in \mathbb{R}^n$ é CL de $v_1, v_2, \dots, v_n$ pois $\vec{0} = 0\,v_1 + 0\,v_2 + \cdots + 0\,v_n$
  $(\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0)$
- v_1 é CL de $v_1, v_2, \dots, v_n$ pois $v_1 = 1\,v_1 + 0\,v_2 + \cdots + 0\,v_n$
  $(\alpha_1 = 1$ e $\alpha_2 = \cdots = \alpha_n = 0)$
- Em geral, cada v_i é CL de $v_1, v_2, \dots, v_n$ $(\alpha_i = 1$ e $\alpha_j = 0$ se $j \neq i)$

Determinação da comb. linear de vetores - exemplos

- 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,

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

Determinação da comb. linear de vetores - exemplo

- Logo se $b = \alpha_1 v_1 + \alpha_2 v_2$ então (α_1, α_2) é solução do sistema linear $[\, v_1 \;\; v_2 \mid b \,]$
- Aplicando o método de Gauss a este sistema, obtém-se (verifique!),

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

- Como o sistema é possível podemos escrever b como CL de v_1 e v_2 com coeficientes α_1 e α_2. Para determinar esses coeficientes aplicamos a fase ascendente do método de eliminação de Gauss, obtendo-se,

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

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

Determinação de combinações lineares de vetores - resumo

Pretende-se decidir se $b \in \mathbb{R}^m$ é combinação linear de $v_1, \ldots, v_n \in \mathbb{R}^m$

Designando $A = [\, v_1 \; v_2 \; \cdots \; v_n \,]$ e aplicando a fase descendente do método de Gauss à matriz ampliada $[A|b] = [\, v_1 \; v_2 \; \cdots \; v_n \mid b \,]$ do sistema $Ax = b$ vem

$$[A|b] = [\, v_1 \; v_2 \; \cdots \; v_n \mid b \,] \;\to\; \cdots \;\to [A'|b'] \;\; \text{(com } A' \text{ em escada).}$$

Tem-se então o seguinte:

- Se $[\, A \mid b \,]$ for impossível, então b não é combinação linear de $v_1, \ldots, v_n$;
- Se $[\, A \mid b \,]$ for possível, b é combinação linear de $v_1, \ldots, v_n$, tendo-se:
  - Se $[\, A \mid b \,]$ for PD, b escreve-se como combinação linear de $v_1, \ldots, v_n$ de uma única forma;
  - Se $[\, A \mid b \,]$ for PI, b escreve-se como combinação linear de $v_1, \ldots, v_n$ de infinitas maneiras distintas;

Observação

Para escrever a CL aplica-se a fase ascendente do método de Gauss à matriz $[A'|b']$ (caso o sistema seja possível). Cada solução $(\alpha_1, \ldots, \alpha_n)$ deste sistema dá origem a uma CL $b = \alpha_1 v_1 + \cdots + \alpha_n v_n$.

Determinação de comb. linear de vetores - exemplos

- Justifique que $c = (0, 0, 1)$ não é CL de $v_1 = (2, 2, 1)$ e $v_2 = (2, 3, 1)$.
  De facto, aplicando a fase descendente ao sistema $[\,A \mid c\,] = [\,v_1\ v_2 \mid c\,]$ vem,

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

  Como o sistema é IMP, c não é CL de v_1 e v_2.

- Justifique que $b = (5, 1)$ é CL de $v_1 = (1, 1)$, $v_2 = (1, 0)$ e $v_3 = (1, -2)$ de infinitas maneiras distintas e escreva essas CLs.
  De facto, aplicando a fase descendente ao sistema $[\,A \mid b\,] = [\,v_1\ v_2\ v_3 \mid b\,]$ vem,

$$[A|b] = [v_1\ v_2\ v_3|b] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 1 & 0 & -2 & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 0 & -1 & -3 & -4 \end{array}\right] = [A'|b']$$

  Como o sistema é PI, b é CL de v_1, v_2 e v_3 de infinitas formas distintas. Para determinar essas CLs aplica-se a fase ascendente:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 0 & -1 & -3 & -4 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & -2 & 1 \\ 0 & 1 & 3 & 4 \end{array}\right] \rightsquigarrow \begin{cases} \alpha_1 = 1 + 2\alpha_3 \\ \alpha_2 = 4 - 3\alpha_3. \end{cases}$$

  Obtém-se uma infinidade de CLs consoante o valor do coeficiente $\alpha_3 \in \mathbb{R}$:

$$\begin{aligned} b = \begin{bmatrix} 5 \\ 1 \end{bmatrix} &= (1 + 2\alpha_3) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (4 - 3\alpha_3) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ -2 \end{bmatrix} \\ &= (1 + 2\alpha_3)v_1 + (4 - 3\alpha_3)v_2 + \alpha_3 v_3. \end{aligned}$$

Conceitos de espaço gerado e espaço das colunas

Definições de 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,

$$\langle v_1, \ldots, v_n \rangle = \{ b \in \mathbb{R}^m \ :\ b = \alpha_1 v_1 + \cdots + \alpha_n v_n,\ \text{com } \alpha_1, \ldots, \alpha_n \in \mathbb{R} \}.$$

- 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,

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

Pelos resultados do slide 76 tem-se

$$\mathcal{C}(A) = \langle v_1, \ldots, v_n \rangle = \left\{ b \in \mathbb{R}^m \colon [\,A \mid b\,] \text{ é possível} \right\}.$$

Nas condições das definições anteriores tem-se o seguinte resultado que nos dá o segundo subespaço vetorial fundamental associado a uma matriz.

Teorema

Tem-se que $\mathcal{C}(A) = \langle v_1, \ldots, v_n \rangle$ define um subespaço vetorial de $\mathbb{R}^m$.

Espaço gerado e das colunas - exemplos dos slides 74 e 77 revisitados

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

$$\begin{aligned} \langle v_1, v_2 \rangle = \mathcal{C}(A) &= \{b \in \mathbb{R}^3 \ : \ Ax = b \text{ é possível}\} \\ &= \{b \in \mathbb{R}^3 \ : \ [A|b] \text{ é possível}\} \end{aligned}$$

Aplicando o método de eliminação de Gauss à matriz $[A|b]$ tem-se

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

Espaço gerado e das colunas - exemplos dos slides 74 e 77 revisitados (cont.)

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

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

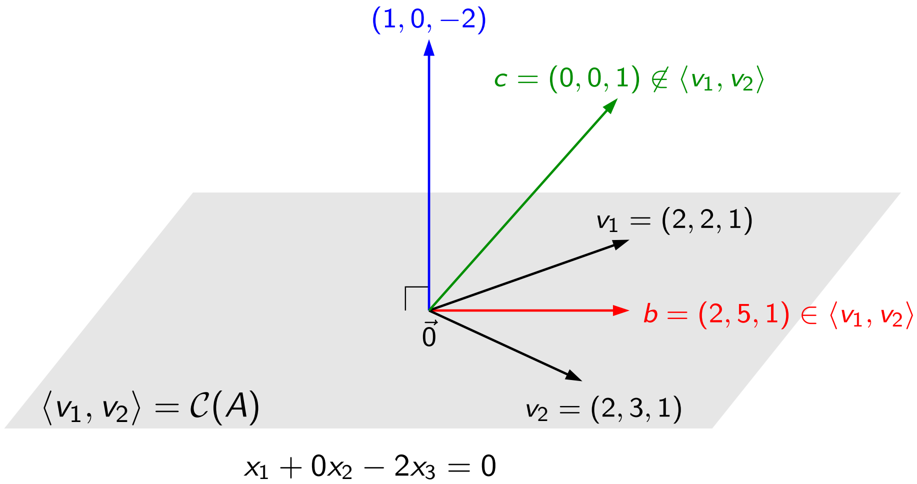
que define o plano de $\mathbb{R}^3$ de equação cartesiana $x_1 + 0x_2 - 2x_3 = 0$, que passa na origem e tem vetor normal $(1, 0, -2)$.

Os vetores que são CL de v_1 e v_2 são os vetores do espaço gerado $\langle v_1, v_2 \rangle$, isto é, os vetores $b = (b_1, b_2, b_3)$ que satisfazem a equação do plano $x_1 + 0x_2 - 2x_3 = 0$.

Por exemplo:

  - O vetor $b = (2, 5, 1) = -2v_1 + 3v_2$ (ver o slide 75) 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 verfica pois satisfaz a equação $x_1 + 0x_2 - 2x_3 = 0$.
  - Já o vetor $c = (0, 0, 1)$ do slide 77 não é CL de v_1 e v_2 (como vimos) e portanto não pertence ao espaço gerado $\langle v_1, v_2 \rangle = \mathcal{C}(A)$, o que se verifica pois não satisfaz a equação $x_1 + 0x_2 - 2x_3 = 0$.

Espaço gerado e das colunas - interp. geom. dos exemplos dos slides 74 e 77



Algoritmo para determinar o espaço das colunas / espaço gerado

- `Input:` Matriz quadrada $A_{m\times n} = [\, v_1 \;\; v_2 \;\; \cdots \;\; v_n \,]$
- `Objectivo:` Determinar $\mathcal{C}(A) = \langle v_1, v_2, \ldots, v_n \rangle$
- Considerar o vetor genérico $b = (b_1, \ldots, b_m) \in \mathbb{R}^m$ e aplicar a fase descendente do método de Gauss à matriz ampliada do sistema $Ax = b$:

$$[A|b] \;\to\; \cdots \;\to [A'|b'] \quad (\text{com } A' \text{ em escada})$$

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

  - se A' não tem linhas nulas, $[A|b]$ é possível para qualquer $b \in \mathbb{R}^m$ e portanto,

$$\mathcal{C}(A) \;=\; \mathbb{R}^m,$$

  ou seja, $v_1, \ldots, v_n$ geram $\mathbb{R}^m$.

  - se A' tem linhas nulas com índices $i_1, i_2, \ldots, i_k$, então

$$\mathcal{C}(A) = \{(b_1, \ldots, b_m) \;:\; b'_{i_1} = b'_{i_2} = \cdots = b'_{i_k} = 0\} \neq \mathbb{R}^m,$$

  ou seja, obtém-se um sistema de equações definidoras para $\mathcal{C}(A) = \langle v_1, v_2, \ldots, v_n \rangle$, anulando as componentes do vetor b' que estão associadas às linhas nulas da matriz em escada A'.

(In)dependência linear

Definição de (in)dependência linear([9])

- Um conjunto formado por um único vetor $\{v\}$ diz-se linearmente dependente se $v = \vec{0}$ e diz-se linearmente independente se $v \neq \vec{0}$.
- Um conjunto com dois ou mais vetores $\{v_1, \ldots, v_k\}$ diz-se linearmente dependente se pelo menos um dos vetores v_i for combinação linear dos restantes vetores do conjunto. Caso contrário, diz-se linearmente independente.

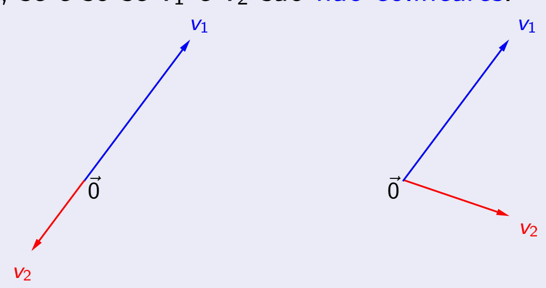
- Um conjunto de vetores é portanto linearmente independente quando não existem relações de linearidade entre esses vetores, no sentido em que nenhum desses vetores se pode escrever como CL dos restantes vetores.
- No caso do conjunto conter apenas 2 vetores, isso significa que nenhum dos 2 vetores é múltiplo do outro vetor.

[9] A definição que está no *Texto de Apoio* é a definição mais usual que aparece na literatura, mas é menos intuitiva e envolve o conceito de combinação linear nula.

(In)dependência linear

Consequências

- $\{v_1, v_2\}$ é linearmente independente se e só se v_1 e v_2 não são múltiplos um do outro, isto é, se e só se v_1 e v_2 são não colineares.



$\{v_1, v_2\}$ l.d. $\qquad$ $\{v_1, v_2\}$ l.i.

- Um conjunto de vetores que contenha um conjunto linearmente dependente é ainda linearmente dependente. Em particular, um conjunto de vetores que contenha o vetor nulo ou vetores múltiplos entre si é linearmente dependente.
- Reciprocamente, um subconjunto não vazio de um conjunto linearmente independente de vetores é ainda linearmente independente.

(In)dependência linear

Exemplos

- $\{v_1\} = \{(0,0,0)\}$ é lin. dept. pois $v_1 = \vec{0}$.
- $\{v_1\} = \{(1,2,3)\}$ é lin. indept. pois $v_1 \neq \vec{0}$.
- $\{v_1, v_2\} = \{(0,0,0),(1,2,3)\}$ é lin. dept. pois $v_1 = 0\, v_2$ (conjuntos de vetores que contenham o vetor nulo são lin. dept.).
- $\{v_1, v_2\} = \{(1,2,3),(2,4,6)\}$ é lin. dept. pois $v_2 = 2\, v_1$, isto é, são colineares.
- $\{v_1, v_2\} = \{(1,2,3),(0,0,1)\}$ é lin. indept. pois nenhum dos vetores é múltiplo do outro, isto é, são não colineares.
- $\{v_1, v_2, v_3\} = \{(1,2,3),(2,4,6),(0,0,1)\}$ é lin. dept. pois $v_2 = 2v_1 + 0v_3$ (conjuntos de vetores que contenham vetores múltiplos entre si são lin. dept.).
- $\{v_1, v_2, v_3, v_4\} = \{(1,2,0,1),(1,1,-1,0),(0,1,1,1),(1,0,0,-1)\}$ (?).
- $\{v_1, v_2, v_4\} = \{(1,2,0,1),(0,1,1,1),(1,0,0,-1)\}$ (?).

Exemplos (cont.)

Consideremos agora o penúltimo exemplo do slide 85,

$$\{v_1, v_2, v_3, v_4\} = \{(1,2,0,1),(1,1,-1,0),(0,1,1,1),(1,0,0,-1)\}.$$

Aplicando a fase descendente à matriz $A = [\, v_1 \ v_2 \ v_3 \ v_4 \,]$ obtém-se

$$A = [\, v_1 \ v_2 \mid v_3 \ v_4 \,] \to \cdots \to A' = \left[\begin{array}{cc|cc} 1 & 1 & 0 & 1 \\ 0 & -1 & 1 & -2 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

"Ignorando" a última coluna, conclui-se que v_3 é CL de v_1 e v_2, porque o sistema $[\, v_1 \ v_2 \mid v_3 \,]$ é possível (uma vez que a terceira coluna de A' não tem pivot). Logo existem $\alpha_1, \alpha_2 \in \mathbb{R}$ tais que

$$v_3 = \alpha_1 v_1 + \alpha_2 v_2.$$

Esta CL pode ser estendida ao vetor v_4 multiplicando-o pelo coeficiente nulo:

$$v_3 = \alpha_1 v_1 + \alpha_2 v_2 + 0\, v_4.$$

Logo um dos vetores de $\{v_1, v_2, v_3, v_4\}$ é CL dos restantes 3 vetores do conjunto e portanto $\{v_1, v_2, v_3, v_4\}$ é linearmente dependente.
Note-se que a CL $v_3 = \alpha_1 v_1 + \alpha_2 v_2 + 0\, v_4 \Leftrightarrow \alpha_1 v_1 + \alpha_2 v_2 - v_3 + 0\, v_4 = \vec{0}$, o que implica que o sistema homogéneo $[\, A \mid \vec{0} \,] = [\, v_1 \ v_2 \ v_3 v_4 \mid \vec{0} \,]$ é indeterminado, uma vez contém a solução não trivial $(\alpha_1, \alpha_2, -1, 0) \neq \vec{0}$.

Exemplos (cont.)

Consideremos agora o último exemplo do slide 85,

$$\{v_1, v_2, v_4\} = \{(1,2,0,1), (0,1,1,1), (1,0,0,-1)\}.$$

Aplicando a fase descendente à matriz $A = [\, v_1 \;\; v_2 \;\; v_4 \,]$ obtém-se

$$A = [\, v_1 \;\; v_2 \;\; v_4 \,] \to \cdots \to A' = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Neste caso todas as colunas de A' têm pivot, isto é, $car(A) = 3 = n$.

Portanto o sistema homogéno $[\, A \,|\, \vec{0} \,]$ só tem a solução trivial $\vec{0}$ o que implica que nenhum dos vetores pode ser CL dos restantes 2 vetores (caso contrário iriam existir soluções não triviais deste sistema como ocorreu no exemplo do slide anterior). Logo $\{v_1, v_2, v_4\}$ é linearmente independente.

Observação

Mais geralmente, considerando $A_{m\times n} = [\, v_1 \;\; v_2 \;\; \cdots \;\; v_n \,] \to \cdots \to A'$, pode-se mostrar que cada vetor de $\{v_1 \;\; v_2, \ldots, v_n\}$ associado a uma coluna sem pivot em A' é CL dos restantes vetores do conjunto. Reciprocamente, se todas as colunas de A' tiverem pivot, o sistema $Ax = \vec{0}$ é determinado e pode-se mostrar que nenhum dos vetores de $\{v_1 \;\; v_2, \ldots, v_n\}$ é CL dos restantes vetores do conjunto.

Caracterização da (in)dependência linear

A partir da observação do slide anterior obtém-se imediatamente o seguinte critério baseado no método de Gauss para decidir sobre a independência linear de um conjunto de vetores.

Critério para decidir a independência linear de um conjunto de vetores

Sejam $v_1, \ldots, v_n \in \mathbb{R}^m$ e apliquemos a fase descendente do método de Gauss à matriz A formada por estes vetores,

$$A_{m\times n} = [\, v_1 \quad \cdots \quad v_n \,] \to \cdots \to A',$$

com A' matriz em escada. Têm-se as seguintes equivalências:

$$\{v_1, \ldots, v_n\} \text{ é linearmente independente}$$

$$\Updownarrow$$

$$\text{Todas as colunas de } A' \text{ têm pivot}$$

$$\Updownarrow$$

$$\text{car}(A) = n$$

$$\Updownarrow$$

$$\mathcal{N}(A) = \{\vec{0}\}$$

Independência linear

Uma vez que $\text{car}(A_{m\times n}) \leq \min\{m, n\}$, tem-se seguinte resultado.

Cardinalidade máxima de um conjunto linearmente independente

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

Exercícios na aula

- Sejam $v_1 = (1, \alpha, 1)$, $v_2 = (0, 1, -1)$ e $v_3 = (\alpha, 3, 3)$.
  Decida sobre a independência linear de $\{v_1, v_2, v_3\}$ em função de α.
- Decida sobre a independência linear de
  $\{(1, 2, 3), (4, 5, 6), (7, 8, 9), (10, 11, 12)\}$.

Solução: $\alpha \neq -3, 2$ Qualquer conjunto com mais que 3 vetores em $\mathbb{R}^3$ é l.d.

Base e dimensão de um subespaço vetorial

Definição de base

Sejam V subespaço vetorial de $\mathbb{R}^m$ e $v_1, \ldots, v_n \in V$. Diz-se que $\{v_1, \ldots, v_n\}$ é *base* de V se as seguintes condições forem verificadas:

- $\{v_1, \ldots, v_n\}$ é linearmente independente
- $\langle v_1, \ldots, v_n\rangle = V$, isto é, $v_1, \ldots, v_n$ geram de V.

- Intuitivamente uma base de um subespaço vetorial V é um subconjunto de vetores de V que não contém vetores redundantes (i.e., nenhum vetor do subconjunto é CL dos restantes vetores do subconjunto) e tal que todo o vetor de V é combinação linear dos vetores da base.
- Pode-se mostrar que quaisquer duas bases de um subespaço vetorial V possuem o mesmo número de vetores. Faz portanto sentido o seguinte.

Definição de dimensão

Chama-se dimensão de um subespaço vetorial V e denota-se $\dim V$, ao número de vetores de uma qualquer base de V.

Convenciona-se que $\{\}$ é a base do subespaço mininal $\{\vec{0}\}$, tendo-se portanto $\dim\{\vec{0}\} = 0$, uma vez que esta base não contém vetores.

Base canónica e dimensão do subespaço maximal $\mathbb{R}^m$

Sejam $e_1 = (1,0,0)$, $e_2 = (0,1,0)$ e $e_3 = (0,0,1)$ (colunas da matriz identidade de ordem 3, I_3). Vejamos que $\{e_1, e_2, e_3\}$ é base de $\mathbb{R}^3$

- Para qualquer $b = (b_1, b_2, b_3) \in \mathbb{R}^3$, tem-se

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

  Logo todo o vetor $b \in \mathbb{R}^3$ é CL de e_1, e_2 e e_3 e portanto $\langle e_1, e_2, e_3\rangle = \mathbb{R}^3$.
- $\{e_1, e_2, e_3\}$ é linearmente independente pois $[\, e_1 \;\; e_2 \;\; e_3 \,]$ é a matriz identidade I_3, que já está em escada e tem todas as colunas com pivot.

Logo $\{e_1, e_2, e_2\}$ é base de $\mathbb{R}^3$, sendo designada por base canónica de $\mathbb{R}^3$. Em particular $\dim \mathbb{R}^3 = 3$ (número de vetores da base). Mais geralmente, tem-se:

Base canónica e dimensão de $\mathbb{R}^m$

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

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

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

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

Caracterização das bases de $\mathbb{R}^m$

- Vimos no slide anterior que $\mathbb{R}^m$ admite a base canónica formada pelos vetores que constituem as m colunas da matriz identidade, e em particular que $\dim \mathbb{R}^m = m$. Logo qualquer base de $\mathbb{R}^m$ tem que possuir m vetores.
- Por outro lado, uma base tem que ser um conjunto linearmente independente.

O próximo resultado mostra que estas duas condições são também suficientes.

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

Sejam $v_1, \ldots, v_m \in \mathbb{R}^m$. Tem-se que

$$\{v_1, \ldots, v_m\} \text{ é base de } \mathbb{R}^m \quad \Leftrightarrow \quad \{v_1, \ldots, v_m\} \text{ é linearmente independente.}$$

Por outras palavras, as bases de $\mathbb{R}^m$ são os conjuntos linearmente independentes formados por m vetores de $\mathbb{R}^m$.

De facto, considerando $A_{m\times m} = [\, v_1 \;\; \cdots \;\; v_m \,] \to \cdots \to A'$ com A' em escada, tem-se que as m colunas de A' têm pivot porque $\{v_1, \ldots, v_m\}$ é l.i. Logo todas as linhas A' também têm pivot porque o número de linhas de A' = colunas de A'. Logo A' não possui linhas nulas e portanto o sistema $[\, A \mid b \,]$ é possível para qualquer $b \in \mathbb{R}^m$, obtendo-se (ver também o slide 82),

$$\langle v_1, \ldots, v_m\rangle = \mathcal{C}(A) = \{b \in \mathbb{R}^m \;:\; [\, A \mid b \,] \;\; \text{é possível}\} = \mathbb{R}^m.$$

Logo $\{v_1, \ldots, v_m\}$ verifica as duas condições da definição de base do slide 90.

Bases de $\mathbb{R}^3$

Exercícios na aula

Quais dos seguintes conjuntos de vetores definem bases de $\mathbb{R}^3$?

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

Nos casos em que não define uma base identifique a(s) condição (ões) da definição de base do slide 90 que falha(m).

Solução: Apenas o terceiro. Não gera $\mathbb{R}^3$ — falham as duas — não é l.i.

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, $\quad V = \mathcal{N}(A), \quad$ 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, $\quad V = \mathcal{C}(A), \quad$ 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 / CS sistema linear homogéneo

Algoritmo

`Input:` Matriz A do tipo $m \times n$
`Objectivo:` Base para $\mathcal{N}(A)$ / CS de um sistema linear homogéneo

- Determinar $\mathcal{N}(A)$ resolvendo o sistema homogéneo $[\,A\,|\,\vec{0}\,]$.
  Seja k o número de variáveis livres do sistema.
- Se $k = 0$ 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 $[\,A\,|\,\vec{0}\,]$ obtidas deste modo constitui uma base para $\mathcal{N}(A)$.

Em particular,

$$\boxed{\dim \mathcal{N}(A) = \text{número de variáveis livres} = n - \operatorname{car}(A)}$$

Base para o espaço nulo de uma matriz / CS sistema homogéneo

Exercícios na aula

Indique uma base e a dimensão dos seguintes subespaços vetoriais:

- O espaço nulo da matriz $A = \begin{bmatrix} 1 & 0 & 2 & 4 \\ 1 & 2 & 0 & 2 \\ -1 & -1 & -1 & -3 \end{bmatrix}$.
- O hiperplano de $\mathbb{R}^4$,

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

- **TPC**: calcular uma base para o espaço nulo do slide 93.

Resolução dos exercícios na aula - base do espaço nulo de A

Vamos determinar uma base do espaço nulo da matriz $A = \begin{bmatrix} 1 & 0 & 2 & 4 \\ 1 & 2 & 0 & 2 \\ -1 & -1 & -1 & -3 \end{bmatrix}$.

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

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

$$\rightsquigarrow \begin{cases} x_1 \qquad 2x_3 + 4x_4 = 0 \\ \quad x_2 - x_3 - \ x_4 = 0 \\ \qquad\qquad\quad 0 = 0 \end{cases}.$$

Como existem variáveis livres (x_3 e x_4) tem-se $\mathcal{N}(A) \neq \{\vec{0}\}$, 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 linearmente independente. Logo $\{v_1, v_2\}$ é base de $\mathcal{N}(A)$ (ver o slide 90), tendo-se $\dim \mathcal{N}(A) = 2$.

Algumas observações

- O primeiro vetor da base do slide anterior, $v_1 = (-2, 1, 1, 0)$, corresponde à solução do sistema homogéneo $[\,A \mid \vec{0}\,]$ considerando a variável livre $x_3 = 1$ e a variável livre $x_4 = 0$. De facto,

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

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

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

- 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 $[\,A \mid \vec{0}\,]$ possua variáveis livres) e conduz sempre a bases de $\mathcal{N}(A)$, não sendo necessário provar que o conjunto é linearmente independente.
- Se o sistema $[\,A \mid \vec{0}\,]$ não possuir variáveis livres então é determinado, tendo-se $\mathcal{N}(A) = \{\vec{0}\}$. Nesse caso a base de $\mathcal{N}(A)$ é $\{\}$ e $\dim \mathcal{N}(A) = 0$.

Resolução dos exercícios na aula - base do hiperplano V

Pretende-se determinar uma base para o hiperplano de $\mathbb{R}^4$ do slide 96,

$$V = \{(x_1, x_2, x_3, x_4) : x_1 - 2x_2 + x_3 - 5x_4 = 0\} = \mathcal{N}([\,1 \quad -2 \quad 1 \quad -5\,]).$$

- Tem-se,

$$\begin{aligned} V &= \{(x_1, x_2, x_3, x_4) : x_1 = 2x_2 - x_3 + 5x_4,\ x_2, x_3, x_4 \in \mathbb{R}\} \\ &= \{(2x_2 - x_3 + 5x_4, x_2, x_3, x_4) :\ x_2, x_3, x_4 \in \mathbb{R}\} \\ &= \{(2x_2, x_2, 0, 0) + (-x_3, 0, x_3, 0) + (5x_4, 0, 0, x_4) :\ x_2, x_3, x_4 \in \mathbb{R}\} \\ &= \underbrace{\{x_2(2,1,0,0) + x_3(-1,0,1,0) + x_4(5,0,0,1) :\ x_2, x_3, x_4 \in \mathbb{R}\}}_{\text{todas as somas de múltiplos de } (2,1,0,0),\ (-1,0,1,0) \text{ e } (5,0,0,1)} \\ &= \langle (2,1,0,0), (-1,0,1,0), (5,0,0,1) \rangle. \end{aligned}$$

  Logo $(2,1,0,0), (-1,0,1,0), (5,0,0,1)$ geram V.
- Note-se que $(2,1,0,0)$ corrresponde à solução em que $x_2 = 1$ e $x_3 = x_4 = 0$, $(-1,0,1,0)$ à solução com $x_3 = 1$ e $x_2 = x_4 = 0$ e finalmente, $(5,0,0,1)$ à solução com $x_4 = 1$ e $x_2 = x_3 = 0$.
- O conjunto das soluções do espaço nulo em que alternadamente, uma das variáveis livres é não nula e as restantes variáveis livres são nulas, é sempre linearmente independente (confirme a independência linear).

Logo $\{(2,1,0,0), (-1,0,1,0), (5,0,0,1)\}$ é base de V. Em particular, $\dim V = 3$ $=$número de vetores da base $=$ número de variáveis livres.

Exercício na aula - base para espaço das colunas

Vamos determinar uma base para o espaço nulo das colunas da matriz

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

Recordemos que $\mathcal{C}(A) = \langle v_1, v_2, v_3, v_4 \rangle = \{b = (b_1, b_2, b_3) \ :\ [\,A\,|\,b\,] \text{ é possível}\}$.
Aplicando a fase descendente à matriz $[\,A\,|\,b\,]$ obtém-se,

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

Logo para o sistema $[\,A\,|\,b\,]$ ser possível, $b_3' = b_3 + \frac{3}{2}b_1 - \frac{1}{2}b_2 = 0$ e portanto

$$\mathcal{C}(A) = \langle v_1, v_2, v_3, v_4 \rangle = \left\{ b = (b_1, b_2, b_3) \ :\ b_3 + \tfrac{3}{2}b_1 - \tfrac{1}{2}b_2 = 0 \right\}.$$

Uma vez que as que as operações elementares que são efetuadas no método de Gauss dependem apenas dos pivots e as colunas sem pivot não influenciam a discussão do sistema $[\,A\,|\,b\,]$, podemos eliminar a terceira e quarta colunas de A (associadas às colunas sem pivot em A'), sem alterar o espaço das colunas, isto é,

$$\mathcal{C}(A) = \langle v_1, v_2, v_3, v_4 \rangle = \left\{ b = (b_1, b_2, b_3) \ :\ b_3 + \tfrac{3}{2}b_1 - \tfrac{1}{2}b_2 = 0 \right\} = \langle v_1, v_2 \rangle.$$

Como $\{v_1, v_2\}$ é linearmente independente, porque está associado às colunas com pivot em A', e gera $\mathcal{C}(A)$, conclui-se que $\{v_1, v_2\}$ é base de $\mathcal{C}(A) = \langle v_1, v_2, v_3, v_4 \rangle$.
Em particular, $\dim \mathcal{C}(A) = \text{car}(A) = 2$ (número de pivots em A').

Base para o espaço das colunas / espaço gerado por um conjunto de vetores

Algoritmo

`Input:` $A = [\, v_1 \ \cdots \ v_n \,]$, 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:
$$A \to \cdots \to A' \text{ com } A' \text{ escada.}$$
- Para obtermos uma base de $\mathcal{C}(A) = \langle v_1, \ldots, v_n \rangle$ consideramos o subconjunto das colunas de A associado às colunas com pivot em A'.

Em particular, tem-se

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

- Voltando ao exemplo do slide anterior, tem-se que $\{v_1, v_2\}$ é base de $\mathcal{C}(A) = \langle v_1, v_2, v_3, v_4 \rangle$, porque v_1 e v_2 são as colunas de A associadas às colunas com pivot em A' (1^a e 2^a colunas de A').
- Em particular, $\langle v_1, v_2, v_3, v_4 \rangle = \langle v_1, v_2 \rangle$, o que significa que os geradores v_3 e v_4, associados às colunas sem pivot em A', são redundantes.
- A característica de uma matriz A é muitas vezes definida como $\dim \mathcal{C}(A)$.

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

- Seja A matriz do tipo $m \times n$. Pelos resultados do slide 95 e do slide anterior, tem-se
  - $\dim \mathcal{N}(A) = n - \text{car}(A)$
  - $\dim \mathcal{C}(A) = \text{car}(A)$
- Daqui resulta imediatamente a seguinte resultado que estabelece uma relação entre as dimensões dos dois subespaços fundamentais associados à matriz A.

Proposição

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

	subespaços vetoriais	dimensão
$\mathbb{R}^2$	$\{\vec{0}\}$	0
	retas que passam na origem	1
	$\mathbb{R}^2$	2
$\mathbb{R}^3$	$\{\vec{0}\}$	0
	retas que passam na origem	1
	planos que passam na origem	2
	$\mathbb{R}^3$	3

Em geral, têm-se as seguintes caracterizações dos subespaços minimal e maximal de $\mathbb{R}^m$ (m arbitrário), em termos das suas dimensões:

- $V = \{\vec{0}\}$ (subespaço minimal) $\Leftrightarrow$ $\dim V = 0$
- $V = \mathbb{R}^m$ (subespaço maximal) $\Leftrightarrow$ $\dim V = m$

Inclusão / igualdade de subespaços vetoriais e dimensão

- Pelo quadro do slide anterior, se um subespaço vetorial U estiver contido num subespaço vetorial V de dimensão um (reta) em $\mathbb{R}^2$ ou em $\mathbb{R}^3$, então $U = \{\vec{0}\}$ se $\dim U = 0$, ou U define uma reta se $\dim U = 1$, tendo-se nesse caso $U = V$
- Analogamente, se U estiver contido num subespaço vetorial V de dimensão dois (plano) em $\mathbb{R}^3$, então $U = \{\vec{0}\}$ se $\dim U = 0$, U é uma reta se $\dim U = 1$, ou U define um plano se $\dim U = 2$, tendo-se nesse caso $U = V$

Mais geralmente, tem-se o seguinte resultado.

Teorema

Sejam U e V subespaços vetoriais de $\mathbb{R}^m$ com $U \subset V$. Então $\dim U \leq \dim V$, tendo-se $U = V$ se e só se $\dim U = \dim V$

Para podemos usar o resultado anterior vamos começar ver como é que podemos mostrar que um subsepaço dado por geradores está contido noutro subespaço vetorial.

Vetor pertence ao espaço nulo / colunas de uma matriz

Recordatória

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

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

Exercício na aula

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,

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

é possível.

Inclusão entre espaços gerados por vetores

Proposição

Sejam $u_1, \ldots, u_k \in \mathbb{R}^m$, $U = \langle u_1, \ldots, u_k \rangle$ e V subespaço vetorial de $\mathbb{R}^m$. Então

$$U = \langle u_1, \ldots, u_k \rangle \subset V \quad \Leftrightarrow \quad \begin{cases} u_1 \in V, \\ \quad \vdots \\ u_k \in V. \end{cases}.$$

Em particular,

- Se $V = \mathcal{N}(A)$, então

$$U \subset V \quad \Leftrightarrow \quad u_i \in \mathcal{N}(A),\ i = 1, \ldots, k \quad \Leftrightarrow \quad A\,u_i = \vec{0},\ i = 1, \ldots, k.$$

- Se $V = \mathcal{C}(A) = \langle v_1, v_2 \ldots, v_n \rangle$, então

$$\begin{aligned} U \subset V \quad &\Leftrightarrow \quad u_i \in \mathcal{C}(A) = \langle v_1, v_2 \ldots, v_n \rangle,\ i = 1, \ldots, k \\ &\Leftrightarrow \quad [\,A \mid u_i\,] \text{ é possível},\ i = 1, \ldots, k \\ &\Leftrightarrow \quad [\,v_1 \ \cdots \ v_n \mid u_1\,],\ \ldots,\ [\,v_1 \ \cdots \ v_n \mid u_k\,] \text{ todos possíveis} \\ &\Leftrightarrow \quad [\,v_1 \ \cdots \ v_n \mid u_1 \ \cdots \ u_k\,] \quad \text{possível} \end{aligned}$$

Igualdade entre espaços gerados por vetores

Exercício na aula

Considere $U = \langle u_1, u_2 \rangle$ e $V = \langle v_1, v_2 \rangle$, em que $u_1 = (1, 0, 2)$, $u_2 = (2, 1, 1)$, $v_2 = (-1, -1, 1)$ e $v_2 = (0, -1, 3)$. Mostre que $U = V$.

- Primeira abordagem: provar as inclusões $U \subset V$ e $V \subset U$.
  - Para mostrar que $U \subset V$, isto é, que $\langle u_1, u_2 \rangle \subset \langle v_1, v_2 \rangle$, vamos provar que $[\, v_1 \;\; v_2 \mid u_1 \;\; u_2 \,]$ é possível (ver o slide 106). De facto,

$$[v_1 \;\; v_2 | u_1 \;\; u_2] = \left[\begin{array}{rr|rr} -1 & 0 & 1 & 2 \\ -1 & -1 & 0 & 1 \\ 1 & 3 & 2 & 1 \end{array}\right] \to \cdots \to \left[\begin{array}{rr|rr} -1 & 0 & 1 & 2 \\ 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

  - Analogamente se mostra que $V \subset U$, provando que $[\, u_1 \;\; u_2 \mid v_1 \;\; v_2 \,]$ é possível, o que fica como exercício.
- Segunda abordagem: mostrar que $U \subset V$ (já vimos) e que $\dim U = \dim V$.
  - Vejamos que $\dim U = \dim V$: tem-se $U = \langle u_1, u_2 \rangle = \mathcal{C}(A)$ com $A = [\, u_1 \;\; u_2 \,]$. Logo $\dim U = \dim \mathcal{C}(A) = \operatorname{car}(A) = 2$ (verifique). Analogamente, $\dim V = \dim \mathcal{C}(B) = \operatorname{car}(B) = 2$, com $B = [\, v_1 \;\; v_2 \,]$. Alternativamente, $U = \langle u_1, u_2 \rangle$ com $\{u_1, u_2\}$ linearmente independente (vetores não colineares), logo $\{u_1, u_2\}$ é base de U e portanto $\dim U = 2$ porque é o número de vetores da base (e analogamente para V).

Consequências da dimensão (cont.)

- Vimos anteriormente que as bases de $\mathbb{R}^n$ são os conjuntos linearmente independententes formados por n vetores de $\mathbb{R}^n$.
- Temos uma caracterização análoga para qualquer subespaço vetorial V cuja dimensão se conhece!

Teorema

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

Nos exercícios podemos aplicar este resultado usando uma formulação um pouco distinta:

Observação

Sejam V é um 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 se verificarem as seguintes 3 condições:

- $v_1, \ldots, v_k \in V$.
- $\{v_1, \ldots, v_k\}$ é linearmente independente.
- $\dim V = k$.

Consequências da dimensão (exercício)

Exercício na aula

Considere $v_1 = (-2, 1, 0, 1)$, $v_2 = (-1, 0, -1, 1)$ e

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

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

Pela observação do slide anterior basta ver as seguintes condições:

- $v_1, v_2 \in \mathcal{N}(A)$. De facto, $Av_1 = \vec{0}$ e $Av_2 = \vec{0}$ (verifique!).
- $\{v_1, v_2\}$ é linearmente independente. De facto, v_1 e v_2 são não colineares.
- $\dim \mathcal{N}(A) = 2$. De facto, a matriz em escada A' obtida a partir de A tem duas colunas sem pivot (verifique!)

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

Ortogonalidade entre vetores

- 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

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

Ortogonalidade entre um vetor e um subespaço vetorial

Definição de vetor ortogonal a um subespaço vetorial

Sejam $u \in \mathbb{R}^m$ e V subespaço vetorial de $\mathbb{R}^m$.
Diz-se que u é *ortogonal* a V e denota-se $u \perp V$ se $u \perp v$, $\forall v \in V$, isto é, se $u \cdot v = 0$, $\forall v \in V$

Por outras palavras, u é ortogonal ao subespaço vetorial V se for ortogonal a todos os vetores de V.

Vejamos um exemplo:

- Consideremos o plano $V = \{(x_1, x_2, x_3) \colon x_1 + 2x_2 + 3x_3 = 0\}$. Tem-se,
$$V = \{(x_1, x_2, x_3) \colon (x_1, x_2, x_3) \cdot (1, 2, 3) = 0\},$$
o que mostra os vetores de V são os vetores $\mathbb{R}^3$ que são ortogonais a $(1, 2, 3)$ e portanto $(1, 2, 3) \perp V$.
- Este resultado pode ser generalizado para qualquer plano que passe na origem, $V = \{(x_1, x_2, x_3) \colon a\,x_1 + b\,x_2 + c\,x_3 = 0\}$. De facto,
$$V = \{(x_1, x_2, x_3) \colon (x_1, x_2, x_3) \cdot (a, b, c) = 0\},$$
tendo-se $(a, b, c) \perp V$, razão pela qual o vetor dos coeficientes (a, b, c) é designado por vetor normal ao plano V.

Vetor ortogonal a um subespaço dado por geradores

Proposição

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

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

Por outras palavras, para mostrar que $u \perp V$ basta mostrar que u é ortogonal a um conjunto de geradores de V (ou, em particular, a uma base de V).

Demonstração

- Se $u \perp V$ então $u \perp v$, $\forall v \in V$, e em particular, $u \perp v_1, \dots, u \perp v_n$
- Reciprocamente suponhamos que $u \perp v_1, \dots, u \perp v_n$ e seja $b \in V$. Como $V = \langle v_1 \dots, v_n\rangle$, existem $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tais que $b = \alpha_1 v_1 + \cdots + \alpha_n v_n$ e tem-se
$$\begin{aligned} u \cdot b &= u \cdot (\alpha_1 v_1 + \cdots + \alpha_n v_n) = u \cdot (\alpha_1 v_1) + \cdots + u \cdot (\alpha_n v_n) \\ &= \alpha_1 (\underbrace{u \cdot v_1}_{=0}) + \cdots + \alpha_n (\underbrace{u \cdot v_n}_{=0}) = 0 \end{aligned}$$

Logo $u \perp b$, $\forall b \in V$ e portanto $u \perp V$.

Exemplo na aula

Sejam $v_1 = (1, 2, -1)$, $v_2 = (2, 0, 2)$, $V = \langle v_1, v_2 \rangle$ e $b = (1, -1, -1)$.
Tem-se:

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

Como $b \perp v_1$ e $b \perp v_2$, conclui-se pelo resultado do slide anterior 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,

$$V^\perp = \{u \in \mathbb{R}^m \;:\; u \perp V\}$$

- 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 $\mathcal{C}(A)^\perp = \langle v_1, v_2 \rangle^\perp$.
Dado $b = (b_1, b_2, b_3) \in \mathbb{R}^3$ tem-se:

$$b \perp \mathcal{C}(A) = \langle v_1, v_2 \rangle \quad \Leftrightarrow \quad \left\{ \begin{array}{l} v_1 \cdot b = 0 \\ v_2 \cdot b = 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} v_1^T b = 0 \\ \\ v_2^T b = 0 \end{array} \right.$$

$$\Leftrightarrow \left\{ \begin{array}{l} \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = 0 \\ \\ \begin{bmatrix} 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = 0 \end{array} \right.$$

$$\Leftrightarrow \begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow A^T b = \vec{0}.$$

Obtivemos portanto a seguinte relação,

$$b \perp \mathcal{C}(A) = \langle v_1, v_2 \rangle \Leftrightarrow A^T b = \vec{0} \Leftrightarrow b \in \mathcal{N}(A^T).$$

Logo, $\mathcal{C}(A)^\perp = \left\{ b \in \mathbb{R}^3 \;:\; b \perp \mathcal{C}(A) \right\} = \mathcal{N}(A^T)$.

Exemplo (cont.)

Calculando $\mathcal{N}(A^T)$ obtém-se (confirme!),

$$\begin{aligned} \mathcal{N}(A^T) &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 \;:\; x_1 = -x_3,\; x_2 = x_3,\; x_3 \in \mathbb{R}\} \\ &= \{(-x_3, x_3, , x_3) \;:\; x_3 \in \mathbb{R}\} = \langle (-1, 1, 1) \rangle. \end{aligned}$$

Note-se que,

$$\mathcal{C}(A) = \left\{ b : [\, A \,|\, b \,] \text{ é possível} \right\} = \left\{ (b_1, b_2, b_3) \;:\; -b_1 + b_2 + b_3 = 0 \right\},$$

como se pode verificar aplicando a fase descendente à matriz $[\, A \,|\, b \,]$:

$$[\, A \,|\, b \,] = \left[\begin{array}{rr|c} 1 & 2 & b_1 \\ 2 & 0 & b_2 \\ -1 & 2 & b_3 \end{array} \right] \rightarrow \cdots \rightarrow \left[\begin{array}{rr|c} 1 & 2 & b_1 \\ 0 & -4 & b_2 - 2b_1 \\ 0 & 0 & b_3 + b_2 - b_1 \end{array} \right].$$

Constata-se assim que $\mathcal{C}(A) = \langle v_1, v_2 \rangle$ é o plano de $\mathbb{R}^3$ que passa na origem com equação cartesiana $-x_1 + x_2 + x_3 = 0$, sendo que o complemento ortogonal deste plano, $\mathcal{C}(A)^\perp = \mathcal{N}(A^T) = \langle (-1, 1, 1) \rangle$, define a reta que passa na origem, perpendicular ao plano $\mathcal{C}(A)$, uma vez que tem a direção $(-1, 1, 1)$ do vetor normal ao plano $-x_1 + x_2 + x_3 = 0$.

Note-se ainda que $\dim \mathcal{C}(A) + \dim \mathcal{C}(A)^\perp = 2 + 1 = 3 = \dim \mathbb{R}^3$.

Uma relação fundamental

As relações estabelecidas no slide 115 podem ser generalizadas para o espaço das colunas de uma matriz arbitrária A. Mais precisamente, tem-se o seguinte:

- Dados $v_1, \ldots, v_n \in \mathbb{R}^m$, $A = [\, v_1 \quad \cdots \quad v_n \,]$ e $b \in \mathbb{R}^m$ tem-se,

$$b \perp \mathcal{C}(A) = \langle v_1, \ldots, v_n \rangle \;\Leftrightarrow\; A^T b = \vec{0} \;\Leftrightarrow\; b \in \mathcal{N}(A^T),$$

donde resulta a relação fundamental:

Complemento ortogonal do espaço de colunas / espaço gerado

Se $A = [\, v_1 \quad \cdots \quad v_n \,]$ então

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

O complemento ortogonal é um subespaço vetorial

- Uma vez que qualquer subespaço vetorial $V \subset \mathbb{R}^m$ admite uma base $\{v_1, \ldots, v_n\}$, podemos escrever,

$$V = \langle v_1, \ldots, v_n \rangle = \mathcal{C}(A)$$

com $A = [\, v_1 \quad \cdots \quad v_n \,]_{m \times n}$ é a matriz da base, tendo-se

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

Logo o complemento ortogonal de um subespaço V de $\mathbb{R}^m$ é também um subespaço vetorial de $\mathbb{R}^m$, uma vez que pode ser definido como o espaço nulo de uma matriz do tipo $n \times m$

- Por outro lado usando a relação $\mathrm{car}(A) = \mathrm{car}(A^T)$, obtém-se,

$$\begin{aligned} \dim V^\perp &= \dim \mathcal{C}(A)^\perp = \dim \mathcal{N}(A^T) \\ &= m - \mathrm{car}(A^T) = m - \mathrm{car}(A) = m - \dim V, \end{aligned}$$

atendendo a que m é o número de colunas de A^T.

Propriedades do complemento ortogonal

Teorema

Seja V um subespaço vetorial de $\mathbb{R}^m$. Então

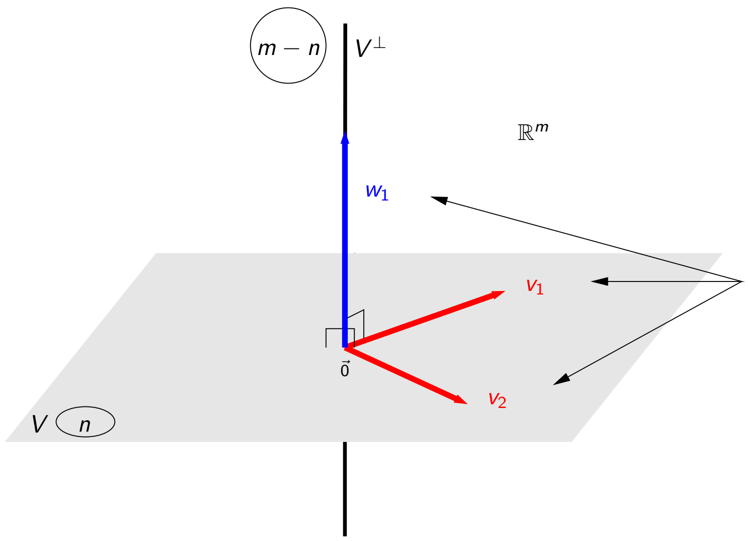
- $V \cap V^\perp = \{\vec{0}\}$
- $\dim V + \dim V^\perp = \dim \mathbb{R}^m = m$
- $(V^\perp)^\perp = V$
- Se $\{v_1, \dots, v_n\}$ é base de V e $\{w_1, \dots, w_{m-n}\}$ é base de $V^\perp$, então

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

Note-se que 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 $(V^\perp)^\perp = V$ implica que se $W = V^\perp$ então $W^\perp = V$.

Ilustração das propriedades do complemento ortogonal



Neste exemplo:

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

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

$\dim V = n = 2$ (V plano)

$\dim V^\perp = m - n = 1$ ($V^\perp$ 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

Observação: para qualquer $m \geq 2$ tem-se

- $\dim(\mathbb{R}^m)^\perp = m - \dim \mathbb{R}^m = 0$ e portanto $(\mathbb{R}^m)^\perp = \{\vec{0}\}$.
- $\dim\{\vec{0}\}^\perp = m - \dim\{\vec{0}\} = m$ e portanto $\{\vec{0}\}^\perp = \mathbb{R}^m$.

Tem-se portanto que
- subespaço maximal$^\perp$ = subespaço minimal
- subespaço minimal$^\perp$ = subespaço maximal

Caso dos subespaços vetoriais dados por equações

- Vimos no slide slide 119 que se $W = V^\perp$ então $W^\perp = V$. Logo, como $\mathcal{N}(A^T) = \mathcal{C}(A)^\perp$, conclui-se que $\mathcal{N}(A^T)^\perp = \mathcal{C}(A)$. Substituindo na relação anterior A^T por B tem-se $A = (A^T)^T = B^T$, obtendo-se $\mathcal{N}(B)^\perp = \mathcal{C}(B^T)$ (e podemos substituir, obviamente, B por A...)

  Têm-se portanto as duas fórmulas abaixo que permitem calcular o complemento ortogonal de um subespaço vetorial dado por geradores, $V = \mathcal{C}(A)$ e / ou de um subespaço vetorial definido por um sistema de equações homogéneas, $V = \mathcal{N}(A)$:

Complemento ortogonal de subespaços definidos por matrizes

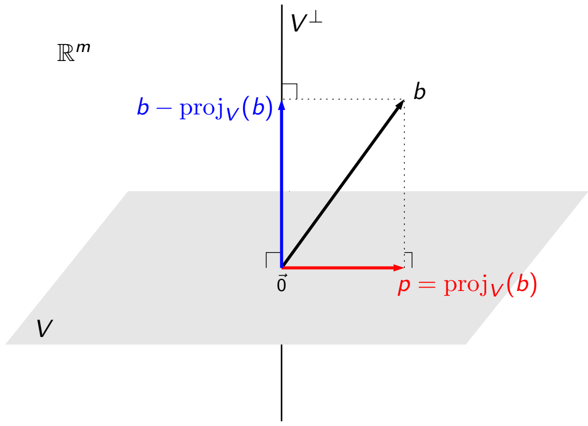
V	$V^\perp$
$\mathcal{C}(A)$	$\mathcal{N}(A^T)$
$\mathcal{N}(A)$	$\mathcal{C}(A^T)$

O complemento ortogonal do espaço das colunas [resp. espaço nulo] de uma matriz é o espaço nulo [resp. espaço das colunas] dessa matriz transposta.

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, $(b - p) \in V^\perp$.

Exercício na aula (cont.)

Tem-se:

- $p = (0,2,2) \in V = \langle v_1, v_2 \rangle = \mathcal{C}(A)$, onde $A = [\, v_1 \;\; v_2 \,]$, 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, $p = \mathrm{proj}_V(b)$.

Casos triviais: projeção ortogonal 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:

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

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

- $\mathrm{proj}_{\{\vec{0}\}}(b) = \vec{0}$ para todo o $b \in \mathbb{R}^m$.

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

E 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.

Projeção ortogonal sobre uma reta

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

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

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,

$$\begin{aligned} v^T(b-p) = 0 \quad &\Leftrightarrow \quad v^T(b - \alpha v) = 0 \Leftrightarrow v^T b - \alpha v^T v = 0 \\ &\Leftrightarrow \quad \alpha v^T v = v^T b \Leftrightarrow \alpha = \frac{v^T b}{v^T v} = \frac{v \cdot b}{v \cdot v}.([10]) \end{aligned}$$

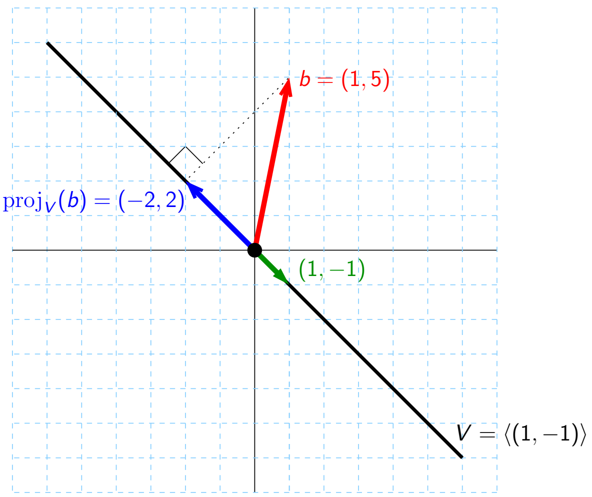
Logo $p = \alpha v = \dfrac{v^T b}{v^T v} v = \dfrac{v \cdot b}{v \cdot v} v.$ □

[10] Note-se que $v \cdot v = \|v\|^2 \neq 0$, onde $\|v\|$ é o comprimento (norma) do vetor $v \neq 0$ (como veremos mais adiante).

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,

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

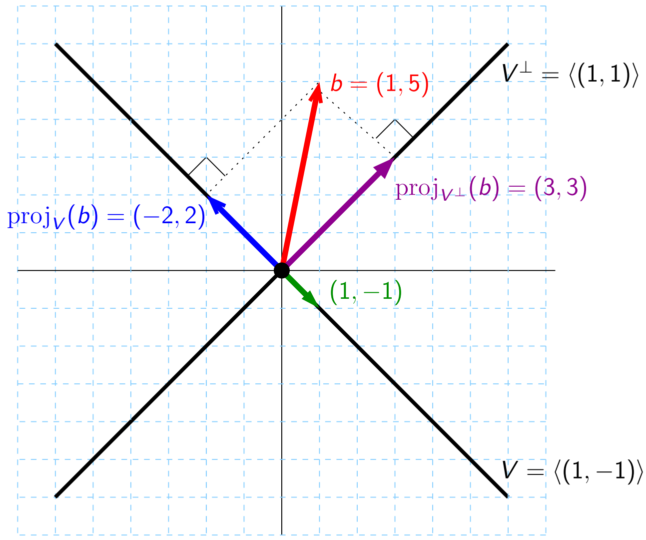
Como $b = p + (b - p)$ obtivemos a seguinte decomposição (única) de b segundo V e $V^\perp$:

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

Exemplo do slide 128 revisitado

Considerando novamente $V = \langle (1,-1) \rangle$ e $b = (1,5)$, obtém-se a decomposição descrita no slide anterior,

$$b = (1,5) = \mathrm{proj}_V(b) + \mathrm{proj}_{V^\perp}(b) = (-2,2) + (3,3).$$



TPC: verifique que $V^\perp = \langle (1,1) \rangle$.

Aplicação: projeção sobre um espaço de codimensão um

Seja V subespaço vetorial de $\mathbb{R}^m$ de codimensão um, isto é, tal que $\dim V = \dim \mathbb{R}^m - 1 = m - 1$. Nessa altura $\dim V^\perp = m - \dim V = 1$ e portanto $V^\perp$ é uma reta tendo-se $V^\perp = \langle w \rangle$ para algum vetor diretor $w \neq \vec{0}$. Pela decomposição do slide 130 e pela fórmula da projeção sobre uma reta obtém-se

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

Vejamos um exemplo. Sejam $V = \{(x_1, x_2, x_3) \colon x_1 + 2x_2 + 3x_3 = 0\}$ que define um plano de dimensão 2 e $b = (1,1,1)$. Pretende-se calcular $\mathrm{proj}_V(b)$.

Tem-se $V = \mathcal{N}(A)$ com $A = [\,1 \ \ 2 \ \ 3\,]$. Tem-se (ver o slide slide 122),

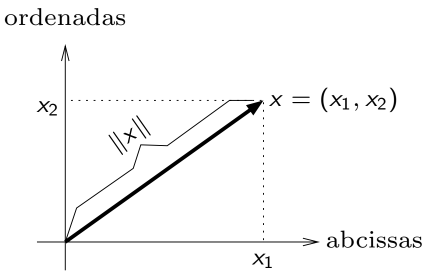
$$V^\perp = \mathcal{N}(A)^\perp = \mathcal{C}(A^T) = \mathcal{C}\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right) = \langle (1,2,3) \rangle.$$

Logo $V^\perp$ é uma reta com vetor diretor $(1,2,3)$ e tem-se

$$\mathrm{proj}_{V^\perp}(b) = \mathrm{proj}_{\langle (1,2,3) \rangle}((1,1,1)) = \frac{(1,2,3)\cdot(1,1,1)}{(1,2,3)\cdot(1,2,3)}(1,2,3) = \frac{3}{7}(1,2,3).$$

Logo $\mathrm{proj}_V(b) = b - \mathrm{proj}_{V^\perp}(b) = (1,1,1) - \frac{3}{7}(1,2,3) = \frac{1}{7}(4,1,-2)$.

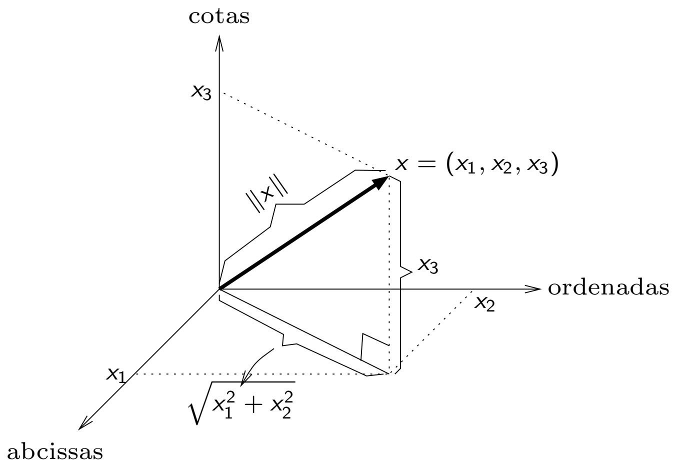
Interlúdio: norma de um vetor do plano



$\|x\|$ 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 e distância entre vetores

Em geral, se $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ define-se a sua norma de forma análoga:

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

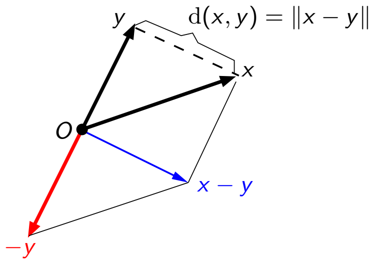
Por exemplo,

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

.

A partir da norma define-se a *distância (euclideana)* entre $x, y \in \mathbb{R}^n$, por

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



Por exemplo,

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

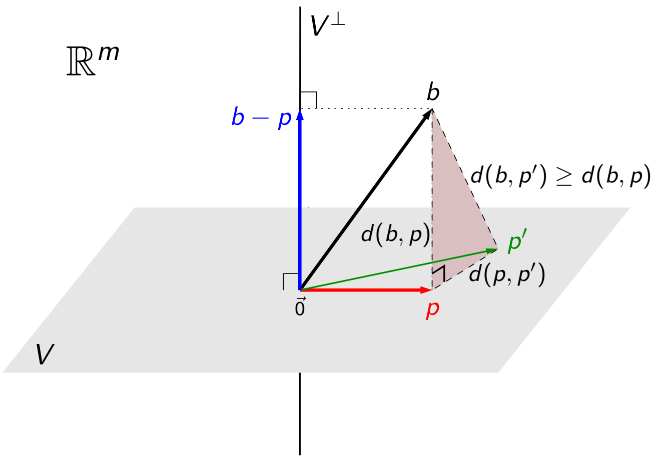
A projeção ortogonal é o vetor à menor distância...

Teorema

Sejam $b \in \mathbb{R}^m$ e V subespaço vetorial de $\mathbb{R}^m$. Então a $\mathrm{proj}_V(b)$ é o vetor de V à menor distância de b

Ideia da demonstração: pelo teorema de Pitágoras com $p = \mathrm{proj}_V(b)$, tem-se

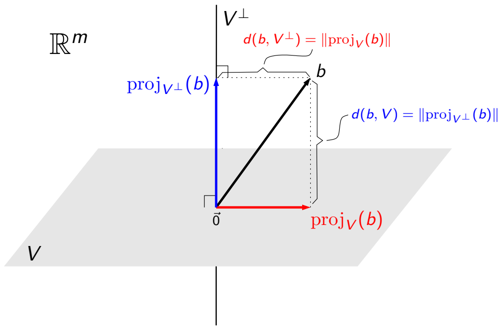
$$d(b,p')^2 = d(b,p)^2 + \underbrace{d(p,p')^2}_{\geq 0} \geq d(b,p)^2 \Rightarrow d(b,p') \geq d(b,p), \quad \forall p' \in V.$$



Distância de um vetor a um subespaço vetorial

Denotando por $d(b, V)$ a distância do vetor b ao subespaço V, que se define como a distância de b ao vetor de V mais próximo de b, tem-se:

- $d(b, V) = d(b, \text{proj}_V(b)) = \|b - \text{proj}_V(b)\| = \|\text{proj}_{V^\perp}(b)\|$.
- $d(b, V^\perp) = d(b, \text{proj}_{V^\perp}(b)) = \|b - \text{proj}_{V^\perp}(b)\| = \|\text{proj}_V(b)\|$.
- $d(b, V)^2 + d(v, V^\perp)^2 = \|b\|^2$ (pelo teorema de Pitágoras).



Observação

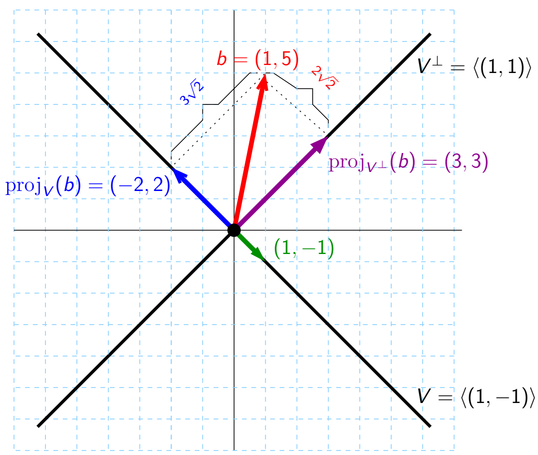
Note-se que a distância de b ao subespaço V ou $V^\perp$ é nula se e só se o vetor pertencer a esse subespaço. De facto,

- $d(b, V) = 0 \Leftrightarrow \text{proj}_{V^\perp}(b) = \vec{0} \Leftrightarrow b = \text{proj}_V(b) \Leftrightarrow b \in V$.
- $d(b, V^\perp) = 0 \Leftrightarrow \text{proj}_V(b) = \vec{0} \Leftrightarrow b = \text{proj}_{V^\perp}(b) \Leftrightarrow b \in V^\perp$.

Distância de um vetor a um subespaço vetorial - exemplo do slide 128

Considerando novamente $V = \langle(1, -1)\rangle$ e $b = (1, 5)$ do slide 128 tem-se:

- $d(b, V) = \|b - \text{proj}_V(b)\| = \|\text{proj}_{V^\perp}(b)\| = \|(3, 3)\| = 3\sqrt{2}$.
- $d(b, V^\perp) = \|b - \text{proj}_{V^\perp}(b)\| = \|\text{proj}_V(b)\| = \|(-2, 2)\| = 2\sqrt{2}$.
- $d(b, V)^2 + d(v, V^\perp)^2 = (3\sqrt{2})^2 + (2\sqrt{2})^2 = 26 = \|(1, 5)\|^2 = \|b\|^2$.



Equações normais

Sejam $V \subset \mathbb{R}^m$ subespaço vetorial de $\dim n$, $\{v_1, \ldots, v_n\}$ base de V e $A = [\, v_1 \; v_2 \; \cdots \; v_n \,]_{m\times n}$ a respetiva matriz da base. Em particular, $V = \mathcal{C}(A)$ e $\text{car}(A) = \dim \mathcal{C}(A) = \dim V = n$. Seja $b \in \mathbb{R}^m$. Por definição de projeção ortogonal, $p = \text{proj}_V(b)$ verifica:

(i) $p \in V = \langle v_1, \ldots, v_n\rangle = \mathcal{C}(A)$, isto é, $Ax = p$ é possível e portanto existe uma solução $\bar{x} = (\alpha_1, \ldots, \alpha_n)$ tal que $p = A\,\bar{x}$.([11])

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

Tem-se então,

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

O vetor $\bar{x}$ é portanto a solução do sistema de equações lineares,

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

que se designa por sistema das equações normais. Uma vez que A é a matriz de uma base de V, o sistema é PD como veremos a seguir. Tem-se então $\text{proj}_V(b) = p = A\,\bar{x}$ com $\bar{x}$ solução do sistema de equações normais.

[11] Note-se que $p = \alpha_1 v_1 + \cdots + \alpha_n v_n$.

Método das equações normais

Algoritmo

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

- Determinar uma base para V, $\{v_1, \ldots, v_n\}$.
  Seja $A = [\, v_1 \; \cdots v_n \,]$ a matriz da base.
- Determinar a solução única $\bar{x}$ das equações normais, $A^T A\,x = A^T b$.
- Tem-se então $\text{proj}_V(b) = A\,\bar{x}$.

Exercício na aula

- Determinar a projeção ortogonal de $b = (1,0,4)$ sobre o subespaço vetorial $V = \big\langle (1,0,1), (1,1,2) \big\rangle$ utilizando o método das equações normais e indicar as distâncias de b a V e a $V^{\perp}$.
- **TPC**: recalcular a projeção de $b = (1,0,4)$ sobre V começando por calcular a projeção sobre $V^{\perp}$ (reta).

Exercício na aula (resolução)

- $\{v_1, v_2\} = \{(1,0,1),(1,1,2)\}$ é base de $V = \langle (1,0,1),(1,1,2)\rangle$ (justifique!). Seja $A = [\, v_1 \;\; v_2 \,]$ a matriz da base.
- Tem-se:

$$A^T A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix}.$$

$$A^T b = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \\ 9 \end{bmatrix}.$$

A solução (única) do sistema $A^T A\, x = A^T b$ é $\bar{x} = (1,1)$. De facto,

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

- Logo $\operatorname{proj}_V(b) = A\,\bar{x} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$.

$d(b, V) = \|\operatorname{proj}_{V^\perp}(b)\| = \|b - \operatorname{proj}_V(b)\| = \|(1,0,4)-(2,1,3)\| = \|(-1,-1,1)\| = \sqrt{3}$.
$d(b, V^\perp) = \|\operatorname{proj}_V(b)\| = \|(2,1,3)\| = \sqrt{14}$.

Matriz de projeção

Recordemos que no método das equações normais $\operatorname{proj}_V(b) = A\,\bar{x}$ onde A é a matriz de uma base $\{v_1, \ldots, v_n\}$ de V e $\bar{x}$ a solução do sistema das equações normais, $A^T A\, x = A^T b$, isto é, verifica a relação

$$A^T A\,\bar{x} = A^T b.$$

Uma vez que A é a matriz de uma base pode-se mostrar que $A^T A$ é invertível. Multiplicando à esquerda por $(A^T A)^{-1}$ ambos os membros da igualdade anterior conclui-se que $\bar{x} = (A^T A)^{-1} A^T b$ e portanto que

$$\boxed{\operatorname{proj}_V(b) = A\,\bar{x} = A\,(A^T A)^{-1} A^T b = Pb,}$$

com,

$$P = A(A^T A)^{-1} A^T.$$

A matriz P não depende da escolha da base de V e designa-se por matriz de projeção sobre V, uma vez que $\operatorname{proj}_V(b) = P\,b$ para todo o $b \in \mathbb{R}^m$.

Propriedades da matriz de projeção

- $P^T = P$ (simétrica).
- $P^2 = P$ (idempotente).

(Ver o exercício 25.10 da sebenta).

Matriz de projeção (cont.)

Observações

- Se P é a matriz de projeção sobre $V \subset \mathbb{R}^m$ então $I_m - P$ é a matriz de projeção sobre $V^\perp$, onde I_m denota a matriz identidade de ordem m, tendo-se para todo o $b \in \mathbb{R}^m$,

$$\text{proj}_{V^\perp}(b) = (I-P)b = b - Pb = b - \text{proj}_V(b).$$

- Se $V = \langle v \rangle = \mathcal{C}(v)$, é uma reta a fórmula da matriz de projeção simplifica-se obtendo-se uma expressão muito elegante:

$$P = v(v^T v)^{-1} v^T = (v^T v)^{-1} v v^T = \frac{v v^T}{v^T v},$$

ou seja, a matriz de projeção sobre a reta $\langle v \rangle$ é, $P = \dfrac{v\,v^T}{v^T v}$.

Exercício na aula

Determinar as matrizes de projeção sobre V e $V^\perp$ onde $V = \langle (1,0,1),(1,1,2) \rangle$ é o subespaço do slide 140 e calcular $\text{proj}_V(1,0,4)$ e $\text{proj}_{V^\perp}(1,0,4)$

Exercício na aula (resolução)

Mnemónica para calcular a inversa de uma matriz 2×2

"Switch diagonally, negate the wings and divide by a cross" (*):

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

Note-se que a matriz anterior é invertível se e só se $ad - bc \neq 0$, valor que se designa por determinante da matriz e será estudado a seguir.
(*)https://www.dam.brown.edu/people/mchb/la/matrix_algebra.pdf

- Pelos resultados do slide 141, $A^T A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix}$,

com A matriz da base de V. Usando a fórmula acima para obter $(A^T A)^{-1}$ vem,

$$\begin{aligned} P &= A(A^T A)^{-1}A^T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix}^{-1}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix}\frac{1}{2\times 6 - 3\times 3}\begin{bmatrix} 6 & -3 \\ -3 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 3 & -3 & 0 \\ -1 & 2 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}. \end{aligned}$$

Exercício na aula (resolução - cont.)

- Constata-se imediatamente que a matriz de projeção P sobre V é simétrica, isto é, $P^T = P$. Verifique que P é também idempotente, isto é, que verifica $P^2 = P$.
- A matriz de projeção sobre $V^\perp$ é $I - P$ onde I é a matriz identidade, isto é,

$$I - P = \frac{1}{3}\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}.$$

- Observemos que calculando $V^\perp = \mathcal{C}(A)^\perp = \mathcal{N}(A^T)$, se tem $V^\perp = \langle w \rangle$ com $w = (-1, -1, 1)$ (verifique) e portanto a matriz de projeção sobre $V^\perp$ pode ser alternativamente reobtida, pela fórmula do slide 143, como $\dfrac{w\,w^T}{w^T w}$. De facto,

$$\frac{w\,w^T}{w^T w} = \frac{\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}\begin{bmatrix} -1 & -1 & 1 \end{bmatrix}}{\begin{bmatrix} -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}} = \frac{\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}}{3}.$$

- Finalmente, designando $b = (1, 0, 4)$ tem-se,

$$\begin{aligned} \mathrm{proj}_V(b) &= Pb = \frac{1}{3}\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 6 \\ 3 \\ 9 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \in V, \\ \mathrm{proj}_{V^\perp}(b) &= (I - P)b = b - Pb = (-1, -1, 1) \in V^\perp. \end{aligned}$$

Conjunto ortogonal e ortonormado de vetores

Definição de conjunto ortogonal e ortonormado de vetores

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

- $\{v_1, \ldots, v_k\}$ diz-se um conjunto *ortogonal* se os vetores forem 2 a 2 perpendiculares entre si, isto é, se $v_i \cdot v_j = 0$, $\forall i, j = 1, \ldots, k$, $i \neq j$.
- Se além disso, todos os vetores forem unitários, isto é, $\|v_i\| = 1$, $\forall i$, então $\{v_1, \ldots, v_k\}$ diz-se um conjunto *ortonormado*.

Exemplos

- $\{v_1, v_2, v_3\} = \{(0, 1, 1), (1, 2, -2), (4, -1, 1)\}$ é um conjunto ortogonal de vetores de $\mathbb{R}^3$.De facto,

$$\begin{aligned} v_1 \cdot v_2 &= (0, 1, 1) \cdot (1, 2, -2) = 0, \quad \text{isto é,} \quad v_1 \perp v_2, \\ v_1 \cdot v_3 &= (0, 1, 1) \cdot (4, -1, 1) = 0, \quad \text{isto é,} \quad v_1 \perp v_3, \\ v_2 \cdot v_3 &= (1, 2, -2) \cdot (4, -1, 1) = 0, \quad \text{isto é,} \quad v_2 \perp v_3. \end{aligned}$$

Portanto os vetores são ortogonais 2 a 2 entre si.

- A base canónica de $\mathbb{R}^m$ é um conjunto ortonormado (verifique!).

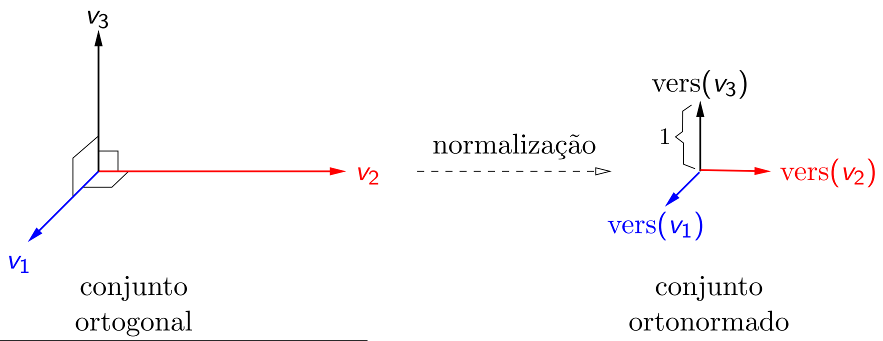
Normalização de vetores. . .

Observação

Se $\{v_1, \ldots, v_k\}$ for um conjunto *ortogonal de vetores não nulos* de $\mathbb{R}^m$, então

$$\{\text{vers}(v_1), \ldots, \text{vers}(v_k)\} = \left\{ \frac{v_1}{\|v_1\|}, \ldots, \frac{v_k}{\|v_k\|} \right\}$$

é um conjunto ortonormado de vetores de $\mathbb{R}^m$. ([12])



[12]O versor de um vetor $x \neq \vec{0}$, $\text{vers}(x) = \dfrac{x}{\|x\|}$, é o único vetor unitário que tem a mesma direção e sentido que x.

Exemplo de normalização. . .

Exemplo

Voltando ao conjunto ortogonal $\{(0, 1, 1), (1, 2, -2), (4, -1, 1)\}$ do exemplo do slide 144, concluímos que o conjunto de vetores

$$\{\text{vers}(0, 1, 1), \text{vers}(1, 2, -2), \text{vers}(4, -1, 1)\} =$$

$$\left\{ \frac{(0, 1, 1)}{\sqrt{2}}, \frac{(1, 2, -2)}{3}, \frac{(4, -1, 1)}{3\sqrt{2}} \right\}$$

é um conjunto ortonormado de vetores de $\mathbb{R}^3$.

Base ortogonal/ortonormada de um subespaço vetorial

Definição

Uma base ortogonal/ortonormada de um subespaço vetorial V é uma base $\{v_1, \ldots, v_n\}$ de V que é simultaneamente um conjunto ortogonal/ortonormado.

Observações

- Pode-se mostrar que um conjunto ortogonal de vetores não nulos $\{v_1, \ldots, v_n\}$ é sempre linearmente independente.
- Do ponto anterior resulta imediatamente que um conjunto ortogonal de geradores não nulos de um subespaco vetorial V, isto é, $V = \langle v_1, \ldots, v_n \rangle$, com $v_i \neq \vec{0}$ e $v_i \perp v_j$ se $i \neq j$, define uma base ortogonal de V.
- Em particular, qualquer conjunto ortogonal de m vetores não nulos de $\mathbb{R}^m$, $\{v_1, \ldots, v_m\}$, define uma base ortogonal de $\mathbb{R}^m$.
- A base canónica $\{e_1, \ldots, e_m\}$ de $\mathbb{R}^m$ é o exemplo mais importante de base ortonormada (b.o.n) de $\mathbb{R}^m$.
- As bases ortogonais são importantes pois vão-nos permitir calcular a projeção ortogonal de forma imediata, como veremos no próximo slide.

Projeção sobre um subespaço vetorial munido de uma base ortogonal

Teorema

Seja $\{v_1, \ldots, v_k\}$ uma base ortogonal de V. Para todo o $b \in \mathbb{R}^m$ tem-se

$$\text{proj}_V(b) = \text{proj}_{v_1}(b) + \cdots + \text{proj}_{v_k}(b) = \frac{b \cdot v_1}{v_1 \cdot v_1} v_1 + \cdots + \frac{b \cdot v_k}{v_k \cdot v_k} v_k$$

Muito importante: o resultado é falso se a base não for ortogonal (!!!)

Exemplo na aula

Calcular $\text{proj}_V(b)$ onde $V = \langle v_1, v_2 \rangle = \langle (0,1,1), (1,2,-2) \rangle$ e $b = (-1,0,4)$.

Como $v_1 \cdot v_2 = 0$ (verifique) $\{v_1, v_2\}$ é um conjunto ortogonal de geradores não nulos de V e portanto define uma base ortogonal de $V = \langle v_1, v_2 \rangle$, tendo-se

$$\begin{aligned}
\text{proj}_V(b) &= \text{proj}_{\langle (0,1,1),(1,2,-2) \rangle}(b) = \text{proj}_{(0,1,1)}(b) + \text{proj}_{(1,2,-2)}(b) \\
&= \tfrac{(-1,0,4)\cdot(0,1,1)}{(0,1,1)\cdot(0,1,1)}(0,1,1) + \tfrac{(-1,0,4)\cdot(1,2,-2)}{(1,2,-2)\cdot(1,2,-2)}(1,2,-2) \\
&= \tfrac{4}{2}(0,1,1) + \tfrac{-9}{9}(1,2,-2) = (-1,0,4).
\end{aligned}$$

Note-se que neste exercício $\text{proj}_V(b) = b$. Qual o significado dessa relação?

Vamos ver a seguir como obter bases ortogonais para subespaços vetoriais. . .

Construção de uma base ortogonal de um subespaço - ideia

Seja $\{u_1, \ldots, u_n\}$ base de V. Em particular, $V = \langle u_1, \ldots, u_n \rangle$ e dim $V = n$. Define-se:

$$v_1 = u_1 \in V$$

$$v_2 = \text{proj}_{\langle v_1 \rangle^\perp}(u_2) = u_2 - \text{proj}_{v_1}(u_2) = u_2 - \frac{v_1 \cdot u_2}{v_1 \cdot v_1} v_1 \Rightarrow \begin{cases} v_2 \perp v_1 & \text{(por construção)} \\ v_2 \in V & (v_2 \text{ é CL de } u_1, u_2) \\ v_2 \neq \vec{0} & \text{(senão } u_2 \text{ múltiplo de } u_1) \end{cases}$$

$$\begin{aligned} v_3 &= \text{proj}_{\langle v_1, v_2 \rangle^\perp}(u_3) = v_3 - \text{proj}_{\langle v_1, v_2 \rangle}(u_3) \\ &= u_3 - \text{proj}_{v_1}(u_3) - \text{proj}_{v_2}(u_3) = u_3 - \frac{v_1 \cdot u_3}{v_1 \cdot v_1} v_1 - \frac{v_2 \cdot u_3}{v_2 \cdot v_2} v_2 \Rightarrow \begin{cases} v_3 \perp v_1, v_2 \\ v_3 \in V \\ v_3 \neq \vec{0} \end{cases}. \end{aligned}$$

$$\vdots$$

$$\begin{aligned} v_n &= u_n - \text{proj}_{\langle v_1, v_2, \ldots, v_{n-1} \rangle}(u_n) = u_n - \text{proj}_{v_1}(u_n) - \text{proj}_{v_2}(u_n) - \cdots - \text{proj}_{v_{n-1}}(u_n) \\ &= u_n - \frac{v_1 \cdot u_n}{v_1 \cdot v_1} v_1 - \frac{v_2 \cdot u_n}{v_2 \cdot v_2} v_2 - \cdots - \frac{v_{n-1} \cdot u_n}{v_{n-1} \cdot v_{n-1}} v_{n-1} \Rightarrow \begin{cases} v_n \perp v_1, v_2, \ldots, v_{n-1} \\ v_n \in V \\ v_n \neq \vec{0} \end{cases} \end{aligned}$$

Pode-se mostrar analogamente que $\{v_1, \ldots, v_n\}$ verifica

- $\{v_1, \ldots, v_n\}$ é um conjunto ortogonal de vetores não nulos logo lin. indep.,
- $v_1, \ldots, v_n \in V$ e
- dim $V = n$.

e portanto define uma base de V (ver o slide 108) que é ortogonal.

Método de ortogonalização de Gram-Schmidt

A partir da ideia descrita no slide anterior obtém-se o seguinte método que permite ortogonalizar qualquer base de um subespaço vetorial.

Algoritmo

`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 ortogonalização de Gram-Schmidt

Observação

- 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 ortonormada de V, $\left\{ \frac{v_1}{\|v_1\|}, \ldots, \frac{v_n}{\|v_n\|} \right\}$.

Exercício na aula

- Justifique que $\{u_1, u_2, u_3\} = \{(1, -1, 1), (1, 0, 1), (1, 1, 2)\}$ é base de $\mathbb{R}^3$.
- A partir da base anterior obtenha uma base ortogonal $\{v_1, v_2, v_3\}$. de $\mathbb{R}^3$ usando o método de Gram-Schmidt.
- Transforme a base ortogonal anterior numa base ortonormada de $\mathbb{R}^3$.

Exercício na aula (resolução)

$\{u_1, u_2, u_3\}$ define uma base de $\mathbb{R}^3$ pois é um conjunto linearmente independente formado por 3 vetores ($\dim \mathbb{R}^3 = 3$). De facto, aplicando o método de Gauss,

$$[\, u_1 \quad u_2 \quad u_3 \,] = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{(escada)}$$

verifica-se que todas as colunas da matriz em escada obtida a partir de $[\, u_1 \quad u_2 \quad u_3 \,]$ têm pivot. No entanto esta base não é ortogonal pois $u_1 \cdot u_2 = 2 \neq 0$ (por exemplo). Aplicando o método de Gram-Schmidt à base anterior 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 a base ortogonal de $\mathbb{R}^3$ anterior, 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

- Determinar a projeção ortogonal de $b = (1, 1, 3)$ sobre $V = \big\langle (1, -1, 1), (1, 0, 1) \big\rangle$ ortogonalizando uma base de V.
- Estender a base ortogonal de V da alínea anterior a 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:

$$\begin{aligned} v_1 &= u_1 = (1, -1, 1) \\ v_2 &= u_2 - \text{proj}_{v_1}(u_2) = u_2 - \frac{u_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \frac{1}{3}(1, 2, 1) \rightsquigarrow (1, 2, 1) \end{aligned}$$

  Uma base ortogonal de V é portanto $\{v_1, v_2\} = \{(1, -1, 1), (1, 2, 1)\}$, tendo-se (ver o teorema do slide 150),

$$\begin{aligned} \text{proj}_V(b) &= \text{proj}_{v_1}(b) + \text{proj}_{v_2}(b) = \frac{v_1 \cdot b}{v_1 \cdot v_1} v_1 + \frac{v_2 \cdot b}{v_2 \cdot v_2} v_2 \\ &= \frac{3}{3}(1, -1, 1) + \frac{6}{6}(1, 2, 1) = (2, 1, 2). \end{aligned}$$

- 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.

Conceito de determinante

- Seja $A = [\, v_1 \;\; v_2 \;\; \cdots v_n \,]$ matriz quadrada de ordem n. Vamos associar a A um valor real, designado por *determinante de A* e denotado por $\det A$ ou por $|A|$, verificando certas propriedades, entre as quais a propriedade

$$\boxed{A \text{ é invertível} \quad \Leftrightarrow \quad \det A \neq 0}$$

Recorde que $A_{n\times n} = [\, v_1 \;\; \cdots \;\; v_n \,]$ é invertível se e só se $\mathrm{car}(A) = n$ se e só se $\{v_1, \ldots, v_n\}$ é linearmente independente se e só se $\{v_1, \ldots, v_n\}$ é base de $\mathbb{R}^n$.

  - Se $n = 1$, define-se,
$$\det [a] = a$$
  - Se $n = 2$, define-se
$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

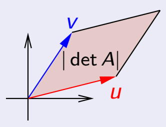
Determinante de matrizes 2×2

- Por exemplo, se $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

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

- Note-se que A é invertível uma vez que as colunas de A são vetores linearmente independentes e portanto $\mathrm{car}(A) = 2$.

Interpretação geométrica do determinante de matrizes 2×2



O valor absoluto do determinante de uma matriz

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = [\, u \;\; v \,], \text{ com } u = (a, c) \text{ e } v = (b, d),$$

corresponde à área do paralelogramo definido por u e v.
Note que a área do paralelogramo é não nula apenas quando os vetores u e v são não colineares, isto é, A é invertível!

Determinante de matrizes 3×3: regra de Sarrus

$$\det(A) = \det \left[\begin{array}{ccc} a & b & c \\ d & e & f \\ g & h & i \end{array}\right] \begin{array}{cc} a & b \\ d & e \\ g & h \end{array}$$

$$= {\color{red} aei + bfg + cdh} - ({\color{blue} ceg + afh + bdi})$$

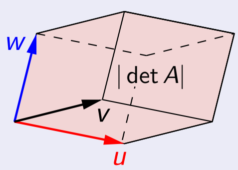
Por exemplo,

$$\det A = \det \begin{bmatrix} 1 & 0 & -2 \\ 2 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} \begin{array}{cc} 1 & 0 \\ 2 & 1 \\ -1 & 3 \end{array} = {\color{red} 1 \times 1 \times 1 + 0 \times 1 \times (-1) + (-2) \times 2 \times 3}$$
$$- ({\color{blue} (-2) \times 1 \times (-1) + 1 \times 1 \times 3 + 0 \times 2 \times 1}) = {\color{red} 1 + 0 - 12} - ({\color{blue} 2 + 3 + 0}) = -16 \neq 0$$

Logo, A é invertível, isto é, as colunas de A são linearmente independentes e portanto definem uma base de $\mathbb{R}^3$.

Determinantes 3x3 (cont.)

Interpretação geométrica do determinante de matrizes 3×3



O valor absoluto do determinante de uma 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.

Note-se que o volume do paralelipípedo é não nulo se e só se o paralelipípedo é não degenerado se e só se u, v, w são não complanares, ou seja, se e só se $\{u, v, w\}$ é linearmente independente, isto é, A é invertível.

- Por exemplo, o paralelipípedo definido pelas 3 colunas da matriz A do slide anterior, $u = (1, 2, -1)$, $v = (0, 1, 3)$ e $w = (-2, 1, 1)$, tem volume 16.
- A regra de Sarrus só se aplica às matrizes 3x3 (!)
- *E o caso das matrizes $n \times n$ com $n \geq 4$?*

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 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 det. 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 det. 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

Consideremos a matriz $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 2 & 3 \\ -1 & 2 & 1 \end{bmatrix}$.

- Ao longo da 2ª linha tem-se

$$\begin{aligned} \det A &= a_{21}\Delta_{21} + a_{22}\Delta_{22} + a_{23}\Delta_{23} = 0(-1)^{2+1}\det\begin{bmatrix} -2 & 3 \\ 2 & 1 \end{bmatrix} + \\ & 2(-1)^{2+2}\det\begin{bmatrix} 1 & 3 \\ -1 & 1 \end{bmatrix} + 3(-1)^{2+3}\det\begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} \\ &= 0 + 2\times(1+3) - 3\times 0 = 8. \end{aligned}$$

- Ao longo da 1ª coluna tem-se

$$\begin{aligned} \det A &= a_{11}\Delta_{11} + a_{21}\Delta_{21} + a_{31}\Delta_{31} = 1(-1)^{1+1}\det\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} + \\ & 0(-1)^{2+1}\det\begin{bmatrix} -2 & 3 \\ 2 & 1 \end{bmatrix} + (-1)(-1)^{3+1}\det\begin{bmatrix} -2 & 3 \\ 2 & 3 \end{bmatrix} \\ &= 1\times(2-6) + 0 + (-1)\times(-6-6) = 8. \end{aligned}$$

Regra de Laplace: exemplo 4×4

Exercício na aula

Calcular o determinante de

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

(sol. $\det A = -6$)

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 anteriores, usando a regra de Sarrus no primeiro e a regra de Laplace no segundo.

Consequências da regra de Laplace

Tem-se (ver o exercício 27 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,

- $\det\big(\operatorname{diag}(a_1, \dots, a_n)\big) = \det \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix} = a_1 a_2 \cdots a_n.$
- $\det(\alpha I_n) = \det\big(\operatorname{diag}(\alpha, \dots, \alpha)\big) = \alpha^n.$
- $\det I_n = 1.$

Propriedades do determinante

Proposição

Sejam A e B matrizes quadradas de ordem n e $\alpha \in \mathbb{R}$. Tem-se:

- $\det(AB) = \det A \det B$.
- $\det(A^T) = \det A$.
- $\det(\alpha A) = \alpha^n \det A$!!!
- A é invertível se e só se $\det A \neq 0$ e nessa altura

$$\boxed{\det(A^{-1}) = \frac{1}{\det A}}$$

- Para ver o terceiro ponto basta notar que pelo primeiro ponto se tem,

$$\det(\alpha A) = \det(\alpha(I_n A)) = \det((\alpha I_n) A) = \det(\alpha I_n) \det A = \alpha^n \det A.$$

- Para obter $\det A^{-1}$ basta notar que pelo primeiro ponto se tem,

$$1 = \det I_n = \det(A A^{-1}) = \det A \det(A^{-1}).$$

- Atenção: em geral, $\det(A + B) \neq \det A + \det B$!!!

Efeito das operações elementares do método de Gauss sobre o determinante

- O "apagador", isto é, adicionar um múltiplo de uma linha a outra linha ($L_j \to L_j + \alpha L_i$) não afeta o determinante.
- Trocar duas linhas entre si ($L_i \leftrightarrow L_j$) troca o sinal do determinante.
- Multiplicar uma linha por um escalar $\alpha \neq 0$ ($L_i \to \alpha L_i$) multiplica o determinante por α.

Esquematicamente,

$$\det \begin{bmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_j + \alpha L_i \\ \vdots \\ L_n \end{bmatrix} = \det \begin{bmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_j \\ \vdots \\ L_n \end{bmatrix}, \quad \det \begin{bmatrix} L_1 \\ \vdots \\ L_j \\ \vdots \\ L_i \\ \vdots \\ L_n \end{bmatrix} = -\det \begin{bmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_j \\ \vdots \\ L_n \end{bmatrix}, \quad \det \begin{bmatrix} L_1 \\ \vdots \\ \alpha L_i \\ \vdots \\ L_n \end{bmatrix} = \alpha \det \begin{bmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_n \end{bmatrix}.$$

Cálculo do determinante pelo método de Gauss: exemplos

- Calcular o seguinte determinante usando o método de Gauss

$$\begin{aligned}
\det \begin{bmatrix} 0 & 1 & 3 \\ 1 & 1 & 2 \\ 2 & 1 & -1 \end{bmatrix} &= -\det \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 2 & 1 & -1 \end{bmatrix} \quad (L_1 \leftrightarrow L_2) \\
&= -\det \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & -1 & -5 \end{bmatrix} \quad (L_3 - 2L_1 \to L_3) \\
&= -\det \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{bmatrix} \quad (L_3 + L_2 \to L_3) \\
&= -(1 \times 1 \times (-2)) = 2 \quad (\text{determ. de matriz triangular})
\end{aligned}$$

- Note que se no método de Gauss multiplicarmos uma linha da matriz por um escalar $\alpha \neq 0$ temos que multiplicar o determinante da matriz resultante por $1/\alpha$ para não alterar o valor do determinante!
- Por exemplo, na matriz abaixo multiplicou-se a primeira linha por $\frac{1}{2}$ pelo que teve que se multiplicar o determinante da matriz resultante por 2:

$$\det \begin{bmatrix} 2 & 2 & 4 \\ 3 & 5 & 0 \\ -2 & 3 & -4 \end{bmatrix} = 2 \det \begin{bmatrix} 1 & 1 & 2 \\ 3 & 5 & 0 \\ -2 & 3 & -4 \end{bmatrix}.$$

Método "híbrido"

- Podemos usar o método de Gauss para "limpar" uma coluna (com excepção do respectivo pivot), para depois se aplicar a regra de Laplace ao longo dessa coluna...

No seguinte exemplo foram envolvidos os 3 métodos dados anteriormente para obter o valor do determinante:

$$\det\begin{bmatrix} 1 & 1 & 2 & 0 \\ -1 & 1 & 1 & 2 \\ 2 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \det\begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 2 & 3 & 2 \\ 0 & -3 & -3 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \text{ (método de Gauss)}$$

$$= 1 \times (-1)^{1+1} \det\begin{bmatrix} 2 & 3 & 2 \\ -3 & -3 & 1 \\ 0 & -1 & 1 \end{bmatrix} \text{ (Laplace na 1}^{\underline{a}}\text{ col.)}$$

$$= \det\begin{bmatrix} 2 & 3 & 2 \\ -3 & -3 & 1 \\ 0 & -1 & 1 \end{bmatrix}\begin{matrix} 2 & 3 \\ -3 & -3 \\ 0 & -1 \end{matrix} = 11 \text{ (regra de Sarrus)}$$

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,

$$Av = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 2\, v$$

Valores e vetores próprios de uma matriz

Em geral, para determinar os vetores próprios de uma matriz é necessário começar por determinar os seus valores próprios !

Para isso vamos recordar algumas relações que nos vão ser úteis.

Observação

Seja A matriz quadrada de ordem n. As seguintes afirmações são equivalentes:

- O sistema homogéneo $Ax = \vec{0}$ é indeterminado.
- $\dim \mathcal{N}(A) > 0$.
- $\text{car}(A) < n$.
- A não é invertível.
- $\det(A) = 0$.

Como determinar os valores próprios de uma matriz ?

Tem-se:

- $\lambda \in \mathbb{R}$ é valor próprio de uma matriz A se e só se existe um vetor próprio $v \neq \vec{0}$ tal que $Av = \lambda v$, isto é, $Av - \lambda v = \vec{0}$, ou seja, $(A - \lambda I)v = \vec{0}$.
- A condição anterior significa que o sistema homogéneo $(A - \lambda I)x = \vec{0}$ admite uma solução $v \neq \vec{0}$ e portanto que é indeterminado.
- Pela observação do slide anterior aplicada à matriz $(A - \lambda I)$ conclui-se que:

$$\boxed{\lambda \in \mathbb{R} \text{ é valor próprio de } A \quad \Leftrightarrow \quad \det(A - \lambda I) = 0}$$

Exemplo

Consideremos novamente a matriz $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ do exemplo do slide 170.

Tem-se que $\lambda = 2$ é valor próprio de A, como visto no slide 170, uma vez que

$$\det(A - 2I) = \begin{vmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{vmatrix} = 0 \quad \text{(possui uma linha de zeros).}$$

Polinómio característico de uma matriz

Polinómio característico e multiplicidade algébrica

- A expressão $\det(A - \lambda I)$ com $A_{n\times n}$ define um polinómio de grau n na variável λ, que se designa por polinómio característico de A e se denota por $p_A(\lambda)$.
- Pelas conclusões do slide anterior os valores próprios de A são as raízes reais e complexas do polinómio característico $p_A(\lambda)$.
- A *multiplicidade algébrica* de um valor próprio λ, que se denota por m.a.(λ), é o número de vezes que λ aparece repetido como raíz na factorização de $p_A(\lambda)$.

Exemplo do slide 170 revisitado

Consideremos novamente a matriz do slide 170, $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

- Tem-se,

$$
\begin{aligned}
p_A(\lambda) &= \det(A - \lambda I) = \det\left(\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & & \\ & \lambda & \\ & & \lambda \end{bmatrix}\right) \\
&= \underbrace{\det\begin{bmatrix} 1-\lambda & 1 & 0 \\ 0 & 2-\lambda & 0 \\ 0 & 1 & 1-\lambda \end{bmatrix}}_{\text{Laplace na } 1^{\underline{a}} \text{ coluna}} \\
&= (-1)^{1+1}(1-\lambda)\det\begin{bmatrix} 2-\lambda & 0 \\ 1 & 1-\lambda \end{bmatrix} + 0 + 0 \\
&= (1-\lambda)(2-\lambda)(1-\lambda) = (1-\lambda)^2(2-\lambda).
\end{aligned}
$$

- $p_A(\lambda)$ admite portanto a raíz dupla $\lambda = 1$ uma vez que aparece repetida 2 vezes na factorização do polinómio e a raíz simples $\lambda = 2$.
- Logo A admite os valores próprios distintos, $\lambda = 1$ com multiplicidade algébrica 2 ($m.a.(1) = 2$) e $\lambda = 2$ com multiplicidade algébrica 1 ($m.a.(2) = 1$).

Subespaço próprio de uma matriz

- Seja λ é um valor próprio de A de ordem n e $v \in \mathbb{R}^n$, $v \neq \vec{0}$([14]). Tem-se:

$$\begin{aligned} v \text{ é vetor próprio de } A \text{ associado a } \lambda \quad &\Leftrightarrow \quad Av = \lambda v \\ &\Leftrightarrow \quad (A - \lambda I)v = \vec{0} \text{ (ver o slide 172)} \\ &\Leftrightarrow \quad v \in \mathcal{N}(A - \lambda I). \end{aligned}$$

Tem portanto sentido a seguinte definição.

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 de $\mathbb{R}^n$,

$$E(\lambda) = \mathcal{N}(A - \lambda I)$$

A dimensão de $E(\lambda)$ designa-se por multiplicidade geométrica de λ e denota-se por $m.g.(\lambda)$, isto é, $m.g.(\lambda) = \dim E(\lambda)$.

Os vetores próprios de A associados ao valor próprio λ são os vetores não nulos do subespaço próprio $E(\lambda)$.

[14] Note-se que $\vec{0}$ nunca é vetor próprio de uma matriz !

Exemplo do slide 170 revisitado

Consideremos novamente a matriz $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. Vimos no slide 174 que A admite os valores próprios $\lambda = 1$ e $\lambda = 2$. Vamos calcular os respectivos subespaços próprios $E(1)$ e $E(2)$.

- Tem-se $E(1) = \mathcal{N}(A - 1\,I) = \mathcal{N}(A - I)$. Aplicando o método de Gauss,

$$\left[\, A - I \,|\, \vec{0} \,\right] = \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \text{ e portanto,}$$

$$\begin{aligned} \mathcal{N}(A - I) &= \{(x_1, x_2, x_3) : x_2 = 0, x_1, x_3 \in \mathbb{R}\} \\ &= \{(x_1, 0, x_3) : x_1, x_3 \in \mathbb{R}\} = \langle (1,0,0), (0,0,1) \rangle. \end{aligned}$$

- 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 $m.g.(1) = \dim E(1) = 2$.
- Os vetores próprios de A associados ao valor próprio $\lambda = 1$ são os vetores não nulos de $E(1)$. Por exemplo, tomando $x_1 = 1$ e $x_3 = -2$ obtém-se o vetor próprio $(1, 0, 2)$ de A associado ao valor próprio $\lambda = 1$.
- Geometricamente $E(1)$ define o plano de $\mathbb{R}^3$ que passa na origem com vetores diretores $(1,0,0)$ e $(0,0,1)$.

Exemplo (cont.)

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 $\left[\, A - 2I \mid \vec{0} \,\right]$ obtém-se,

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

- Logo, $E(2) = \mathcal{N}(A - 2I) = \{(x_3, x_3, x_3) : x_3 \in \mathbb{R}\} = \langle (1,1,1) \rangle$.
- Uma base para $E(2)$ é portanto $\{(1,1,1)\}$, tendo-se m.g.$(2) = \dim E(2) = 1$.
- Os vetores próprios de A associados ao valor próprio $\lambda = 2$ são portanto os vetores da reta que passa na origem com vetor diretor $(1,1,1)$, com excepção da origem.
- A informação, dita espectral sobre a matriz A pode ser organizada numa tabela

λ	$m.a.(\lambda)$	$m.g.(\lambda)$	base de $E(\lambda)$
1	2	2	$\{(1,0,0),(0,0,1)\}$
2	1	1	$\{(1,1,1)\}$

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 α é valor próprio de A
  $\longrightarrow$ Mostrar que $\det(A - \alpha I) = 0$.

- 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 não nulos de $E(\lambda) = \mathcal{N}(A - \lambda I)$.
  A multiplicidade geométrica de λ é m.g.$(\lambda) = \dim E(\lambda)$.

Propriedades dos valores próprios

Proposição

Seja A uma matriz quadrada de ordem n. Então:

- Para todo o valor próprio λ de A tem-se

$$\boxed{1 \leq m.g.(\lambda) \leq m.a.(\lambda)}$$

- 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, $tr(A)$, 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 1

Consideremos a matriz D do exercício 29.3 e a respectiva informação espectral,

$$D = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 5 \end{bmatrix}$$

λ	$m.a.(\lambda)$	$m.g.(\lambda)$	base de $E(\lambda)$
1	2	1	$\{(1,0,0)\}$
6	1	1	$\{(1,5,10)\}$

(ver as soluções da sebenta de exercícios). Constata-se que D possui 2 valores próprios distintos $\lambda_1 = 1$ e $\lambda_2 = 6$, tais que:

- $1 = m.g.(1) \leq m.a.(1) = 2$
  $1 = m.g.(6) \leq m.a.(6) = 1$.
- $m.a.(1) + m.a.(6) = 2 + 1 = 3 = n$ (ordem da matriz D).
- A soma dos valores próprios de D contando com repetições (m.a.), $1 + 1 + 6 = 8$, coincide com o traço de D, $tr(D) = 1 + 2 + 5 = 8$ (soma das entradas da diagonal principal, a vermelho na matriz).
- O produto dos valores próprios de D, contando com repetições (m.a.), $1 \times 1 \times 6$ coincide com $det(D) = 6$ (verifique).
- Como $\lambda = 0$ não é valor próprio a matriz D é invertível.

Exemplo 2

Consideremos a matriz $A = \begin{bmatrix} 2 & 2 & 6 \\ 0 & 4 & 6 \\ 1 & 8 & 1 \end{bmatrix}$. Tem-se:

- $v = (1, 1, 1)$ é vetor próprio de A associado ao valor próprio $\lambda = 10$. De facto,

$$Av = \begin{bmatrix} 2 & 2 & 6 \\ 0 & 4 & 6 \\ 1 & 8 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 10 \\ 10 \end{bmatrix} = 10 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 10v.$$

- $\lambda = 2$ é valor próprio de A. De facto, $\det(A - 2I) = \begin{vmatrix} 0 & 2 & 6 \\ 0 & 2 & 6 \\ 1 & 8 & -1 \end{vmatrix} = 0$,

  uma vez que $A - 2I$ que possui a primeira e segunda linhas múltiplas entre si.
- Logo A possui os valores próprios $\lambda_1 = 10$ e $\lambda_2 = 2$. Uma vez que A é uma matriz de ordem 3, A possui um terceiro valor próprio λ_3 (que pode ser repetido). Como a soma dos valores próprios de A é igual ao traço de A, ou seja, é igual à soma dos elementos da diagonal principal de A, tem-se

$$\lambda_1 + \lambda_2 + \lambda_3 = 2 + 4 + 1 = 7,$$

  e portanto $\lambda_3 = 7 - \lambda_1 - \lambda_2 = 7 - (10 + 2) = -5$.
- Logo A admite valores próprios distintos $\lambda_1 = 10$, $\lambda_2 = 2$ e $\lambda_3 = -5$. Como A tem ordem 3, não podem existir mais valores próprios. Logo os 3 valores próprios têm multiplicidade algébrica um, uma vez que não há valores próprios repetidos.
- Logo $m.g.(10) = m.g.(2) = m.g.(-5) = 1$ uma vez que $1 \leq m.g. \leq m.a.$

Exemplo 2 (cont.)

- Vamos reobter os valores próprios de A calculando as raízes do polinómio característico $p_A(\lambda) = \det(A - \lambda I)$.
- Aplicando a regra de Laplace ao longo da 1ª coluna da matriz $A - \lambda I$ obtém-se,

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 2 & 6 \\ 0 & 4-\lambda & 6 \\ 1 & 8 & 1-\lambda \end{vmatrix} \\ &= (2-\lambda)(-1)^{1+1}\big((4-\lambda)(1-\lambda) - 48\big) + 0 + 1(-1)^{3+1}\big(12 - 6(4-\lambda)\big) \\ &= (2-\lambda)(\lambda^2 - 5\lambda - 44) - 12 + 6\lambda = (2-\lambda)(\lambda^2 - 5\lambda - 44) - 6(2-\lambda) \\ &= (2-\lambda)(\lambda^2 - 5\lambda - 50) = (2-\lambda)(\lambda - 10)(\lambda + 5). \end{aligned}$$

- Na factorização do polinómio $\lambda^2 - 5\lambda - 50 = (\lambda - 10)(\lambda + 5)$ da última igualdade usou-se a fórmula resolvente para determinar as raízes $\alpha = 10$ e $\beta = -5$ de $\lambda^2 - 5\lambda - 50$([15]).
- Conclui-se novamente que A admite 3 valores próprios distintos, $\lambda_1 = 10$, $\lambda_2 = 2$ e $\lambda_3 = -5$, todos com multiplicidade algébrica um, porque na factorização de $p_A(\lambda)$ não há raízes repetidas.

[15] Recorde que se um polinómio de 2º grau $p(x) = ax^2 + bx + c$ admite raízes α e β então pode factorizar-se como $p(x) = a(x - \alpha)(x - \beta)$.

Valores próprios e independência linear

Sejam v_1 e v_2 vetores próprios colineares de A associados a valores próprios λ_1 e λ_2, isto é $v_2 = \alpha v_1$ com $\alpha \neq 0$, tais que $Av_1 = \lambda_1 v_1$ e $Av_2 = \lambda v_2$. Ora, podemos obter Av_2 de duas formas distintas:

$$Av_2 = \lambda_2 v_2 \quad \text{e} \quad Av_2 = A(\alpha v_1) = \alpha A v_1 = \alpha \lambda_1 v_1 = \lambda_1(\alpha v_1) = \lambda_1 v_2.$$

Logo $\lambda_2 v_2 = \lambda_1 v_2$, ou seja, $(\lambda_2 - \lambda_1)v_2 = \vec{0}$. Como $v_2 \neq \vec{0}$, tem-se $\lambda_2 = \lambda_1$.

Logo se dois 2 vetores próprios estão associados a valores próprios distintos então não podem ser colineares e portanto definem um conjunto linearmente independente. Mais geralmente tem-se o seguinte resultado.

Teorema

Sejam A matriz de ordem n e $\lambda_1, \dots, \lambda_k$ k valores próprios distintos de A. Tem-se:

- Um conjunto $\{v_1, \dots, v_k\}$ formado por k vetores próprios de A associados aos k valores próprios distintos $\lambda_1, \dots, \lambda_k$ ($Av_i = \lambda_i v_i$, $i = 1, \dots, k$), é linearmente independente.
- Mais geralmente, um conjunto obtido reunindo bases dos k subespaços próprios $E(\lambda_1), \dots, E(\lambda_k)$ define um conjunto linearmente independente de $\mathbb{R}^n$, que contém $m.g.(\lambda_1) + \cdots + m.g(\lambda_k)$ vetores próprios de A.

Bases próprias de $\mathbb{R}^n$ associadas a uma matriz

Como consequência do teorema do slide anterior tem-se o seguinte.

Critérios para a existência de bases próprias de $\mathbb{R}^n$

Sejam A matriz de ordem n e $\lambda_1, \dots, \lambda_k$ os valores próprios distintos de A.

- Se $k = n$, isto é, se A admite n valores próprios distintos, então existe uma base de $\mathbb{R}^n$ formada por vetores próprios de A. Mais precisamente, qualquer conjunto formado por n vetores próprios de A associados aos n valores distintos de A define uma base própria de $\mathbb{R}^n$ associada à matriz A.
- Se $k < n$ então existe uma base de $\mathbb{R}^n$ formada por vetores próprios de A, (obtida reunindo bases dos subesp. próprios $E(\lambda_1), \dots, E(\lambda_k)$) se e só se

$$m.g.(\lambda_1) + \cdots + m.g.(\lambda_k) = n.$$

Critério alternativo para a existência de bases próprias de $\mathbb{R}^n$

Como $m.a.(\lambda_1) + \cdots + m.a.(\lambda_k) = n$ e $m.g.(\lambda_i) \leq m.a.(\lambda_i)$ para todo o i, a condição anterior $m.g.(\lambda_1) + \cdots + m.g.(\lambda_k) = n$ é equivalente à condição,

$$m.g.(\lambda_i) = m.a.(\lambda_i), \qquad i = 1, \dots, k.$$

Bases próprias de $\mathbb{R}^n$ associadas a uma matriz - exemplos

- Consideremos novamente a matriz do slide 170 e a respectiva informação espectral (ver o slide 174),

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

λ	$m.a.(\lambda)$	$m.g.(\lambda)$	base de $E(\lambda)$
1	2	2	$\{(1,0,0),(0,0,1)\}$
2	1	1	$\{(1,1,1)\}$

Os valores próprios distintos de A, são portanto $\lambda_1 = 1$ e $\lambda_2 = 2$, tendo-se $m.g.(1) = m.a.(1) = 2$ e $m.g.(2) = m.a.(2) = 1$. Logo existe uma base de $\mathbb{R}^3$ formada por vetores próprios de A, que é obtida reunindo bases dos subespaços próprios $E(1)$ e $E(2)$, como por exemplo, a base própria

$$\{(1,0,0),(0,0,1),(1,1,1)\}.$$

- Como a matriz $A = \begin{bmatrix} 2 & 2 & 6 \\ 0 & 4 & 6 \\ 1 & 8 & 1 \end{bmatrix}$ do slide 181 é uma matriz de ordem 3 que admite 3 valores próprios distintos ($\lambda_1 = -5$, $\lambda_2 = 5$ e $\lambda_3 = 10$), obtém-se uma base de $\mathbb{R}^3$ formada por 3 vetores próprios de A considerando 3 vetores próprios associados a estes 3 valores próprios. Deixa-se como exercício indicar tal base.
- A matriz $D = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 5 \end{bmatrix}$ do exercício 29.3 possui 2 valores próprios distintos, $\lambda_1 = 1$ e $\lambda_2 = 6$, tendo-se $m.g.(1) = 1 < m.a.(1) = 2$ como vimos slide 180. Como $m.g.(1) \neq m.a.(1)$, não existe uma base de $\mathbb{R}^3$ formada por vetores próprios de D. Note-se que neste caso $m.g.(1) + m.g(6) = 1 + 1 = 2 < 3 = n$.

Bases próprias e diagonalização de matrizes

- Seja A uma matriz quadrada de ordem n.
- Dadas matrizes $P = [\, v_1 \ \cdots \ v_n \,]$ invertível e $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ diagonal, tem-se

$$\begin{aligned} AP &= A[\, v_1 \ v_2 \ \cdots \ v_n \,] = [\, Av_1 \ Av_2 \ \cdots \ Av_n \,], \\ PD &= [\, v_1 \ v_2 \ \cdots \ v_n \,] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\, \lambda_1 v_1 \ \lambda_2 v_2 \ \cdots \ \lambda_n v_n \,]. \end{aligned}$$

- Logo $AP = PD$ se só se $Av_i = \lambda_i v_i$, $i = 1, \ldots, n$, isto é, 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 $\lambda_1, \ldots, \lambda_n$.
- Multiplicando ambos os membros à esquerda por P^{-1}, obtém-se

$$AP = PD \quad \Leftrightarrow \quad P^{-1}AP = P^{-1}PD \quad \Leftrightarrow \quad P^{-1}AP = D,$$

dizendo-se então que A é diagonalizável, como veremos a seguir...

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 $P^{-1}AP = D$.

A matriz P designa-se por matriz de diagonalização para A.

Pelas considerações do slide anterior tem-se o seguinte.

Teorema

Seja A uma matriz quadrada de ordem n. Tem-se:

1. A é diagonalizável se e só se existe uma base de $\mathbb{R}^n$ formada por vetores próprios de A.
2. Se $\{v_1, \ldots, v_n\}$ é uma base de $\mathbb{R}^n$ formada por vetores próprios de A então a matriz desta base própria $P = [\, v_1 \ \cdots \ v_n \,]$ é uma matriz de diagonalização para A tendo-se,

$$P^{-1}AP = D = \text{diag}(\lambda_1, \ldots, \lambda_n),$$

onde $\lambda_1, \ldots, \lambda_n$ são os valores próprios (possivelmente repetidos) associados aos vetores próprios $v_1, \ldots, v_n$, respectivamente.

Exemplo do slide 170 revisitado

- Consideremos a matriz $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ do slide 170 cujos valores próprios distintos são $\lambda = 1$ e $\lambda = 2$ com $m.a.(1) = 2$ e $m.a.(2) = 1$.
- Vimos no slide 185 que $\{u_1, u_2, u_3\} = \{(1, 0, 0,), (0, 0, 1), (1, 1, 1)\}$ é uma base de $\mathbb{R}^3$ formada por vetores próprios de A, obtida reunindo a base $\{(1, 0, 0), (0, 0, 1)\}$ de $E(1)$ com a base $\{(1, 1, 1)\}$ de $E(2)$.
- Então 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.
- De facto, calculando a inversa de P tem-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 do slide 180 revisitado

- Consideremos novamente a matriz $D = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 5 \end{bmatrix}$ do slide 180 cujos valores próprios distintos são $\lambda = 1$ e $\lambda = 6$, tendo-se $m.a.(1) = 2 > m.g.(1) = 1$ e $m.a.(6) = m.g.(6) = 1$.
- 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.
- Note-se que a cardinalidade (número de vetores) máxima de um conjunto linearmente independente formado por vetores próprios de D é $m.g.(1) + m.g.(6) = 2 < 3 = \dim \mathbb{R}^3$.

Reconstrução de matrizes diagonalizáveis

Observações

- A partir da informação sobre os valores e vetores próprios (informação espectral) de uma matriz diagonalizável podemos reconstruir essa matriz.
- Mais precisamente, se $P = [\, v_1 \;\; \cdots \;\; v_n \,]$ é a matriz de uma base de $\mathbb{R}^n$ formada por vetores próprios de uma matriz diagonalizável A e $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ é a matriz diagonal que contém os correspondentes valores próprios ($Av_i = \lambda_i v_i$, $i = 1, \ldots, n$), então

$$\begin{aligned} P^{-1}AP = D \quad &\Leftrightarrow \quad P(P^{-1}AP)P^{-1} = PDP^{-1} \\ &\Leftrightarrow \quad (PP^{-1})A(PP^{-1}) = PDP^{-1} \\ &\Leftrightarrow \quad A = PDP^{-1} \end{aligned}$$

- Logo a matriz diagonalizável A pode ser reobtida a partir das matrizes P, formada por vetores próprios de A, e D matriz diagonal que contém os respectivos valores próprios, isto é, a partir da informação espectral de A.
- Notemos que se trocarmos a ordem dos vetores próprios na matriz de diagonalização P temos que trocar a ordem dos correspondentes valores próprios na matriz diagonal D.

Reconstrução de matrizes diagonalizáveis

Exercício na aula

Determinar a matriz A a partir da informação espectral dada na seguinte tabela

λ	$m.a.(\lambda)$	$m.g.(\lambda)$	base de $E(\lambda)$
-3	2	2	$\{(0,1,0),(1,1,-1)\}$
2	1	1	$\{(1,-1,0)\}$

Resolução:

- A é diagonalizável pois $m.g(-3) = m.a.(-3)$ e $m.g(2) = m.a.(2)$.
- Reunindo as bases dos subespaços próprios de $E(-3)$ e $E(2)$ dadas na tabela, obtém-se a base de $\mathbb{R}^3$ formada por vetores próprios de A,
$$\{v_1, v_2, v_3\} = \{\underbrace{(0,1,0)\,,\,(1,1,-1)}_{\text{base de } E(-3)}\,,\,\underbrace{(1,-1,0)}_{\text{base de } E(2)}\},$$
associados aos valores próprios $\lambda_1 = -3$, $\lambda_2 = -3$ e $\lambda_3 = 2$, respectivamente.
- Definindo $P = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$ e $D = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$ e calculando P^{-1} obtém-se por fim matriz $A = PDP^{-1}$, isto é,
$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & -1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 5 \\ -5 & -3 & -5 \\ 0 & 0 & -3 \end{bmatrix}.$$

Aplicação: cálculo de potências de matrizes diagonalizáveis

- Seja A uma matriz diagonalizável de ordem n e P matriz de diagonalização para A, tal que
$$P^{-1}AP = D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$
A partir da relação anterior obtém-se $A = PDP^{-1}$ (ver o slide 190) e portanto
$$\begin{aligned} A^\ell &= (PDP^{-1})^\ell = \underbrace{(PDP^{-1})(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})}_{\ell \text{ vezes}} \\ &= PD\underbrace{(P^{-1}P)}_{I_n}D\underbrace{(P^{-1}P)}_{I_n}D\cdots D\underbrace{(P^{-1}P)}_{I_n}DP^{-1} \\ &= PD^\ell P^{-1}. \end{aligned}$$
- Atendendo a que a potência de matrizes diagonais vem dada simplesmente por,
$$D^\ell = (\mathrm{diag}(\lambda_1, \ldots, \lambda_n))^\ell = \mathrm{diag}(\lambda_1^\ell, \ldots, \lambda_n^\ell),$$
obtém-se finalmente
$$A^\ell = P\,\mathrm{diag}(\lambda_1^\ell, \ldots, \lambda_n^\ell)\,P^{-1}.$$

Potências de matrizes diagonalizáveis - exercício

Exercício na aula

Determinar uma matriz de diagonalização P para

$$A = \begin{bmatrix} 11 & 6 \\ -18 & -10 \end{bmatrix}$$

e aplicar os cálculos descritos no slide anterior para calcular A^{10}.

Resolução: Calculando os valores próprios de A e os respectivos subespaços próprios obtém-se a matriz de diagonalização $P = \begin{bmatrix} -2 & -1 \\ 3 & 2 \end{bmatrix}$ associada à matriz diagonal $D = \text{diag}(2, -1)$ (verifique). A matriz inversa de P é (ver a mnemónica do slide 144),

$$P^{-1} = \frac{1}{-4+3} \begin{bmatrix} 2 & 1 \\ -3 & -2 \end{bmatrix} = \begin{bmatrix} -2 & -1 \\ 3 & 2 \end{bmatrix}.$$

Tem-se então $D^{10} = \left(\text{diag}(2, -1)\right)^{10} = \text{diag}\left(2^{10}, (-1)^{10}\right) = \text{diag}(1024, 1)$, e portanto

$$\begin{aligned} A^{10} = PD^{10}P^{-1} &= \begin{bmatrix} -2 & -1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1024 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & -1 \\ 3 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 4093 & 2046 \\ -6138 & -3068 \end{bmatrix}. \end{aligned}$$

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
Tomate	8000 m^3/ha	40 h/ha	300 €/ha
Trigo		20 h/ha	200 €/ha
Disponibilidades	$\leq$ 320000 m^3	$\leq$ 2000 h	max
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,

$$\begin{array}{lll}
\max & z = 300\,x + 200\,y & \\
\text{s.a} & x + y \leq 80 & (\text{T}) \\
& 8000\,x \leq 320000 & (\text{A}) \\
& 40\,x + 20\,y \leq 2000 & (\text{MO}) \\
& x, y \geq 0 &
\end{array}$$

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.

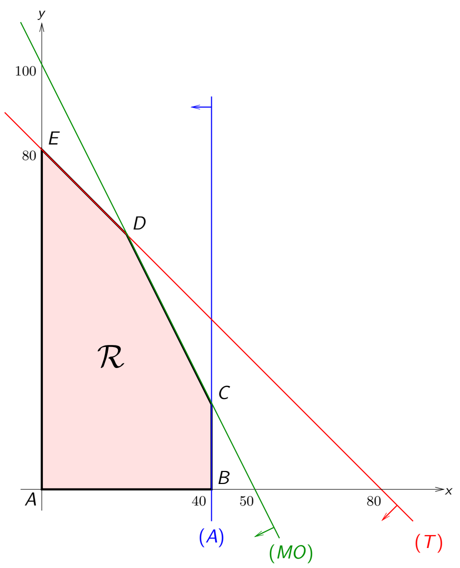
Solução e região admissível de um problema em PL

- 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 Problema 1 por 8000 e a terceira por 20 (ver o slide 196), 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 portanto 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.

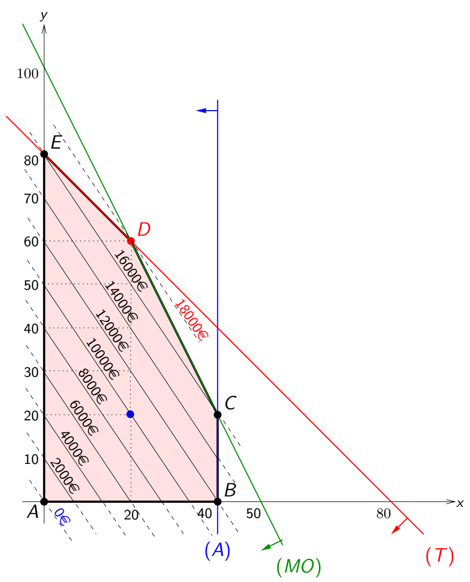
- Por exemplo, cultivar 20 hectares de tomate e 20 hectares de trigo corresponde à solução admissível $(20,20) \in \mathcal{R}$ e gera uma receita de $k = 300x20 + 200x20 = 10000$€.
- O conjunto das soluções admissíveis que geram a mesma receita que a solução $(20,20)$ é o conjunto

$$\{(x,y) \in \mathcal{R}\ :\ 300\,x + 200\,y = 10000\},$$

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, ou seja, 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

- Veremos mais adiante que se a região admissível $\mathcal{R}$ for não vazia e limitada, então pelo menos uma solução ótima do problema de PL ocorre num vértice do polígono de admissibilidade $\mathcal{R}$, "bastando" por isso enumerar todos os seus vértices e determinar o(s) vértice(s), onde a função objectivo atinge o valor máximo (ou mínimo).
- 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

$$C: \quad \begin{cases} x = 40 \\ 2x + y = 100 \end{cases} \Leftrightarrow \begin{cases} x = 40 \\ y = 20 \end{cases}$$

- 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€

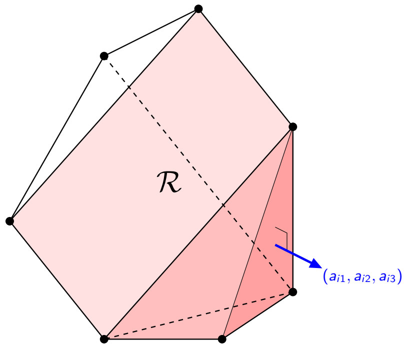
Formulação de um problema de PL: caso geral

- 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 z designada por função objetivo (f.o.), satisfazendo um conjunto de restrições funcionais (restrições lineares) (1),. . . , (m) e de sinal (m+1):

$$\begin{array}{lll} \text{max ou min} & z = c_1\, x_1 + c_2\, x_2 + \cdots + c_k\, x_k & \text{(f.o.)} \\ s.a & a_{11}\, x_1 + a_{12}\, x_2 + \cdots + a_{1k}\, x_k \geq, \leq \text{ ou } = b_1 & (1) \\ & a_{21}\, x_1 + a_{22}\, x_2 + \cdots + a_{2k}\, x_k \geq, \leq \text{ ou } = b_2 & (2) \\ & \vdots & \\ & a_{m1}\, x_1 + a_{m2}\, x_2 + \cdots + a_{mk}\, x_k \geq, \leq \text{ ou } = b_m & (m) \\ & x_1 \geq 0, \leq 0 \text{ ou livre}, \ldots, x_k \geq 0, \leq 0 \text{ ou livre} & (m+1) \end{array}$$

- 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 $\mathcal{R}$.

Poliedros e combinações convexas



Exemplo de um POLIEDRO DE ADMISSIBILIDADE em $\mathbb{R}^3$ (problema de PL com 3 variáveis de decisão), obtido como interseção de 7 semi-espaços em $\mathbb{R}^3$ definidos por 7 restrições lineares do tipo

$a_{i1}\, x_1 \;+\; a_{i2}\, x_2 \;+\; a_{i3}\, x_3 \;\geq$ ou $\leq\;\; b_i,\;\; i = 1, \ldots, 7.$

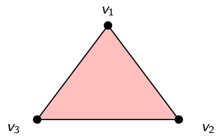
Os 7 (hiper)planos de suporte que contêm as 7 facetas do poliedro são definidos pelas equações lineares

$a_{i1}\, x_1 \;+\; a_{i2}\, x_2 \;+\; a_{i3}\, x_3 \;=\; b_i,\;\; i = 1, \ldots, 7,$

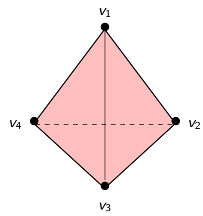
com cada vetor (a_{i1}, a_{i2}, a_{i3}) normal a uma faceta do poliedro.

Uma COMBINAÇÃO CONVEXA de $v_1, \ldots, v_k \in \mathbb{R}^n$ é uma combinação linear da forma $\alpha_1 v_1 + \cdots + \alpha_k v_k$ com $\alpha_1, \ldots, \alpha_k \geq 0$ e $\alpha_1 + \cdots + \alpha_k = 1$.

Cada ponto do segmento de reta é uma combinação convexa de v_1 e v_2.



Cada ponto do polígono (triângulo) é uma combinação convexa de v_1, v_2 e v_3.



Cada ponto do poliedro (tetraedro) é uma combinação convexa de v_1, v_2, v_3 e v_4.

Vértices da região admissível e solução ótima

Vimos no Problema 1 que uma solução ótima era atingida num vértice da região admissível. Nesta secção vamos generalizar essa propriedade para o problema genérico de um problema de PL,

$$\begin{array}{ll} \max\ (\text{ou } \mathit{min}) & z = c_1\, x_1 + c_2\, x_2 + \cdots + c_k\, x_k \\ s.a & (x_1, x_2, \ldots, x_k) \in \mathcal{R}, \end{array}$$

com $c_1, \ldots, c_k \in \mathbb{R}$ e onde $\mathcal{R}$ é a região admissível descrita no slide 202.

Teorema

Se $\mathcal{R}$ for limitada e não vazia tem-se:

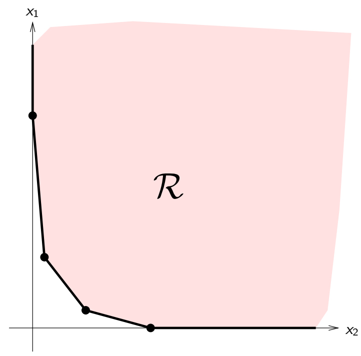
- Existe um vértice de $\mathcal{R}$ que é solução ótima do problema de PL anterior.
- Se k vértices de $\mathcal{R}$, $v_1, \ldots, v_k$, são soluções ótimas do problema de PL anterior então qualquer combinação convexa destes k vértices é também solução ótima do mesmo problema de PL.

O teorema anterior reduz o problema de determinar uma solução ótima de um problema de PL com região admissível limitada e não vazia, ao problema de identificar os vértices dessa região admissível (que são em número finito) e determinar o(s) vértice(s) onde a função objectivo atinge o maior ou menor valor, consoante o problema seja de maximização ou minimização.

Caso da região admissível não limitada

Observações

- Se a região admissível de um problema de PL for não limitada (como por exemplo na região da figura abaixo) pode não existir um vértice onde ocorra uma solução ótima.
- Por exemplo, no exercício 32.3 da sebenta de exercícios existe um vértice onde ocorre o mínimo da f.o., mas não existe um vértice onde ocorra o máximo (que é $+\infty$).



Formulação de um problema de PL na forma *standard*

Definição de formulação 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.

- Iremos ver que os vértices da região admissível do problema de PL original vão corresponder a um certo tipo especial de soluções, ditas soluções básicas admissíveis (s.b.a), da região admissível do problema de PL convertido para a forma *standard*.
- No que se segue iremos considerar apenas problemas de PL em que as respetivas variáveis de decisão $x_1, \ldots, x_k$ tomam valores não negativos, isto é, com restrições de sinal do tipo $x_1, \ldots, x_k \geq 0$.

Conversão de inequações lineares para equações lineares

As seguintes equivalências permitem converter cada restrição funcional definida por uma inequação linear numa restrição definida por uma equação linear acrescentando variáveis auxiliares ditas variáveis de folga que também tomam valores não negativos:

$$
\begin{aligned}
a_1x_1 + a_2x_2 + \ldots + a_kx_k \leq b \quad &\Leftrightarrow \quad \overbrace{b - (a_1x_1 + a_2x_2 + \ldots + a_kx_k)}^{f} \geq 0 \\
&\Leftrightarrow \quad \begin{cases} f &= \; b - (a_1x_1 + a_2x_2 + \ldots + a_kx_k) \\ f &\geq \; 0 \end{cases} \\
&\Leftrightarrow \quad \begin{cases} a_1x_1 + a_2x_2 + \ldots + a_kx_k + f &= \; b \\ f &\geq \; 0 \end{cases}
\end{aligned}
$$

e

$$
\begin{aligned}
a'_1x_1 + a'_2x_2 + \ldots + a'_kx_k \geq b' \quad &\Leftrightarrow \quad \overbrace{a'_1x_1 + a'_2x_2 + \ldots + a'_kx_k - b'}^{f'} \geq 0 \\
&\Leftrightarrow \quad \begin{cases} f' &= \; a'_1x_1 + a'_2x_2 + \ldots + a'_kx_k - b' \\ f' &\geq \; 0 \end{cases} \\
&\Leftrightarrow \quad \begin{cases} a'_1x_1 + a'_2x_2 + \ldots + a'_kx_k - f' &= \; b' \\ f' &\geq \; 0. \end{cases}.
\end{aligned}
$$

Conversão de um problema de PL para a forma *standard*

Regras de conversão

- Para cada restrição funcional do tipo,

$$a_1x_1 + a_2x_2 + \cdots + a_kx_k \leq b$$

acrescentamos uma nova variável f e substituímos essa restrição pelas restrições linear e de sinal,

$$a_1x_1 \; + \; a_2x_2 \; + \; \cdots \; + \; a_kx_n + f \; = b, \qquad f \geq 0;$$

- Para cada restrição funcional do tipo,

$$a'_1x_1 \; + \; a'_2x_2 \; + \; \cdots \; + \; a'_{ik}x_k \geq b',$$

acrescentamos uma nova variável f' e substituímos essa restrição pelas restrições linear e de sinal,

$$a'_1x_1 \; + \; a'_2x_2 \; + \; \cdots \; + \; a'_kx_n - f' \; = \; b', \qquad f' \geq 0;$$

- As restrições funcionais do tipo, $a''_1x_1 \; + \; a''_2x_2 \; + \; \cdots \; + \; a''_kx_k = b''$ e a função objetivo ficam **inalteradas**.

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		

Construção do modelo matemático

Variáveis de decisão

Temos 3 variáveis x_A, x_B e x_C que representam, respetivamente, as quantidades, em toneladas, de fertilizante dos tipos A, B e C a produzir mensalmente.

Função objetivo

A função objetivo (f.o.) traduz a quantidade de resíduos tóxicos gerados mensalmente que se pretende minimizar:

$$z = 50x_A + 40X_B + 60x_C.$$

Restrições funcionais

- A **capacidade mensal de produção** de fertilizantes é de 15000 t: $x_A + x_B + x_C \leq 15000$.
- A **produção mensal de fertilizante A** deve ser pelo menos de 5000 t: $x_A \geq 5000$.
- O **lucro mensal** deve ser pelo menos 100000 €: $10x_A + 5x_B + 10x_C \geq 100000$.
- a **produção mensal de fertilizantes** deve representar pelo menos 80 % da capacidade mensal de produção: $x_A + x_B + x_C \geq 15000 \times 0.80 = 12000$.

Restrições de sinal

Pela sua natureza as variáveis não podem tomar valores negativos: $x_A, x_B, x_C \geq 0$.

Formulação do Problema 2 em PL

O Problema 2 pode então ser formulado em PL como,

$$
\begin{array}{lllr}
\min & z = 50x_A + 40X_B + 60x_C & & \\
\text{s.a} & x_A + x_B + x_C \leq 15000 & & (1) \\
 & x_A \geq 5000 & & (2) \\
 & 10x_A + 5x_B + 10x_C \geq 100000 & & (3) \\
 & x_A + x_B + x_C \geq 15000 \times 0.80 = 12000 & & (4) \\
 & x_A,\ x_B,\ x_C \geq 0 & &
\end{array}
$$

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 208 (ver também o slide 207).

Veremos adiante como identificar os vértices da região admissível de um problema de PL determinando um certo tipo de soluções, ditas básicas admissíveis, da região admissível do problema na forma *standard*.

Formulação do Problema 2 na forma *standard*

$$
\begin{array}{llr}
\min & z = 50x_A + 40X_B + 60x_C & \\
\text{s.a} & x_A + x_B + x_C + f_1 = 15000 & (1) \\
 & x_A - f_2 = 5000 & (2) \\
 & 10x_A + 5x_B + 10x_C - f_3 = 100000 & (3) \\
 & x_A + x_B + x_C - f_4 = 12000 & (4) \\
 & x_A,\ x_B,\ x_C,\ f_1,\ f_2,\ f_3,\ f_4 \geq 0 &
\end{array}
$$

em que as variáveis do problema x_A, x_B e x_C já foram descritas no slide anterior e 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.

Soluções básicas admissíveis (s.b.a.)

Seja $A_{m\times n}$ tal que $\mathrm{car}(A) = m$ e $n > m$. Em particular, $\mathcal{C}(A) = \mathbb{R}^m$ e portanto o sistema $\left[A \,|\, b\right]$ é possível indeterminado para todo $b \in \mathbb{R}^m$. Define-se a região

$$\mathcal{F} = \{\bar{x} \in \mathbb{R}^n \;:\; A\,\bar{x} = b,\; \bar{x} \geq 0\},$$

onde $\bar{x} \geq 0$ significa que todas as componentes de $\bar{x}$ são não negativas.

Definição de solução básica admissível

- Uma solução básica é uma solução $x_{i_1,\ldots,i_m} \in \mathbb{R}^n$ obtida escolhendo um conjunto linearmente independente de m colunas $i_1,\ldots,i_m$ de A, resolvendo o sistema em ordem apenas às m correspondentes variáveis $x_{i_1},\ldots,x_{i_m}$ e fazendo as restantes $n-m$ variáveis iguais a zero.
- Se na solução básica $x_{i_1,\ldots,i_m}$ todas as componentes forem não negativas, $x_{i_1,\ldots,i_m}$ diz-se uma solução básica admissível (s.b.a.) de $\mathcal{F}$. Caso contrário, $x_{i_1,\ldots,i_m} \notin \mathcal{F}$ e diz-se solução básica não admissível (s.b.n.a.)

Como existem $\binom{n}{m}$ formas distintas de escolher m colunas de um conjunto de n, o número de soluções básicas (admissíveis e não admissíveis) não pode exceder $\binom{n}{m}$ (note-se que apenas escolhas de conjuntos linearmente independentes com m colunas dão origem a soluções básicas).

"Toy example"

Exercício na aula

Determinar as soluções básicas admissíveis de $\mathcal{F} = \{\bar{x} \in \mathbb{R}^3 : A\bar{x} = b,\; \bar{x} \geq 0\}$, com $\left[A \,|\, b\right] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 1 & -2 & -5 & -2 \end{array}\right]$.

Resolução: A matriz A tem $n = 3$ colunas e $m = 2$ linhas, tendo-se $\mathrm{car}(A) = m = 2$ e existem $\binom{3}{2} = 3$ maneiras distintas de escolher 2 colunas em 3:

- Resolvendo o sistema apenas com as colunas 1 e 2 de A (que são linearmente independentes), vem $\left[\begin{array}{cc|c} 1 & 1 & 4 \\ 1 & -2 & -2 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 2 \end{array}\right]$.
  Logo $x_1 = x_2 = 2$. Fazendo $x_3 = 0$ obtém-se a s.b.a. $x_{1,2} = (2,2,0) \in \mathcal{F}$.
- Resolvendo o sistema apenas com as colunas 1 e 3 de A (que são linearmente independentes), vem $\left[\begin{array}{cc|c} 1 & 1 & 4 \\ 1 & -5 & -2 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 1 \end{array}\right]$.
  Logo $x_1 = 3$ e $x_3 = 1$. Fazendo $x_2 = 0$ obtém-se a s.b.a. $x_{1,3} = (3,0,1) \in \mathcal{F}$.
- Resolvendo o sistema apenas com as colunas 2 e 3 de A (que são linearmente independentes), vem $\left[\begin{array}{cc|c} 1 & 1 & 4 \\ -2 & -5 & -2 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & 6 \\ 0 & 1 & -2 \end{array}\right]$.
  Logo $x_2 = 6$ e $x_3 = -2$. Fazendo $x_1 = 0$ obtém-se a s.b.n.a. $x_{2,3} = (0,6,-2) \notin \mathcal{F}$ (uma vez que possui uma componente negativa).

Logo $\mathcal{F}$ possui as s.b.a. $x_{1,2} = (2,2,0)$ e $x_{1,3} = (3,0,1)$ e a s.b.n.a. $x_{2,3} = (0,6,-2)$.

Vértices da região admissível e s.b.a.

O seguinte resultado torna evidente a importância do conceito de s.b.a.

Teorema

Consideremos um problema de PL com k variáveis de decisão $x_1, \ldots, x_k$ (problema original)

$$\begin{array}{ll} \min & z = c_1x_1 + c_2x_2 + \cdots + c_kx_k \\ \text{s.a} & x = (x_1, x_2, \ldots, x_k) \in \mathcal{R} \end{array}$$

e o correspondente problema na forma *standard* (ao qual se acrescentaram s variáveis de folga $f_1, \ldots, f_s$):

$$\begin{array}{ll} \min & z = c_1x_1 + c_2x_2 + \cdots + c_kx_k \\ \text{s.a} & \bar{x} = (x_1, \ldots, x_k, f_1, \ldots, f_s) \in \mathcal{F}. \end{array}$$

Para cada vetor $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$ denotamos por $\bar{x} = (x_1, \ldots, x_k, f_1, \ldots, f_s) \in \mathbb{R}^{k+s}$ o vetor que se obtém adicionando as respetivas folgas $f_1, \ldots, f_s$. Tem-se então que:

$$\begin{array}{c} x = (x_1, \ldots, x_k) \\ \text{é vértice de } \mathcal{R} \end{array} \quad \Leftrightarrow \quad \begin{array}{c} \bar{x} = (x_1, \ldots, x_k, f_1, \ldots, f_s) \\ \text{é s.b.a. de } \mathcal{F} \end{array}$$

Exemplo: Problema 1 revisitado

Consideremos novamente a formulação do Problema 1 do slide 196 mas com a região admissível $\mathcal{R}$ definida pelas restrições lineares já simplificadas do slide 197 e com as variáveis de decisão designadas por x_1, x_2 em vez de x, y:

Problema original

$$\begin{array}{lrl} \max & z = & 300\,x_1 + 200\,x_2 \\ \text{s.a} & & x_1 + x_2 \leq 80 \\ & & 2\,x_1 + x_2 \leq 100 \\ & & x_1 \quad\;\; \leq 40 \\ & & x_1,\ x_2 \;\; \geq 0 \end{array}$$

$--\!\!\rightarrow$

Problema na forma *standard*

$$\begin{array}{lrl} \max & z = & 300\,x_1 + 200\,x_2 \\ \text{s.a} & & x_1 + x_2 + f_1 \qquad\qquad = 80 \\ & & 2\,x_1 + x_2 \qquad + f_2 \qquad = 100 \\ & & x_! \qquad\qquad\qquad + f_3 = 40 \\ & & x_1,\ x_2 \geq 0 \end{array}$$

Temos a correspondência,

$$x = (x_1, x_2) \longrightarrow \bar{x} = (x_1, x_2, f_1, f_2, f_3) \text{ em que } \left\{ \begin{array}{lcl} f_1 & = & 80 - (x_1 + x_2) \\ f_2 & = & 100 - (2x_1 + x_2) \\ f_3 & = & 40 - x_1 \end{array} \right.$$

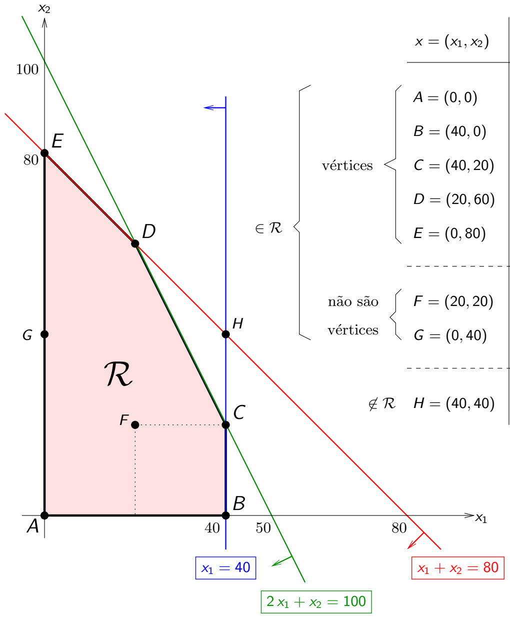
que associa a cada vértice da região admissível $\mathcal{R}$ do problema original uma s.b.a. da região admissível $\mathcal{F}$ do problema na forma standard e vice-versa.

Considerando, por exemplo, o vértice $D = (20, 60) \in \mathcal{R}$ (ver o slide 201), obtém-se

$$D = (20, 60) \;\longrightarrow\; \bar{D} = (20, 60, 0, 0, 20) \text{ s.b.a. de } \mathcal{F}.$$

De facto, calculando as folgas, tem-se $f_1 = 80 - (x_1 + x_2) = 80 - (20 + 60) = 0$, $f_2 = 100 - (2x_1 + x_2) = 100 - (40 + 60) = 0$ e $f_3 = 40 - x_1 = 40 - 20 = 20$.

Vértices e correspondentes s.b.a. do Problema 1, etc. . .



		$x = (x_1, x_2)$	$\bar{x} = (x_1, x_2, f_1, f_2, f_3)$		
$\in \mathcal{R}$	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:

$f_1 = 80 - (x_1 + x_2)$

$f_2 = 100 - (2\,x_1 + x_2)$

$f_3 = 40 - x_2$

Note-se que todas as s.b.a. têm 2=5-3 zeros e todas as suas componentes são não negativas. No próximo slide vamos dar um critério para verificar se um dado vetor é s.b.a. da região admissível $\mathcal{F}$ de um problema na forma *standard*.

Como reconhecer soluções básicas admissíveis ?

Critério para identificar uma s.b.a

Consideremos $A_{m\times n}$ tal que $\text{car}(A) = m$ e $n > m$ e

$$\mathcal{F} = \{\bar{x} \in \mathbb{R}^n \ : \ A\bar{x} = b, \ \bar{x} \geq 0\}.$$

Pode-se mostrar que $\bar{x} \in \mathbb{R}^n$ é uma solução básica admissível de $\mathcal{F}$ se e só se verificar as seguintes condições:

- $\bar{x}$ é solução do sistema linear $\left[A \,|\, b\right]$, isto é, verifica $A\,\bar{x} = b$.
- $\bar{x} \geq 0$, isto é, todas as suas componentes são não negativas.
- O número de componentes nulas de $\bar{x}$ é superior ou igual a $n - m$ (número de variáveis - número de equações).
- As colunas de A associadas às componentes não nulas de $\bar{x}$ formam um conjunto de vetores linearmente independente.

Note que se $\bar{x}$ verificar todas as condições excepto a segunda então $\bar{x}$ é solução básica não admissível (s.b.n.a.)

Exemplo

Exercício na aula

Considere o problema de PL (adaptado do Problema 2 do slide 212):

$$
\begin{array}{lllll}
\min & z = 5x_1 + 4X_2 + 6x_C & & & \\
\text{s.a} & \quad x_1 + \;\; x_2 + \quad x_3 & & \leq & 125 \\
& \quad x_1 + \;\; x_2 + \quad x_3 & & \geq & 120 \\
& \quad x_1 & & \geq & 50 \\
& 10x_1 + 5x_2 + 10x_3 & & = & 1000 \\
& x_1, x_2, x_3 & & \geq & 0
\end{array}
$$

- Converta o problema à forma standard.
- Mostre que $(50, 50, 25)$ é vértice da região admissível.

Resolução do exercício - forma *standard*

Aplicando as regras de conversão do slide 208 podemos formular o problema do slide anterior na forma *standard* como

$$
\begin{array}{lllll}
\min & z = 5x_1 + 4x_2 + 6x_3 & & & \\
\text{s.a} & \quad x_1 + \;\; x_2 + \quad x_3 + f_1 & & = & 125 \\
& \quad x_1 + \;\; x_2 + \quad x_3 \quad - f_2 & & = & 120 \\
& \quad x_1 \qquad\qquad\qquad\qquad - f_3 & & = & 50 \\
& 10x_1 + 5x_2 + 10x_3 & & = & 1000 \\
& x_1, x_2, x_3, f_1, f_2, f_3 & & \geq & 0
\end{array}
$$

em que

- $f_1 = 125 - (x_1 + x_2 + x_3)$,
- $f_2 = x_1 + x_2 + x_3 - 120$,
- $f_3 = x_1 - 50$.

Resolução do exercício - mostrar que é vértice

Calculando as folgas associadas ao vetor $x = (x_1, x_2, x_3) = (50, 50, 25)$ obtém-se o vetor $\bar{x} = (x_1, x_2, x_3, f_1, f_2, f_3) = (50, 50, 25, 0, 5, 0)$. De facto,

- $f_1 = 125 - (x_1 + x_2 + x_3) = 125 - (50 + 50 + 25) = 0$
- $f_2 = x_1 + x_2 + x_3 - 120 = 125 - 120 = 5$
- $f_3 = x_1 - 50 = 50 - 50 = 0.$

Tem-se que $x = (50, 50, 25)$ é vértice de $\mathcal{R}$ se e só se $\bar{x} = (50, 50, 25, 0, 5, 0)$ for s.b.a. da região admissível $\mathcal{F}$ do problema na forma *standard*.

Para mostrar que $\bar{x} = (50, 50, 25, 0, 5, 0)$ é s.b.a. de $\mathcal{F}$ começamos por escrever $\mathcal{F}$ em notação matricial,

$$\mathcal{F} = \{\bar{x} = (x_1, x_2, x_3, f_1, f_2, f_3) \in \mathbb{R}^6 \ : \ A\,\bar{x} = b, \ \bar{x} \geq 0\},$$

com

$$[A \mid b] = \left[\begin{array}{cccccc|c} 1 & 1 & 1 & 1 & 0 & 0 & 125 \\ 1 & 1 & 1 & 0 & -1 & 0 & 120 \\ 1 & 0 & 0 & 0 & 0 & -1 & 50 \\ 10 & 5 & 10 & 0 & 0 & 0 & 1000 \end{array}\right]$$

matriz ampliada do sistema linear que define a região admissível na forma *standard* (ver o slide anterior) e em que $\bar{x} \geq 0$ significa $x_1, x_2, x_3, f_1, f_2, f_3 \geq 0$.

Resolução do exercício - mostrar que é vértice (cont.)

Pelo slide 217 tem-se então que $\bar{x} = (50, 50, 25, 0, 5, 0)$ é s.b.a. de $\mathcal{F}$ se e só se verificar as 4 condições seguintes:

- $\bar{x}$ é solução de $[A \mid b]$, isto é, verifica $A\,\bar{x} = b$. De facto,

$$A\,\bar{x} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 \\ 10 & 5 & 10 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 50 \\ 50 \\ 25 \\ 0 \\ 5 \\ 0 \end{bmatrix} = \begin{bmatrix} 125 \\ 120 \\ 50 \\ 1000 \end{bmatrix} = b.$$

- $\bar{x} \geq 0$, isto é, todas as suas componentes são não negativas, o que se verifica.
- O número de componentes nulas de $\bar{x}$ é superior ou igual a $n - m$ com n número de variáveis (contando com variáveis de folga) e m número de restrições funcionais. De facto, $\bar{x}$ possui 2 componentes nulas verificando-se $2 \geq n - m = 6 - 4$ (6 variáveis e 4 restrições).
- As colunas de A associadas às componentes não nulas de $\bar{x}$ formam um conjunto de vetores linearmente independente (ver o próximo slide).

Resolução do exercício - mostrar que é vértice (concl.)

As colunas de A associadas às componentes não nulas de $\bar{x} = (50, 50, 25, 0, 5, 0)$ são a 1^{a}, 2^{a}, 3^{a} e 5^{a} colunas de A.

Logo basta mostrar que a matriz B constituída por estas 4 colunas tem característica 4, ou equivalentemente, que $\det B \neq 0$ (note-se que apenas se pode usar o determinante nos casos em que a matriz é quadrada).

De facto,

$$B = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 10 & 5 & 10 & 0 \end{bmatrix} \rightarrow \cdots \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} = B'.$$

Como todas as colunas da matriz em escada B' têm pivot o conjunto das colunas de A associadas às 4 componentes não nulas de $\bar{x}$ é linearmente independente.

Verificámos as 4 condições. Logo $\bar{x}$ é s.b.a. de $\mathcal{F}$ e portanto x é vértice de $\mathcal{R}$.

TPC: verifique que $\det(B) \neq 0$.