MIEEICOM
2018/19

Álgebra Linear e Geometria Analítica EE

Salvatore Cosentino
Departamento de Matemática e Aplicações - Universidade do Minho
Campus de Gualtar - 4710 Braga - PORTUGAL
gab CG - Edifício 6 - 3.48, tel 253 604086
e-mail scosentino@math.uminho.pt
url `http://w3.math.uminho.pt/~scosentino`

13 de Dezembro de 2018

Resumo

This is not a book! These are notes written for personal use while preparing lectures on "Álgebra Linear e Geometria Analítica" for students of FIS and MIENGFIS during the a.y.'s 2012/13 and 2013/14, and then for students of MIEEICOM during the a.y. 2018/19. They are rather informal and certainly contain mistakes (indeed, they are constantly actualized). I tried to be as synthetic as I could, without missing the observations that I consider important.

Most probably I will not lecture all I wrote, and did not write all I plan to lecture. So, I included sketched or even empty paragraphs, about material that I think should/could be lectured within the same course, given enough time.

References contain some introductory manuals that I like, some classics, books where I have learnt things in the past century, recent books which I find interesting. Almost all material can be found in [Ap69], [La97] and [Ax15]. Good material and further references can easily be found in the web, for example in Scholarpedia, in Wikipedia, or in the MIT OpenCoureWare.

Everything about the course may be found in my web pages

`http://w3.math.uminho.pt/~scosentino/salteaching.html`

The notation is as follows:

e.g. means EXEMPLI GRATIA, that is, "for example", and is used to introduce important or (I hope!) interesting examples.

ex: means "exercise", to be solved at home or in the classroom.

ref: means "references", places where you can find and study what follows inside each section.

Black paragraphs form the main text.

Blue paragraphs deal with applications and ideas relevant in physics, engineering or other sciences. They are the main reason why all this maths is worth studying for you. Some of them will only be understood and appreciated much later.

Red paragraphs (mostly written in english) are more advanced or non trivial facts and results which may be skipped in a first (and also second) reading.

$\square$ indicates the end of a proof.

Pictures were made with *Grapher* on my MacBook, or taken from Wikipedia, or produced with `Matlab` or `Mathematica®8` .

Conteúdo

1 Vetores

21 set 2018

ref: [Ap69] *Vol 1*, **12.1-4** ; [La97] *Ch. I*, **1-2**

A linguagem da filosofia. "... Signor Sarsi, la cosa non istà così. La filosofia è scritta in questo grandissimo libro che continuamente ci sta aperto innanzi agli occhi (io dico l'universo), ma non si può intendere se prima non s'impara a intender la lingua, e conoscer i caratteri, ne' quali è scritto. Egli è scritto in lingua matematica, e i caratteri son triangoli, cerchi, ed altre figure geometriche, senza i quali mezi è impossibile a intenderne umanamente parola; senza questi è un aggirarsi vanamente per un oscuro laberinto." [1]

A reta real. Fixada uma origem (ou seja, um ponto 0), um unidade de medida (ou seja, a distância entre 0 e 1) e uma orientação (ou seja, uma direção "positiva"), é possivel representar cada ponto de uma reta com um número real $x \in \mathbb{R}$. Vice-versa, ao número $x \in \mathbb{R}$ corresponde o ponto da reta colocado a uma distância $\sqrt{x^2}$ da origem, na direção positiva se $x > 0$ e negativa se $x < 0$.

O plano cartesiano. O *plano cartesiano*[2] $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$ é o conjunto dos *pontos* $\mathbf{r} = (x, y)$, com "coordenadas (cartesianas)" $x, y \in \mathbb{R}$ (chamadas *abcissa* e *ordenada*, respetivamente). A *origem* é o ponto $\mathbf{0} := (0, 0)$. O ponto $\mathbf{r} = (x, y)$ pode ser pensado como o *vetor* (i.e. o segmento orientado, a seta) entre a origem e o ponto $\mathbf{r}$. A *soma* dos vetores $\mathbf{r} = (x, y)$ e $\mathbf{r}' = (x', y')$ é o vetor

$$\mathbf{r} + \mathbf{r}' := (x + x', y + y'),$$

que representa uma diagonal do paralelogramo de lados $\mathbf{r}$ e $\mathbf{r}'$. Por exemplo, $(1, 2) + (3, 4) = (4, 6)$. O *produto* do número/escalar $\lambda \in \mathbb{R}$ pelo vetor $\mathbf{r} = (x, y)$ é o vetor

$$\lambda\mathbf{r} := (\lambda x, \lambda y)$$

que representa uma dilatação/contração (e uma inversão se $\lambda < 0$) de razão λ do vetor $\mathbf{r}$. Por exemplo, $3(1, 2) = (3, 6)$, e $-\frac{1}{2}(10, 12) = (-5, -6)$. Cada vetor pode ser representado de maneira única como soma

$$\mathbf{r} = (x, y) = x\mathbf{i} + y\mathbf{j},$$

onde $\mathbf{i} := (1, 0)$ e $\mathbf{j} := (0, 1)$ denotam os vetores da base canónica.

Lugares geométricos (pontos, retas, circunferências, parábolas, ...) podem ser descritos/definidos por equações algébricas, ditas "equações cartesianas".

ex: Descreva as coordenadas cartesianas dos pontos da reta que passa por $(1, 2)$ e $(-1, 3)$.

ex: Descreva as coordenadas cartesianas do triângulo de vértices $(0, 0)$, $(1, 0)$ e $(0, 2)$.

ex: Esboce os lugares geométricos definidos pelas equações

$$xy = 1 \qquad y = 2x - 7 \qquad (x + 1)^2 + (y - 3)^2 = 9 \qquad x - 2y^2 = 3$$

$$\begin{cases} x + y = 1 \\ x - 1 = 1 \end{cases} \qquad \begin{cases} x + y = 3 \\ -2x - 2y = -6 \end{cases} \qquad \begin{cases} x + y = 1 \\ 3x + 3y = 1 \end{cases}$$

ex: Escboce os lugares geométricos definidos pelas seguintes desigualdades

$$x - y \leq 1 \qquad \begin{cases} 0 \leq x \leq 1 \\ 0 \leq y \leq 1 \end{cases} \qquad \begin{cases} x + y \leq 1 \\ x - y \leq 1 \end{cases}$$

[1] Galileo Galilei, *Il Saggiatore*, 1623.

[2] René Descartes, *La Géométrie* [em *Discourse de la Méthode*, 1637].

ex: Determine umas desigualdades (cartesianas) que definem os pontos do paralelogramo de lados $(2,1)$ e $(3,5)$.

O espaço, o espaço-tempo e o espaço de fases da física newtoniana. O espaço onde acontece a física newtoniana é o *espaço 3-dimensional* $\mathbb{R}^3 := \mathbb{R}\times\mathbb{R}\times\mathbb{R}$. A posição de uma partícula é um ponto/vetor

$$\mathbf{r} = (x,y,z) := x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \in \mathbb{R}^3$$

onde $\mathbf{i} := (1,0,0)$, $\mathbf{j} := (0,1,0)$ e $\mathbf{k} := (0,0,1)$ denotam os vetores da base canónica. Ou seja, uma receita para deslocar um ponto material da *origem* $\mathbf{0} := (0,0,0)$ até a posição $\mathbf{r} = (x,y,z)$ consiste em fazer x "passos" na direção do vetor $\mathbf{i}$, depois y "passos" na direção do vetor $\mathbf{j}$ e enfim z "passos" na direção do vetor $\mathbf{k}$ (mas a ordem é indifferente!).

A *lei horária/trajetória*, de uma partícula é uma função $t \mapsto \mathbf{r}(t)$ que associa a cada tempo t num intervalo $I \subset \mathbb{R}$ a posição $\mathbf{r}(t) = (x(t), y(t), z(t)) \in \mathbb{R}^3$ da partícula no instante t. A *velocidade* da partícula no instante t é o vetor $\mathbf{v}(t) := \dot{\mathbf{r}}(t) = (\dot{x}(t), \dot{y}(t), \dot{z}(t))$ formado pelas derivadas das coordenadas. A *aceleração* da partícula no instante t é o vetor $\mathbf{a}(t) := \dot{\mathbf{v}}(t) = \ddot{\mathbf{r}}(t) = (\ddot{x}(t), \ddot{y}(t), \ddot{z}(t))$ formado pelas segundas derivadas das coordenadas. A aceleração de uma partícula de massa $m > 0$ num referencial inercial é determinada pela equação de Newton[3]

$$m\,\mathbf{a}(t) = \mathbf{F}(\mathbf{r}(t), \dot{\mathbf{r}}(t))$$

onde $\mathbf{F} : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ é um campo de forças.

Por exemplo, a lei horária do movimento retilíneo uniforme é $\mathbf{r}(t) = \mathbf{a} + \mathbf{v}t$ onde $\mathbf{a}$ é a posição inicial e $\mathbf{v}$ a velocidade, um vetor constante.

O *espaço-tempo* da física newtoniana é o produto cartesiano $\mathbb{R}\times\mathbb{R}^3 \approx \mathbb{R}^4$, o espaço dos *eventos* $(t,x,y,z) \in \mathbb{R}^4$, onde $\mathbf{r} = (x,y,z) \in \mathbb{R}^3$ representa uma posição num referencial inercial, e $t \in \mathbb{R}$ é o *tempo absoluto*.

O *estado* de uma partícula, a informação necessária e suficiente para resolver a equação de Newton e portanto determinar a sua trajetória futura (e passada), é um ponto $(\mathbf{r}, \mathbf{p}) \in \mathbb{R}^3\times\mathbb{R}^3 \approx \mathbb{R}^6$ do *espaço dos estados/de fases*, onde $\mathbf{r}$ é a posição e $\mathbf{p} := m\mathbf{v}$ é o *momento (linear)*.

ex: A lei horária da queda livre, próximo da superfície da terra, é $\mathbf{r}(t) = \mathbf{a} + \mathbf{v}t + (0,0,-g/2)t^2$, onde $\mathbf{a}$ é a posição inicial, $\mathbf{v}$ a velocidade inicial e $g \simeq 9.8$ m/s^2 a aceleração gravitacional. Calcule a velocidade $\dot{\mathbf{r}}(t)$ e a aceleração $\ddot{\mathbf{r}}(t)$.

ex: Determine a "dimensão" do espaço de fases de um sistema composto por 8 planetas (como, por exemplo, Mercúrio, Vênus, Terra, Marte, Júpiter, Saturno, Urano, Netuno) e de um sistema composto por 6×10^{23} moléculas.

Reações químicas. O estado de uma reação química

$$aA + bB + cC + \dots \quad\to\quad xX + yY + zZ + \dots$$

entre os n reagentes $A, B, C, \dots$ e os m produtos $X, Y, Z, \dots$ é descrito usando as concentrações $[A]$, $[B]$, $[C]$, ..., $[X]$, $[Y]$, $[Z]$, ..., e portanto $n+m$ números.

O espaço vetorial $\mathbb{R}^n$. O *espaço vetorial real* de *dimensão* n é o conjunto

$$\mathbb{R}^n := \underbrace{\mathbb{R}\times\mathbb{R}\times\cdots\times\mathbb{R}}_{n \text{ vezes}}$$

das n-uplas $\mathbf{x} = (x_1, x_2, \dots, x_n)$ de números reais, ditas *vetores* ou *pontos*, munido das operações *adição* $+ : \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}^n$, definida por

$$\boxed{\mathbf{x}, \mathbf{y} \mapsto \mathbf{x} + \mathbf{y} := (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)}$$

[3] Isaac Newton, *Philosophiæ Naturalis Principia Mathematica*, 1687.

e *multiplicação por um escalar* $\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, definida por

$$\boxed{\lambda, \mathbf{x} \mapsto \lambda\mathbf{x} := (\lambda x_1, \lambda x_2, \dots, \lambda x_n)}$$

O *vetor nulo/origem* é o vetor $\mathbf{0} := (0, 0, \dots, 0)$ (também denotado pela letra O), tal que $\mathbf{x} + \mathbf{0} = \mathbf{x}$ para todo $\mathbf{x} \in \mathbb{R}^n$. O *simétrico* do vetor $\mathbf{x} = (x_1, x_2, \dots, x_n)$ é o vetor $-\mathbf{x} := (-1)\mathbf{x} = (-x_1, -x_2, \dots, -x_n)$, tal que $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$. Isto justifica a notação $\mathbf{x} - \mathbf{y} := \mathbf{x} + (-\mathbf{y})$. A soma de vetores é comutativa, ou seja, $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$, e associativa, ou seja, $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$ (e portanto as parêntesis são inúteis). Em particular, a ordem dos vetores numa soma (finita) $\mathbf{x} + \mathbf{y} + \cdots + \mathbf{z}$ é irrelevante. É também evidente que $\lambda(\mu\mathbf{x}) = (\lambda\mu)\mathbf{x}$, e que valem as propriedades distributivas $(\lambda + \mu)\mathbf{x} = \lambda\mathbf{x} + \mu\mathbf{x}$ e $\lambda(\mathbf{x} + \mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}$. Finalmente, a multiplicação pelo escalar 1 é a identidade, i.e. $1\mathbf{x} = \mathbf{x}$, e a multiplicação pelo escalar "zero" produz o vetor nulo, i.e. $0\mathbf{x} = \mathbf{0}$.
.

A *combinação linear* dos vetores $\mathbf{v}_1$, $\mathbf{v}_2$, $\dots$, $\mathbf{v}_k \in \mathbb{R}^n$ com *coeficientes* λ_1, λ_2, $\dots$, $\lambda_k \in \mathbb{R}$ é o vetor

$$\sum_{i=1}^{k} \lambda_i \mathbf{v}_i := \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_k\mathbf{v}_k \,.$$

A *base canónica* de $\mathbb{R}^n$ é o conjunto ordenado dos vetores

$$\mathbf{e}_1 = (1, 0, \dots, 0) \quad \mathbf{e}_2 = (0, 1, 0, \dots, 0) \qquad \dots \qquad \mathbf{e}_n = (0, \dots, 0, 1)$$

assim que cada vetor $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ é uma combinação linear única

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$$

dos vetores da base canónica. O número x_k é chamado k-ésima *coordenada*, ou *componente*, do vetor $\mathbf{x}$ (relativamente à base canónica).

Um *subespaço vetorial* de $\mathbb{R}^n$ é um subconjunto não vazio $\mathbf{V} \subset \mathbb{R}^n$ que contém a origem, i.e. $\mathbf{0} \in \mathbf{V}$, e que contém todas as combinações lineares dos seus vetores, i.e. tal que se $\mathbf{x}, \mathbf{y} \in \mathbf{V}$ então também $\lambda\mathbf{x} + \mu\mathbf{y} \in \mathbf{V}$ para todos $\lambda, \mu \in \mathbb{R}$.

No *plano* $\mathbb{R}^2$ os pontos costumam ser denotados por $\mathbf{r} = (x, y)$, e no *espaço (3-dimensional)* $\mathbb{R}^3$ por $\mathbf{r} = (x, y, z)$.

ex: Calcule

$$(1, 2, 3) + (2, 3, 4) \qquad\qquad 6 \cdot (-1, -6, 0) \qquad\qquad (1, -1) - (3, 2)$$

ex: Calcule e esboce os pontos $A + B$, $A - B$, $2A - 3B$ e $-A + \frac{1}{2}B$ quando

$$A = (1, 2) \;\text{ e }\; B = (-1, 1) \qquad \text{ou} \qquad A = (0, 1, 7) \;\text{ e }\; B = (-2, 3, 0)$$

ex: [Ap69] **12.4**.

Vetores aplicados. Um *vetor aplicado/geométrico* (uma força, uma velocidade, ...) é um segmento orientado $\vec{AB}$ entre um ponto de aplicação $A \in \mathbb{R}^n$ e um ponto final $B \in \mathbb{R}^n$. Dois vetores aplicados $\vec{AB}$ e $\vec{CD}$ são *paralelos* se $B - A = \lambda(D - C)$ com $\lambda \in \mathbb{R}$, e são *equivalentes* (e portanto definem o mesmo "vetor" $\mathbf{x} = B - A$) se $B - A = D - C$.

ex: Mostre que cada vetor aplicado é equivalente a um vetor aplicado na origem.

ex: Diga se são paralelos ou equivalentes $\vec{AB}$ e $\vec{CD}$ quando

$$A = (1, 2) \qquad B = (-1, 1) \qquad C = (2, 3) \qquad D = (4, , 4)$$

$$A = (0, 1, \pi) \qquad B = (-2, 3, 0) \qquad C = (1, 0, -\pi) \qquad D = (2, 3, 0)$$

ex: Determine $D \in \mathbb{R}^n$ de maneira tal que $\vec{AB}$ e $\vec{CD}$ sejam equivalentes quando

$$A = (1,2) \qquad B = (-1,1) \qquad C = (2,3)$$

$$A = (0,1,\pi) \qquad B = (-2,3,0) \qquad C = (0,0,0)$$

Composição de forças. Se duas força $\mathbf{F}$ e $\mathbf{G}$ atuam sobre uma partícula colocada num certo ponto do espaço, então a "resultante" é uma força $\mathbf{F} + \mathbf{G}$.

Translações e homotetias. A soma de vetores e a multiplicação por um escalar, as duas operações algébricas definidas no espaço vetorial $\mathbb{R}^n$, descrevem transformações geométricas que são translações e homotetias.

Uma *translação* do espaço $\mathbb{R}^n$ é uma transformação $T_\mathbf{a} : \mathbb{R}^n \to \mathbb{R}^n$ definida por

$$\mathbf{x} \mapsto T_\mathbf{a}(\mathbf{x}) := \mathbf{x} + \mathbf{a}\,,$$

com $\mathbf{a} \in \mathbb{R}^n$. A imagem de um subconjunto $X \subset \mathbb{R}^n$ é também denotada $\mathbf{a} + X := T_\mathbf{a}(X)$.

A *homotetia* de razão $\lambda \neq 0$ é a transformação $M_\lambda : \mathbb{R}^n \to \mathbb{R}^n$ definida por

$$\mathbf{x} \mapsto M_\lambda(\mathbf{x}) := \lambda \mathbf{x}\,.$$

Representa uma dilatação quando $\lambda > 1$, e uma contração quando $\lambda < 1$. A imagem de um subconjunto $X \subset \mathbb{R}^n$ é também denotada $\lambda X := M_\lambda(X)$.

A composição de uma homotetia e uma translação é a transformação

$$\mathbf{x} \mapsto T_\mathbf{a} \circ M_\lambda(\mathbf{x}) = \lambda \mathbf{x} + \mathbf{a}\,.$$

Em particular, é possível definir uma *homotetia* de centro $\mathbf{c} \in \mathbb{R}^n$ e razão $\lambda \neq 0$ como sendo a transformação $H_{\mathbf{c},\lambda} : \mathbb{R}^n \to \mathbb{R}^n$

$$\mathbf{x} \mapsto H_{\mathbf{c},\lambda}(\mathbf{x}) := \mathbf{c} + \lambda(\mathbf{x} - \mathbf{c})\,.$$

Observe que o centro é um ponto fixo de uma homotetia, i.e. $H_{\mathbf{c},\lambda}(\mathbf{c}) = \mathbf{c}$.

ex: Mostre que $T_\mathbf{a} \circ T_\mathbf{b} = T_{\mathbf{a}+\mathbf{b}}$ e deduza que $T_\mathbf{a} \circ T_{-\mathbf{a}} = T_{-\mathbf{a}} \circ T_\mathbf{a} = \mathbf{1}$.

ex: Mostre que $\forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ existe uma translação $T_\mathbf{a}$ tal que $T_\mathbf{a}(\mathbf{x}) = \mathbf{y}$.

ex: Descreva o efeito das transformações M_λ, com $\lambda < 1$ e $\lambda > 1$, no plano $\mathbb{R}^2$.

ex: Determine e compare as transformações compostas $T_\mathbf{a} \circ M_\lambda$ e $M_\lambda \circ T_\mathbf{a}$. São iguais?

Graus Celsius, Fahrenheit e Kelvin. A temperatura pode ser medida em graus Celsius (C), Fahrenheit (F) e Kelvin (K), e

$$F = 1.8 \cdot C + 32 \qquad K = (F + 459.67)/1.8$$

ex: Determine a relação entre graus Kelvin e Celsius.

ex: Determine a relação entre um grau Kelvin e um grau Fahrenheit.

Invariância galileiana/sistemas inerciais. Seja $\mathbf{r} \in \mathbb{R}^3$ a posição de uma partícula num referencial inercial $\mathcal{R}$. Num referencial $\mathcal{R}'$ em movimento retilíneo uniforme com velocidade (constante) $\mathbf{V} \in \mathbb{R}^3$ e origem $\mathbf{R}$ no instante $t = 0$, a posição da partícula é dada pela "transformação de Galileo" [LL78]

$$\mathbf{r}' = \mathbf{r} - (\mathbf{R} + \mathbf{V}t)\,. \tag{1.1}$$

ex: Verifique que a diferença $\mathbf{r}_1 - \mathbf{r}_2$ entre as posições de duas partículas é invariante para as transformações (1.1), ou seja, $\mathbf{r}'_1 - \mathbf{r}'_2 = \mathbf{r}_1 - \mathbf{r}_2$. Verifique que a aceleração $\mathbf{a} := \ddot{\mathbf{r}}$ da trajetória $t \mapsto \mathbf{r}(t)$ de uma partícula também é invariante para as transformações (1.1), ou seja, a aceleração não depende do sistema inercial no qual é calculada. Deduza que se a força entre duas partículas apenas depende da diferença entre as posições (interação gravitacional, interação elétrica, ...), então a lei de Newton $m\ddot{\mathbf{r}} = \mathbf{F}$ é invariante.

ex: Mostre que o momento linear $\mathbf{p} := m\dot{\mathbf{r}}$ transforma segundo a lei

$$\mathbf{p}' = \mathbf{p} + m\mathbf{V} .$$

Centro de massas. O *centro de massas* do sistema de partículas de massas $m_1, m_2, \dots, m_N$ colocadas nos pontos $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N \in \mathbb{R}^3$ é

$$\mathbf{R} := \frac{1}{M} \sum_{k=1}^{N} m_k \mathbf{r}_k = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \cdots + m_N\mathbf{r}_N}{M}$$

onde $M := m_1 + m_2 + \cdots + m_N$ é a massa total do sistema.

ex: Três massas unitárias são colocadas nos vértices de um triângulo no plano. Caracterize o centro de massa.

2 Produto escalar, norma e distância

28 set 2018

ref: [Ap69] *Vol 1*, **12.5-11** ; [La97] *Ch. I*, **3-4**

Módulo e distância na reta real. O *módulo*, ou *valor absoluto*, do número real $x \in \mathbb{R}$ é

$$|x| := \max\{x, -x\} = \begin{cases} x & \text{se } x \geq 0 \\ -x & \text{se } x < 0 \end{cases} .$$

A *distância* entre os pontos x e y da reta real $\mathbb{R}$ é

$$\mathrm{d}(x,y) := |x - y|$$

O módulo e a distância satisfazem as desigualdades do triângulo

$$|x+y| \leq |x| + |y| \qquad \text{e} \qquad \mathrm{d}(x,y) \leq \mathrm{d}(x,z) + \mathrm{d}(y,z) .$$

ex: Mostre que as desigualdades podem ser estritas.

O plano euclidiano segundo Descartes. A geometria euclidiana do plano (distâncias, ângulos, paralelismo e perpendicularidade, ...) pode ser deduzida a partir da noção algébrica de *produto escalar/interno*

$$\mathbf{r} \cdot \mathbf{r}' := xx' + yy' .$$

Os vetores $\mathbf{r} = (x, y)$ e $\mathbf{r}' = (x', y')$ são *perpendiculares/ortogonais* quando $\mathbf{r} \cdot \mathbf{r}' = 0$, i.e. quando $xx' + yy' = 0$. O *comprimento*, ou *norma*, do vetor $\mathbf{r} = (x, y)$ é dado pelo teorema de Pitágoras:

$$\|\mathbf{r}\| := \sqrt{\mathbf{r} \cdot \mathbf{r}} = \sqrt{x^2 + y^2} .$$

A *distância* entre os pontos $\mathbf{r} = (x, y)$ e $\mathbf{r}' = (x', y')$ é

$$\mathrm{d}(\mathbf{r}, \mathbf{r}') := \|\mathbf{r} - \mathbf{r}'\| = \sqrt{(x - x')^2 + (y - y')^2} .$$

ex: Determine um vetor ortogonal ao vetor $(2, 3)$.

ex: Calcule o comprimento do vetor $(3, 4)$.

ex: Calcule a distância entre os pontos $(2, 1)$ e $(1, 2)$.

Produto escalar euclidiano. O *produto escalar/interno (euclidiano)* em $\mathbb{R}^n$ é a função $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definida por

$$\boxed{\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n}$$

Usando o *símbolo de Kroneker* δ_{ij}, definido por $\delta_{ii} = 1$ e $\delta_{ij} = 0$ se $i \neq j$, o produto escalar Euclidiano entre os vetores $\mathbf{x}$ e $\mathbf{y}$ é

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i,j} \delta_{ij} x_i y_j .$$

O produto escalar é "comutativo/simétrico", i.e.

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x} ,$$

"bilinear" (ou seja, linear em cada uma das duas variáveis), i.e.

$$(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z} , \quad \mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z} \quad \text{e} \quad (\lambda \mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (\lambda \mathbf{y}) = \lambda (\mathbf{x} \cdot \mathbf{y}) ,$$

e "positivo", i.e.

$$\mathbf{x} \cdot \mathbf{x} \geq 0 , \qquad \text{e} \qquad \mathbf{x} \cdot \mathbf{x} = 0 \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{0} .$$

Os vetores $\mathbf{x}$ e $\mathbf{y}$ são ditos *ortogonais/perpendiculares* quando $\mathbf{x} \cdot \mathbf{y} = 0$. (notação $\mathbf{x} \perp \mathbf{y}$)

ex: Mostre que o produto interno é simétrico, bilinear e positivo.

ex: Verifique que os vetores da base canónica são ortogonais dois a dois, i.e. $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ se $i \neq j$.

ex: Se $\mathbf{v}$ é ortogonal a todos os vetores $\mathbf{x} \in \mathbb{R}^n$, então $\mathbf{v} = \mathbf{0}$.

ex: Calcule o produto interno entre $\mathbf{x} = (1, 2)$ e $\mathbf{y} = (-1, 1)$, e entre $\mathbf{x} = (0, 1, \pi)$ e $\mathbf{y} = (-2, 3, 0)$.

ex: Determine se são ortogonais $\mathbf{x} = (1, 2)$ e $\mathbf{y} = (-1, 1)$, ou $\mathbf{x} = (0, 1, \pi)$ e $\mathbf{y} = (-2, 3, 0)$.

ex: Se $\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{z}$ então $\mathbf{y} = \mathbf{z}$?

ex: [Ap69] **12.8**.

Norma euclidiana. A *norma (euclidiana)* (ou seja, o *comprimento*) do vetor $\mathbf{x} \in \mathbb{R}^n$ é o número não-negativo

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

A norma é "(positivamente) homogénea", i.e.

$$\|\lambda \mathbf{x}\| = |\lambda| \, \|\mathbf{x}\| \, ,$$

e "positiva", i.e.

$$\|\mathbf{x}\| \geq 0 \, , \qquad \text{e} \qquad \|\mathbf{x}\| = 0 \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{0} \, .$$

Um vetor $\mathbf{x}$ é dito *unitário* se $\|\mathbf{x}\| = 1$.

ex: Mostre que a norma é homogénea e positiva.

ex: Verifique que os vetores da base canónica são unitários, i.e. $\|\mathbf{e}_i\| = 1$.

ex: Verifique que se $\mathbf{v} \neq 0$ então $\mathbf{u} = \mathbf{v}/\|\mathbf{v}\|$ é um vetor unitário paralelo a $\mathbf{v}$.

ex: Mostre que

$$\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 = 4 \, \mathbf{x} \cdot \mathbf{y} \, ,$$

e deduza que $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x} - \mathbf{y}\|$ sse $\mathbf{x} \cdot \mathbf{y} = 0$.

ex: Prove o *teorema de Pitágoras*: se $\mathbf{x}$ e $\mathbf{y}$ são ortogonais então

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \, .$$

ex: Verifique (e interprete) a *identidade do paralelogramo*

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2 \, \|\mathbf{x}\|^2 + 2 \, \|\mathbf{y}\|^2 \, .$$

ex: Sejam $\mathbf{x}$ e $\mathbf{y}$ dois vetores não paralelos (em particular, não nulos), e $\lambda > 0$ um escalar positivo. Então

$$\frac{\|\lambda \mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|\lambda \mathbf{y}\|}{\|\mathbf{y}\|} = \frac{\|\lambda(\mathbf{x} + \mathbf{y})\|}{\|\mathbf{x} + \mathbf{y}\|} = \lambda \, .$$

Fixado um ponto, por exemplo a origem O, desenhe os triângulos OAB e $OA'B'$, onde $A = O + \mathbf{x}$, $A' = O + \lambda \mathbf{x}$, $B = O + (\mathbf{x} + \mathbf{y})$ e $B' = O + \lambda(\mathbf{x} + \mathbf{y})$, e deduza/reconheça o *teorema de Tales* (ou *teorema da interseção*).

ex: Verifique que o produto interno euclidiano pode ser deduzido da norma usando a *identidade de polarização*

$$\mathbf{x}\cdot\mathbf{y} = \tfrac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\right) ,$$

ou

$$\mathbf{x}\cdot\mathbf{y} = \tfrac{1}{2}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\right) .$$

ex: Calcule a norma dos vetores $\mathbf{x} = (1-1,1)$, $\mathbf{y} = (-1,1)$ e $\mathbf{z} = (1,2,3,4)$.

Projeção e componente. Seja $\mathbf{v} \in \mathbb{R}^n$ um vetor não nulo. Cada vetor $\mathbf{x} \in \mathbb{R}^n$ pode ser representado de maneira única como soma

$$\mathbf{x} = \lambda\mathbf{v} + \mathbf{w}$$

de um vetor $\lambda\mathbf{v}$ proporcional a $\mathbf{v}$ e um vetor $\mathbf{w}$ ortogonal a $\mathbf{v}$. De fato, a condição de ortogonalidade $\mathbf{w} \perp \mathbf{v}$, ou seja, $(\mathbf{x}-\lambda\mathbf{v})\cdot\mathbf{v} = 0$, obriga a escolher

$$\lambda = \frac{\mathbf{x}\cdot\mathbf{v}}{\|\mathbf{v}\|^2} .$$

O vetor $\lambda\mathbf{v}$ é dito *projeção (ortogonal)* do vetor $\mathbf{x}$ sobre (a reta definida pel)o vetor $\mathbf{v}$, e o coeficiente λ é dito *componente* de $\mathbf{x}$ ao longo de $\mathbf{v}$. Em particular, a componente de $\mathbf{x}$ ao longo de um vetor unitário $\mathbf{u}$ é o produto escalar $\mathbf{x}\cdot\mathbf{u}$. Por exemplo, a componente do vetor $\mathbf{x}$ sobre o vetor $\mathbf{e}_k$ da base canónica é $x_k = \mathbf{x}\cdot\mathbf{e}_k$, a k-ésima coordenada de $\mathbf{x}$.

ex: Calcule a componente de $\mathbf{x} = (1,2)$ ao longo de $\mathbf{v} = (-1,1)$, e a projeção de $\mathbf{x} = (0,1,\pi)$ sobre $\mathbf{v} = (-2,3,0)$.

ex: [Ap69] **12.11**.

Desigualdade de Schwarz, ângulos e desigualdade do triângulo. Sejam $\mathbf{x},\mathbf{y}$ dois vetores do espaço euclidiano $\mathbb{R}^n$. Se $\mathbf{x}$ ou $\mathbf{y}$ é nulo, então o produto escalar $\mathbf{x}\cdot\mathbf{y}$ é também nulo. Se, por outro lado, $\mathbf{y} \neq 0$ e $\lambda\mathbf{y}$ é a projeção de $\mathbf{x}$ sobre $\mathbf{y}$, então a positividade da norma de $\mathbf{x}-\lambda\mathbf{y}$ diz que

$$0 \leq \|\mathbf{x}-\lambda\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - |\mathbf{x}\cdot\mathbf{y}|^2/\|\mathbf{y}\|^2 .$$

Isto prova a

Teorema 2.1 (desigualdade de Schwarz)**.** *O módulo do produto escalar entre cada dois vetores* $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$ *é limitado por*

$$\boxed{|\mathbf{x}\cdot\mathbf{y}| \leq \|\mathbf{x}\|\,\|\mathbf{y}\|} \tag{2.1}$$

e a igualdade verifica-se sse os vetores $\mathbf{x}$ *e* $\mathbf{y}$ *são proporcionais.*

Uma consequência é que a norma satisfaz a *desigualdade do triângulo* (ou *subaditividade*)

$$\boxed{\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|}$$

A desigualdade de Schwarz também implica que, se $\mathbf{x}$ e $\mathbf{y}$ são vetores não nulos, então

$$-1 \leq \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} \leq 1 .$$

Isto permite definir o *ângulo* entre os vetores não nulos $\mathbf{x}$ e $\mathbf{y}$ de $\mathbb{R}^n$ como sendo o único $\theta \in [0,\pi]$ tal que

$$\boxed{\cos\theta = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}}$$

ex: Sejam $\mathbf{x}$ e $\mathbf{y}$ dois vetores não nulos, e considere os vetores unitários $\mathbf{u} = \mathbf{x}/\|\mathbf{x}\|$ e $\mathbf{v} = \mathbf{y}/\|\mathbf{y}\|$. Calcule $\|\mathbf{u} \pm \mathbf{v}\|^2$ e mostre que $-1 \leq \mathbf{u} \cdot \mathbf{v} \leq 1$. Deduza a desigualdade de Schwarz (2.1).

ex: Prove a desigualdade do triângulo (calcule $\|\mathbf{x} + \mathbf{y}\|^2$ e use a desigualdade de Schwarz).

ex: Mostre que se θ é o ângulo entre $\mathbf{x}$ e $\mathbf{y}$ então

$$\|\mathbf{x} \pm \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \pm 2\|\mathbf{x}\|\|\mathbf{y}\| \cos\theta$$

ex: Calcule o coseno do ângulo entre $\mathbf{x} = (1, 2)$ e $\mathbf{y} = (-1, 1)$. Calcule o coseno do ângulo entre $\mathbf{x} = (0, 1, \pi)$ e $\mathbf{y} = (-2, 3, 0)$.

ex: Calcule o coseno dos ângulos do triângulo de vértices $A = (1, 1)$, $B = (-1, 3)$ e $C = (0, 2)$. Calcule o coseno dos ângulos do triângulo de vértices $A = (1, 2, 5)$, $B = (2, 1, 2)$ e $C = (0, 3, 0)$.

ex: Determine um vetor ortogonal ao vetor $(1, -1)$, e um vetor ortogonal ao vetor $(1, 3, 6)$.

ex: Determine a família dos vetores (ou seja, todos os vetores) de $\mathbb{R}^2$ ortogonais ao vetor (a, b). Determine a família dos vetores de $\mathbb{R}^3$ ortogonais ao vetor (a, b, c).

ex: [Ap69] **12.8**.

ex: [Ap69] **12.11**.

Distância euclidiana. A *distância (euclidiana)* entre os pontos $\mathbf{x}$ e $\mathbf{y}$ de $\mathbb{R}^n$ é o número não-negativo

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

Observe que $\mathrm{d}(\mathbf{x}, \mathbf{y}) = 0$ sse $\mathbf{x} = \mathbf{y}$. A distância satisfaz a *desigualdade do triângulo* (o comprimento de um lado de um triângulo é inferior a soma dos comprimentos dos outros dois lados)

$$\boxed{\mathrm{d}(\mathbf{x}, \mathbf{y}) \leq \mathrm{d}(\mathbf{x}, \mathbf{z}) + \mathrm{d}(\mathbf{z}, \mathbf{y})}$$

ex: Prove a desigualdade do triângulo (use a desigualdade homónima da norma).

ex: Prove que $\mathrm{d}(\lambda\mathbf{x}, \lambda\mathbf{y}) = |\lambda|\, \mathrm{d}(\mathbf{x}, \mathbf{y})$.

ex: Calcule a distância entre $\mathbf{x} = (1, 2)$ e $\mathbf{y} = (-1, 1)$, e a distância entre $\mathbf{x} = (0, 1, \pi)$ e $\mathbf{y} = (-2, 3, 0)$.

Trabalho e energia cinética. O *trabalho (mecânico)* infinitesimal que realiza um campo de forças $\mathbf{F}$ ao deslocar uma partícula (ao longo do segmento) do ponto $\mathbf{r}$ ao ponto $\mathbf{r} + d\mathbf{r}$ (se o deslocamento $d\mathbf{r}$ é "pequeno" então o campo pode ser considerado constante) é $dT := \mathbf{F} \cdot d\mathbf{r}$. Usando a equação de Newton $\mathbf{F} = \dot{\mathbf{p}}$, com $\mathbf{p} = m\mathbf{v}$, e considerando m constante, temos que

$$dT = \mathbf{F} \cdot d\mathbf{r} = m\frac{d\mathbf{v}}{dt} \cdot \mathbf{v}dt = m\,\mathbf{v} \cdot d\mathbf{v} = m\,\tfrac{1}{2}\, d(\mathbf{v} \cdot \mathbf{v})\,.$$

Portanto $dT = dK$, ou seja, o trabalho infinitesimal é igual a variação infinitesimal da *energia cinética*

$$K := \tfrac{1}{2}m\|\mathbf{v}\|^2\,.$$

Ao integrar, temos o *teorema trabalho-energia*: o trabalho $T := \int_{t_0}^{t_1} \mathbf{F} \cdot d\mathbf{r}$ realizado sobre uma partícula por um campo de forças $\mathbf{F}$ é igual a variação $\Delta K := K(t_1) - K(t_0)$ da energia cinética.

Bolas e esferas. A *bola aberta* e a *bola fechada* (ou *círculo*, se $n = 2$) de centro $\mathbf{x} \in \mathbb{R}^n$ e raio $r > 0$ são

$$\boxed{B_r(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^n \text{ t.q. } \|\mathbf{y} - \mathbf{x}\| < r\} \qquad \text{e} \qquad \overline{B_r(\mathbf{x})} := \{\mathbf{y} \in \mathbb{R}^n \text{ t.q. } \|\mathbf{y} - \mathbf{x}\| \leq r\}}$$

respetivamente. A *esfera* (ou *circunferência*, se $n = 2$) de centro $\mathbf{x} \in \mathbb{R}^n$ e raio $r > 0$ é

$$\boxed{S_r(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^n \text{ t.q. } \|\mathbf{y} - \mathbf{x}\| = r\}}$$

Em particular, a *esfera unitária* de dimensão $n - 1$ é

$$\mathbb{S}^{n-1} := S_1(\mathbf{0}) = \{\mathbf{x} \in \mathbb{R}^n \text{ t.q. } \|\mathbf{x}\| = 1\}\,.$$

ex: Determine uma equação cartesiana da circunferência de centro $(2, -1)$ e raio 7.

ex: Diga quando (ou seja, para quais valores dos raios r e r' dependendo dos centros $\mathbf{x}$ e $\mathbf{x}'$) a interseção $B_r(\mathbf{x}) \cap B_{r'}(\mathbf{x}')$ é $\neq \emptyset$.

ex: Verifique que $B_r(\mathbf{x}) = T_{\mathbf{x}}\left(M_{\mathbf{r}}\left(B_1(\mathbf{0})\right)\right)$ e $S_r(\mathbf{x}) = T_{\mathbf{x}}\left(M_{\mathbf{r}}\left(\mathbb{S}^{n-1}\right)\right)$.

Centroide. O *centroide* do sistema de pontos $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N \in \mathbb{R}^n$ é o ponto

$$\mathbf{C} := \frac{1}{N}\left(\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_N\right)$$

(o centro de massa de um sistema de partículas de massas unitárias colocadas nas posições $\mathbf{x}_1, \mathbf{x}_2, \dots \mathbf{x}_N$). Observe que em dimensão $n = 1$ o centróide da coleção de números $x_1, x_2, \dots, x_N$ é a *média aritmética* $\overline{x} := (x_1 + x_2, \cdots + x_N)/N$.

ex: Mostre que o centroide é o ponto $\mathbf{y}$ que minimiza a função

$$\mathcal{E}(\mathbf{y}) = \sum_{k=1}^{N} \|\mathbf{x}_k - \mathbf{y}\|^2\,.$$

ex: Calcule o centroide do sistema composto pelos pontos $(0, 1)$, $(2, 2)$ e $(3, 0)$ do plano.

ex: Mostre que o centroide de 3 pontos do plano, A, B e C, é a interseção dos segmentos que unem os vértices do triângulo ABC aos pontos médios dos lados opostos.

3 Retas e planos

ref: [Ap69] *Vol 1*, **13.1-8** ; [La97] *Ch. I*, **5-6**

5 out 2018

Curvas e superfícies. Curvas, superfícies e outros subconjuntos de $\mathbb{R}^n$ podem ser definidos de forma *paramétrica*, ou seja, como imagens $A = f(S) = \{f(s) \text{ com } s \in S\}$ de funções $f : S \to \mathbb{R}^n$ definidas em espaços de "parâmetros" $S = [0,1], \mathbb{R}, \mathbb{R}^2, \dots$ (por exemplo, uma "trajectória" $t \mapsto \mathbf{r}(t) \in \mathbb{R}^3$, onde t é o "tempo"), ou de forma *cartesiana*, ou seja, como "lugares geométricos" $B = \{\mathbf{x} \in \mathbb{R}^n \text{ t.q. } f_1(\mathbf{x}) = 0,\ f_2(\mathbf{x}) = 0, \dots\}$ dos pontos de $\mathbb{R}^n$ onde as funções $f_1 : \mathbb{R}^n \to \mathbb{R}$, $f_2 : \mathbb{R}^n \to \mathbb{R}, \dots$ se anulam.

ex: Descreva e esboce os seguintes subconjuntos de $\mathbb{R}^2$ ou $\mathbb{R}^3$.

$$\{(x,y) \in \mathbb{R}^2 \text{ t.q. } x^2 + y^2 = 2\} \qquad \{(x,y) \in \mathbb{R}^2 \text{ t.q. } xy = 0\}$$

$$\{(1,1) + t(0,3) \text{ com } t \in \mathbb{R}\} \qquad \{(0,4,0) + t(2,3,4) \text{ com } t \in \mathbb{R}\}$$

$$\{(\cos t, \sin t) \text{ com } t \in [0,2\pi]\} \qquad \{(\cos t, \sin t, s) \text{ com } t \in [0,2\pi]\,,\ s \in \mathbb{R}\}$$

$$\{(x,y) \in \mathbb{R}^3 \text{ t.q. } x^2 + y^2 + z^2 < \pi\} \qquad \{(t,|t|) \text{ com } t \in [-1,1]\} \qquad \{(t,t^2) \text{ com } t \in \mathbb{R}\}$$

$$\{(x,y) \in \mathbb{R}^2 \text{ t.q. } \langle (1,2),(x,y)\rangle = 0\} \qquad \{(x,y) \in \mathbb{R}^2 \text{ t.q. } 2x - 3y = 0\}$$

$$\{(x,y) \in \mathbb{R}^2 \text{ t.q. } \langle (3,-1),(x,y)\rangle = 2\} \qquad \{(x,y) \in \mathbb{R}^2 \text{ t.q. } 2x - 3y = 1\}$$

$$\{(x,y) \in \mathbb{R}^2 \text{ t.q. } -3x + y = 0 \text{ e } x - 7y = 0\} \qquad \{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } x + y + z = 0\}$$

$$\{(x,y) \in \mathbb{R}^2 \text{ t.q. } x + y = 2 \text{ e } 2x - y = 1\} \qquad \{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } x + y + z = 1\}$$

$$\{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } 2x - 3y - z = 0 \text{ e } x + y + 11z = 3\}$$

Partícula livre. A trajetória $t \mapsto \mathbf{r}(t) \in \mathbb{R}^3$ de uma partícula livre de massa m num referencial inercial é modelada pela equação de Newton

$$\frac{d}{dt}(m\mathbf{v}) = 0\,, \qquad \text{ou seja, se } m \text{ é constante,} \qquad m\mathbf{a} = 0\,,$$

onde $\mathbf{v}(t) := \dot{\mathbf{r}}(t)$ denota a velocidade da partícula no instante t, e $\mathbf{a}(t) := \ddot{\mathbf{r}}(t)$ denota a aceleração da partícula no instante t. Em particular, o "momento linear" $\mathbf{p} := m\mathbf{v}$, é uma constante do movimento, de acordo com o princípio de inércia de Galileo[4] ou a primeira lei de Newton[5]. As soluções da equação de Newton da partícula livre são as retas afins

$$\mathbf{r}(t) = \mathbf{s} + \mathbf{v}t$$

onde $\mathbf{s}, \mathbf{v} \in \mathbb{R}^3$ são a posição inicial e a velocidade (inicial), respetivamente.

ex: Determine a trajetória de uma partícula livre que passa, no instante $t_0 = 0$, pela posição $\mathbf{r}(0) = (3,2,1)$ com velocidade $\dot{\mathbf{r}}(0) = (1,2,3)$.

ex: Determine a trajetória de uma partícula livre que passa pela posição $\mathbf{r}(0) = (0,0,0)$ no instante $t_0 = 0$ e pela posição $\mathbf{r}(2) = (1,1,1)$ no instante $t_1 = 2$. Calcule a sua "velocidade escalar", ou seja, a norma $v = \|\mathbf{v}\|$.

---

[4] "...il mobile durasse a muoversi tanto quanto durasse la lunghezza di quella superficie, né erta né china; se tale spazio fusse interminato, il moto in esso sarebbe parimenti senza termine, cioè perpetuo."
. [Galileo Galilei, *Dialogo sopra i due massimi sistemi del mondo*, 1623]

[5] "Corpus omne perseverare in statu suo quiescendi vel movendi uniformiter in directum, nisi quatenus a viribus impressis cogitur statum illum mutare." [Isaac Newton, *Philosophiæ Naturalis Principia Mathematica*, 1687]

Retas. Um vetor não nulo $\mathbf{v} \in \mathbb{R}^n$ define/gera uma reta $\mathbb{R}\mathbf{v} := \{t\mathbf{v} \text{ com } t \in \mathbb{R}\}$ passando pela origem, o subespaço vetorial de $\mathbb{R}^n$ formado pelos vetores proporcionais a $\mathbf{v}$. Dois vetores não nulos e proporcionais geram a mesma reta. De fato, se $\mathbf{w} = \lambda\mathbf{v}$ com $\lambda \neq 0$, então $t\mathbf{v} = s\mathbf{w}$ quando $t = s\lambda$, e portanto $\mathbb{R}\mathbf{v} = \mathbb{R}\mathbf{w}$ (as retas diferem apenas pela parametrização, ou seja, a velocidade).

A *reta (afim)* paralela ao vetor não nulo $\mathbf{v} \in \mathbb{R}^n$ que passa pelo ponto $\mathbf{a} \in \mathbb{R}^n$ é

$$\boxed{\mathbf{a} + \mathbb{R}\mathbf{v} := \{\mathbf{a} + t\mathbf{v} \text{ com } t \in \mathbb{R}\}}$$

($\mathbf{v}$ é dito *vetor direccional* da reta, e pode ser pensado como a velocidade da trajetória $\mathbf{r}(t) = \mathbf{a}+t\mathbf{v}$). Duas retas afins $\mathbf{a}+\mathbb{R}\mathbf{v}$ e $\mathbf{b}+\mathbb{R}\mathbf{w}$ são ditas *paralelas* quando $\mathbf{w}$ é proporcional a $\mathbf{v}$ (ou seja, $\mathbf{w} = \lambda\mathbf{v}$), e *ortogonais* quando $\mathbf{w}$ é ortogonal a $\mathbf{v}$ (ou seja, $\mathbf{w} \cdot \mathbf{v} = 0$).

No plano $\mathbb{R}^2$, é possível eliminar o parâmetro t das equações $(x(t), y(t))) = \mathbf{a} + t\mathbf{v}$, e deduzir uma equação cartesiana da reta: por exemplo, se $\mathbf{a} = (a, b)$ e $\mathbf{v} = (v, w)$, então

$$\boxed{\mathbf{a} + \mathbb{R}\mathbf{v} = \{(x, y) \in \mathbb{R}^2 \text{ t.q. } w(x - a) - v(y - b) = 0\}}$$

Um vetor não nulo $\mathbf{n} \in \mathbb{R}^2$ define uma *reta normal* $\mathbf{n}^\perp := \{\mathbf{x} \in \mathbb{R}^2 \text{ t.q. } \mathbf{x} \cdot \mathbf{n} = 0\}$, subespaço vetorial de $\mathbb{R}^2$ formado pelos vetores ortogonais a $\mathbf{n}$. A reta perpendicular/normal ao vetor $\mathbf{n} \in \mathbb{R}^2$ que passa pelo ponto $\mathbf{a} \in \mathbb{R}^2$ é

$$\boxed{\mathbf{a} + \mathbf{n}^\perp := \{\mathbf{x} \in \mathbb{R}^2 \text{ t.q. } (\mathbf{x} - \mathbf{a}) \cdot \mathbf{n} = 0\}}$$

($\mathbf{n}$ é dito vetor *normal* à reta). Por exemplo, se $\mathbf{a} = (a, b)$ e $\mathbf{n} = (m, n)$, então

$$\boxed{\mathbf{a} + \mathbf{n}^\perp = \{(x, y) \in \mathbb{R}^2 \text{ t.q. } m(x - a) + n(y - b) = 0\}}$$

ex: Mostre que a reta que passa pelo pontos $\mathbf{a}$ e $\mathbf{b}$ de $\mathbb{R}^n$ é

$$\{\mathbf{a} + t\,(\mathbf{b} - \mathbf{a}) \text{ com } t \in \mathbb{R}\}$$

e portanto

$$\{t_1\mathbf{a} + t_2\mathbf{b} \text{ com } t_1, t_2 \in \mathbb{R} \text{ t.q. } t_1 + t_2 = 1\}$$

ex: Mostre que se $L = \mathbf{a} + \mathbb{R}\mathbf{v}$ e $L' = \mathbf{a}' + \mathbb{R}\mathbf{v}'$ são duas retas paralelas, então existe um vetor $\mathbf{r}$ tal que $L' = \mathbf{r} + L$.

ex: Determine uma equação paramétrica da reta

que passa pelo ponto $(2, 3)$ e é paralela ao vetor $(-1, 2)$

que passa pelo ponto $(5, 1, -2)$ e é paralela ao vetor $(3, -7, 2)$

que passa pelos pontos $(3, 3)$ e $(-1, -1)$

que passa pelos pontos $(0, 3, 4)$ e $(8, 3, 2)$

$$\{(x, y) \in \mathbb{R}^2 \text{ t.q. } 2x - 3y = 5\}$$

$$\{(x, y) \in \mathbb{R}^2 \text{ t.q. } -x + 7y = 0\}$$

ex: Determine uma equação cartesiana da reta

que passa pelo ponto $(5, -1)$ e é paralela ao vetor $(-6, 2)$

que passa pelos pontos $(0, 1)$ e $(-3, 4)$

que passa pelo ponto $(0, 0)$ e é perpendicular ao vetor $(-2, -3)$

que passa pelo ponto $(2, 1)$ e é perpendicular ao vetor $(9, 3)$

$$(-2, 3) + t(5, 1)$$

ex: Calcule o (coseno do) ângulo entre as retas

$$x - y = 0 \quad \text{e} \quad -x + y = -7$$

$$x + y = 1 \quad \text{e} \quad x - 2y = -4$$

ex: Determine um vetor normal à reta

$$\text{que passa pelos pontos } (3,0) \text{ e } (2,1)$$

$$5x - 3y = 2$$

ex: Determine $P \in \mathbb{R}^2$ e $\mathbf{v} \in \mathbb{R}^2$ tais que

$$\{(x,y) \in \mathbb{R}^2 \ \text{ t.q. } \ x + 2y = -1\} = \{P + t\mathbf{v} \ \text{ com } \ t \in \mathbb{R}\}$$

ex: As retas

$$\{(x,y) \in \mathbb{R}^2 \text{ t.q. } \ 2x - 3y = 5\} \quad \text{e} \quad \{(x,y) \in \mathbb{R}^2 \text{ t.q. } \ 3x - 2y = 5\}$$

$$\{(x,y) \in \mathbb{R}^2 \text{ t.q. } \ x + 7y = 3\} \quad \text{e} \quad \{(x,y) \in \mathbb{R}^2 \text{ t.q. } \ -2x - 14y = 0\}$$

são paralelas? São perpendiculares?

ex: Determine as interseções entre as retas

$$x - 2y = 1 \quad \text{e} \quad -2x + 4y = 3$$

$$3x + 5y = 0 \quad \text{e} \quad x - y = -1$$

$$(3,1) + t(1,3) \quad \text{e} \quad (0,1) + t(-1,-2)$$

ex: Determine a família das retas paralelas ao vetor $\mathbf{v} = (a,b)$ do plano.

ex: Determine a família das retas que passam pelo ponto (a,b) do plano.

ex: Verifique que homotetias e translações enviam cada reta numa reta paralela.

ex: [Ap69] **13.5**.

Segmentos. O *segmento (afim)* entre os pontos $\mathbf{x}$ e $\mathbf{y}$ de $\mathbb{R}^n$ é

$$\boxed{\overline{\mathbf{xy}} = \{t\mathbf{x} + (1-t)\mathbf{y} \ \text{ com } \ t \in [0,1]\}}$$

ex: Determine o segmento entre

$$\text{os pontos } (2,4) \text{ e } (-3,-8)$$

$$\text{os pontos } (1,0,1) \text{ e } (0,1,0)$$

ex: Determine a família dos pontos de $\mathbb{R}^n$ equidistantes de $\mathbf{x}$ e $\mathbf{y}$ (no plano, este conjunto é chamado *mediatriz* do segmento $\overline{\mathbf{xy}}$).

Ternos pitagóricos, método da corda de Diofanto. Uma solução inteira (i.e. com $X, Y, Z \in \mathbb{Z}$) da equação

$$X^2 + Y^2 = Z^2$$

é uma solução racional (i.e. com $x, y \in \mathbb{Q}$) de

$$x^2 + y^2 = 1$$

(basta dividir por $Z \neq 0$), ou seja, um ponto racional $(x,y) \in \mathbb{Q}^2$ da circunferência unitária $\mathbb{S} := \{(x,y) \in \mathbb{R}^2 \ \text{ t.q. } \ x^2 + y^2 = 1\} \subset \mathbb{R}^2$.

ex: Determine a (outra) interseção entre uma reta que passa pelo ponto $(-1, 0)$ e a circunferência unitária $\mathbb{S} := \{(x, y) \in \mathbb{R}^2 \text{ t.q. } x^2 + y^2 = 1\}$.

ex: Mostre que quando o declive d da reta é racional, ou seja $d = U/V$ com $U, V \in \mathbb{Z}$, a interseção determina uma solução inteira de $X^2 + Y^2 = Z^2$. Verifique que esta solução corresponde à solução de Euclides

$$X = (U^2 - V^2)W\,, \qquad Y = 2UVW\,, \qquad Z = (U^2 + V^2)W\,,$$

com $U, V, W \in \mathbb{Z}$.

Distância entre um ponto e uma reta. Seja $\mathbf{r}(t) = \mathbf{a} + t\mathbf{v}$, com $t \in \mathbb{R}$, a representação paramétrica dos pontos de uma reta em $\mathbb{R}^n$ passando pelo ponto $\mathbf{a}$ e paralela ao vetor não nulo $\mathbf{v}$. O quadrado da distância entre $\mathbf{r}(t)$ e a origem (ou qualquer outro ponto, basta mudar a origem do referencial!) é um polinómio de segundo grau no tempo t,

$$\|\mathbf{r}(t)\|^2 = t^2\,\|\mathbf{v}\|^2 + 2t\,\mathbf{a}\cdot\mathbf{v} + \|\mathbf{a}\|^2\,.$$

Esta função assume um mínimo quando o tempo é igual a $\bar{t} = -\mathbf{a}\cdot\mathbf{v}/\|\mathbf{v}\|^2$. Portanto o ponto da reta mais próximo da origem é

$$\mathbf{r}(\bar{t}) = \mathbf{a} - \frac{\mathbf{a}\cdot\mathbf{v}}{\|\mathbf{v}\|}\,\mathbf{v}\,,$$

o ponto onde $\mathbf{r}(\bar{t})$ é perpendicular ao vetor $\mathbf{v}$.

Em particular, no plano euclidiano, a distância do ponto $\mathbf{x} \in \mathbb{R}^2$ à reta $R = \mathbf{a} + \mathbf{n}^\perp \subset \mathbb{R}^2$ é a norma da projeção de $\mathbf{x} - \mathbf{a}$ sobre o vetor normal $\mathbf{n}$, i.e.

$$\boxed{\mathrm{d}(\mathbf{r}, R) = \frac{|(\mathbf{x} - \mathbf{a})\cdot\mathbf{n})|}{\|\mathbf{n}\|}}$$

Se a reta é $R = \{(x, y) \in \mathbb{R}^2 \text{ t.q. } mx + ny + c = 0\}$, então

$$\boxed{\mathrm{d}((x, y), R) = \frac{|mx + ny + c|}{\sqrt{m^2 + n^2}}}$$

ex: Mostre que $\|\mathbf{a}+t\mathbf{v}\| \geq \|\mathbf{a}\|$ para cada tempo $t \in \mathbb{R}$ sse $\mathbf{a}\cdot\mathbf{v} = 0$ (calcule o quadrado da norma de $\mathbf{a} + t\mathbf{v}$). Dê uma interpretação geométrica.

ex: Mostre que a norma de cada ponto $\mathbf{r} \in \mathbf{a}+\mathbf{n}^\perp$ da reta que passa por $\mathbf{a} \in \mathbb{R}^2$ e é perpendicular ao vetor não nulo $\mathbf{n} \in \mathbb{R}^2$ verifica

$$\|\mathbf{r}\| \geq d = \frac{|\mathbf{a}\cdot\mathbf{n}|}{\|\mathbf{n}\|}$$

e que $\|\mathbf{r}\| = d$ sse $\mathbf{r}$ é a projeção de $\mathbf{a}$ sobre $\mathbf{n}$, ou seja $\mathbf{r} = \frac{|\mathbf{a}\cdot\mathbf{n}|}{\|\mathbf{n}\|^2}\mathbf{n}$ (observe que a equação da reta é $\mathbf{r}\cdot\mathbf{n} = \mathbf{a}\cdot\mathbf{n}$ e use a desigualdade de Schwarz).

ex: Calcule a distância entre

o ponto $(8, -3)$ e a reta $(1, 0) + t(3, 3)$

o ponto $(2, 4)$ e a reta $x - y = 0$

o ponto $(11, -33)$ e a reta $y = 0$

ex: [Ap69] **13.5**.

Planos. Dois vetores $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ "linearmente independentes" (i.e. não paralelos) geram um *plano* $\mathbb{R}\mathbf{v} + \mathbb{R}\mathbf{w} := \{t\mathbf{v} + s\mathbf{w} \text{ com } (t,s) \in \mathbb{R}^2\} \subset \mathbb{R}^n$, subespaço vetorial de $\mathbb{R}^n$. O *plano (afim)* gerado pelos vetores $\mathbf{v}$ e $\mathbf{w}$ que passa pelo ponto $\mathbf{a} \in \mathbb{R}^n$ é

$$\boxed{\mathbf{a} + (\mathbb{R}\mathbf{v} + \mathbb{R}\mathbf{w}) := \{\mathbf{a} + t\mathbf{v} + s\mathbf{w} \;\text{ com }\; (t,s) \in \mathbb{R}^2\}}$$

No espaço $\mathbb{R}^3$, é possível eliminar os parâmetros t e s e deduzir uma equação cartesiana do plano, da forma

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

Um vetor não nulo $\mathbf{n} \in \mathbb{R}^3$ define um plano normal $\mathbf{n}^\perp := \{\mathbf{x} \in \mathbb{R}^3 \text{ t.q. } \mathbf{x} \cdot \mathbf{n} = 0\}$, subespaço vetorial de $\mathbb{R}^3$. O plano ortogonal/perpendicular/normal ao vetor não nulo $\mathbf{n} \in \mathbb{R}^3$ que passa pelo ponto $\mathbf{a} \in \mathbb{R}^3$ é

$$\boxed{\mathbf{a} + \mathbf{n}^\perp := \{\mathbf{x} \in \mathbb{R}^3 \;\text{ t.q. }\; (\mathbf{x} - \mathbf{a}) \cdot \mathbf{n} = 0\}}$$

O vetor $\mathbf{n}$ (que é definido a menos de um múltiplo $\neq 0$) é dito *vetor normal* ao plano. O ângulo entre dois planos de $\mathbb{R}^3$ pode ser definido como sendo o ângulo entre dois vetores normais aos planos.

ex: Mostre que o plano que passa pelo três pontos $\mathbf{a}$, $\mathbf{b}$ e $\mathbf{c}$ de $\mathbb{R}^n$, com $\mathbf{b}-\mathbf{a}$ e $\mathbf{c}-\mathbf{a}$ não paralelos, é

$$\{\mathbf{a} + t(\mathbf{b} - \mathbf{a}) + s(\mathbf{c} - \mathbf{a}) \;\text{ com }\; (t,s) \in \mathbb{R}^2\}\,,$$

e portanto

$$\{t_1\mathbf{a} + t_2\mathbf{b} + t_3\mathbf{c} \;\text{ com }\; t_1, t_2, t_3 \in \mathbb{R} \;\text{ t.q. }\; t_1 + t_2 + t_3 = 1\}$$

ex: Mostre que uma equação cartesiana do plano perpendicular ao vetor $\mathbf{n} = (m, n, p) \in \mathbb{R}^3$ que passa pelo ponto $\mathbf{a} = (a, b, c) \in \mathbb{R}^3$ é

$$m(x - a) + n(y - b) + p(z - c) = 0$$

ex: Determine uma equação paramétrica do plano

que passa pelo ponto $(5, 1, -2)$ e é gerado pelos vetores $(3, -7, 2)$ e $(-1.0, -1)$

que passa pelos pontos $(0, 3, 4)$, $(0, 5, 0)$ e $(8, 3, 2)$

$\{(x, y, z) \in \mathbb{R}^3 \text{ t.q. } x + y + z = 1\}$

$\{(x, y, z) \in \mathbb{R}^3 \text{ t.q. } z = 0\}$

ex: Determine uma equação cartesiana do plano

que passa pelo ponto $(-1, 1, 11)$ e é gerado pelos vetores $(1, 0, 0)$ e $(0, 1, 0)$

que passa pelo ponto $(0, 0, 0)$ e é gerado pelos vetores $(3, -7, 2)$ e $(-1.0, -1)$

que passa pelos pontos $(3, 3, 3)$ e é paralelo ao plano $x + y + z = 0$

que passa pelos pontos $(0, 3, 4)$, $(0, 5, 0)$ e $(8, 3, 2)$

que passa pelo ponto $(0, 0, 0)$ e é perpendicular ao vetor $(-2, -3, -4)$

que passa pelo ponto $(2, 1, 0)$ e é perpendicular ao vetor $(9, 3, 0)$

ex: Calcule o (coseno do) ângulo entre os planos

$$x - y + z = 0 \quad \text{e} \quad -x + 3y + 5z = -7$$

$$x - z = 2 \quad \text{e} \quad x - y = -3$$

ex: Determine um vetor normal ao plano

$$\text{que passa pelos pontos } (0,0,0)\,,\ (1,0,0) \text{ e } (0,1,0)$$

$$y+z=1$$

ex: Determine as intersecções entre os planos

$$x+2y+3z=-1 \qquad \text{e} \qquad -2x+4y-z=3$$

$$3x-5y=0 \qquad \text{e} \qquad x+y+z=1$$

ex: [Ap69] **13.8** e **13.17**.

Distância entre um ponto e um plano em $\mathbb{R}^3$. A distância do ponto $\mathbf{x} \in \mathbb{R}^3$ ao plano $P = \{\mathbf{x} \in \mathbb{R}^3 \text{ t.q. } (\mathbf{x}-\mathbf{a})\cdot\mathbf{n} = 0\} \subset \mathbb{R}^3$ é a norma da projeção de $\mathbf{x}-\mathbf{a}$ sobre o vetor normal $\mathbf{n}$, i.e.

$$\boxed{\mathrm{d}(\mathbf{x}, P) = \frac{|(\mathbf{x}-\mathbf{a})\cdot\mathbf{n}|}{\|\mathbf{n}\|}}$$

Se o plano é $P = \{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } mx+ny+pz+c=0\}$, então

$$\boxed{\mathrm{d}((x,y,z), P) = \frac{|mx+ny+pz+c|}{\sqrt{m^2+n^2+p^2}}}$$

ex: Calcule a distância entre

$$\text{o ponto } (2,4,1) \text{ e o plano } x+y+z=0$$

$$\text{o ponto } (5,7,1) \text{ e o plano que passa por } (2,0,3) \text{ e é normal ao vetor } (1,1,0)$$

$$\text{o ponto } (15,11,17) \text{ e o plano } xy$$

ex: [Ap69] **13.17**.

4 Subespaços, bases e dimensão

ref: [Ap69] *Vol 1*, **12.12-17** ; [La97] *Ch. III*, **1-5**

12 out 2018

Subespaços e geradores. Um *subespaço vetorial/linear* de $\mathbb{R}^n$ é um subconjunto não vazio $\mathbf{V} \subset \mathbb{R}^n$ tal que $\mathbf{v} + \mathbf{v}' \in \mathbf{V}$ e $\lambda \mathbf{v} \in \mathbf{V}$ para todos os $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$ e $\lambda \in \mathbb{R}$. Em particular, todo subespaço contem o vetor nulo $\mathbf{0}$, e o subespaço minimal é o subespaço trivial $\{\mathbf{0}\}$. O subespaço maximal é o próprio $\mathbb{R}^n$.

Seja $S = \{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \dots\}$ um subconjunto de $\mathbb{R}^n$, finito ou não. O conjunto das "combinações lineares" (finitas) dos elementos de S, i.e.

$$\boxed{\mathrm{Span}(S) := \{\lambda_1 \mathbf{s}_1 + \lambda_2 \mathbf{s}_2 + \cdots + \lambda_k \mathbf{s}_k \quad \text{com} \quad \mathbf{s}_i \in S\,,\ \lambda_i \in \mathbb{R}\}}$$

é um subsespaço vetorial de $\mathbb{R}^n$, dito subespaço *gerado* por S.

e.g. **Retas.** O subespaço gerado por um vetor não nulo $\mathbf{v} \in \mathbb{R}^n$ é a reta $\mathbb{R}\mathbf{v}$.

e.g. **Planos.** O subespaço gerado por dois vetores não nulos $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ é o plano $\mathbb{R}\mathbf{v} + \mathbb{R}\mathbf{w}$, se $\mathbf{v}$ e $\mathbf{w}$ não são paralelos, ou a reta $\mathbb{R}\mathbf{v}$ se $\mathbf{w} = \lambda \mathbf{v}$.

ex: Determine o subespaço gerado por

$$(3,-2) \text{ e } (-6,4) \qquad \text{em } \mathbb{R}^2$$
$$(1,1) \text{ e } (1,-1) \qquad \text{em } \mathbb{R}^2$$
$$(1,0,0) \text{ e } (0,1,0) \qquad \text{em } \mathbb{R}^3$$
$$(1,0,0)\,,\ (0,1,0) \text{ e } (1,-1,0) \qquad \text{em } \mathbb{R}^3$$
$$\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k \qquad \text{em } \mathbb{R}^n$$

ex: Diga se são subespaços vetoriais os seguintes subconjuntos de $\mathbb{R}^2$ ou $\mathbb{R}^3$

$$\{(t,3t) \ \text{com} \ t \in \mathbb{R}\} \qquad \{(2t,3t) \ \text{com} \ t \in [0,1]\} \qquad \{(t-1, 2+3t) \ \text{com} \ t \in \mathbb{R}\}$$
$$\{(t, 3s-t, s) \ \text{com} \ (t,s) \in \mathbb{R}^2\} \qquad \{(1-t, t+s, 5) \ \text{com} \ (t,s) \in \mathbb{R}^2\}$$
$$\{(x,y) \in \mathbb{R}^2 \ \text{t.q.} \ x - 2y = 0\} \qquad \{(x,y) \in \mathbb{R}^2 \ \text{t.q.} \ x - 2y = 1\}$$
$$\{(x,y,z) \in \mathbb{R}^3 \ \text{t.q.} \ x - y + z = 0 \ \text{e} \ -2x + y - z = 1\}$$
$$\{(x,y,z) \in \mathbb{R}^3 \ \text{t.q.} \ x + y + z = 0 \ \text{e} \ x - y - 3z = 0\}$$

ex: Dado $\mathbf{n} \in \mathbb{R}^n$, o conjunto

$$\mathbf{n}^\perp := \{\mathbf{x} \in \mathbb{R}^n \ \text{t.q.} \ \mathbf{n} \cdot \mathbf{x} = 0\}$$

é um subespaço linear de $\mathbb{R}^n$. Dados $\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_m \in \mathbb{R}^n$, o conjunto

$$\bigcap_{k=1}^{m} \mathbf{n}_k{}^\perp = \{\mathbf{x} \in \mathbb{R}^n \ \text{t.q.} \ \mathbf{n}_i \cdot \mathbf{x} = 0 \quad \forall i = 1, 2, \dots, m\}$$

é um subespaço linear de $\mathbb{R}^n$. Em geral, dado um subconjunto $N \subset \mathbb{R}^n$ (não necessariamente um subespaço), o conjunto

$$N^\perp := \{\mathbf{x} \in \mathbb{R}^n \ \text{t.q.} \ \mathbf{n} \cdot \mathbf{x} = 0 \quad \forall \mathbf{n} \in N\}$$

é um subespaço linear de $\mathbb{R}^n$

ex: O conjunto

$$\{\mathbf{x} \in \mathbb{R}^n \ \text{t.q.} \ \mathbf{n} \cdot \mathbf{x} = 1\}$$

é um subespaço linear de $\mathbb{R}^n$?

ex: [Ap69] **12.15**.

Dependência linear. Um vetor $\mathbf{v} \in \mathbb{R}^n$ é *(linearmente) dependente* dos vetores $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m$ (ou do conjunto finito $S = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\} \subset \mathbb{R}^n$), se $\mathbf{v} \in \mathrm{Span}(\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\})$, ou seja, se $\mathbf{v}$ é uma combinação linear

$$\mathbf{v} = \lambda_1 \mathbf{s}_1 + \lambda_2 \mathbf{s}_2 + \dots + \lambda_m \mathbf{s}_m$$

dos vetores $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m$, ou seja, se

$$\lambda \mathbf{v} + \lambda_1 \mathbf{s}_1 + \lambda_2 \mathbf{s}_2 + \dots + \lambda_m \mathbf{s}_m = 0$$

com $\lambda \neq 0$. Um vetor dependente do conjunto vazio $\emptyset$ é um vetor nulo.

ex: O vetor $\mathbf{v}$ é dependente do vetor $\mathbf{w}$ sse é proporcional a $\mathbf{w}$.

ex: Cada $\mathbf{s}_k$, com $1 \leq k \leq m$, é dependente dos vetores $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m$.

ex: Se $\mathbf{v}$ é dependente dos $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m$ mas não é dependente dos $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{m-1}$, então $\mathbf{s}_m$ é dependente dos $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{m-1}, \mathbf{v}$.

ex: Se $\mathbf{v}$ é dependente dos $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m$ e cada $\mathbf{s}_k$, com $1 \leq k \leq m$, é dependente dos $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_p$, então $\mathbf{v}$ é dependente dos $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_p$.

Conjuntos livres/linearmente independentes. O conjunto finito $S = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\} \subset \mathbb{R}^n$ é dito *dependente* se um dos vetores é uma combinação linear dos outros, e.g. $\mathbf{s}_k = \sum_{i \neq k} \lambda_i \mathbf{s}_i$.

O conjunto finito S é *livre/(linearmente) independente* (ou os vetores $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m$ são *livres/independentes*) se não é dependente, ou seja, se nenhum dos $\mathbf{s}_k$ é dependente dos outros, e portanto se S gera o vetor nulo $\mathbf{0}$ duma única maneira, i.e. se

$$\boxed{\lambda_1 \mathbf{s}_1 + \lambda_2 \mathbf{s}_2 + \dots + \lambda_m \mathbf{s}_m = \mathbf{0} \qquad \Rightarrow \qquad \lambda_1 = \lambda_2 = \dots = \lambda_n = 0}$$

(ou seja, se a única soluçã do sistema homogéneo acima, de n equações em m incógnitas, apenas admite a solução trivial), e portanto se S gera cada vetor de $\mathrm{Span}(S)$ duma única maneira, i.e. se cada vetor $\mathbf{v} \in \mathrm{Span}(S)$ admite uma única representação $\mathbf{v} = \lambda_1 \mathbf{s}_1 + \lambda_2 \mathbf{s}_2 + \dots + \lambda_m \mathbf{s}_m$.

ex: Os vetores $\mathbf{v}$ e $\mathbf{w}$ são dependentes sse são paralelos.

ex: Um conjunto que contém o vetor nulo não é independente (assuma que $\mathbf{s}_1 = \mathbf{0}$, e observe que $\mathbf{0} = 1\mathbf{s}_1 + 0\mathbf{s}_2 + \dots + 0\mathbf{s}_m$).

ex: Se $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m$ são independentes mas $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m, \mathbf{v}$ são dependentes, então $\mathbf{v}$ é dependente dos $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m$.

ex: Cada conjunto finito $S = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\} \subset \mathbb{R}^n$ contém um subconjunto (que pode ser vazio) de vetores do qual cada $\mathbf{s}_k$ é dependente.

ex: Se $S = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\} \subset \mathbb{R}^n$ é um conjunto independente de m vetores, então todo o conjunto de $m+1$ vetores do espaço $\mathrm{Span}(S)$ gerado por S é dependente (prova por indução).

ex: Mostre que os vetores (a, b) e (c, d) de $\mathbb{R}^2$ são independentes sse

$$ad - bc \neq 0\,.$$

ex: Verifique se os seguintes conjuntos de vetores são linearmente independentes:

$$(1,2) \qquad (-1,2) \qquad \text{em } \mathbb{R}^2$$

$$(7,-7/3) \qquad (-1,3) \qquad \text{em } \mathbb{R}^2$$

$$(17,4) \qquad (23,-6) \qquad (-8,99) \qquad \text{em } \mathbb{R}^2$$

$$(1,1,0) \qquad (0,1,1) \qquad (1,0,1) \qquad \text{em } \mathbb{R}^3$$

$$(\sqrt{3},1,0) \qquad (1,\sqrt{3},1) \qquad (0,1,\sqrt{3}) \qquad \text{em } \mathbb{R}^3$$

$$(\sqrt{2},1,0) \qquad (1,\sqrt{2},1) \qquad (0,1,\sqrt{2}) \qquad \text{em } \mathbb{R}^3$$

ex: Verifique se

$$(7,-1) \text{ é combinação linear de } (3,8) \text{ e } (-1,0)$$

$$(1,5,3) \text{ é combinação linear de } (0,1,2) \text{ e } (-1,0,5)$$

ex: Determine os valores de t para os quais os vetores

$$(t,1,0) \qquad (1,t,1) \qquad (0,1,t)$$

são dependentes.

ex: Dê 3 vetores independentes de $\mathbb{R}^3$.

ex: [Ap69] **12.15.**

Bases e dimensão. Uma *base* de $\mathbb{R}^n$ é um conjunto livre de geradores de $\mathbb{R}^n$, ou seja, um conjunto que gera $\mathbb{R}^n$ duma única maneira. Uma base de $\mathbb{R}^n$ é portanto um conjunto $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\} \subset \mathbb{R}^n$ tal que cada vetor $\mathbf{x} \in \mathbb{R}^n$ admite uma única representação $\mathbf{x} = \lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2 + \cdots + \lambda_n\mathbf{b}_n$ como combinação linear dos vetores de $\mathcal{B}$. Os coeficientes λ_k são ditos coordenadas/componentes de $\mathbf{x}$ relativamente à base $\mathcal{B}$.

A *base canónica* de $\mathbb{R}^n$ é o conjunto formado pelos vetores $\mathbf{e}_1 = (1,0,\dots,0)$, $\mathbf{e}_2 = (0,1,0,\dots,0)$, ... , $\mathbf{e}_n = (0,\dots,0,1)$. É tautológico que a base canónica é uma base de $\mathbb{R}^n$.

Teorema 4.1. *Toda a base de $\mathbb{R}^n$ é composta de n vetores. Qualquer conjunto de n vetores linearmente independentes é uma base de $\mathbb{R}^n$. Qualquer conjunto linearmente independente é um subconjunto de uma base de $\mathbb{R}^n$.*

A prova é elementar mas comprida, e pode ser consultada em [Ap69] ou [Wa91].

A cardinalidade de uma (e portanto de qualquer) base é chamada *dimensão* do espaço. Em particular, a dimensão do espaço $\mathbb{R}^n$ é n.

ex: Os vetores $(1,1)$ e $(1,-1)$ formam uma base de $\mathbb{R}^2$?

ex: Determine uma base de $\mathbb{R}^2$ contendo o vetor $(2,3)$.

ex: Determine uma base de $\mathbb{R}^3$ contendo os vetores $(0,1,1)$ e $(1,1,1)$.

ex: Verifique se os vetores $(1,0)$ e $(1,1)$ formam uma base de $\mathbb{R}^2$.

ex: Verifique se os vetores $(0,1,2)$, $(-1,0,3)$ e $(1,1,-1)$ formam uma base de $\mathbb{R}^3$.

ex: Calcule as coordenadas do vetor $(1,1)$ relativamente a base formada pelos vetores $(1,2)$ e $(-1,2)$.

ex: Verifique se os vetores $(1,0,0,0)$, $(1,1,0,0)$, $(1,1,1,0)$ e $(1,1,1,1)$ formam uma base de $\mathbb{R}^4$.

ex: [Ap69] **12.15**.

Conjuntos e bases ortogonais e ortonormados. O conjunto de vetores não nulos $S = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\} \subset \mathbb{R}^n$ é *ortogonal* se $\mathbf{s}_i \cdot \mathbf{s}_j = 0$ quando $i \neq j$.

Teorema 4.2. *Um sistema ortogonal de vetores não nulos é linearmente independente.*

Demonstração. Seja $\lambda_1 \mathbf{s}_1 + \lambda_2 \mathbf{s}_2 + \dots + \lambda_m \mathbf{s}_m = \mathbf{0}$. Se calculamos o produto escalar com os vetores $\mathbf{s}_k$'s do sistema ortogonal, temos

$$\begin{aligned} 0 &= \mathbf{s}_k \cdot (\lambda_1 \mathbf{s}_1 + \lambda_2 \mathbf{s}_2 + \dots + \lambda_m \mathbf{s}_m) \\ &= \lambda_k \, \|\mathbf{s}_k\|^2 \end{aligned}$$

(pois todos os outros produtos são nulos pela ortogonalidade). Sendo os vetores $\mathbf{s}_k$ não nulos, isto implica que todos os coeficientes λ_k são nulos. □

Se o vetor $\mathbf{x} \in \mathbb{R}^n$ é gerado pelo sistema ortogonal de vetores não nulos $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\}$, ou seja, $\mathbf{x} = \sum_{k=1}^m x_k \mathbf{s}_k$, então os coeficientes x_k são as projeções de $\mathbf{x}$ sobre $\mathbf{s}_k$, i.e.

$$\boxed{x_k = \frac{\mathbf{x} \cdot \mathbf{s}_k}{\|\mathbf{s}_k\|^2}}$$

Um conjunto ortogonal composto de vetores unitários (i.e. tais que $\|\mathbf{s}_i\| = 1 \ \forall i = 1, 2, \dots, m$) é dito *ortonormado*. Em particular, se $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\}$ é uma base ortonormada, então cada $\mathbf{x} \in \mathbb{R}^n$ é

$$\boxed{\mathbf{x} = \sum_{k=1}^n x_k \mathbf{s}_k = \sum_{k=1}^n (\mathbf{x} \cdot \mathbf{s}_k) \, \mathbf{s}_k}$$

e.g. **Base canónica.** A base canónica de $\mathbb{R}^n$ é ortonormada.

ex: Verifique se o conjunto formado pelos vetores $(0, \sqrt{3}/2, 1/2)$, $(0, -1/2, \sqrt{3}/2)$ e $(1,0,0)$ é ortogonal e ortonormado.

ex: Determine uma base ortogonal de $\mathbb{R}^2$ contendo o vetor $(1,1)$.

ex: Determine uma base ortonormada de $\mathbb{R}^2$ contendo o vetor $(1/\sqrt{2}, 1/\sqrt{2})$.

ex: [Ap69] **12.15**.

5 Produto vetorial, área e volume

ref: [Ap69] *Vol 1*, **13.9-17**

19 out 2018

Independência no plano e determinante. Os vetores $\mathbf{r} = (a, c)$ e $\mathbf{r}' = (b, d)$ são independentes sse

$$x \begin{pmatrix} a \\ c \end{pmatrix} + y \begin{pmatrix} b \\ d \end{pmatrix} = \mathbf{0}$$

implica $x = 0$ e $y = 0$, ou seja, sse a única solução do "sistema linear homogéneo"

$$\begin{cases} ax + by & = & 0 \\ cx + dy & = & 0 \end{cases}$$

é a solução trivial. Ao retirar b vezes a segunda equação de d vezes a primeira equação, e depois ao retirar c vezes a primeira equação de a vezes a segunda equação, temos que

$$\begin{cases} ax + by & = & 0 \\ cx + dy & = & 0 \end{cases} \qquad \Rightarrow \qquad \begin{cases} (ad - bc)\, x & = & 0 \\ (ad - bc)\, y & = & 0 \end{cases}$$

O *determinante* da matriz 2×2 cujas colunas são as componentes dos vetores (a, c) e (b, d) (ou cujas linhas são as componentes dos vetores (a, b) e (c, d)) de $\mathbb{R}^2$ é

$$\boxed{\operatorname{Det} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} := ad - bc}$$

(observe que trocar linhas com colunas não altera o determinante!). Portanto, os vetores (a, c) e (b, d) são independentes sse $\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| \neq 0$.

ex: Diga se os vetores $(-1, 4)$ e $(3, -12)$ são independentes.

ex: Diga se os vetores $(5, 7)$ e $(2, 9)$ são independentes.

Determinante e área. O *paralelogramo* gerado/definido pelos vetores $\mathbf{x} = (a, b)$ e $\mathbf{y} = (c, d)$ de $\mathbb{R}^2$ é o conjunto $P = \{t\mathbf{x} + s\mathbf{y} \quad \text{com } 0 \leq t, s \leq 1\} \subset \mathbb{R}^2$. A sua área é igual ao módulo do determinante da matriz $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, i.e.

$$\boxed{\text{Área}\,(P) = \left|\operatorname{Det} \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right| = |ad - bc|}$$

ex: Prove a fórmula acima para a área do paralelogramo (retire da área do retângulo de base $a + c$ e altura $b + d$ as áreas dos retângulos e triângulos que sobram ...).

ex: Calcule a área do paralelogramo definido pelos vetores $(0, 1)$ e $(1, 1)$, e do paralelogramo definido pelos vetores $(5, -2)$ e $(-3, 1)$.

ex: Calcule a área do triângulo de vértices $(3, 2)$, $(6, -4)$ e $(8, 8)$.

Produto vetorial. O *produto vetorial* (ou *externo*, em inglês *cross product*) no espaço $\mathbb{R}^3$ (munido da orientação definida pela base canónica $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$) é a operação $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$\boxed{\mathbf{r}, \mathbf{r}' \mapsto \mathbf{r} \times \mathbf{r}' := (yz' - zy'\,,\ -xz' + zx'\,,\ xy' - yx')}$$

onde $\mathbf{r} = (x, y, z)$ e $\mathbf{r}' = (x', y', z')$. O produto vetorial é distributivo sobre a adição e compatível com a multiplicação escalar, ou seja, é bilinear, i.e.

$$(\lambda\mathbf{r} + \mu\mathbf{r}') \times \mathbf{r}'' = \lambda(\mathbf{r} \times \mathbf{r}'') + \mu(\mathbf{r}' \times \mathbf{r}'')$$

$$\mathbf{r} \times (\lambda \mathbf{r}' + \mu \mathbf{r}'') = \lambda (\mathbf{r} \times \mathbf{r}') + \mu (\mathbf{r} \times \mathbf{r}'')$$

e "anti-comutativo", i.e.

$$\mathbf{r} \times \mathbf{r}' = -\mathbf{r}' \times \mathbf{r}$$

O produto vetorial satisfaz a *identidade de Jacobi*

$$\boxed{\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}} \tag{5.1}$$

e a *identidade de Lagrange*

$$\boxed{\|\mathbf{r} \times \mathbf{r}'\|^2 = \|\mathbf{r}\|^2 \, \|\mathbf{r}'\|^2 - (\mathbf{r} \cdot \mathbf{r}')^2} \tag{5.2}$$

(as demosntrações são cálculos elementares).

ex: Mostre que o produto vetorial é bilinear e anti-comutativo.

ex: $\mathbf{r} \times \mathbf{r}' = \mathbf{0}$ sse $\mathbf{r}$ e $\mathbf{r}'$ são dependentes, i.e. proporcionais (use a identidade de Lagrange e a desigualdade de Schwarz).

ex: O vetor $\mathbf{r} \times \mathbf{r}'$ é ortogonal ao subespaço vetorial $\mathbb{R}\mathbf{r} + \mathbb{R}\mathbf{r}'$ gerado por $\mathbf{r}$ e $\mathbf{r}'$ (calcule $\mathbf{r} \cdot (\mathbf{r} \times \mathbf{r}')$ e $\mathbf{r}' \cdot (\mathbf{r} \times \mathbf{r}')$ ...).

ex: Calcule os produtos vetoriais entre os vetores da base canónica de $\mathbb{R}^3$, $\mathbf{i} = (1,0,0)$, $\mathbf{j} = (0,1,0)$ e $\mathbf{k} = (0,0,1)$, e verifique que

$$\mathbf{i} \times \mathbf{j} = \mathbf{k} \qquad \mathbf{j} \times \mathbf{k} = \mathbf{i} \qquad \text{e} \qquad \mathbf{k} \times \mathbf{i} = \mathbf{j}$$

ex: O produto vetorial não é associativo! Por exemplo, $\mathbf{i} \times (\mathbf{i} \times \mathbf{j}) \neq (\mathbf{i} \times \mathbf{i}) \times \mathbf{j}$.

ex: Verifique a *fórmula de Lagrange*

$$\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a}\,(\mathbf{c} \cdot \mathbf{b}) - \mathbf{b}\,(\mathbf{c} \cdot \mathbf{a})$$

ex: Calcule $\mathbf{r} \times \mathbf{r}'$ quando

$$\mathbf{r} = (1,-1,1) \qquad \text{e} \qquad \mathbf{r}' = (2,-2,2)$$

$$\mathbf{r} = (-2,-1,3) \qquad \text{e} \qquad \mathbf{r}' = (\pi,-\pi,0)$$

Produto vetorial e determinante. Uma representação formal do produto vetorial entre os vetores $\mathbf{r} = (x,y,z)$ e $\mathbf{r}' = (x',y',z')$ é

$$\boxed{\mathbf{r} \times \mathbf{r}' = \text{“Det} \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & z \\ x' & y' & z' \end{pmatrix} \text{”} := \mathbf{i} \begin{vmatrix} y & z \\ y' & z' \end{vmatrix} - \mathbf{j} \begin{vmatrix} x & z \\ x' & z' \end{vmatrix} + \mathbf{k} \begin{vmatrix} x & y \\ x' & y' \end{vmatrix}}$$

ex: Calcule $\mathbf{r} \times \mathbf{r}'$ quando

$$\mathbf{r} = (3,-2,8) \qquad \text{e} \qquad \mathbf{r}' = (1,1,1)$$

$$\mathbf{r} = (-\pi,e,10) \qquad \text{e} \qquad \mathbf{r}' = (7,5,3)$$

ex: [Ap69] **13.11**.

Produto vetorial e área. Usando a identidade de Lagrange é imediato ver que que norma do produto vetorial $\mathbf{r} \times \mathbf{r}'$ é

$$\boxed{\|\mathbf{r} \times \mathbf{r}'\| = \|\mathbf{r}\| \, \|\mathbf{r}'\| \, |\sin\theta|}$$

onde θ é o ângulo entre os vetores $\mathbf{r}$ e $\mathbf{r}'$. Portanto, o comprimento do produto vetorial $\mathbf{r} \times \mathbf{r}'$ é a área do paralelogramo $\{t\mathbf{r} + s\mathbf{r}' \quad \text{com } 0 \leq t,s \leq 1\} \subset \mathbb{R}^3$ definido pelos vetores $\mathbf{r}$ e $\mathbf{r}'$.

ex: Prove a fórmula acima (use a identidade de Lagrange e a definição de θ).

ex: $\|\mathbf{r} \times \mathbf{r}'\| = \|\mathbf{r}\|\|\mathbf{r}'\|$ sse $\mathbf{r}$ e $\mathbf{r}'$ são ortogonais.

ex: Calcule a área do paralelogramo de lados $\vec{OP}$ e $\vec{OQ}$, onde $P = (2, 4, -1)$ e $Q = (1, -3, 1)$.

ex: Calcule a área do triângulo de vértices $(1, 2, 0)$, $(2, 3, 4)$ e $(-1, 0, 0)$.

Produto vetorial e vetor normal. Se $\mathbf{u}$ e $\mathbf{v}$ são dois vetores linearmente independentes de $\mathbb{R}^3$, então $\mathbf{n} = \mathbf{u} \times \mathbf{v}$ é um vetor não nulo ortogonal ao plano $\mathbb{R}\mathbf{u} + \mathbb{R}\mathbf{v}$, e o conjunto $\{\mathbf{u}, \mathbf{v}, \mathbf{n}\}$ é uma base de $\mathbb{R}^3$. Em particular, se $\mathbf{u}$ e $\mathbf{v}$ são vetores unitários e ortogonais, então também $\mathbf{n} = \mathbf{u} \times \mathbf{v}$ é unitário, e o conjunto $\{\mathbf{u}, \mathbf{v}, \mathbf{n}\}$ é uma base ortonormada de $\mathbb{R}^3$. Neste caso valem as relações

$$(\mathbf{u} \times \mathbf{v}) \times \mathbf{u} = \mathbf{v} \qquad \text{e} \qquad \mathbf{v} \times (\mathbf{u} \times \mathbf{v}) = \mathbf{u}$$

(que podem ser provadas usando a fórmula de Lagrange).

ex: O plano gerado pelos vetores linearmente independentes $\mathbf{u}$ e $\mathbf{v}$ no espaço $\mathbb{R}^3$ é

$$\mathbb{R}\mathbf{u} + \mathbb{R}\mathbf{v} = (\mathbf{u} \times \mathbf{v})^{\perp} = \{\mathbf{r} \in \mathbb{R}^3 \text{ t.q. } \mathbf{r} \cdot (\mathbf{u} \times \mathbf{v}) = 0\}$$

ex: Determine um vetor normal aos vetores $(2, 3, -1)$ e $(5, 2, 4)$.

ex: Determine uma base de $\mathbb{R}^3$ contendo os vetores $(0, 1, 1)$ e $(1, 1, 1)$.

ex: Se $\mathbf{a}$ e $\mathbf{b}$ são dois vetores ortogonais e unitários de $\mathbb{R}^3$, então $\{\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}\}$ é uma base ortonormada de $\mathbb{R}^3$.

ex: Determine um vetor normal ao plano que passa pelos pontos $(0, 1, 0)$, $(1, 1, 0)$ e $(0, 2, 3)$

ex: Determine uma equação cartesiana do plano

$$\text{gerado pelos vetores } (-3, 1, 2) \text{ e } (1, 5, -2)$$

$$\text{que passa pelos pontos } (0, 0, 0)\,,\ (1, 0, 0) \text{ e } (0, 1, 0)$$

$$\{(1 + t + s, t - s, 5t) \quad \text{com} \ \ (t, s) \in \mathbb{R}^2\}$$

ex: [Ap69] **13.11**.

Produto misto/triplo escalar e determinante. O *produto misto* dos vetores $\mathbf{r}$, $\mathbf{r}'$ e $\mathbf{r}''$ de $\mathbb{R}^3$ é o escalar $\mathbf{r} \cdot (\mathbf{r}' \times \mathbf{r}'')$. É também igual ao *determinante* da matriz 3×3 cujas linhas são as componentes dos três vetores, i.e.

$$\boxed{\mathbf{r} \cdot (\mathbf{r}' \times \mathbf{r}'') = \mathrm{Det} \begin{pmatrix} x & y & z \\ x' & y' & z' \\ x'' & y'' & z'' \end{pmatrix} := x \begin{vmatrix} y' & z' \\ y'' & z'' \end{vmatrix} - y \begin{vmatrix} x' & z' \\ x'' & z'' \end{vmatrix} + z \begin{vmatrix} x' & y' \\ x'' & z'' \end{vmatrix}}$$

Os vetores $\mathbf{a}$, $\mathbf{b}$ e $\mathbf{c}$ são independentes (e portanto formam uma base de $\mathbb{R}^3$) sse $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) \neq 0$.

ex: Calcule o produto misto $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ quando

$$\mathbf{a} = (1, 1, 0) \qquad \mathbf{b} = (1, 3, 1) \qquad \mathbf{c} = (0, 1, 1)$$

$$\mathbf{a} = (-2, 5, 1) \qquad \mathbf{b} = (0, 3, 0) \qquad \mathbf{c} = (6, 7, -3)$$

ex: Diga se são independentes

$$(7,2,3) \qquad (-1,-5,3) \qquad \text{e} \qquad (0,1,-3)$$
$$(1,0,1) \qquad (0,1,0) \qquad \text{e} \qquad (1,1,1)$$

ex: Determine os valores de t para os quais os vetores

$$(t,1,0) \qquad (1,t,1) \qquad (0,1,t)$$

são dependentes.

ex: [Ap69] **13.14**.

Determinantes e volumes no espaço. O *paralelepípedo* gerado/definido pelos vetores $\mathbf{a}$, $\mathbf{b}$ e $\mathbf{c}$ de $\mathbb{R}^3$ é o conjunto $\{t\mathbf{a}+s\mathbf{b}+u\mathbf{c} \text{ com } 0 \leq t,s,u \leq 1\}$. O seu volume é igual ao módulo do produto misto $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$, i.e.

$$\boxed{\text{Volume}\left(\{t\mathbf{a}+s\mathbf{b}+u\mathbf{c} \text{ com } 0 \leq t,s,u \leq 1\}\right) = |\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})| = \left|\text{Det}\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}\right|}$$

ex: Verifique que
$$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \|\mathbf{a}\|\|\mathbf{b}\|\|\mathbf{c}\|\sin(\theta)\cos(\phi)$$
onde θ é o ângulo entre $\mathbf{b}$ e $\mathbf{c}$, e ϕ é o ângulo entre $\mathbf{a}$ e $\mathbf{b}\times\mathbf{c}$.

ex: Calcule o volume do paralelepípedo definido pelos vetores $\mathbf{i}+\mathbf{j}$, $\mathbf{j}+\mathbf{k}$ e $\mathbf{k}+\mathbf{i}$.

ex: Calcule o volume do paralelepípedo gerado pelos vetores

$$(3,3,1) \qquad (2,1,2) \qquad (5,1,1)$$
$$(0,0,1) \qquad (5,7,-3) \qquad (-9,0,0)$$

Regra de Cramer. Resolver o sistema de três equações lineares

$$\begin{cases} a_1x+b_1y+c_1z &= d_1 \\ a_2x+b_2y+c_2z &= d_2 \\ a_3x+b_3y+c_3z &= d_3 \end{cases} \qquad \text{ou seja} \qquad x\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + y\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} + z\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

nas três incógnitas x, y e z, significa representar o vetor $D=(d_1,d_2,d_3)$ como combinação linear

$$xA+yB+zC=D$$

dos vetores $A=(a_1,a_2,a_3)$, $B=(b_1,b_2,b_3)$ e $C=(c_1,c_2,c_3)$ com coeficientes x, y e z. Ao multiplicar esta representação por $B\times C$, ou por $C\times A$, ou por $A\times B$, resulta que

$$\begin{aligned} x\,A\cdot(B\times C) &= D\cdot(B\times C)\,, \\ y\,B\cdot(C\times A) &= D\cdot(C\times A)\,, \\ z\,C\cdot(A\times B) &= D\cdot(A\times B)\,. \end{aligned}$$

Portanto, se o produto misto $A\cdot(B\times D)\neq 0$ (i.e. se o determinante da matriz 3×3 cujas colunas são os vetores A, B e C é diferente de zero, ou seja, se os três vetores são independentes), então o sistema admite una única solução (x,y,z), dada pela *regra de Cramer*:

$$x = \frac{D\cdot(B\times C)}{A\cdot(B\times C)}\,, \qquad y = \frac{C\cdot(D\times A)}{A\cdot(B\times C)}\,, \qquad z = \frac{A\cdot(B\times D)}{A\cdot(B\times C)}$$

Observe que o denominador é o determinante da matriz 3×3 cujas colunas são os vetores A, B e C, e o numerador da i-ésima coordenada é obtido ao substituir, nesta matriz, a i-ésima coluna com o vetor D.

ex: Resolva os sistemas

$$\left\{\begin{array}{lcl} x+y+z & = & 6 \\ 2x+y+3z & = & 9 \\ 3x-y-5z & = & -17 \end{array}\right. \qquad \left\{\begin{array}{lcl} x+2y-3z & = & 4 \\ 2x-3y+z & = & 1 \\ 3x-y-2z & = & 5 \end{array}\right.$$

Força magnética. A *força de Lorentz* que experimenta uma partícula com carga eléctrica q e velocidade $\mathbf{v}$ num campo elétrico $\mathbf{E}$ e magnético $\mathbf{B}$ é (nas unidades do S.I.)

$$\mathbf{F} = q\,(\mathbf{E} + \mathbf{v}\times\mathbf{B})$$

ex: Mostre que num referencial inercial em que o campo elétrico é nulo, i.e. $\mathbf{E} = \mathbf{0}$, e portanto a única força é força magnética $q\mathbf{v}\times\mathbf{B}$, a energia cinética $K := \frac{1}{2}m\|\mathbf{v}\|^2$ é conservada, calculando a derivada

$$\frac{d}{dt}\left(\tfrac{1}{2}m\|\mathbf{v}\|^2\right) = m\dot{\mathbf{v}}\cdot\mathbf{v}$$

e utilizando a equação de Newton $m\dot{\mathbf{v}} = \mathbf{F}$.

Momento angular e torque. O *momento angular* (relativo à origem do referencial) de uma partícula de massa $m > 0$ colocada na posição $\mathbf{r}\in\mathbb{R}^3$ com momento linear $\mathbf{p} = m\dot{\mathbf{r}}$, é o produto vetorial

$$\mathbf{L} := \mathbf{r}\times\mathbf{p}$$

A derivada do momento angular de uma partícula sujeita à lei de Newton $\dot{\mathbf{p}} = \mathbf{F}$ é igual ao *binário* (ou *torque*) $\mathbf{T} := \mathbf{r}\times\mathbf{F}$, pois

$$\begin{aligned}\dot{\mathbf{L}} &= \dot{\mathbf{r}}\times\mathbf{p} + \mathbf{r}\times\dot{\mathbf{p}} \\ &= \mathbf{r}\times\mathbf{F}\,.\end{aligned}$$

sendo $\dot{\mathbf{r}}\times\mathbf{p} = \mathbf{0}$.

ex: O momento angular do sistema de n partículas de massas m_i, colocadas nas posições $\mathbf{r}_i\in\mathbb{R}^3$ com momentos lineares $\mathbf{p}_i := m_i\dot{\mathbf{r}}_i$, com $i = 1, 2, \dots, n$, é

$$\mathbf{L} := \sum_{i=1}^{n}\mathbf{r}_i\times\mathbf{p}_i$$

Sejam $\mathbf{R} := \frac{1}{M}\sum_{i=1}^n m_i\mathbf{r}_i$, com $M := \sum_{i=1}^n m_i$, o centro de massa do sistema, e $\mathbf{P} := M\dot{\mathbf{R}} = M\sum_{i=1}^n m_i\dot{\mathbf{r}}_i$ o momento linear do centro de massa. Mostre que

$$\mathbf{L} = \mathbf{R}\times\mathbf{P} \,+\, \mathbf{L}'$$

onde $\mathbf{L}' := \sum_{i=1}^n \mathbf{r}'_i\times\mathbf{p}'_i$, com $\mathbf{r} = \mathbf{R} + \mathbf{r}'$, é o momento angular relativo ao centro de massa.

Angular momentum of Kepler orbits and Hamilton's theorem. If two point masses with positions $\mathbf{r}_1$ and $\mathbf{r}_2$ move around their center of masses according to Newton's gravitational law, then the difference vector $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ obeys the law

$$\ddot{\mathbf{r}} = -\frac{\mathbf{r}}{\|\mathbf{r}\|^3} \tag{5.3}$$

(if we set to one both the reduced mass of the system and the universal gravitational constant). If $\mathbf{v} = \dot{\mathbf{r}}$ denotes the velocity vector, then a computation shows that the angular momentum

$$\mathbf{L} := \mathbf{r}\times\mathbf{v}$$

is a constant of the motion. If at some (initial) time the vectors $\mathbf{r}$ and $\mathbf{v}$ are not parallel, then $\mathbf{L}\neq\mathbf{0}$. We may therefore choose a reference Cartesian system in which $\mathbf{L} = \ell\,\mathbf{k}$ for some $\ell > 0$, and write the position vector as $\mathbf{r}(t) = r\cos(\theta)\,\mathbf{i} + r\sin(\theta)\,\mathbf{j}$ for some time-dependent angle θ and lenght $r = \|\mathbf{r}\|$. A computation shows that

$$\ell = r^2\dot{\theta}\,, \tag{5.4}$$

expressing *Kepler second law*[6], according to which "the position vector from the sun to a planet sweeps out equal areas in equal times". If we eliminate dt from Newton equation (5.3) (written for $\dot{\mathbf{v}}$) and equation (5.4), we get

$$d\mathbf{v}/d\theta = -\ell^{-1}(\cos(\theta)\,\mathbf{i} + \sin(\theta)\,\mathbf{j})$$

and, after integration,

$$\mathbf{v} = \mathbf{v}_0 \,+\, \ell^{-1}(\sin(\theta)\,\mathbf{i} - \,\cos(\theta)\,\mathbf{j})\,,$$

for some constant vector $\mathbf{v}_0$ (on the $z = 0$ plane). Thus, "the velocity vector moves along a circle orthogonal to the angular velocity vector with radius inversely proportional to its length" [7] [8].

[6] J. Kepler, *Astronomia Nova*, Prague, 1609.

[7] W.R. Hamilton, The hodograph or a new method of expressing in symbolic language the Newtonian law of attraction, *Proc. Roy. Irish Acad.* **3** (1846), 344-353.

[8] J. Milnor, On the geometry of the Kepler problem. *Amer. Math. Monthly* **90** (1983), 353-365.

6 Números complexos*

ref: [Ap69] *Vol 1*, **9.1-5**

Very short history. Complex numbers were invented/discovered in the XVI century as a "sophistic" device/trick to solve cubic equations like $x^3 + px + q = 0$. Today, they enter in our formulation of most fundamental laws of Nature, as for example in the *Schrödinger equation*

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi + V\psi$$

or in *Feynman path integrals* of quantum field theory

$$\int_{\text{paths}} e^{i\mathcal{S}[x]/\hbar}\,\mathcal{D}x\,.$$

O corpo dos números complexos. O corpo dos *números complexos* é $\mathbb{C} := \mathbb{R}(i)$, onde $i^2 = -1$. Ou seja, é o conjunto $\mathbb{C} \approx \mathbb{R}^2$ dos números/pontos $z = x + iy \approx (x, y)$, com $x, y \in \mathbb{R}$, munido das operações binárias "soma", definida por

$$(x_1 + iy_1) + (x_2 + iy_2) := (x_1 + x_2) + i\,(y_1 + y_2) \tag{6.1}$$

(que corresponde à soma dos vetores $z_1 \approx (x_1, y_1)$ e $z_2 \approx (x_2, y_2)$ do plano $\mathbb{R}^2$), e "multiplicação", definida por

$$(x_1 + iy_1)\cdot(x_2 + iy_2) := (x_1x_2 - y_1y_2) + i\,(x_1y_2 + x_2y_1)\,. \tag{6.2}$$

Em particular, se $i := 0 + i\cdot 1 \in \mathbb{C}$, então $i\cdot i = -1$, ou seja, $\pm i$ são (as únicas) "raízes quadradas de -1". De fato (e esta é a origem das fórmulas acima, que não devem ser decoradas!), somas e multiplicações de números complexos podem ser manipuladas como as correspondentes operações entre números reais (ou seja, usando as propriedades associativas, comutativas e distributivas), e depois substituindo $i\cdot i$ por -1.

O conjunto $\mathbb{C}$, munido da operação $+$ definida em (6.1), é um grupo (abeliano) aditivo, cujo elemento neutro é $0 := 0 + i0$. Somar um número complexo z, i.e. fazer $w \mapsto w + z$, corresponde a uma translação no plano complexo $\mathbb{C} \approx \mathbb{R}^2$.

Todo $z = x + iy \neq 0$ admite um inverso multiplicativo, um número complexo $1/z$ tal que $z\cdot(1/z) = (1/z)\cdot z = 1$, dado por

$$\frac{1}{z} = \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2}\,,$$

como é fácil verificar (observe que se $z \neq 0$ então $x^2 + y^2 > 0$). De consequência, o conjunto $\mathbb{C}^\times := \mathbb{C}\backslash\{0\}$, munido da operação $\cdot$ definida em (6.2), é um grupo (abeliano) multiplicativo, o grupo multiplicativo dos números complexos invertíveis, cujo elemento neutro é $1 := 1 + i0$ (o significado geométrico da multiplicação deve esperar mais um pouco).

Potências. As potências inteiras de um número complexo são definidas por recorrência: $z^{n+1} := z\cdot z^n$, se $n \geq 1$, sendo $z^0 := 1$. Se $z \in \mathbb{C}^\times$, então as potências negativas são definidas por $z^{-n} := (1/z)^n$.

Conjugação. O *conjugado* de $z = x + iy$ é

$$\boxed{\overline{z} := x - iy\,,}$$

ou seja, a imagem do ponto $x + iy \approx (x, y)$ pela reflexão na reta $y = 0$ do plano $\mathbb{R}^2 \approx \mathbb{C}$. A conjugação respeita soma e produtos, ou seja, verifica

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} \qquad \text{e} \qquad \overline{z_1\cdot z_2} = \overline{z_1}\cdot\overline{z_2}$$

(a segunda identidade não é óbvia, mas um milagre que relaciona multiplicação e geometria do plano). Observe também que a conjugação é uma involução, ou seja, $\overline{\overline{z}} = z$.

Os números reais

$$x = \Re(z) := \frac{z+\overline{z}}{2} \qquad \text{e} \qquad y = \Im(z) := \frac{z-\overline{z}}{2i}$$

são ditos *parte real* e *parte imaginária* do número complexo $z = x+iy$, respetivamente. Observe que $z = \overline{z}$ sse z é real, i.e. sse $\Im(z) = 0$.

Módulo. A conjugação permite definir $N(z) := z\overline{z} = x^2+y^2$, que é um número real não-negativo (o “módulo” de z no sentido da teoria de números), e portanto o *módulo*, ou *valor absoluto*, de $z = x+iy$,

$$\boxed{|z| := \sqrt{z\overline{z}} = \sqrt{x^2+y^2}}$$

que é a norma euclidiana do vetor $(x,y) \in \mathbb{R}^2$. Em particular, $|z| = 0$ sse $z = 0$. O valor absoluto é multiplicativo, ou seja, $|zw| = |z|\,|w|$, e $|z/w| = |z|/|w|$ se $w \neq 0$. O inverso multiplicativo de um número complexo $z \neq 0$ é então

$$1/z = \overline{z}/|z|^2 .$$

ex: Calcule

$$(2+i3)+(3-i2) \qquad (1-i)\cdot(2-i) \qquad (1+i)+(1-i)\cdot(2-i5)$$

ex: Represente na forma $x+iy$ os seguintes números complexos

$$1/i \qquad \frac{2-i}{1+i} \qquad \frac{1-i}{1+i}\cdot\frac{i}{2+i} \qquad (1-i3)^2 \qquad i^{17} \qquad (2\pm i)^3$$

ex: Interprete, e prove, a seguinte identidade entre números reais:

$$(a^2+b^2)(c^2+d^2) = (ac-bd)^2+(ad+bc)^2 .$$

ex: Resolva as seguintes equações

$$z^2-2z+2=0 \qquad z^2+z+1=0$$

Representação polar. A *representação polar* do número complexo $z = x+iy \approx (x,y) \in \mathbb{R}^2$ é

$$\boxed{z = \rho\, e^{i\theta}}$$

onde $\rho = |z| = \sqrt{x^2+y^2} \geq 0$ é o módulo de z, $\theta = \arg(z) \in \mathbb{R}/2\pi\mathbb{Z}$ é um *argumento* de z, ou seja, um ângulo tal que $x = \rho\cos(\theta)$ e $y = \rho\sin(\theta)$ (logo definido a menos de múltiplos inteiros de 2π), e o número complexo $e^{i\theta} \in \mathbb{S} := \{|z| = 1\}$ é (provisoriamente) definido pela *fórmula de Euler*

$$\boxed{e^{i\theta} := \cos(\theta) + i\sin(\theta) .} \tag{6.3}$$

Pode ser útil escolher um valor do argumento, e chamar *argumento principal* de um número z o único argumento que satisfaz $\mathrm{Arg}(z) \in (-\pi,\pi]$.

Interpretação geométrica da multiplicação. Se $z_1 = \rho_1 e^{i\theta_1}$ e $z_2 = \rho_2 e^{i\theta_2}$, então as fórmulas de adição para seno e coseno mostram que

$$\boxed{z_1 z_2 = \rho_1 \rho_2 e^{i(\theta_1+\theta_2)}}$$

e, se $z_2 \neq 0$,

$$\boxed{\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2} e^{i(\theta_1-\theta_2)} .}$$

Estas fórmulas revelam o significado geométrico da multiplicação entre números complexos. Uma primeira consequência é que o inverso do número complexo $z = \rho e^{i\theta}$, com $\rho > 0$, é $z^{-1} = \rho^{-1} e^{-i\theta}$. Outra é que a multiplicação por $z = \rho e^{i\theta} \neq 0$, no plano $\mathbb{C} \approx \mathbb{R}^2$, ou seja, a transformação $w \mapsto zw$, corresponde a uma homotetia $w \mapsto \rho w$ de razão $|z| = \rho > 0$ (uma dilatação ou contração se $\rho \neq 1$) e uma rotação $w \mapsto e^{i\theta} w$ de um ângulo θ. Em particular, a multiplicação por um número complexo de módulo um, ou seja, da forma $e^{i\theta}$, corresponde a uma rotação. Por exemplo, a multiplicação por $i = e^{i\pi/2}$ é uma "raiz quadrada" da rotação $z \mapsto e^{i\pi} z = -z$ de um ângulo π, i.e. uma rotação de um ângulo $\pi/2$.

Raízes. Se $n = 1, 2, 3, \ldots$, então cada número complexo $w \neq 0$ possui n *raízes n-ésimas*, i.e. n números complexos z que resolvem

$$z^n = w .$$

De fato, as raízes n-ésimas de $w = \rho e^{i\theta}$, com $\rho \neq 0$, são os números

$$z_k = \sqrt[n]{\rho}\, e^{i(\theta+2\pi k)/n}$$

com $k = 0, 1, 2, \ldots, n-1$. Os pontos z_k formam os vértices de um polígono regular de n lados, inscrito na circunferência de raio $\sqrt[n]{\rho}$ e centro 0. Em particular, os números complexos

$$\zeta_k := e^{i2\pi k/n} ,$$

com $k = 0, 1, 2, \ldots, n-1$, que resolvem $(\zeta_k)^n = 1$ e portanto pertencem a circunferência unitária, são chamados *raízes n-ésimas da unidade*. Observe que $\zeta_k = (\zeta_1)^k$, onde $\zeta_1 = e^{i2\pi/n}$ é uma raiz "primitiva".

ex: Represente na forma polar os seguintes números complexos:

$$-i \qquad i-1 \qquad 1+i$$

ex: Use a representação polar e a fórmula de Euler (6.3) para provar a *fórmula de de Moivre*

$$(\cos(\theta) + i\sin(\theta))^n = \cos(n\theta) + i\sin(n\theta) .$$

Deduza as fórmulas

$$\cos(n\theta) = \ldots \qquad \text{e} \qquad \sin(n\theta) = \ldots$$

ex: Verifique que o conjugado de $z = \rho e^{i\theta}$ é $\overline{z} = \rho e^{-i\theta}$.

ex: Calcule

$$e^{i\pi} \qquad e^{-i\pi/2} \qquad \sqrt{i} \qquad \sqrt{-i} \qquad \sqrt{1+i} \qquad \sqrt[4]{i}$$

ex: Resolva as equações $z^3 = 1$, $z^5 = 1$ e $z^3 = 81$.

ex: Verifique que $(1+z+z^2+\cdots+z^n)(1-z) = 1-z^{n+1}$, e portanto, se $z \neq 1$,

$$1+z+z^2+\cdots+z^n = \frac{1-z^{n+1}}{1-z}$$

Considere $z = e^{i\theta}$ com $\theta \neq 2\pi\mathbb{Z}$ e real, calcule a parte real e deduza

$$1+\cos(\theta)+\cos(2\theta)+\cdots+\cos(n\theta) = \frac{1}{2}+\frac{\sin((n+1/2)\theta)}{2\,\sin(\theta/2)}$$

Mostre que se ω é uma raiz n-ésima não trivial da unidade (ou seja, $\omega^n = 1$ e $\omega \neq 1$) então

$$1+\omega+\omega^2+\omega^3+\ldots+\omega^{n-1} = 0\,.$$

Desigualdades. É evidente que a parte real e a parte imaginária de um número complexo são limitadas pelo módulo, i.e.

$$-|z| \leq \Re(z) \leq |z| \qquad \text{e} \qquad -|z| \leq \Im(z) \leq |z| \tag{6.4}$$

Por outro lado, um cálculo direto mostra que $|z \pm w|^2 = |z|^2+|w|^2 \pm 2\,\Re(z\overline{w})$. Usando as (6.4), deduzimos a *desigualdade do triângulo*

$$\boxed{|z+w| \leq |z|+|w|} \tag{6.5}$$

Norma e métrica euclidiana. A desigualdade do triângulo diz que $|z|$ é uma norma, e portanto

$$d(z,w) := |z-w|$$

é uma métrica no plano complexo. Ou seja, é positiva quando $z \neq w$, nula sse $z = w$, e satisfaz a desigualdade do triângulo

$$\mathrm{d}(z,w) \leq \mathrm{d}(z,p)+\mathrm{d}(p,w)\,.$$

De fato, como já observado, é a métrica euclidiana de $\mathbb{C} \approx \mathbb{R}^2$, definida pelo produto escalar euclidiano $\langle (x,y),(x',y')\rangle = xx'+yy' = \Re(z\overline{z'})$.

ex: Diga quando valem as igualdades nas (6.4) e (6.5).

ex: Mostre que

$$|z+w|^2+|z-w|^2 = 2\left(|z|^2+|w|^2\right)$$

ex: Prove a desigualdade

$$|z-w| \geq ||z|-|w||$$

ex: Seja $\|z\|_\infty := \max\{|\Re(z)|, |\Im(z)\}$. Mostre que

$$\|z\|_\infty \leq |z| \leq \sqrt{2}\,\|z\|_\infty \tag{6.6}$$

Discos e bolas. O *disco* (ou *bola*) aberto de raio $r > 0$ e centro $p \in \mathbb{C}$ é

$$D_r(p) := \{z \in \mathbb{C} \ \text{ t.q. } \ |z-p| < r\}\,.$$

Particularmente importante é o *disco unitário* $\mathbf{D} := D_1(0)$, formado pelos números complexos de módulo $|z| < 1$.

ex: Esboce os lugares do plano definidos pelas seguintes equações ou desigualdades:

$$|z-i|=2 \qquad 0<\Re(z)<1 \qquad \Im(z)>0 \qquad 4/z=\overline{z}$$

$$|z-i|>|z+i| \qquad |z-1|+|z+1|=3 \qquad |z+1|-|z-1|=1$$

ex: Mostre que se $z \in D_r(p)$, então $D_{r'}(z) \subset D_r(p)$ se $r' < r - \mathrm{d}(z,p)$.

Circunferências. A circunferência (ou círculo) $C_\rho(a)$ de centro $a \in \mathbb{C}$ e raio $\rho > 0$ é o lugar dos pontos que satisfazem $|z-a| = \rho$. Ao calcular os quadrados, temos $(z-a)(\overline{z}-\overline{a}) = \rho^2$, ou seja,

$$z\overline{z} - \overline{a}z - a\overline{z} + b = 0\,. \tag{6.7}$$

onde $b := |a|^2 - \rho^2 \in \mathbb{R}$. Por exemplo, a circunferência unitária $\mathbb{S} := C_1(0)$ é definida pela equação cartesiana $|z|^2 = 1$.

Retas. A equação paramétrica de uma reta passando pelo ponto $p \in \mathbb{C}$ com velocidade $v \in \mathbb{C}^\times$ é $z(t) = p + vt$, com $t \in \mathbb{R}$. Ao resolver para t, a condição "t real" traduz-se $\Im((z-p)/v) = 0$. Portanto, uma reta no pano complexo $\mathbb{C} \approx \mathbb{R}^2$ pode ser definida por uma equação cartesiana da forma $\Im(\alpha z + \beta) = 0$, onde $\alpha = 1/v \in \mathbb{C}^\times$ e $\beta = -p/v \in \mathbb{C}$ são parâmetros complexos. Uma equação cartesiana da reta é portanto

$$\overline{a}z + a\overline{z} + b = 0\,, \tag{6.8}$$

onde $a = -\overline{\alpha}/i = iv/|v|^2 \in \mathbb{C}^\times$ e $b = 2\,\Im(\beta) \in \mathbb{R}$. Por exemplo, umas equações cartesianas do eixo real $\mathbb{R}$ e do eixo imaginário $i\mathbb{R}$ são $z - \overline{z} = 0$ e $z + \overline{z} = 0$, respetivamente.

ex: Determine uma equação cartesiana da reta que passa por $z_0 = 1+i$ com velocidade $v = 2-i$.

ex: Determine uma equação cartesiana da circunferência de centro $1+i$ e raio 3.

ex: Mostre que a equação $\Re\left((z-a)/(z-b)\right) = 0$ define uma circunferência cujo diâmetro é o segmento entre a e b.

Exponencial complexo e funções trigonométricas. A *função exponencial* $\exp(z) := e^z$, é a função inteira $\exp : \mathbb{C} \to \mathbb{C}$ definida pela série de potências

$$\boxed{e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \dots}$$

ex: Verifique a fórmula de adição $e^{z+w} = e^z e^w$, e deduza que $e^z \neq 0$ para todo o $z \in \mathbb{C}$.

ex: Verifique que e^z é igual à sua derivada, ou seja, $\exp'(z) = \exp(z)$.

ex: Mostre que, se $\theta \in \mathbb{R}$, então o conjugado de $e^{i\theta}$ é $e^{-i\theta}$, e portanto $|e^{i\theta}| = 1$. Defina as funções reais de variável real "cos" e "sin" usando a fórmula de Euler $e^{i\theta} = \cos\theta + i\sin\theta$, ou seja,

$$\cos(\theta) := \frac{e^{i\theta} + e^{-i\theta}}{2} \qquad \text{e} \qquad \sin(\theta) := \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

e deduza as suas expansões em série de potências em torno de 0.

ex: Deduza que, se $\alpha, \theta \in \mathbb{R}$,

$$e^{\alpha + i\theta} = e^\alpha\left(\cos\theta + i\sin\theta\right)$$

Oscilador harmónico e sobreposições. A função $z(t) = e^{i\omega t}$ descreve um ponto que percorre o círculo unitário $\mathbb{S}^1 := \{z \in \mathbb{C} \text{ t.q. } |z| = 1\}$ do plano complexo no sentido anti-horário com "frequência angular" $\omega > 0$ (e portanto com período $\tau = 2\pi/\omega$ e frequência $\nu = \omega/(2\pi)$).

ex: Verifique qua a função $z(t) = e^{i\omega t}$ satisfaz a equação diferencial linear

$$\dot{z} = i\omega z \,.$$

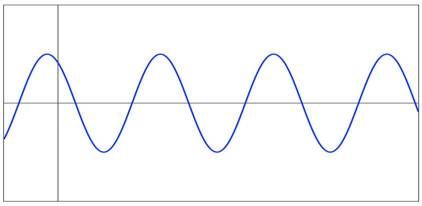
ex: Verifique que $z_\pm(t) = e^{\pm i\omega t}$, e portanto uma sobreposição $z(t) = az_+(t) + bz_-(t)$ com coeficientes $a, b \in \mathbb{C}$, satisfazem a equação do oscilador harmónico (ou lei de Hooke)

$$\ddot{z} = -\omega^2 z \,.$$

Deduza que a parte real e a parte imaginária de $z(t)$, por exemplo

$$q(t) = \rho\, \cos(\omega t + \varphi) \qquad \text{e} \qquad \rho\, \sin(\omega t + \varphi)\,,$$

são soluções reais do oscilador harmónico $\ddot{q} = -\omega^2 q$. Identifique as condições iniciais $q(0)$ e $\dot{q}(0)$ em quanto funções de $\rho e^{i\varphi}$.



Oscilação $q(t) = \rho \cos(\omega t + \varphi)$.

ex: Observe que a sobreposição das oscilações $z_1(t) = e^{i\omega_1 t}$ e $z_2(t) = e^{i\omega_2 t}$,

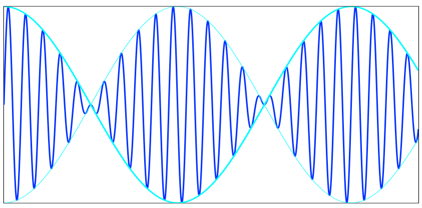
$$z(t) = e^{i\omega_1 t} + e^{i\omega_2 t}\,,$$

é máxima quando $\omega_1 t = \omega_2 t$ (módulo 2π), e mínima quando $\omega_1 t - \omega_2 t = \pi$ (módulo 2π).

ex: Observe que, se $\omega_\pm = \omega \pm \varepsilon$, então a sobreposição das duas oscilações $z_\pm(t) = e^{i\omega_\pm t}$ pode ser representada como

$$z(t) = z_+(t) + z_-(t) = e^{i\omega t}\left(e^{i\varepsilon t} + e^{-i\varepsilon t}\right) = 2e^{i\omega t}\cos(\varepsilon t)$$

Em particular, se $|\varepsilon| \ll |\omega|$, então a sobreposição consiste numa modulação lentai, com período $2\pi/\varepsilon \gg 2\pi/\omega$ (chamada *batimentos*, em inglês, *beats*), da frequência fundamental $\omega \simeq \omega_\pm$.



Sobreposição $z(t) = \Im\left(e^{i0.95\cdot t} + e^{i1.05\cdot t}\right) = \sin(0.95 \cdot t) + \sin(1.05 \cdot t)$.

7 Espaços lineares

ref: [Ap69] *Vol. 2*, **1.1-10** ; [La97] *Ch. III*

26 out 2018

Espaços lineares/vetoriais. Um *espaço linear/vetorial real* (ou seja, sobre o corpo $\mathbb{R}$ dos números reais) é um conjunto $\mathbf{V}$ munido de duas operações:
a "adição" : $\mathbf{V} \times \mathbf{V} \to \mathbf{V}$

$$\mathbf{v}\,,\, \mathbf{w} \mapsto \mathbf{v} + \mathbf{w}\,,$$

que satisfaz os axiomas

EL1 *(propriedade associativa)* $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$,

EL2 *(propriedade comutativa)* $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$,

EL3 *(existência do elemento neutro)* existe $\mathbf{0} \in \mathbf{V}$, tal que $\mathbf{0} + \mathbf{v} = \mathbf{v}$

EL4 *(existência do simétrico)* $\forall \mathbf{v} \in \mathbf{V}$ existe $-\mathbf{v} \in \mathbf{V}$ tal que $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.

e a "multiplicação por escalares/números" : $\mathbb{R} \times \mathbf{V} \to \mathbf{V}$

$$\lambda\,,\, \mathbf{v} \mapsto \lambda\mathbf{v} \in \mathbf{V}$$

que satisfaz os axiomas

EL5 *(propriedade associativa)* $(\lambda\mu)\mathbf{v} = \lambda(\mu\mathbf{v})$,

EL6 *(propriedade distributiva para a adição em $\mathbb{R}$)* $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$,

EL7 *(propriedade distributiva para a adição em $\mathbf{V}$)* $\lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$,

EL8 *(existência do elemento neutro)* $1\mathbf{v} = \mathbf{v}$,

Um *isomorfismo* entre os espaços lineares $\mathbf{V}$ e $\mathbf{V}'$ é uma aplicação bijectiva $f : \mathbf{V} \to \mathbf{V}'$, $\mathbf{v} \leftrightarrow \mathbf{v}'$, que respeita as operações, i.e. tal que se $\mathbf{v} \leftrightarrow f(\mathbf{v}) = \mathbf{v}'$ e $\mathbf{w} \leftrightarrow f(\mathbf{w}) = \mathbf{w}'$, então

$$\mathbf{v} + \mathbf{w} \leftrightarrow \mathbf{v}' + \mathbf{w}' \qquad \text{e} \qquad \lambda\mathbf{v} \leftrightarrow \lambda\mathbf{v}' \qquad \forall \lambda \in \mathbb{R}\,.$$

Se substituimos o corpo $\mathbb{R}$ dos números reais pelo corpo $\mathbb{C}$ dos números complexos, obtemos a definição de *espaço linear/vetorial complexo*.

ex: Verifique que $\mathbb{R}$, munido das operações usuais "+" e "·", é um espaço vetorial real.

ex: Verifique que $\mathbb{R}^n$, munido das operações "adição" e "produto por um escalar" definidas no exercício 1 do capítulo 1, é um espaço vetorial real.

ex: Mostre que o elemento neutro $\mathbf{0}$ é único.

ex: Mostre que o simétrico $-\mathbf{v}$ de cada $\mathbf{v} \in \mathbf{V}$ e único.

ex: Mostre que $0\mathbf{v} = \mathbf{0}$ para todo $\mathbf{v} \in \mathbf{V}$ e que $\lambda\mathbf{0} = \mathbf{0}$ para todo $\lambda \in \mathbb{R}$.

ex: Mostre que $\lambda\mathbf{v} = \mathbf{0}$ implica $\mathbf{v} = \mathbf{0}$ ou $\lambda = 0$.

ex: Mostre que $\lambda\mathbf{v} = \mu\mathbf{v}$ implica $\mathbf{v} = \mathbf{0}$ ou $\lambda = \mu$.

ex: Mostre que $\lambda\mathbf{v} = \lambda\mathbf{w}$ implica $\mathbf{v} = \mathbf{w}$ ou $\lambda = 0$.

ex: [Ap69] **15.5**.

O espaço linear complexo $\mathbb{C}^n$. O *espaço vetorial complexo* de *dimensão* n é o espaço

$$\mathbb{C}^n := \underbrace{\mathbb{C} \times \mathbb{C} \times \cdots \times \mathbb{C}}_{n \text{ vezes}}$$

das n-uplas $\mathbf{z} = (z_1, z_2, \dots, z_n)$ de números complexos, munido das operações *adição*, definida por

$$\boxed{\mathbf{z}, \mathbf{w} \mapsto \mathbf{z} + \mathbf{w} := (z_1 + w_1, z_2 + w_2, \dots, z_n + w_n)}$$

e *multiplicação por um escalar*, definida por

$$\boxed{\lambda, \mathbf{z} \mapsto \lambda\mathbf{z} := (\lambda z_1, \lambda z_2, \dots, \lambda z_n)}$$

Espaço afim. Um *espaço afim* modelado sobre o espaço vetorial $\mathbf{V}$ é um conjunto $\mathcal{A}$ munido de uma aplicação $\mathcal{A} \times \mathcal{A} \to \mathbf{V}$

$$P, Q \mapsto \vec{PQ}$$

que satisfaz os axiomas

EA1 para cada $P \in \mathcal{A}$ e cada $\mathbf{v} \in \mathbf{V}$ existe um único $Q \in \mathcal{A}$ tal que $\vec{PQ} = \mathbf{v}$

EA2 para quaisquer $P, Q, R \in \mathcal{A}$,

$$\vec{PQ} + \vec{QR} = \vec{PR}$$

ex: Verifique que o conjunto $\mathbb{R}^n$ munido da aplicação $P, Q \mapsto Q - P$ é um espaço afim modelado sobre o espaço vetorial real $\mathbb{R}^n$.

Espaços de funções. Os espaços inrteressantes em análise, em física e em engenharia, são espaços de funções, chamados "espaços funcionais".

Sejam X um conjunto e $\mathbb{R}^X = \mathcal{F}(X, \mathbb{R})$ ou $\mathbb{C}^X = \mathcal{F}(X, \mathbb{C})$ os espaços das funções reais ou complexas definidas em X, cujos elementos são denotados por $x \mapsto f(x)$. Os espaços $\mathcal{F}(X, \mathbb{R})$ e $\mathcal{F}(X, \mathbb{C})$, munidos das operações "adição" e "produto por um escalar" definidas por

$$(f + g)(x) := f(x) + g(x) \qquad \text{e} \qquad (\lambda f)(x) := \lambda f(x)\,,$$

são espaços vetoriais reais e complexos, respetivamente. Também interessantes são os 'campos vetoriais", funções com valores num espaço vetorial como $\mathbb{R}^m$ (campos de força, de velocidade, campo eletro-magnético, ...

Engenheiros e físicos estão por exemplo interessados em "sinais" $f(t)$ (a intensidade de uma onda de som, onde t é o tempo), ou "funções de onda" $\psi(\mathbf{r}, t)$, em mecânica quânticas, ou outros "campos" $u(\mathbf{r}, t)$ (um deslocamento, um campo de velocidades, o campo eletro-magnético, ...) que resolvem certas equações diferenciais parciais como a equação de onda, de calor, de Laplace, de Schrödinger, ...).

e.g. **Polinómios.** Sejam $\mathrm{Pol}(\mathbb{R}) := \mathbb{R}[t]$ o espaço dos polinómios $p(t) = a_n t^n + \cdots + a_1 t + a_0$ com coeficientes reais na variável t, e $\mathrm{Pol}_{\leq n}(\mathbb{R})$ o espaço dos polinómios reais de grau $\leq n$. $\mathrm{Pol}(\mathbb{R})$ e $\mathrm{Pol}_{\leq n}(\mathbb{R})$ são espaços lineares. Por outro lado, o conjunto dos polinómios de grau n não é um espaço linear (porque?).

e.g. **Funções contínuas e diferenciáveis.** São espaços vetoriais, reais ou complexos, o espaço $\mathcal{C}^0(\mathbb{R})$ das funções contínuas $f : \mathbb{R} \to \mathbb{R}$ ou $\mathbb{C}$, os espaços $\mathcal{C}^k(\mathbb{R})$ das funções com k derivadas contínuas, o espaço $\mathcal{C}^\infty(\mathbb{R})$ das funções infinitamente deriváveis. As inclusões sendo

$$\mathcal{C}^\infty(\mathbb{R}) \subset \cdots \subset \mathcal{C}^{k+1}(\mathbb{R}) \subset \mathcal{C}^k(\mathbb{R}) \subset \cdots \subset \mathcal{C}^0(\mathbb{R})$$

e.g. Sucessões. Seja $\mathbb{R}^\infty := \mathbb{R}^\mathbb{N} = \mathcal{F}(\mathbb{N}, \mathbb{R})$ o espaço das sucessões reais $\mathbf{x} = (x_1, x_2, \dots, x_n, \dots)$, com $x_n \in \mathbb{R}$. Descreva a sua estrutura de espaço linear real. Mostre que o espaço b das sucessões limitadas, o espaço c das sucessões convergentes, o espaço c_0 das sucessões que convergem para 0, e o espaço ℓ das sucessões com suporte compacto (i.e. tais que $x_n = 0$ se n é suficientemente grande) são subespaços vetoriais de $\mathbb{R}^\infty$, e que

$$\ell \subset c_0 \subset c \subset b \subset \mathbb{R}^\infty .$$

Amostragem/ADC. Um sinal contínuo $f(t)$ podes ser observado apenas em múltiplos inteiros de um tempo de amostragem $\tau > 0$, e transformado num sinal discreto $x[n] := f(n\tau)$, com $n \in \mathbb{Z}$ ou $\mathbb{N}$, que é uma sequência.

ex: [Ap69] **15.5**.

Superposition principle for linear ODEs. Consider a "linear homogeneous" ordinary differential equation, for example with constant coefficients and of second order like

$$\ddot{y} + \alpha\dot{y} + \beta y = 0 . \tag{7.1}$$

A superposition $ay_1(t) + by_2(t)$ of any two solutions, $y_1(t)$ and $y_2(t)$, is still a solution. Therefore, the space $\mathbf{H}$ of solutions of (7.1) is a linear space (a subspace of the infinite-dimensional space $\mathcal{C}^2(\mathbb{R})$ of twice differentiable functions on the real line). Actually, as you will see in analysis, this space is two-dimensional, so that $\mathbf{H} \approx \mathbb{R}^2$, and therefore any solution is a superposition

$$y(t) = c_+ \, y_+(t) + c_- \, y_-(t)$$

of two basic independent solutions $y_\pm(t)$, forming a base of $\mathbf{H}$. We now add a time-dependent force $f(t)$, and consider the linear ordinary differential equation

$$\ddot{x} + \alpha\dot{x} + \beta x = f(t) . \tag{7.2}$$

The difference $x_1(t) - x_2(t)$ between any two solutions of (7.2) is a solution of (7.1). Therefore, if $z(t)$ is any (particular) solution of (7.2), then the space of all solutions of (7.2) is the affine space $z + \mathbf{H}$.

Subespaços e geradores. Seja $\mathbf{V}$ um espaço vetorial real ou complexo. Um subconjunto não vazio $W \subset \mathbf{V}$ que é também um espaço vetorial (ou seja, tal que $w + w' \in W$ e $\lambda w \in W$ para todos os $w, w' \in W$ e $\lambda \in \mathbb{R}$) é dito *subespaço (linear/vetorial)* de $\mathbf{V}$. Se $S \subset \mathbf{V}$ é um subconjunto de $\mathbf{V}$, o conjunto $\mathrm{Span}(S)$ das combinações lineares finitas

$$\lambda_1\mathbf{s}_1 + \lambda_2\mathbf{s}_2 + \cdots + \lambda_m\mathbf{s}_m \qquad \mathbf{s}_i \in S \, , \, \lambda_i \in \mathbb{R} \text{ ou } \mathbb{C}$$

é um subsespaço de $\mathbf{V}$, dito subespaço *gerado* por S. O espaço linear $\mathbf{V}$ tem "dimensão finita" se admite um conjunto finito de geradores.

Se X e Y são subespaços de $\mathbf{V}$, então a "soma"

$$X + Y := \{\mathbf{x} + \mathbf{y} \text{ com } \mathbf{x} \in X, \, \mathbf{y} \in Y\}$$

é um subespaço de $\mathbf{V}$. Se cada vetor $\mathbf{v} \in X + Y$ admite uma única representação $\mathbf{v} = \mathbf{x} + \mathbf{y}$ com $\mathbf{x} \in X$ e $\mathbf{y} \in Y$, então a soma é chamada *soma direta*, e denotada por $X \oplus Y$.

ex: Se $X_1, X_2, \dots$ são subespaços de $\mathbf{V}$, então

$$\bigcap_i X_i := \{\mathbf{x} \text{ t.q. } \mathbf{x} \in X_i \, \forall i\}$$

é um subespaço de $\mathbf{V}$.

ex: Dados os números $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$, o conjunto dos vetores $\mathbf{x} \in \mathbb{R}^n$ tais que

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0$$

é um subespaço linear de $\mathbb{R}^n$.

ex: Se V é um subespaço do espaço euclidiano $\mathbb{R}^n$, então

$$V^\perp := \{\mathbf{x} \in \mathbb{R}^n \text{ t.q. } \mathbf{x} \cdot \mathbf{v} = 0 \quad \forall \mathbf{v} \in V\}$$

é um subespaço de $\mathbb{R}^n$.

ex: Verifique que $\mathcal{C}(\mathbb{R}, \mathbb{R})$ é um subespaço de $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

ex: Verifique que $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ é um subespaço de $\mathcal{C}^{k+1}(\mathbb{R}, \mathbb{R})$, que é um subespaço de $\mathcal{C}^k(\mathbb{R}, \mathbb{R})$, que é um subespaço de $\mathcal{C}(\mathbb{R}, \mathbb{R})$.

ex: Verifique que $\mathrm{Pol}_{\leq n}(\mathbb{R})$ é um supespaço de $\mathrm{Pol}(\mathbb{R})$, que é um subespaço de $\mathcal{C}(\mathbb{R}, \mathbb{R})$. O espaço $\mathrm{Pol}_n(\mathbb{R})$ dos polinómios de grau (exatamente) n é um suberspaço vetorial de $\mathrm{Pol}(\mathbb{R})$?

ex: Determine os subespaços gerados por

$$(1,2) \qquad (-1,-2) \qquad \text{em } \mathbb{R}^2$$

$$(0,1,0) \qquad (0,0,1) \qquad \text{em } \mathbb{R}^3$$

$$\mathbf{i} \quad \mathbf{j} \qquad \mathbf{i}+\mathbf{j} \qquad \text{em } \mathbb{R}^3$$

$$\{1, t, t^2, \dots, t^n, \dots\} \qquad \text{em } \mathbb{R}^{\mathbb{R}}$$

$$t \qquad t^2 \qquad \text{em } \mathrm{Pol}(\mathbb{R})$$

$$e^t \qquad e^{-t} \qquad \text{em } \mathcal{F}(\mathbb{R}, \mathbb{R})$$

ex: O conjunto b das sucessões $(x_n)_{n\in\mathbb{N}}$ limitadas é um subespaço do espaço das sucessões reais $\mathbb{R}^\infty := \mathbb{R}^{\mathbb{N}}$? E o conjunto c das sucessões convergentes? E o conjunto c_0 das sucessões convergentes tais que $x_n \to 0$?

ex: O espaço das funções não negativas, i.e. tais que $f(t) \geq 0$ $\forall t$, é um subespaço do espaço $\mathcal{F}(\mathbb{R}, \mathbb{R})$?

ex: Os conjuntos das funções pares e ímpares, definidos por

$$\mathcal{F}_\pm(\mathbb{R}, \mathbb{R}) = \{f \in \mathcal{F}(\mathbb{R}, \mathbb{R}) \quad \text{t.q. } f(t) = \pm f(-t) \quad \forall t \in \mathbb{R}\},$$

são subespaços de $\mathcal{F}(\mathbb{R}, \mathbb{R})$? Mostre que cada $f \in \mathcal{F}(\mathbb{R}, \mathbb{R})$ é uma soma $f = f_+ + f_-$ com $f\pm \in \mathcal{F}_\pm(\mathbb{R}, \mathbb{R})$.

ex: [Ap69] **15.9**.

Conjuntos livres/linearmente independentes. Seja $\mathbf{V}$ um espaço linear. O conjunto $S \subset \mathbf{V}$ é *livre/(linearmente) independente* se gera cada vetor de $\mathrm{Span}(S)$ duma única maneira, ou seja, se gera o vetor nulo $\mathbf{0}$ duma única maneira, i.e. se quaisquer que sejam os elementos distintos $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m \in S$,

$$\boxed{\lambda_1 \mathbf{s}_1 + \lambda_2 \mathbf{s}_2 + \cdots + \lambda_m \mathbf{s}_m = \mathbf{0} \qquad \Longrightarrow \qquad \lambda_i = 0 \quad \forall i = 1, 2, \dots, m}$$

Caso contrário, o conjunto é dito *(linearmente) dependente*.

ex: Verifique se os seguintes conjuntos de vetores são linearmente independentes

$$(1,2) \qquad (-1,2) \qquad \text{em } \mathbb{R}^2$$
$$(1,1,0) \qquad (0,1,1) \qquad (1,0,1) \qquad \text{em } \mathbb{R}^3$$
$$(1,2,3) \qquad (2,3,4) \qquad (3,4,5) \qquad \text{em } \mathbb{R}^3$$
$$(1,0,0,0) \qquad (1,1,0,0) \qquad (1,1,1,0) \qquad (1,1,1,1) \qquad \text{em } \mathbb{R}^4$$
$$\cos t \qquad \sin t \qquad \text{em } \mathcal{F}(\mathbb{R},\mathbb{R})$$
$$\cos^2 t \qquad \sin^2 t \qquad 1/2 \qquad \text{em } \mathcal{F}(\mathbb{R},\mathbb{R})$$
$$1 \qquad t \qquad t^2 \qquad \text{em } \mathrm{Pol}(\mathbb{R})$$
$$(1,-1,1,-1,1,-1,\dots) \qquad (1,1,1,1,1,1,\dots) \qquad \text{em } \mathbb{R}^\infty$$

Bases e dimensão. Seja $\mathbf{V}$ um espaço linear de dimensão finita, real ou complexo. Uma *base* de $\mathbf{V}$ é um conjunto livre de geradores de $\mathbf{V}$, ou seja, um conjunto ordenado $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ de vetores tal que cada $\mathbf{v} \in \mathbf{V}$ admite uma e uma única representação

$$\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n$$

Os números v_i são as *componentes*. do vetor $\mathbf{v}$ relativamente à base $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$. A *dimensão* $\dim_{\mathbb{R}} \mathbf{V}$ do espaço linear $\mathbf{V}$ (de dimensão finita) é o número de elementos de uma base. Se $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ é uma base ordenada do espaço linear $\mathbf{V}$, então a aplicação

$$\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n \mapsto (v_1, v_2, \dots, v_n)$$

define um isomorfismo $\mathbf{V} \approx \mathbb{R}^n$ ou $\mathbb{C}^n$ (dependendo se o espaço linear é real ou complexo).

ex: Determine uma base e a dimensão dos seguintes subespaços de $\mathbb{R}^2$ ou $\mathbb{R}^3$

$$\{(x,y) \in \mathbb{R}^2 \text{ t.q. } x = y\} \subset \mathbb{R}^2$$
$$\{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } x = y\} \subset \mathbb{R}^3$$
$$\{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } x+y+z = 0\} \subset \mathbb{R}^3$$
$$\{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } x+y = 0 \text{ e } y+z = 0\} \subset \mathbb{R}^3$$
$$\{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } x-2y+z = 0\} \subset \mathbb{R}^3$$
$$\{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } x+y+z = 0 \text{ e } z = 0\} \subset \mathbb{R}^3$$

ex: Verifique que

$$1 \qquad t \qquad t^2 \qquad t^3 \qquad \dots \qquad t^n$$

é uma base do espaço $\mathrm{Pol}_n(\mathbb{R})$ dos polinómios de grau $\leq n$. Qual a dimensão de $\mathrm{Pol}_n(\mathbb{R})$?

ex: Determine as coordenadas do polinómio $f(t) = (1-t)^2$ relativamente à base ordenada $(1, t, t^2)$ de $\mathrm{Pol}_2(\mathbb{R})$.

ex: [Ap69] **15.9**.

Rational linear independence on the line. The real numbers $x_1, x_2, \dots, x_n$ are said *linear independent over the field of rationals* if the only solution of

$$k_1x_1 + k_2x_2 + \cdots + k_nx_n = 0$$

with $k_i \in \mathbb{Z}$ is the trivial solution $k_1 = k_2 = \cdots = k_n = 0$.

ex: Show that (any finite subset of) the sequence

$$\log 2\,, \quad \log 3\,, \quad \log 5\,, \quad \log 7\,, \quad \dots \quad \log p\,, \quad \dots$$

of natural logarithms of prime numbers is linear independent over the rationals (use the unique factorization of an integer into prime factors) [9].

[9] H. Bohr, 1910.

8 Transformações lineares

ref: [Ap69] *Vol 2*, **2.1-9** ; [La97] *Ch. IV*, **1-4**

2 nov 2018

Linearidade. Se cada kilo de ♡ custa A euros e cada kilo de ♠ custa B euros, então a kilos de ♡ e b kilos de ♠ custam $aA + bB$ euros. Ou seja, a função "preço" P satisfaz

$$P(a\heartsuit + b\spadesuit) = a \cdot P(\heartsuit) + b \cdot P(\spadesuit)$$

Esta propriedade é chamada *linearidade*.

Por outro lado, a superfície e o volume de um cubo de lado 2ℓ são 4 e 8 vezes a superfície e o volume de um cubo de lado ℓ, respetivamente (e esta é uma das razões pela existência de dimensões típicas de animais e plantas, como explicado por D'Arcy Thompson[10]). São funções não lineares.

ex: Dê exemplos de funções lineares e de funções não lineares.

Formas lineares, espaço dual. Seja $\mathbf{V}$ um espaço linear real (ou complexo). Uma função real $f : \mathbf{V} \to \mathbb{R}$ (ou complexa $f : \mathbf{V} \to \mathbb{C}$) é dita *aditiva* se $\forall\, \mathbf{v}, \mathbf{w} \in \mathbf{V}$

$$\boxed{f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w})}$$

e é dita *homogénea* se $\forall\, \mathbf{v} \in \mathbf{V}$ e $\forall\, \lambda \in \mathbb{R}$ (ou $\forall\, \lambda \in \mathbb{C}$)

$$\boxed{f(\lambda \mathbf{v}) = \lambda f(\mathbf{v})}$$

Uma função real $\boldsymbol{\xi} : \mathbf{V} \to \mathbb{R}$ (ou complexa $\boldsymbol{\xi} : \mathbf{V} \to \mathbb{C}$) aditiva e homogénea, ou seja, tal que

$$\boxed{\boldsymbol{\xi}(\lambda \mathbf{v} + \mu \mathbf{w}) = \lambda\, \boldsymbol{\xi}(\mathbf{v}) + \mu\, \boldsymbol{\xi}(\mathbf{w})}$$

é dita *forma linear*, ou *covetor* (ou *funcional linear* quando $\mathbf{V}$ é um espaço de funções). Uma notação simétrica para o valor da forma linear $\boldsymbol{\xi}$ sobre o vetor $\mathbf{v}$ é

$$\langle \boldsymbol{\xi}, \mathbf{v} \rangle := \boldsymbol{\xi}(\mathbf{v})\,.$$

O espaço $\mathbf{V}^* := \mathrm{Hom}_{\mathbb{R}}(\mathbf{V}, \mathbb{R})$ (ou $\mathrm{Hom}_{\mathbb{C}}(\mathbf{V}, \mathbb{C})$) das formas lineares, dito *espaço dual (algébrico)* de $\mathbf{V}$, é um espaço linear real (ou complexo) se a adição e o produto por um escalar são definidos por

$$\langle \boldsymbol{\xi} + \boldsymbol{\eta}, \mathbf{v} \rangle := \langle \boldsymbol{\xi}, \mathbf{v} \rangle + \langle \boldsymbol{\eta}, \mathbf{v} \rangle \qquad \text{e} \qquad \langle \lambda \boldsymbol{\xi}, \mathbf{v} \rangle := \langle \boldsymbol{\xi}, \lambda \mathbf{v} \rangle$$

respetivamente, e a forma nula $\mathbf{0}^* \in \mathbf{V}^*$ é definida por $\langle \mathbf{0}^*, \mathbf{v} \rangle = 0\ \forall\, \mathbf{v} \in \mathbf{V}$.

Se o espaço vetorial $\mathbf{V}$ tem dimensão finita, e se $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ é uma base de $\mathbf{V}$ (por exemplo a base canónica de $\mathbf{V} \approx \mathbb{R}^n$), então cada forma linear $\boldsymbol{\xi} \in \mathbf{V}^*$ é determinada pelos seus valores $\xi_i := \langle \boldsymbol{\xi}, \mathbf{e}_i \rangle$ nos vetores da base, pois

$$\langle \boldsymbol{\xi}, \mathbf{v} \rangle = \langle \boldsymbol{\xi}, v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + \cdots + v_n \mathbf{e}_n \rangle = \xi_1 v_1 + \xi_2 v_2 + \cdots + \xi_n v_n\,.$$

Portanto, também $\mathbf{V}^*$ tem dimensão finita, e uma base de $\mathbf{V}^*$, dita *base dual*, é o conjunto ordenado dos covetores $\mathbf{e}_1^*, \mathbf{e}_2^*, \ldots, \mathbf{e}_n^*$ definidos por [11]

$$\langle \mathbf{e}_i^*, \mathbf{e}_j \rangle = \delta_{ij}$$

As coordenadas da forma linear $\boldsymbol{\xi} = \xi_1 \mathbf{e}_1^* + \xi_2 \mathbf{e}_2^* + \cdots + \xi_n \mathbf{e}_n^*$ relativamente à base dual são os números $\xi_i = \langle \boldsymbol{\xi}, \mathbf{e}_i \rangle$, assim que $\langle \boldsymbol{\xi}, \mathbf{v} \rangle = \sum_i \xi_i v_i$.

Para cada $\mathbf{v} \in \mathbf{V}$, a função $\xi \mapsto \langle \boldsymbol{\xi}, \mathbf{v} \rangle$ é uma forma linear em $\mathbf{V}^*$, e portanto existe um homomorfismo injetivo de $\mathbf{V}$ em $(\mathbf{V}^*)^*$. Se $\mathbf{V}$ tem dimensão finita, então todas as formas lineares $g \in (\mathbf{V}^*)^*$ podem ser representadas como $g(\boldsymbol{\xi}) = \langle \boldsymbol{\xi}, \mathbf{v} \rangle$ para algum $\mathbf{v} \in \mathbf{V}$ (basta definir $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ com $v_i = g(\mathbf{e}_i)$), e portanto o espaço dual do espaço dual é isomorfo a $(\mathbf{V}^*)^* \approx \mathbf{V}$ (mas o isomorfismo não é canónico, depende da escolha de uma base!).

[10] D'Arcy Wentworth Thompson, *On growth and form*, 1917 and 1942.

[11] O *símbolo de Kronecker* é definido por

$$\delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

ex: Mostre que uma função homogénea $f : \mathbb{R} \to \mathbb{R}$ é $f(x) = \lambda x$, com $\lambda = f(1) \in \mathbb{R}$. Deduça que uma função homogénea $f : \mathbb{R} \to \mathbb{R}$ é também aditiva, e portanto linear.

ex: Diga se as seguintes funções $f : \mathbb{R}^n \to \mathbb{R}$ são ou não lineares.

$$x \mapsto 3x \qquad x \mapsto 2x - 1 \qquad x \mapsto \sin(2\pi x)$$

$$(x, y) \mapsto 3x - 5y \qquad (x, y) \mapsto x^2 - xy$$

$$(x, y, z) \mapsto 2x - y + 3z \qquad (x, y, z) \mapsto 2x - y + 3z + 8$$

$$(x, y, z) \mapsto 0 \qquad (x, y, z) \mapsto \sqrt{3}$$

$$(x_1, x_2, \dots, x_n) \mapsto 0 \qquad (x_1, x_2, \dots, x_n) \mapsto 66$$

ex: Mostre que, dados $a_1, a_2, \dots, a_n \in \mathbb{R}$, a aplicação

$$(x_1, x_2, \dots, x_n) \mapsto a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

é uma forma linear em $\mathbb{R}^n$.

Formas lineares no espaço euclidiano $\mathbb{R}^n$. Uma forma linear $\boldsymbol{\xi} \in (\mathbb{R}^n)^*$ é determinada pelos seus valores nos vetores de uma base de $\mathbb{R}^n$. Por exemplo, se $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ é a base canónica de $\mathbb{R}^n$, e $\xi_i := \langle \boldsymbol{\xi}, \mathbf{e}_i \rangle$, com $i = 1, 2, \dots, n$, então o valor da forma $\boldsymbol{\xi} \in (\mathbb{R}^n)^*$ sobre o vetor genérico $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n \in \mathbb{R}^n$ é

$$\begin{aligned} \langle \boldsymbol{\xi}, \mathbf{x} \rangle &= \langle \boldsymbol{\xi}, x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n \rangle \\ &= x_1 \langle \boldsymbol{\xi}, \mathbf{e}_1 \rangle + x_2 \langle \boldsymbol{\xi}, \mathbf{e}_2 \rangle + \cdots + x_n \langle \boldsymbol{\xi}, \mathbf{e}_n \rangle \\ &= x_1\xi_1 + x_2\xi_2 + \cdots + x_n\xi_n \,. \end{aligned}$$

Portanto, se $\boldsymbol{\xi} = (\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{R}^n$, então

$$\langle \boldsymbol{\xi}, \mathbf{x} \rangle = \boldsymbol{\xi} \cdot \mathbf{x}$$

A correspondência que associa a forma $\langle \boldsymbol{\xi}, \mathbf{x} \rangle = \boldsymbol{\xi} \cdot \mathbf{x}$ ao vetor $\boldsymbol{\xi} \in \mathbb{R}^n$ é um isomorfismo $\mathbb{R}^n \approx (\mathbb{R}^n)^*$ entre o espaço euclidiano $\mathbb{R}^n$ e o seu dual $(\mathbb{R}^n)^*$ (que depende da estrutura euclidiana, ou seja, do produto escalar euclidiano!).

Uma outra maneira de representar a forma linear é usando o "produto linhas por colunas"

$$\langle \boldsymbol{\xi}, \mathbf{x} \rangle = \Xi X = \begin{pmatrix} \xi_1 & \xi_2 & \dots & \xi_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_n \end{pmatrix}$$

do vetor linha

$$\Xi = \begin{pmatrix} \xi_1 & \xi_2 & \dots & \xi_n \end{pmatrix}$$

pelo vetor coluna

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_n \end{pmatrix}$$

ex: Determine o vetor $\boldsymbol{\xi} = (\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{R}^n$ que define as seguintes formas lineares

$$x \mapsto 3x \qquad (x, y) \mapsto 0$$

$$(x, y) \mapsto 5x + 9y \qquad (x, y, z) \mapsto -3x + 7y - z$$

$$(x_1, x_2, \dots, x_n) \mapsto 3x_k \qquad \text{com } 0 \le k \le n$$

ex: $f : \mathbb{R}^2 \to \mathbb{R}$ é uma função linear tal que $f(\mathbf{i}) = 5$ e $f(\mathbf{j}) = -2$. Determine $f(x, y)$.

ex: $f : \mathbb{R}^3 \to \mathbb{R}$ é uma função linear tal que $f(\mathbf{i}) = -3$, $f(\mathbf{j}) = 1$ e $f(\mathbf{k}) = 7$. Determine $f(x, y, z)$.

Núcleo e hiperplanos. O *núcleo/espaço nulo* (em inglês *kernel*) da forma linear $\boldsymbol{\xi} \in \mathbf{V}^*$ é o subespaço vetorial

$$\boxed{\mathrm{Ker}(\boldsymbol{\xi}) := \{\mathbf{x} \in \mathbf{V} \text{ t.q. } \langle \boldsymbol{\xi}, \mathbf{x} \rangle = 0\} \subset \mathbf{V}\,.}$$

Se $\boldsymbol{\xi} \neq \mathbf{0}'$ e $\mathbf{v} \in \mathbf{V} \backslash \mathrm{Ker}(\boldsymbol{\xi})$ (i.e. um vetor tal que $\langle \boldsymbol{\xi}, \mathbf{v} \rangle \neq 0$), então cada vetor $\mathbf{x} \in \mathbf{V}$ pode ser representado de uma única maneira como

$$\mathbf{x} = \lambda \mathbf{v} + \mathbf{w}$$

com $\lambda \in \mathbb{R}$ e $\mathbf{w} \in \mathrm{Ker}(\xi)$. De fato, a condição $\mathbf{x} - \lambda \mathbf{v} \in \mathrm{Ker}(\xi)$, ou seja, $\langle \xi, \mathbf{x} - \lambda \mathbf{v} \rangle = 0$, obriga a escolher $\lambda = \langle \boldsymbol{\xi}, \mathbf{x} \rangle / \langle \boldsymbol{\xi}, \mathbf{v} \rangle$. Portanto, o núcleo $\mathrm{Ker}(\boldsymbol{\xi})$ de uma forma $\boldsymbol{\xi} \neq \mathbf{0}'$ é um *hiperplano* do espaço linear $\mathbf{V}$, ou seja, um subespaço linear di "co-dimensão" 1. Em particular, se $\mathbf{V}$ tem dimensão finita,

$$\boxed{\dim_{\mathbb{R}} \mathrm{Ker}(\boldsymbol{\xi}) + 1 = \dim_{\mathbb{R}} \mathbf{V}}$$

O *hiperplano afim* que passa pelo ponto $\mathbf{a} \in \mathbf{V}$ e é paralelo ao hiperplano $\mathrm{Ker}(\boldsymbol{\xi})$ é

$$\mathbf{a} + \mathrm{Ker}(\boldsymbol{\xi}) = \{\mathbf{v} \in \mathbf{V} \text{ t.q. } \langle \boldsymbol{\xi}, \mathbf{v} \rangle = \lambda\} \qquad \text{onde } \lambda = \langle \boldsymbol{\xi}, \mathbf{a} \rangle\ .$$

ex: Mostre que o núcleo de uma forma linear $\boldsymbol{\xi} \in \mathbf{V}^*$ é um subespaço linear de $\mathbf{V}$.

ex: Uma "equação linear"

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b$$

define um hiperplano afim em $\mathbb{R}^n$. Mostre que o hiperplano afim é um subespaço linear de $\mathbb{R}^n$ sse $b = 0$.

Interseções de hiperplanos/sistemas homogéneos. Se $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_m \in \mathbf{V}^*$ são m formas lineares independentes num espaço linear de dimensão finita $\mathbf{V} \approx \mathbb{R}^n$ (em particular, $m \leq n$), então a interseção dos núcleos $\mathbf{W} = \cap_{k=1}^m \mathrm{Ker}(\boldsymbol{\xi}_k)$, ou seja o conjunto dos vetores $\mathbf{x} \in \mathbf{V}$ que resolvem o "sistema homogéneo"

$$\langle \boldsymbol{\xi}_1, \mathbf{x} \rangle = 0\,, \quad \langle \boldsymbol{\xi}_2, \mathbf{x} \rangle = 0\,, \qquad \ldots \qquad \langle \boldsymbol{\xi}_m, \mathbf{x} \rangle = 0\,,$$

é um subespaço vetorial de co-dimensão m, e portanto de dimensão $n-m$. De fato, se completamos o sistema de formas independentes até obter uma base $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_m, \boldsymbol{\xi}_{m+1}, \ldots, \boldsymbol{\xi}_n$ de $\mathbf{V}^*$, e se $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ denota a base dual de $\mathbf{V}$ (assim que $\langle \boldsymbol{\xi}_i, \mathbf{b}_j \rangle = \delta_{ij}$), então nas coordenadas relativas a esta base o sistema homogéneo é

$$x_1 = 0\,, \quad x_2 = 0\,, \qquad \ldots \qquad x_m = 0\,,$$

e as soluções são todos os vetores $\mathbf{x} = \sum_{k=m+1}^n x_k \mathbf{b}_k \in \mathbf{W} \approx \mathbb{R}^{n-m}$.

Integral. O integral

$$f \mapsto I(f) := \int_a^b f(t)\, dt$$

é uma forma linear no espaço $\mathcal{C}^0([a, b])$ das funções contínuas no intervalo $[a, b] \subset \mathbb{R}$. O seu núcleo é o hiperplano das funções com média nula no intervalo. Toda função contínua pode ser representada de forma única como soma $f(x) = c + g(x)$ onde $c = I(f)$ é uma constante (a média de f vezes o comprimento do intervalo) e $g(x) = f(x) - I(f)$ é uma função com média nula.

Delta de Dirac. A *delta de Dirac* (no ponto 0), definida por

$$f \mapsto \delta(f) := \text{“} \int_{-\infty}^{\infty} \delta(t) f(t)\, dt \text{”} \ = f(0)\,,$$

é uma forma linear no espaço $\mathcal{C}^0(\mathbb{R})$ das funções contínuas $f : \mathbb{R} \to \mathbb{R}$. O núcleo de δ é o conjunto das funções (contínuas) tais que $f(0) = 0$, que é um hiperplano do espaço $\mathcal{C}^0(\mathbb{R})$.

Conjuntos convexos. Um subconjunto $C \subset \mathbf{V}$ de um espaço vetorial real é *convexo* se contém o segmento entre cada par de seus pontos, i.e. se

$$\mathbf{x}, \mathbf{y} \in C \qquad \Rightarrow \qquad (1-t)\mathbf{x} + t\mathbf{y} \in C \quad \forall t \in [0,1]$$

O menor convexo que contém (ou seja, a interseção de todos os convexos que contêm) um subconjunto $A \subset \mathbb{R}^n$ é chamado *envoltória/invólucro/fecho convexa/o* de A, e denotado por $\mathrm{Conv}(A)$. Em particular, o menor convexo que contém o conjunto finito de pontos $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m \in \mathbf{V}$ é

$$\mathrm{Conv}(\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}) := \{t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \dots + t_m\mathbf{x}_m \quad \text{com } t_i \geq 0 \text{ e } t_1 + t_2 + \dots + t_m = 1\}$$

ex: As bolas $B_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n \text{ t.q. } \|\mathbf{y} - \mathbf{x}\| < r\}$ e $\overline{B_r(\mathbf{x})} = \{\mathbf{y} \in \mathbb{R}^n \text{ t.q. } \|\mathbf{y} - \mathbf{x}\| \leq r\}$ são convexas.

ex: Dados um vetor $\mathbf{n} \in \mathbb{R}^n$ e um escalar $b \in \mathbb{R}$, os semi-espaços $\{\mathbf{x} \in \mathbb{R}^n \text{ t.q. } \mathbf{n} \cdot \mathbf{x} \geq b\}$ são convexos.

ex: Dada uma forma $\boldsymbol{\xi} \in \mathbf{V}^*$ e um escalar $b \in \mathbb{R}$, os semiespaços $\{\mathbf{v} \in \mathbf{V} \text{ t.q. } \langle \boldsymbol{\xi}, \mathbf{v} \rangle \geq b\}$ são convexos.

ex: A envoltória convexa de três pontos, A, B e C do plano $\mathbb{R}^2$ é o triângulo de vértices A, B e C.

ex: Translações e homotetias preservam os convexos, i.e. se $C \subset \mathbf{V}$ é convexo, então também $T_{\mathbf{a}}(C) = C + \mathbf{a}$ e $H_\lambda(C) = \lambda C$ são convexos, $\forall \mathbf{a} \in \mathbf{V}$ e $\forall \lambda \in \mathbb{R}$.

ex: Se $\boldsymbol{\xi} \in \mathbf{V}^*$, então $\{\mathbf{v} \in \mathbf{V} \text{ t.q. } \boldsymbol{\xi}(\mathbf{v}) < 0\}$ é convexo.

Medidas de probabilidades. Uma *(medida de) probabilidade* num “espaço dos acontecimentos” finito $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ é uma função $\mathbb{P} : 2^\Omega := \{\text{subconjuntos } A \subset \Omega\} \to [0,1]$ aditiva, i.e. tal que

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) \qquad \text{se } A \cap B = \emptyset\,,$$

(a probabilidade do evento “A ou B” é igual à probabilidade do evento A mais a probabilidade do evento B se A e B são eventos mutuamente exclusivos) que verifica $\mathbb{P}(\emptyset) = 0$ (a probabilidade do “evento impossível” é nula) e $\mathbb{P}(\Omega) = 1$ (a probabilidade do “evento certo” é um).

Cada vetor $\mathbf{p} = (p_1, p_2, \dots, p_n) \in \mathbb{R}^n$ cujas coordenadas estão limitadas por $0 \leq p_i \leq 1$ e tal que $p_1 + p_2 + \dots + p_n = 1$ define uma probabilidade $\mathbb{P}$, por meio de

$$\mathbb{P}(A) = \sum_{\omega_i \in A} p_i$$

ou seja, $p_i = \mathbb{P}(\{\omega_i\}$. Portanto, o espaço das medidas de probabilidades em Ω é o fecho convexo dos vetores da base canónica de $\mathbb{R}^n$, chamado *“simplex”*

$$\Delta^{n-1} := \{\mathbf{p} = (p_1, p_2, \dots, p_n) \in \mathbb{R}^n \text{ com } 0 \leq p_i \leq 1 \text{ e } p_1 + p_2 + \dots + p_n = 1\} \subset \mathbb{R}^n\,.$$

Transformações lineares. Uma *transformação/aplicação/operador linear* entre os espaços vetoriais reais (ou complexos) $\mathbf{V}$ e $\mathbf{W}$ é uma função $L : \mathbf{V} \to \mathbf{W}$ aditiva e homogénea, ou seja, tal que

$$L(\mathbf{v} + \mathbf{v}') = L(\mathbf{v}) + L(\mathbf{v}') \qquad \text{e} \qquad L(\lambda \mathbf{v}) = \lambda\, L(\mathbf{v})$$

ou seja, tal que

$$\boxed{L(\lambda \mathbf{v} + \lambda' \mathbf{v}') = \lambda L(\mathbf{v}) + \lambda' L(\mathbf{v}')}$$

$\forall\, \mathbf{v}, \mathbf{v}' \in \mathbf{V}$ e $\forall \lambda, \lambda' \in \mathbb{R}$ (ou $\in \mathbb{C}$). É usual omitir as parêntesis, e denotar a imagem do vetor $\mathbf{v}$ simplesmente por $L\mathbf{v}$.

O espaço $\mathrm{Lin}(\mathbf{V}, \mathbf{W}) := \mathrm{Hom}_{\mathbb{R}}(\mathbf{V}, \mathbf{W})$ (ou $\mathrm{Hom}_{\mathbb{C}}(\mathbf{V}, \mathbf{W})$) das transformações lineares de $\mathbf{V}$ em $\mathbf{W}$ é um espaço linear real (ou complexo) se a adição e a multiplicação por um escalar são definidas por

$$(L + L')(\mathbf{v}) := L(\mathbf{v}) + L'(\mathbf{v}) \qquad\qquad (\lambda L)(\mathbf{v}) := \lambda L(\mathbf{v})$$

Em particular, é um espaço linear o espaço $\mathrm{End}(\mathbf{V}) := \mathrm{Lin}(\mathbf{V}, \mathbf{V})$ dos *endomorfismos* de $\mathbf{V}$. Uma transformação linear bijetiva $L : \mathbf{V} \to \mathbf{W}$ é dita *isomorfismo (linear)* entre os espaços lineares $\mathbf{V}$ e $\mathbf{W}$.

ex: A transformação *identidade* $\mathbf{1}_n : \mathbb{R}^n \to \mathbb{R}^n$, definida por $\mathbf{x} \mapsto \mathbf{x}$, e transformação *nula* $\mathbf{0}_n : \mathbb{R}^n \to \mathbb{R}^n$, definida por $\mathbf{x} \mapsto \mathbf{0}$, são lineares.

ex: Se $L : \mathbf{V} \to \mathbf{W}$ é uma transformação linear, então $L(\mathbf{0}) = \mathbf{0}$ e $L(-\mathbf{v}) = -L(\mathbf{v})$.

ex: Uma transformação linear $L : \mathbf{V} \to \mathbf{W}$ envia retas afins $\mathbf{a} + \mathbb{R}\mathbf{v} \subset \mathbf{V}$ em retas afins $\mathbf{b} + \mathbb{R}\mathbf{w} \subset \mathbf{W}$, se $\mathbf{b} = L(\mathbf{a})$ e $\mathbf{w} = L(\mathbf{v}) \neq \mathbf{0}$, ou em pontos $\mathbf{b}$, se $L(\mathbf{v}) = \mathbf{0}$. Em particular, envia retas $\mathbb{R}\mathbf{v} \subset \mathbf{V}$ passando pela origem em retas $\mathbb{R}\mathbf{w} \subset \mathbf{W}$ passando pela origem.

ex: A inversa de uma transformação linear $L : \mathbf{V} \to \mathbf{W}$ bijetiva é uma transformação linear $L^{-1} : \mathbf{W} \to \mathbf{V}$.

ex: Uma transformação $T : \mathbb{R}^n \to \mathbb{R}^m$, definida por

$$(x_1, x_2, \dots, x_n) \mapsto (T_1(x_1, x_2, \dots, x_n), T_2(x_1, x_2, \dots, x_n), \dots, T_m(x_1, x_2, \dots, x_n))$$

é linear sse todas as suas "coordenadas" $T_i : \mathbb{R}^n \to \mathbb{R}$, com $i = 1, 2, \dots, m$, são lineares.

ex: Diga se as seguintes aplicações de $\mathbb{R}^n$ em $\mathbb{R}^m$ são lineares.

$$(x, y) \mapsto (3x - 5y, x - y) \qquad (x, y) \mapsto (x^2, xy)$$

$$(x, y, z) \mapsto (x, y + z, 2) \qquad (x, y, z) \mapsto (x, y + z, 0)$$

$$(x, y, z) \mapsto (x - y, z + y, 3x + 2y - z) \qquad (x, y, z) \mapsto (1, 2, 3)$$

$$(x_1, x_2, \dots, x_n) \mapsto (0, 0, 1) \qquad (x_1, x_2, \dots, x_n) \mapsto (0, 0, \dots, 0) \in \mathbb{R}^m$$

ex: Diga se as seguintes aplicações de $T : \mathbb{R}^2 \to \mathbb{R}^2$ são lineares.

T transforma cada ponto no seu simétrico em relação à reta $y = 0$

T transforma cada ponto no seu simétrico em relação à reta $y = x$

T transforma o ponto de coordenadas polares (r, θ) no ponto de coordenadas polares$(2r, \theta)$

T transforma o ponto de coordenadas polares (r, θ) no ponto de coordenadas polares $(r, \theta + \pi/2)$

ex: Seja $L : \mathbb{R}^2 \to \mathbb{R}^2$ uma transformação linear tal que $L(1, 1) = (1, 4)$ e $L(2, -1) = (-2, 3)$. Determine $L(5, -1)$ (observe que $1 + 2 \cdot 2 = 5$ e $1 + 2 \cdot (-1) = -1$).

ex: Determine a transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que

$$T(\mathbf{i}) = 2\mathbf{i} + \mathbf{j} \qquad \text{e} \qquad T(\mathbf{j}) = \mathbf{i} - 3\mathbf{j}$$

ex: Se $X \subset \mathbf{V}$ é um subespaço linear de $\mathbf{V}$ e $L : \mathbf{V} \to \mathbf{W}$ é uma transformação linear, então a imagem $L(X)$ é um subespaço linear de $\mathbf{W}$.

ex: As translações de $\mathbb{R}^n$, as transformações $\mathbf{v} \mapsto \mathbf{v} + \mathbf{a}$ com $\mathbf{a} \in \mathbb{R}^n$, são transformações lineares?

ex: As homotetias de $\mathbb{R}^n$, as transformações $\mathbf{v} \mapsto \lambda\mathbf{v}$ com $\lambda \in \mathbb{R}$, são transformações lineares?

ex: [Ap69] **16.4**.

Núcleo e imagem. Seja $L : \mathbf{V} \to \mathbf{W}$ uma transformação linear. O *núcleo/espaço nulo* (em inglês, *kernel*) de L é o subespaço vetorial

$$\boxed{\mathrm{Ker}(L) := \{\mathbf{v} \in \mathbf{V} \ \text{ t.q. } \ L(\mathbf{v}) = \mathbf{0}\} \subset \mathbf{V}}$$

A *imagem* de L é o subespaço vetorial

$$\boxed{\mathrm{Im}(L) := L(\mathbf{V}) = \{L(\mathbf{v}) \ \text{ com } \ \mathbf{v} \in \mathbf{V}\} \subset \mathbf{W}}$$

A dimensão do núcleo é dita *nulidade* de L, e a dimensão da imagem é dita *ordem* de L. Acontece que estes dois números satisfazem um "princípio de conservação".

Teorema 8.1. *Seja $L : \mathbf{V} \to \mathbf{W}$ uma transformação linear. Se $\mathbf{V}$ tem dimensão finita, então também a imagem $L(\mathbf{V})$ tem dimensão finita e*

$$\boxed{\dim_{\mathbb{R}} \mathrm{Ker}(L) + \dim_{\mathbb{R}} \mathrm{Im}(L) = \dim_{\mathbb{R}} \mathbf{V}}$$

Demonstração. Seja $\dim_{\mathbb{R}}(\mathbf{V}) = n$. Se os vetores $\mathbf{v}_1, , \dots, \mathbf{v}_k$ formam uma base de $\mathrm{Ker}(L)$, e se juntamente com os vetores $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ formam uma base de $\mathbf{V}$ (se $k < n$, caso contrário o teorema é trivial), então é imediato verificar que os vetores $\mathbf{w}_1 = L(\mathbf{v}_{k+1})$, ..., $\mathbf{w}_{n-k} = L(\mathbf{v}_n)$ geram $\mathrm{Im}(L)$ e são independentes. □

ex: Mostre que $\mathrm{Ker}(L)$ é um subespaço de $\mathbf{V}$ e que $\mathrm{Im}(L)$ é um subespaço de $\mathbf{W}$.

ex: Determine o núcleo, a imagem, a nulidade e a ordem das seguintes transformações lineares

$$L(x, y) = (x + y, x - y) \qquad L(x, y) = (y, -x) \qquad L(x, y) = (x + y, 3x - 2y)$$

$$L(x, y, z) = (2x, 3y, 0) \qquad L(x, y, z) = (x + y, y + z, z + x)$$

$$L(x, y, z) = (0, 0, 0) \qquad L(x, y, z) = (x, y)$$

ex: Determine o núcleo, a imagem, a nulidade e a ordem da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada ponto no seu simétrico em relação à reta $x = 0$

ex: Determine o núcleo, a imagem, a nulidade e a ordem da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada ponto no seu simétrico em relação à origem.

ex: [Ap69] **16.14**.

Operadores. Seja $\mathbf{V}$ um espaço linear (de dimensão não necessariamente finita!). Um *operador (linear)* em $\mathbf{V}$ é uma transformação linear $A : \mathbf{D} \to \mathbf{V}$ definida num subespaço linear $\mathbf{D} \subset \mathbf{V}$, dito *domínio* do operador A (esta definição é importante em análise, quando $\mathbf{V}$ é um espaço de dimensão infinita e o operador apenas pode ser definido num subespaço próprio de $\mathbf{V}$). Um subespaço linear $\mathbf{W} \subset \mathbf{D}$ é *invariante* se $A(\mathbf{W}) \subset \mathbf{W}$ (ou seja, se $\mathbf{v} \in \mathbf{W}$ implica $A\mathbf{v} \in \mathbf{W}$), e portanto a restrição de A a $\mathbf{W}$ é um endomorfismo $A|_{\mathbf{W}} : \mathbf{W} \to \mathbf{W}$.

Operadores derivação, multiplicação e primitivação. O operador *derivação* envia uma função derivável $f(x)$ na função

$$(Df)(x) := f'(x) .$$

Pode ser pensado como um operador $D : \mathcal{C}^{k+1}(\mathbb{R}) \to \mathcal{C}^k(\mathbb{R})$, ou também como un endomorfismo de $\mathcal{C}^\infty(\mathbb{R})$. O operador *multiplicação* envia uma função $f(x)$ na função

$$(Xf)(x) := x \cdot f(x) .$$

O operador *primitivação* envia uma função integrável (na reta real ou num intervalo da reta) $f(x)$ na função

$$(Pf)(x) := \int_c^x f(t)\, dt$$

Pode ser pensado como um operador $P : \mathcal{C}^k(\mathbb{R}) \to \mathcal{C}^{k+1}(\mathbb{R})$, ou como um endomorfismo de $\mathcal{C}^0(\mathbb{R})$ ou de $\mathcal{C}^\infty(\mathbb{R})$.

ex: Determine o núcleo e a imagem do operador derivação $D : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$,

ex: O subespaço $\mathrm{Pol}(\mathbb{R})$ dos polinómio é um subespaço invariante de $D : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$. O subespaço $\mathrm{Pol}_{\leq n}(\mathbb{R}) \subset \mathrm{Pol}(\mathbb{R})$ dos polinómios de grau $\leq n$ é um subespaço invariante do operador derivação $D : \mathrm{Pol}(\mathbb{R}) \to \mathrm{Pol}(\mathbb{R})$?

ex: Determine o núcleo $\mathrm{Ker}(P)$ e a imagem $\mathrm{Im}(P) = P(\mathcal{C}^0(\mathbb{R}))$.

Composição e álgebra dos endomorfismos. A composição de duas transformações lineares $L : \mathbf{V} \to \mathbf{W}$ e $M : \mathbf{W} \to \mathbf{Z}$, definida por $(ML)(\mathbf{v}) := M(L(\mathbf{v}))$, é uma transformação linear. Em particular, a composição de dois endomorfismos de um espaço vetorial $\mathbf{V}$ é um endomorfismo de $\mathbf{V}$. A n-ésima iterada do endomorfismo $L \in \mathrm{End}(\mathbf{V})$ é o endomorfismo L^n definido indutivamente por

$$L^0 = \mathbf{1}_{\mathbf{V}} , \qquad L^{n+1} = LL^n \qquad \text{se } n \geq 1 .$$

A composição de transformações lineares satisfaz as propriedades distributivas

$$(L + M)N = LN + MN \qquad L(M + N) = LM + LN$$

(que justificam a notação "multiplicativa" LM em vez de $L \circ M$).

A composição não é comutativa! Ou seja, em geral, não há razão para que LM seja igual a ML. Os endomorfismos $L, M \in \mathrm{End}(\mathbf{V})$ *comutam entre si/são permutáveis* se $LM = ML$. A obstrução é o *comutador*, definido por

$$[L, M] := LM - ML ,$$

que é igual a transformação nula sse L e M comutam.

ex: Calcule a composição ML quando

$$L(x, y) = (x + y, x - y) \qquad M(x, y) = 2x - 3y$$

$$L(x, y, z) = (x - y + z, z - y) \qquad M(x, y) = (x, y, x + y)$$

ex: Verifique que cada endomorfismo comuta com si próprio, ou seja, $[L, L] = 0$

ex: Calcule o comutador entre os endomorfismos do plano $E_+(x, y) = (y, 0)$ e $E(x, y) = (x, -y)$.

ex: Se $[L, M] = 0$ e $[M, N] = 0$, é verdade que $[L, N] = 0$?

ex: Determine todos os endomorfismos do plano que comutam com $I(x, y) = (x, y)$.

ex: Calcule o comutador $[D, X]$ entre os operadores derivação e multiplicação definidos em $\mathcal{C}^\infty(\mathbb{R})$.

Transformações lineares invertíveis. A transformação linear $L : \mathbf{V} \to \mathbf{W}$ é *invertível* se é *injetiva* (i.e. $L(\mathbf{x}) \neq L(\mathbf{x}'$ se $\mathbf{x} \neq \mathbf{x}'$, e portanto a equação $L(\mathbf{x}) = \mathbf{y}$ admite uma e uma única solução para cada $\mathbf{y} \in \mathrm{Im}(L)$), $\Leftrightarrow$ existe uma transformação linear $L^{-1} : \mathrm{Im}(L) \to \mathbf{V}$ (a *inversa* de L) tal que $L^{-1}L = \mathbf{1}_\mathbf{V}$ e $LL^{-1} = \mathbf{1}_{\mathrm{Im}(L)}$ $\Leftrightarrow$ o núcleo de L é trivial, i.e. $\mathrm{Ker}(L) = \{\mathbf{0}\}$.

Em particular, se $\mathbf{V}$ tem dimensão finita, a transformação linear $L : \mathbf{V} \to \mathbf{W}$ é invertível $\Leftrightarrow$ $\dim_\mathbb{R} \mathrm{Im}(L) = \dim_\mathbb{R} \mathbf{V}$ $\Leftrightarrow$ L transforma vetores independentes de $\mathbf{V}$ em vetores independentes de $\mathbf{W}$ $\Leftrightarrow$ L transforma bases de $\mathbf{V}$ em bases de $\mathrm{Im}(L)$.

Os *automorfismos* $\mathrm{Aut}(\mathbf{V})$ de um espaço linear $\mathbf{V}$ são os endomorfismos $L : \mathbf{V} \to \mathbf{V}$ invertíveis.

- Diga se L é injetiva e, caso afirmativo, determine a imagem $\mathrm{Im}(L)$ e a transformação inversa L^{-1}.

$$L(x, y) = (x, x) \qquad L(x, y) = (y, x)$$
$$L(x, y) = (x - y, x + y) \qquad L(x, y) = (0, y)$$
$$L(x, y, z) = (x + y, y + z, z + x) \qquad L(x, y, z) = (3x, 2y, z)$$
$$L(x, y, z) = (y, z, 0) \qquad L(x, y, z) = (x + y + z, y, z)$$
$$L(x, y) = (x, 0, y) \qquad L(x, y) = (x - x, x + y, 0)$$

- [Ap69] **16.8**.
- Mostre que se $L : \mathbf{V} \to \mathbf{V}$ é invertível então também L^n é invertível e $(L^n)^{-1} = (L^{-1})^n$.
- Mostre que se $L : \mathbf{V} \to \mathbf{V}$ e $M : \mathbf{V} \to \mathbf{V}$ comutam, então também as inversas L^{-1} e M^{-1} comutam, e $(LM)^n = L^n M^n$.
- Considere, no espaço $\mathcal{C}^\infty(\mathbb{R})$, o operador *derivação*, definido por $(Df)(t) := f'(t)$, e o operador *primitivação*, definido por $(Pf)(t) := \int_0^t f(s)\, ds$. Mostre que $DP = I$ mas $PD \neq I$. Descreva o núcleo e a imagem de PD.
- O operador *multiplicação*, definido por $(Xf)(x) := x\, f(x)$, em $\mathcal{C}^\infty(\mathbb{R})$, é invertível?
- O operador *deslocamento* $\sigma : \mathbb{R}^\mathbb{N} \to \mathbb{R}^\mathbb{N}$, definido por

$$(x_1, x_2, x_3, \dots) \mapsto (x_2, x_3, x_4, \dots)\,,$$

é invertível?

9 Matrizes e transformações lineares

ref: [Ap69] *Vol 2*, **2.10-16** ; [La97] *Ch. IV* , **5**

23 nov 2018

Matrizes. Uma *matriz* real (ou complexa) $m \times n$ é uma tabela

$$\boxed{A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}}$$

de $m \cdot n$ números reais (ou complexos) dispostos em m linhas e n colunas. O número real (ou complexo) a_{ij} é dito *elemento/componente/entrada* ij da matriz A. Os vetores

$$\boxed{A^i = (a_{i1}, a_{i2}, \dots, a_{in}) \in (\mathbb{R}^n)^* \approx \mathbb{R}^n \qquad \text{e} \qquad A_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in \mathbb{R}^m}$$

são ditos *i-ésima linha* e *j-ésima coluna* da matriz A, respetivamente. Em particular, uma matriz com m linhas e apenas uma coluna é um vetor de $\mathbb{R}^m$ representado como um vetor coluna, e uma matriz com n colunaas e apenas uma linha é um vetor de $\mathbb{R}^n$ representado como um vetor linha. Se o número de linhas é igual ao número de colunas, i.e. $n = m$, a matriz é dita "quadrada".

Espaço linear das matrizes. O espaço $\mathrm{Mat}_{m\times n}(\mathbb{R})$ (ou $\mathrm{Mat}_{m\times n}(\mathbb{C})$) das matrizes reais (ou complexas) $m \times n$ é um espaço linear real (ou complexo) se a adição e a multiplicação por um escalar são definidas por

$$A + B := (a_{ij} + b_{ij}) \qquad \text{e} \qquad \lambda A := (\lambda a_{ij})$$

se $A = (a_{ij})$ e $B = (b_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, e $\lambda \in \mathbb{R}$. O elemento neutro, ou vetor nulo, é a "matriz nula" $0 = (0)$, cujas entradas são todas nulas, e que satisfaz $A + 0 = A$ para toda a matriz A. A matriz "oposta" da matriz A é a matriz $-A := (-1)A$, tal que $A + (-A) = 0$. Então, podemos simplificar a notação e escrever $A - B := A + (-B)$.

É claro que dimensão do espaço $\mathrm{Mat}_{n\times n}(\mathbb{R})$ é o produto $m \cdot n$, o número de elementos das matrizes. De fato, uma base é o conjunto das matrizes I_{ij}, que têm 1 na entrada ij e 0 nas outras. Então toda matriz é uma combinação linear única $A = \sum_{ij} a_{ij} I_{ij}$, com coordenadas a_{ij}'s, e esta representação define um isomorfismo $\mathrm{Mat}_{m\times n}(\mathbb{R}) \approx \mathbb{R}^{m\cdot n}$.

ex: Verifique que

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

é uma base do espaço linear $\mathrm{Mat}_{2\times 2}(\mathbb{R})$.

ex: [Ap69] **16.16**.

Álgebra das matrizes. Se $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ e $B = (b_{ij}) \in \mathrm{Mat}_{n\times s}(\mathbb{R})$, o *produto (linhas por colunas)* AB é a matriz $AB = C = (c_{ij}) \in \mathrm{Mat}_{m\times s}(\mathbb{R})$ definida por

$$\boxed{c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj}}$$

Ou seja, o elemento c_{ij} de $C = AB$ é o produto escalar $c_{ij} = A^i \cdot B_j$ da i-ésima linha de A e a j-ésima coluna de B. Observe que o produto AB apenas pode ser defindo quando o número de colunas de A é igual ao número de linhas de B.

A "matriz identidade" é a matriz quadrada $I \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ definida por

$$I = (\delta_{ij}) := \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix} .$$

(também denotada por I_n quando é necessário lembrar a dimensão). É claro que $IA = A$ e $BI = B$ para todas as matrizes A e B (quando o produto faz sentido). O produto é "associativo",

$$\boxed{A(BC) = (AB)C}$$

e satisfaz as "propriedades distributivas" à esquerda e à direita,

$$\boxed{A(B+C) = AB + AC \qquad \text{e} \qquad (A+B)C = AC + BC}$$

ex: Existem matrizes $A \neq 0$ e $B \neq 0$ tais que $AB = 0$?

ex: [Ap69] **16.16**.

Matriz de uma transformação linear. Uma transformação linear $L : \mathbf{V} \to \mathbf{W}$ definida num espaço de dimensão finita $\mathbf{V}$ é determinada pelos seus valores nos vetores de uma base. De fato, se $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ é uma base ordenada de $\mathbf{V} \approx \mathbb{R}^n$, e definimos os vetores $\mathbf{w}_j := L(\mathbf{b}_j) \in \mathbf{W}$, com $j = 1, 2, \dots, n$, então o valor de L sobre o vetor genérico $\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n$ é

$$\begin{aligned} L(\mathbf{x}) &= L(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n) \\ &= x_1 L(\mathbf{b}_1) + x_2 L(\mathbf{b}_2) + \cdots + x_n L(\mathbf{b}_n) \\ &= x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \cdots + x_n\mathbf{w}_n . \end{aligned}$$

Em particular, os vetores $\mathbf{w}_1, \dots, \mathbf{w}_n$ geram $\mathrm{Im}(L) \subset \mathbf{W}$. Se também $\mathbf{W}$ tem dimensão finita, e $\mathcal{C} = (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_m)$ é uma base ordenada de $\mathbf{W} \approx \mathbb{R}^m$, e definimos os números a_{ij} como sendo as coordenadas dos $\mathbf{w}_j$' s relativamente à base $\mathcal{C}$, i.e.

$$L(\mathbf{b}_j) = \mathbf{w}_j = a_{1j}\,\mathbf{c}_1 + a_{2j}\,\mathbf{c}_2 + \cdots + a_{mj}\,\mathbf{c}_m$$

com $j = 1, 2, \dots, n$, então o valor de L sobre o vetor genérico $\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n$ é

$$L(\mathbf{x}) = L\left(\sum_j x_j\mathbf{b}_j\right) = \sum_{j=1}^{m}\left(\sum_{i=1}^{n} a_{ij}\,x_j\right)\mathbf{c}_i ,$$

Portanto, as coordenadas do vetor $\mathbf{y} = L(\mathbf{x})$ na base escolhida $\mathcal{C}$ são

$$\boxed{y_i = \sum_{j=1}^{n} a_{ij}\,x_j \qquad\qquad i = 1, 2, \dots, m .}$$

Os números a_{ij} são os elementos de uma matriz $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, a matriz que representa L nas bases escolhidas.

Transformação linear definida por uma matriz. Em particular, fixadas as bases canónicas dos espaços $\mathbb{R}^n$ e $\mathbb{R}^m$ (ou quaisquer outras bases), uma matriz $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ define uma transformação linear $L_A : \mathbb{R}^n \to \mathbb{R}^m$. Se X e Y denotam os "vetores coluna" (matrizes com apenas uma coluna)

$$X := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \qquad \text{e} \qquad Y := \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{R}^m$$

e

$$A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in \mathrm{Mat}_{m\times n}(\mathbb{R}),$$

então a transformação linear $L_A : \mathbb{R}^n \to \mathbb{R}^m$ é representada pela equação matricial

$$\boxed{X \mapsto Y = L_A(X) := AX}$$

ou seja, explicitamente,

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Caraterística. Seja $L_A : \mathbb{R}^n \to \mathbb{R}^m$ a transformação linear $X \mapsto AX$ definida pela matriz $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$. Se E_1, E_2, $\dots E_n$ denotam os vetores coluna da base canónica de $\mathbb{R}^n$, ou seja,

$$E_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad E_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}. \qquad \dots \qquad E_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

então o produto $AE_k = A_k$ é a k-ésima coluna da matriz A, que representa portanto a imagem $L_A(E_k)$ do vetor E_k. A imagem do vetor genérico X é portanto uma combinação linear

$$AX = x_1A_1 + x_2A_2 + \cdots + x_nA_n$$

das colunas da matriz A. A dimensão de $\mathrm{Im}(L_A) \subset \mathbb{R}^m$, ou seja, o número de colunas linearmente independentes de A, é dita *caraterística* da matriz A, e denotada por $\mathrm{rank}(A) := \dim \mathrm{Im}(L_A)$.

Por outro lado, o vetor $X \in \mathbb{R}^n$ pertence ao espaço nulo $\ker(L_A)$ da transformação linear L_A se $AX = 0$, ou seja, se

$$A^1 \cdot X = 0\,, \quad A^2 \cdot X = 0\,, \qquad \dots, \quad A^m \cdot X = 0\,,$$

onde $A^i = (a_{i1}, a_{i2}, \dots, a_{in}) \in \mathbb{R}^n$ é a i-ésima linha da matriz A. O espaço nulo é portanto o espaço ortogonal ao subespaço vetorial $\mathrm{Span}(A^1, A^2, \dots, A^m) \subset \mathbb{R}^n$ gerado pelas linhas de A, ou seja,

$$\mathrm{Ker}(L_A) = \mathrm{Span}(A^1, A^2, \dots, A^m)^\perp\,.$$

A sua dimensão é igual a $n-k$, se k é o número de linhas linearmente independentes de A. Em particular, sendo $\dim\ker(L_A) + \dim\mathrm{Im}(L_A) = n$, a caraterística da matriz $\mathrm{rank}(A)$ é também igual ao número de linhas linearmente independentes de A.

Produto e composição. Se a matriz $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ define a transformaçõa linear $L_A : \mathbb{R}^n \to \mathbb{R}^m$, e se uma segunda matriz $B = (b_{ij}) \in \mathrm{Mat}_{p\times m}(\mathbb{R})$ define a transformação linear $L_B : \mathbb{R}^m \to \mathbb{R}^p$, então a matriz produto $BA \in \mathrm{Mat}_{p\times n}(\mathbb{R})$ define a composição $L_B L_A : \mathbb{R}^n \to \mathbb{R}^p$. De fato, se as coordenadas de $Y = L_A(X)$ são $y_k = \sum_j a_{kj}\, x_j$ e as coordenadas de $Z = L_B(Y)$ são $z_i = \sum_k b_{ik}\, y_k$, então

$$z_i = \sum_k \sum_j b_{ik}\, a_{kj}\, x_j = \sum_j \left(\sum_k b_{ik}\, a_{kj} \right) x_j \qquad i = 1, 2, \dots, p\,,$$

Em notação matricial,

$$\boxed{\text{se} \quad Y = AX \quad \text{e} \quad Z = BY \quad \text{então} \quad Z = BAX}$$

Ou seja, o produto linhas por colunas representa a composição das transformações lineares.

ex: Determine a matriz que define a transformação

$$L(x,y) = (x-y, 2x-3y) \qquad L(x,y,z) = (3x+y-z, -x+2y+z)$$

$$L(x,y,z) = (3x,3y,3z) \qquad L(x,y) = (x+y, x-y, 2x-7y)$$

$$L(x,y,z) = (x,y) \qquad L(x,y,z) = (x,z)$$

ex: Determine a transformação linear $L : \mathbb{R}^n \to \mathbb{R}^m$ definida pela matriz

$$\begin{pmatrix} 1 & 3 \\ -2 & 0 \end{pmatrix} \qquad \begin{pmatrix} -5 & 0 & -1 \\ 3 & 1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 0 & -1 \\ 0 & 8 \\ -1 & -3 \end{pmatrix} \qquad \begin{pmatrix} 1 & 3 \\ -2 & 0 \end{pmatrix}$$

ex: Determine a matriz 2×2 que define a transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que

transforma cada ponto no seu simétrico em relação à reta $x = 0$

transforma cada ponto no seu simétrico em relação à reta $y = -x$

transforma o ponto de coordenadas polares (r, θ) no ponto de coordenadas polares$(r/2, \theta)$

transforma o ponto de coordenadas polares (r, θ) no ponto de coordenadas polares $(r, \theta - \pi/2)$

ex: [Ap69] **16.12**.

Einstein's notation. It is often important to take care of the distinction between vectors and co-vectors, and then different types of tensors, as well as to shorten formulas and computations. One possibility is the convention introduced by Einstein. We denote vectors using upper indices as $\mathbf{x} = (x^i) \in \mathbb{R}^n$, and denote co-vectors using lower indices as $\boldsymbol{\xi} = (\xi_j) \in (\mathbb{R}^n)^*$. The pairing between a co-vector $\boldsymbol{\xi}$ and a vector $\mathbf{x}$ reads $\langle \boldsymbol{\xi}, \mathbf{x} \rangle = \xi_1 x^1 + \xi_2 x^2 + \cdots + \xi_n x^n$, and is shortened as $\langle \boldsymbol{\xi}, \mathbf{x} \rangle = \xi_i x^i$, Einstein's convention being that a repeated index which appears once as an upper index and once as a lower index implies summation. Using Einstein's sum convention, the coordinates of the image of a linear transformation represented by the matrix $T = (t^i{}_j)$ are given by $y^i = t^i{}_j\, x^j$. The composition of $T = (t^i{}_j)$ followed by $S = (s^i{}_j)$ is then represented by the matrix $ST = (s^i{}_k\, t^k{}_j)$.

Endomorfismos e matrizes quadradas. Fixada uma base (por exemplo, a base canónica de $\mathbb{R}^n$), o espaço linear $\mathrm{End}(\mathbf{V}) := \mathrm{Lin}(\mathbf{V}, \mathbf{V})$ dos *endomorfismos* de um espaço linear $\mathbf{V} \approx \mathbb{R}^n$ é isomorfo ao espaço linear $\mathrm{Mat}_{n\times n}(\mathbb{R})$ das matrizes "quadradas" $n \times n$, os seus elementos sendo as transformações lineares

$$\boxed{X \mapsto Y = AX}$$

ou seja, $x_i \mapsto y_i = \sum_j a_{ij}\, x_j$, com $X \in \mathbb{R}^n$ e $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$. O caso complexo é análogo.

A composição de dois endomorfismos corresponde ao produto de matrizes. Em particular, a k-ésima iterada $L^k = L \circ L \circ \cdots \circ L$ (k vezes) do endomorfismo $L : X \mapsto AX$ é representada pela *potência* k-ésima de A, definida recursivamente por

$$A^0 = I \qquad \text{e} \qquad A^{k+1} = AA^k \qquad \text{se } k \geq 0\,.$$

A matriz quadrada A (ou o endomorfismo $L : X \mapsto AX$) é *nilpotente* se existe um inteiro k tal que $A^k = 0$. A matriz quadrada A é *unipotente* se $A - I$ é nilpotente, e portanto existe um inteiro k tal que $(A - I)^k = 0$.

A *diagonal* da matriz quadrada $A = (a_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é o conjunto ordenado dos elementos "diagonais" $a_{11}, a_{22}, \dots, a_{nn}$. O *traço* (em inglês, *trace*) de A é a soma dos elementos da diagonal,

$$\boxed{\mathrm{tr}(A) := \sum_i a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}}$$

É imediato verificar que $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Uma matriz quadrada é uma *matriz diagonal* se os elementos que não pertencem à diagonal são nulos, ou seja, se é da forma

$$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) := \begin{pmatrix} \lambda_1 & 0 & 0 & \dots \\ 0 & \lambda_2 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

ex: Determine a matriz da transformação "identidade" $\mathbf{x} \mapsto \mathbf{x}$ e da transformação "nula" $\mathbf{x} \mapsto \mathbf{0}$.

ex: Determine a matriz da homotetia $\mathbf{x} \mapsto \lambda\mathbf{x}$, com $\lambda \in \mathbb{R}$, e calcule o seu traço.

ex: Mostre que

$$\mathrm{tr}(AB) = \mathrm{tr}(BA)$$

ex: Calcule A^0, A^1, A^2, A^3, ... , A^k , ... quando

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

ex: Determine as matrizes $A \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$ tais que $A^2 = 0$.

ex: Determine as matrizes $A \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$ tais que $A^2 = I$.

ex: Mostre (a matriz que representa) uma projeção ortogonal $P : \mathbb{R}^n \to \mathbb{R}^n$ sobre um subespaço $\mathbb{R}^m \subset \mathbb{R}^n$ (por exemplo, $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_m, 0, \dots, 0)$) é unipotente.

ex: Mostre qua (a matriz que representa) o operador derivação $Df := f'$ no espaço $\mathrm{Pol}_{\leq n} \approx \mathbb{R}^{n+1}$ dos polinómios de grau $\leq n$ é nilpotente.

Comutador. A composição de transformações lineares, e portanto o produto de matrizes, não são comutativos! Ou seja, em geral, $AB \neq BA$. As matrizes quadradas $A, B \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ (e portanto os endomorfismos que representam), *comutam entre si/são permutáveis* se

$$AB = BA\,.$$

A obstrução é o *comutador*, definido por

$$\boxed{[A,B] := AB - BA}$$

O comutador satisfaz a *identidade de Jacobi*

$$\boxed{[[A,B],C] + [[B,C],A] + [[C,A],B] = 0}$$

ex: Mostre que cada matriz quadrada A comuta com si própria, i.e. $[A,A]=0$.

ex: Mostre que duas matrizes diagonais comutam.

ex: Considere as matrizes 2×2

$$E = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad E_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad E_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Calcule $[E, E_+]$, $[E, E_-]$ e $[E_+, E_-]$.

e.g. Rotações do plano. Uma rotação anti-horária de um ângulo θ envia o ponto de coordenadas polares (ρ, φ) no ponto de coordenadas polares $(\rho, \varphi+\theta)$ (e fixa a origem). Então envia o ponto de coordenadas cartesianas $(r\cos\varphi, r\sin\varphi)$ no ponto de coordenadas cartesianas $(r\cos(\varphi+\theta), r\sin(\varphi+\theta))$. As fórmulas de adição implicam que

$$\begin{pmatrix} r\cos(\varphi+\theta) \\ r\sin(\varphi+\theta) \end{pmatrix} = \begin{pmatrix} r\,(\cos\varphi\cos\theta - \sin\varphi\sin\theta) \\ r\,(\sin\varphi\cos\theta + \cos\varphi\sin\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \cos\theta & \sin\theta \end{pmatrix} \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}$$

e portanto que a rotação é uma transformação linear definida pela matriz

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \cos\theta & \sin\theta \end{pmatrix}.$$

Em particular, uma rotação de um ângulo nulo, ou múltiplo inteiro de 2π, é a transformação identidade, definida pela matriz $R_0 = I$. As fórmulas de adição também mostram que as potências de uma rotação são rotações, e de fato

$$R_\theta^2 = R_{2\theta} \quad R_\theta^3 = R_{3\theta} \qquad \dots \qquad R_\theta^n = R_{n\theta}\,.$$

Também é evidente que $R_\theta R_{-\theta} = I$. Mais em geral, a composição de duas rotações de ângulos θ e ϕ é uma rotação de um ângulo $\theta+\phi$, ou seja,

$$R_\theta R_\phi = R_{\theta+\phi}\,.$$

e.g. Cisalhamentos. Um *cisalhamento* (em inglês, *shear*) horizontal é uma transformação do plano $(x,y) \mapsto (x+\alpha y, y)$, definida pela matriz

$$C_\alpha = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix},$$

com $\alpha \in \mathbb{R}$ (fazer um desenho, para observar que o quadrado unitário de lados $\mathbf{i}$ e $\mathbf{j}$ é enviado no paralelogramo de lados $\mathbf{i}$ e $\alpha\mathbf{i}+\mathbf{j}$). Analogamente é possíve definir cisalhamentos verticais. É evidente que as potências de um cisalhamento são ainda cisalhamentos, e que de fato

$$C_\alpha^2 = \begin{pmatrix} 1 & 2\alpha \\ 0 & 1 \end{pmatrix} \qquad \dots \qquad C_\alpha^n = \begin{pmatrix} 1 & n\alpha \\ 0 & 1 \end{pmatrix}$$

e.g. Reflexões no plano. Uma reflexão $T : \mathbb{R}^2 \to \mathbb{R}^2$ ao longo de uma reta passando pela origem satisfaz a identidade $T^2 = I$. Por exemplo, as matrizes $\pm E$, com

$$E = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} ,$$

definem reflexões nos eixos dos x e dos y, respetivamente. A reflexão numa reta genérica de equação cartesiana $y \cos\theta = x \sin\theta$ (ou seja, com declive $\tan\theta$) pode ser obtida como a composição

$$R_\theta \, E \, R_{-\theta} \, ,$$

ou seja, transformando a reta dada no eixo dos x, aplicando a reflexão no eixo dos x, e depois voltando a transformar o eixo dos x na reta dada.

e.g. Affine transformations. The affine transformations of the plane $\mathbf{r} \mapsto \mathbf{r}' = L\mathbf{r} + \mathbf{a}$, where L is the linear transformation defined by a matrix $A = (a_{ij}) \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$ and $\mathbf{a} = (\alpha, \beta) \in \mathbb{R}^2$, are not linear transformations. Nevertheless, they may be represented with the aid of a 3×3 matrix according to

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & a_{12} & \alpha \\ a_{21} & a_{22} & \beta \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Matrizes transpostas. Seja $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ uma matriz. A *matriz transposta* é a matriz $A^\top = (a^\top_{ij}) \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, definida por $a^\top_{ij} = a_{ji}$ (ou seja, as linhas de $A^\top$ são as colunas de A e vice-versa). É imediato verificar que $(A^\top)^\top = A$ e que $(AB)^\top = B^\top A^\top$ (quando o primeiro produto faz sentido) .

Por exemplo, se X e Y são dois vetores/matrizes coluna de $\mathbb{R}^n$, então $Y^\top$ é um vetor/matriz linha, e o produto linha por coluna $Y^\top X$ é igual ao produto escalar

$$Y^\top X = Y \cdot X \, .$$

Em geral, se X é um vetor coluna de $\mathbb{R}^n$ e Y é um vetor coluna de $\mathbb{R}^m$, e se a matriz $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ define a transformação linear $X \mapsto AX$, então a matriz transposta define a transformação linear $Y \mapsto A^\top Y$ tal que

$$Y \cdot AX = (A^\top Y) \cdot X \, ,$$

pois $Y \cdot AX = Y^\top (AX) = (A^\top Y)^\top X = (A^\top Y) \cdot X$.

Uma matriz quadrada A é dita *simétrica* se $A = A^\top$ e *anti-simétrica* se $A^\top = -A$.

ex: Verifique que $(A^\top)^\top = A$.

ex: Mostre que $\mathrm{tr}(A^\top) = \mathrm{tr}(A)$.

ex: Mostre que, se A é uma matriz quadrada, então $A + A^\top$ é simétrica, e $A - A^\top$ é anti-simétrica. Deduza que cada matriz quadrada pode ser decomposta como soma $A = A_+ + A_-$ de uma matriz simétrica $A_+ = (A + A^\top)/2$ com uma matriz anti-simétrica $A_- = (A - A^\top)/2$.

ex: Mostre que o traço de uma matriz (quadrada) anti-simétrica é nulo.

10 Sistemas lineares

ref: [Ap69] *Vol 1*, **16.17-18** ; [La97] *Ch. II*

30 nov 2018

Peppermint Patty's problems.

ex: "In driving from town A to town D you pass first through town B and then through town C. It is 10 miles farther from A to B than from B to C and 10 miles farther from B to C than from C to D. If it is 390 miles from A do D, how far is it from A to B?"[12]

ex: "A man has a daughter and a son.. The son is three years older than the daughter ... In one year the man will be six times as old as the daughter is now, and in ten years he will be fourteen years older than the combined ages of his children ... What is the man's present age?"

ex: "A man has twenty coins consisting of dimes and quarters[13] ... If the dime were quarters and the quarters were dimes, he would have ninety cents more than he has now ... How many dimes and quarters does he have?"

e.g. Equações lineares na reta. Uma equação linear

$$ax = b$$

na reta real $\mathbb{R}$ (ou na reta complexa $\mathbb{C}$, ou, em geral, num corpo), com $a \neq 0$ (caso contrário é apenas a afirmação $b = 0$), admite uma única solução $x = b/a$.

e.g. Equações lineares no plano. Uma equação linear

$$ax + by = c$$

no plano cartesiano $\mathbb{R}^2$, com $\mathbf{n} = (a, b) \neq (0, 0)$, define uma reta afim $R \subset \mathbb{R}^2$. A equação homogénea associada

$$ax + by = 0$$

define uma reta que passa pela origem, ou seja, um subespaço vetorial $\mathbf{n}^\perp = \mathbb{R}\mathbf{v} \subset \mathbb{R}^2$ de dimensão 1 (por exemplo, com $\mathbf{v} = (b, -a)$). Se $\mathbf{r}_0 = (x_0, y_0)$ é um ponto de R, ou seja, (apenas) uma solução de $ax+by = c$, então o espaço de todas as soluções é $R = \mathbf{r}_0 + \mathbb{R}\mathbf{v}$. Ou seja, as soluções de $ax+by = c$ são dadas por

$$(x, y) = \mathbf{r}_0 + t\mathbf{v} = (x_0, y_0) + t\,(b, -a)$$

ao variar o parâmetro $t \in \mathbb{R}$.

[12] Peppermint Patty, in *Peanuts*, by Charles M. Schulz, December 6th, 1968.
[13] A *dime* is a 10 cents coin, and a *quarter* is a 25 cents coin.

Um sistema de duas equações lineares

$$\begin{cases} ax + by &= c \\ a'x + b'y &= c' \end{cases}$$

descreve a interseção entre duas retas afins $(R_1 \cap R_2) \subset \mathbb{R}^2$. Esta interseção pode ser vazia (retas paralelas e distintas), pode ser uma reta $ax + by = c$ (equações proporcionais/equivalentes), ou pode ser um único ponto. A última possibilidade é o caso genérico, e o sistema é equivalente (eliminando x na segunda equação, se $a \neq 0$) ao sistema "em escada de linhas"

$$\begin{cases} ax + & by &= c \\ & b''y &= c'' \end{cases}$$

com $a \neq 0$ e $b'' \neq 0$, e portanto ao sistema "diagonal"

$$\begin{cases} x & &= \alpha \\ & y &= \beta \end{cases}$$

com $\beta = c''/b''$ e $\alpha = (c - b\beta)/a$.

ex: Resolva, se possível, os seguintes sistemas lineares

$$\begin{cases} x + y &= 0 \\ x - y &= 0 \end{cases} \qquad \begin{cases} x + y &= 20 \\ x - y &= 2 \end{cases} \qquad \begin{cases} 2x - y &= 13 \\ -x + 2y &= 7 \end{cases}$$

e.g. Equações lineares no espaço. Uma equação linear

$$ax + by + cz = d$$

no espaço $\mathbb{R}^3$, com $\mathbf{n} = (a, b, c) \neq (0, 0, 0)$, define um plano afim $P = \{\mathbf{n} \cdot \mathbf{r} = d\} \subset \mathbb{R}^3$. A equação homogénea associada

$$ax + by + cz = 0$$

define o supespaço vetorial $\mathbf{n}^\perp \subset \mathbb{R}^3$.

Um sistema de duas equações lineares

$$\begin{cases} ax + by + cz &= d \\ a'x + b'y + c'z &= d' \end{cases}$$

descreve a interseção entre dois planos afins $(P_1 \cap P_2) \subset \mathbb{R}^3$. Esta interseção pode ser vazia (dois planos paralelos e distintos), pode ser um plano $ax + by + cz = d$ (duas equações proporcionais/equivalentes), ou pode ser uma reta. A última possibilidade é o caso genérico, e o sistema é equivalente (eliminando x na segunda equação, se $a \neq 0$) ao sistema "em escada de linhas"

$$\begin{cases} ax + & by + & cz &= d \\ & b''y + & c''z &= d'' \end{cases}$$

A última variável pode ser pensada como um parâmetro $z = t$ da reta:

$$t \mapsto (\alpha t + \gamma, \beta t + \delta, t)\,.$$

Um sistema de três equações lineares

$$\begin{cases} ax + by + cz &= d \\ a'x + b'y + c'z &= d' \\ a''x + b''y + c''z &= d'' \end{cases}$$

descreve a interseção entre três planos afins $(P_1 \cap P_2 \cup P_3) \subset \mathbb{R}^3$. Esta interseção pode ser vazia (dois planos paralelos e distintos, ou um plano paralelo à reta de interseção entre os outros dois), pode ser um plano $ax + by + cz = d$ (equações proporcionais/equivalentes), pode ser uma reta (sistema equivalente a um sistema de duas equações), ou pode ser um único ponto. A última possibilidade é o caso genérico, e o sistema é equivalente (eliminando x na segunda e na terceira equação, se $a \neq 0$, e depois y na terceira, se $b''' \neq 0$) ao sistema "em escada de linhas"

$$\left\{ \begin{array}{llll} ax+ & by+ & cz & = d \\ & b'''y+ & c'''z & = d''' \\ & & c''''z & = d'''' \end{array} \right.$$

com $a \neq 0$, $b''' \neq 0$ e $c'''' \neq 0$, e portanto ao sistema "diagonal"

$$\left\{ \begin{array}{llll} x & & & = \alpha \\ & y & & = \beta \\ & & z & = \gamma \end{array} \right.$$

com $\gamma = d''''/c''''$, $\beta = (d''' - c'''\gamma)/b'''$ e $\alpha = (d - c\gamma - b\beta)/a$.

ex: Resolva, se possível, os seguintes sistemas lineares

$$\left\{ \begin{array}{rl} 3x - y & = 0 \\ x + y + z & = 1 \end{array} \right. \qquad \left\{ \begin{array}{rcl} x + y + z & = & 1 \\ x + y - z & = & 0 \end{array} \right. \qquad \left\{ \begin{array}{lcl} x & = & 1 \\ x + y & = & 2 \\ x + y + z & = & 3 \end{array} \right.$$

Sistemas lineares. Um *sistema* de m equações lineares nas *incógnitas* $x_1, x_2, \dots, x_n$ é um conjunto de equações

$$\boxed{\begin{array}{lcl} a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1n}\,x_n & = & b_1 \\ a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2n}\,x_n & = & b_2 \\ \ddots & & \vdots \\ a_{m1}\,x_1 + a_{m2}\,x_2 + \cdots + a_{mn}\,x_n & = & b_m \end{array}} \tag{10.1}$$

com a_{ij} e b_k números reais (ou complexos). A matriz $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ é dita *matriz dos coeficientes* do sistema, e os números b_k são chamados *termos independentes*, coordenadas de um vetor $\mathbf{b} = (b_1, b_2, \dots, b_m) \in \mathbb{R}^m$.

Uma *solução* do sistema linear é um vetor $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ cujas coordenadas satisfazem as m equações (10.1). Um sistema linear pode ter uma solução única (sistema *determinado*), ter uma família (uma reta afim, um plano afim, ...) de soluções (sistema *indeterminado*), ou não ter nenhuma solução (sistema *impossível*).

O *sistema homogéneo* correspondente/associado é o sistema

$$\boxed{\begin{array}{lcl} a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1n}\,x_n & = & 0 \\ a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2n}\,x_n & = & 0 \\ \ddots & & \vdots \\ a_{m1}\,x_1 + a_{m2}\,x_2 + \cdots + a_{mn}\,x_n & = & 0 \end{array}} \tag{10.2}$$

onde todos os termos independentes são nulos. Um sistema homogéneo admite pelo menos a *solução trivial* $\mathbf{0} = (0, 0, \dots, 0)$. Se $\mathbf{x} \in \mathbb{R}^n$ é uma solução de um sistema homogéneo, então todos os pontos $t\mathbf{x}$ da reta $\mathbb{R}\mathbf{x}$ também são soluções. Se $\mathbf{x}$ e $\mathbf{y}$ são soluções de um sistema homogéneo, então também a soma $\mathbf{x} + \mathbf{y}$ é solução. Portanto, o espaço das soluções de um sistema homogéneo (10.2) é um espaço linear.

Soluções de um sistema linear. O sistema linear (10.1) é equivalente a $L_A(X) = B$, ou seja,

$$\boxed{AX = B}$$

onde $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ é a matriz dos coeficientes, $B = \mathbf{b}^\top \in \mathbb{R}^m$ e $X = \mathbf{x}^\top \in \mathbb{R}^n$ são vetores coluna, e $L_A : \mathbb{R}^n \to \mathbb{R}^m$ é a transformação linear $L_A(X) = AX$.

O vetor $X = (x_1, x_2, \dots, x_n)^\top \in \mathbb{R}^n$ é solução do sistema $AX = B$ se

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

ou seja, se

$$x_1 A_1 + x_2 A_2 + \cdots + x_n A_n = B$$

onde $A_j = (a_{1j}, a_{2j}, \dots, a_{mj})^\top \in \mathbb{R}^m$ é a j-ésima coluna da matriz A. Portanto, o sistema admite (pelo menos) uma solução (i.e. é *possível*) sse B é uma combinação linear das colunas da matriz A, i.e. sse $B \in \mathrm{Im}(L_A) = \mathrm{Span}(A_1, A_2, \dots, A_n)$. A dimensão $r = \dim \mathrm{Im}(L_A)$ da imagem de L_A, ou seja, o número de colunas linearmente independentes de A, é dita *caraterística* da matriz A, e denotada por $r = \mathrm{rank}(A) := \dim \mathrm{Im}(L_A)$.

O sistema linear homogéneo (10.2) é equivalente a $L_A(X) = 0$, ou seja,

$$\boxed{AX = 0}$$

O vetor $X = (x_1, x_2, \dots, x_n)^\top \in \mathbb{R}^n$ é solução do sistema homogéneo se

$$A^1 \cdot X = 0\,, \quad A^2 \cdot X = 0\,, \qquad \dots\,, \quad A^m \cdot X = 0\,,$$

onde $A^i = (a_{i1}, a_{i2}, \dots, a_{in}) \in \mathbb{R}^n$ é a i-ésima linha da matriz A, ou seja, se é ortogonal ao espaço vetorial $\mathrm{Span}(A^1, A^2, \dots, A^m)$ gerado pelas linhas de A. Portanto, o espaço das soluções do sistema homogéneo é

$$\mathrm{Ker}(L_A) = \mathrm{Span}(A^1, A^2, \dots, A^m)^\perp$$

e a sua dimensão é igual a $n - k$, se k é o número de linhas linearmente independentes de A. Em particular, sendo $\dim \ker(L_A) + \dim \mathrm{Im}(L_A) = n$, a caraterística $r = \mathrm{rank}(A)$ da matriz A é também igual ao número de linhas linearmente independentes de A.

Se X e X' são soluções do sistema $AX = B$, então a diferença $Z = X - X'$ é solução do sistema homogéneo $AZ = 0$. Portanto, se X é uma (apenas uma!) das soluções do sistema linear possível $AX = B$, então o espaço d(e todas)as soluções é o subespaço afim

$$X + \mathrm{Ker}(L_A)$$

de dimensão $k = n - r$. Se os vetores E_1, E_2, $\dots$, E_k formam uma base de $\ker(L_A) \subset \mathbb{R}^n$, então a "solução geral" do sistema possível é

$$X + t_1 E_1 + t_2 E_2 + \cdots + t_k E_k$$

com $t_1, t_2, \dots, t_k$ parámetros reais.

Em particular, a solução do sistema possível $L_A(X) = B$ é única quando $\mathrm{rank}(A) = n$, ou seja, quando o núcleo de L_A é trivial, ou seja, quando L admite uma inversa $L_A^{-1} : \mathrm{Im}(L_A) \to \mathbb{R}^n$. Neste caso, a solução é dada por $X = L_A^{-1}(B)$.

ex: Estude os seguintes sistemas (ou seja, diga se são possíveis e, caso afirmativo, determine o espaço das soluções)

$$\begin{cases} 2x + y & = -1 \\ x - y & = 3 \end{cases} \qquad \begin{cases} x + y & = 11 \\ x - y & = 33 \end{cases}$$

$$\begin{cases} 2x - 3y & = -1 \\ -6x + 9y & = 0 \end{cases} \qquad \begin{cases} x + y - z & = 1 \end{cases}$$

$$\begin{cases} 2x - 5y + 4z & = -3 \\ x - 2y + z & = 5 \\ x - 4y + 6z & = 10 \end{cases} \qquad \begin{cases} 3x + y - 10z & = 1 \\ -2x - 5y + 7z & = 2 \\ x + 3y - z & = 0 \end{cases}$$

ex: [Ap69] **16.20**.

Eliminação de Gauß-Jordan. Considere o sistema linear

$$AX = B\,,$$

com $A = (A_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, $B = (b_1, b_2, \dots, b_m)^\top \in \mathbb{R}^m$ e $X = (x_1, x_2, \dots, x_n)^\top \in \mathbb{R}^n$ vetores-coluna, ou seja,

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

O *método de eliminação de Gauss-Jordan* consiste em efectuar as seguintes "operações elementares" sobre as equações, e portanto sobre as linhas da "matriz ampliada"

$$(A|B) := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix} \in \mathrm{Mat}_{m\times(n+1)}(\mathbb{R})$$

i) trocar a ordem das equações,

ii) multiplicar (todos os termos de) uma equação por um escalar não nulo $\lambda \neq 0$,

iii) somar a uma equação um múltiplo de outra equação,

. . . até obter um sistema equivalente $A'X = B'$, com A' "matriz em escada de linhas", ou seja, da forma

$$\boxed{A' = \begin{pmatrix} \star & * & * & * & * & \dots & * \\ 0 & \star & * & * & * & \dots & * \\ 0 & 0 & 0 & \star & * & \dots & * \\ 0 & 0 & 0 & 0 & \star & \dots & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}}$$

onde os "pivots" $\star$ são os elementos $\neq 0$ mais à esquerda de cada linha (este processo, que transforma a matriz A numa matriz em escada de linhas A', é chamado *condensação*).

É evidente que as operações elementares não mudam a caraterística de uma matriz (pois não alteram o espaço das soluções da equação homogénea). De consequência, a caraterística de uma matriz é igual ao número de linhas não nulas da matriz em escada de linhas equivalente.

ex: Usando operações elementares sobre as linhas, tranforme a matriz A dada numa matriz em escada e calcule a característica de A.

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & 1 & 3 & -2 \\ 2 & 1 & -4 & 3 \\ 2 & 3 & 2 & -1 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 2 & -1 & 2 \\ 2 & 4 & 1 & -2 \end{pmatrix}$$

ex: Resolva os seguintes sistemas lineares usando o método de eliminação de Gauss.

$$\left\{\begin{array}{ll} 3x + 2y + z & = 1 \\ 5x + 3y + 3z & = 2 \\ -x + y + z & = -1 \end{array}\right. \qquad \left\{\begin{array}{ll} 3x + 2y + z & = 1 \\ 2x - 6y + 4z & = 3 \\ x + y + z & = -2 \\ 2x - 5y + 5z & = -1 \end{array}\right.$$

$$\left\{\begin{array}{ll} 2x + y + 4z & = 2 \\ 6x + y & = -10 \\ -x + 2y - 10z & = -4 \end{array}\right. \qquad \left\{\begin{array}{ll} y + z & = 1 \\ x + 2y - z & = 3 \\ x + y + z & = 1 \end{array}\right.$$

$$\begin{cases} x+2y+3z+4w &= 3 \\ 5z+6w &= 0 \\ z+3w &= 1 \\ x-y+8w &= 0 \end{cases} \qquad \begin{cases} -x+y-z &= 1 \\ x+3z &= -3 \\ y-z &= 0 \end{cases}$$

ex: [Ap69] **16.20**.

ex: Dê exemplos de

- um sistema de 2 equações lineares com 2 incógnitas com solução única,
- um sistema de 2 equações lineares com 2 incógnitas sem nenhuma solução,
- um sistema de 3 equações lineares com 3 incógnitas tal que o espaço das soluções seja uma reta afim.
- um sistema de 3 equações lineares com 3 incógnitas com solução única,
- um sistema de 2 equações lineares com 3 incógnitas tal que o espaço das soluções seja um plano afim.
- um sistema de 2 equações lineares com 3 incógnitas tal que o espaço das soluções seja um subespaço vetorial de dimensão 1.

Equações diferenciais lineares com coeficientes constantes. O núcleo do operador derivação $(Df)(t) = f'(t)$ é o espaço de dimensão um das funções constantes $f(t) = c$. As soluções da equação diferencial linear

$$Df = g\,,$$

onde $g(t)$ é uma função integrável dada, são $f(t) = (Pg)(t) + c$, onde $(Pg)(t) = \int_0^t g(s)\,ds$ é uma solução particular, e $c = f(0)$ é uma constante arbitrária, solução geral da equação homogénea $Df = 0$.

Uma equação diferencial linear com coeficientes constantes é uma equação do género

$$a_n D^n f + \cdots + a_1 Df + a_0 = g\,,$$

para a função $f(t)$, onde os a_k são coeficientes (reais ou complexos) e $g(t)$ é uma função dada (uma força quando $n = 2$). Pode ser pensada como $Lf = g$ se L é operador diferencial

$$L := a_n D^n + \cdots + a_1 D + a_0 I\,.$$

O núcleo de L, o espaço das soluções da equação homogénea

$$a_n D^n f + \cdots + a_1 Df + a_0 = 0\,,$$

é um espço vetorial de dimensão n. De fato, $h(t) = e^{zt}$ é uma solução de $Lh = 0$ se z é uma raiz do polinómio caraterístico $a_n z^n + \cdots + a_1 z + a_0 = 0$. No caso genérico, este polinómio tem n raizes complexas $z_1, z_2, \dots, z_n$ (conjugadas em pares se os coeficientes a_k forem reais). Assim, os exponenciais (complexos, produtos de exponenciais reais e funções trigonométricas) $e^{z_k t}$, com $k = 1, 2, \dots, n$, formam uma base de $\ker(L)$. Se $f(t)$ é uma "solução particular" (ou seja, apenas uma!) da equação $Lf = g$, então a "solução geral" (ou seja, todas as soluções) é da forma

$$f(t) + c_1 e^{z_1 t} + c_2 e^{z_2 t} + \cdots + c_n e^{z_n t}\,,$$

com $c_1, c_2, \dots, c_n$ coeficientes arbitrários (determinados pelas condições iniciais).

e.g. Inversão de matrizes 2×2. A matriz 2×2

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

representa o endomorfismo genérico do plano $L(x, y) = (ax + by, cx + dy)$. A transformação L é invertível se para cada vetor $(\alpha, \beta) \in \mathbb{R}^2$ é possível encontrar um vetor $(x, y) \in \mathbb{R}^2$ tal que $L_A(x, y) = (\alpha, \beta)$, ou seja, resolver o sistema linear

$$\left\{ \begin{array}{rcl} ax + by & = & \alpha \\ cx + dy & = & \beta \end{array} \right. \qquad \Rightarrow \qquad \left\{ \begin{array}{rcl} (ad - bc)\, x & = & d\alpha - c\beta \\ (ad - bc)\, y & = & a\beta - c\alpha \end{array} \right.$$

(o segundo sistema é obtido ao retirar b vezes a segunda equação de d vezes a primeira equação, e depois ao retirar c vezes a primeira equação de a vezes a segunda equação). Portanto, a transformação L_A é invertível sse $\mathrm{Det} A := ad - bc \neq 0$, e a sua inversa é a transformação linear

$$L_A^{-1}(\alpha, \beta) = \frac{1}{ad - bc} (d\alpha - c\beta, a\beta - c\alpha)\,,$$

representada pela matriz

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

Automorfismos e matrizes invertíveis. A matriz quadrada $A \in \mathrm{Mat}_{n \times n}(\mathbb{R})$ é *invertível* (ou *não-singular*, ou *regular*) se existe uma matriz quadrada $A^{-1} \in \mathrm{Mat}_{n \times n}(\mathbb{R})$, dita *inversa* de A, tal que

$$\boxed{A^{-1} A = A A^{-1} = I}$$

A transformação linear $L : \mathbb{R}^n \to \mathbb{R}^n$, representada pela equação matricial $X \mapsto Y = AX$, é invertível sse a matriz A é invertível, e a sua inversa é a transformação linear $Y \mapsto X = A^{-1} Y$. Se $A = (a_{ij})$, então as entradas da inversa $A^{-1} = (b_{ij})$ satisfazem as n^2 equações lineares

$$\sum_k b_{ik}\, a_{kj} = \delta_{ij}$$

Se A e B são invertíveis, então também AB é invertível e a sua inversa é

$$\boxed{(AB)^{-1} = B^{-1} A^{-1}}$$

Se A é invertível então também $A^\top$ é invertível e

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

Se $AX = B$ é um sistema de n equações com n incógnitas, e se $A \in \mathrm{Mat}_{n \times n}(\mathbb{R})$ é uma matriz invertível, então o sistema admite uma solução única dada por $X = A^{-1} B$. Isto acontece quando o núcleo da transformação linear $X \mapsto AX$ é trivial, e portanto a caraterística da matriz (o número de linhas ou de colunas linearmente independentes) é n.

ex: Mostre que, se $A, B \in \mathrm{Mat}_{n \times n}(\mathbb{R})$, então $BA = I \Rightarrow AB = I$ (ou seja, uma inversa esquerda é também uma inversa direita, logo uma inversa).

ex: Diga se as seguintes matrizes são invertíveis e, caso afirmativo, calcule a inversa.

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad \begin{pmatrix} 1 & -3 \\ -2 & 6 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \\ 5 & 4 & 3 \end{pmatrix} \quad \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & 0 & 1 \end{pmatrix}$$

ex: [Ap69] **16.20**.

Mudança de bases/coordenadas. Sejam $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ e $\mathcal{B}' = (\mathbf{b}'_1, \mathbf{b}'_2, \dots, \mathbf{b}'_n)$ duas bases do espaço vetorial $\mathbb{R}^n$. Então existe uma matriz invertível $U = (u_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, com inversa $U^{-1} = (v_{ij})$, tal que

$$\mathbf{b}'_j = \sum_i u_{ij}\,\mathbf{b}_i \qquad \text{e} \qquad \mathbf{b}_j = \sum_i v_{ij}\,\mathbf{b}'_i \,.$$

Se (x_i) são as coordenadas do vetor $\mathbf{x} \in \mathbf{V}$ relativamente à base $\mathcal{B}$, então as coordenadas do vetor $\mathbf{x}$ relativamente à base $\mathcal{B}'$ são

$$\boxed{x'_i = \sum_k v_{ij}\,x_j \qquad \text{ou seja} \qquad x_i = \sum_j u_{ij}\,x'_j}$$

pois $\mathbf{x} = \sum_j x_j \mathbf{b}_j = \sum_{j,k} v_{ij} x_j \mathbf{b}'_i$. A matriz U, cujas colunas são as coordenadas dos vetores da base $\mathcal{B}'$ relativamente à base $\mathcal{B}$, é a matriz que realiza a mudança de coordenadas. As matrizes U e U^{-1} podem ser pensadas como matrizes das derivadas parciais, pois $u_{ij} = \partial x_i/\partial x'_j$ e $v_{ij} = \partial x'_i/\partial x_j$. Nesta notação, as fórmulas para a mudança de coordenadas parecem tautológicas (o que explica o valor da notação de Leibniz para as derivadas):

$$x'_i = \sum_j \frac{\partial x'_i}{\partial x_j}\,x_j \qquad \text{e} \qquad x_i = \sum_j \frac{\partial x_i}{\partial x'_j}\,x'_j \,.$$

Na notação matricial, se X e X' são os vetores coluna de coordenadas x_i's e x'_i's respetivamente, a mudança de coordenadas assume a forma

$$X = UX' \qquad \text{ou seja,} \qquad X' = U^{-1}X \,.$$

Seja $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ a matriz de uma transformação $L : \mathbb{R}^n \to \mathbb{R}^m$ relativamente a certas coordenadas de $\mathbb{R}^n$ e $\mathbb{R}^m$, dada por $X \mapsto Y = AX$. Sejam $X' = UX$ e $Y' = VY$ umas mudanças de coordenadas em $\mathbb{R}^n$ e $\mathbb{R}^m$, respetivamente, definidas pelas matrizes invertíveis $U \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ e $V \in \mathrm{Mat}_{m\times m}(\mathbb{R})$. Então a matriz da transformação L relativamente às novas coordenadas é

$$\boxed{A' = V^{-1}AU}$$

De fato, se $y_i = \sum_j a_{ij}\,x_j$, então $y'_i = \sum_k t_{ik}\,y_k = \sum_{k,\ell} t_{ik}\,a_{k\ell}\,x_\ell = \sum_{k,\ell,j} v_{ik}\,a_{k\ell}\,u_{\ell j}\,x'_j$. Em notação matricial,

$$X \mapsto Y = AX \qquad \Rightarrow \qquad X' \mapsto Y' = V^{-1}Y = V^{-1}AX = (V^{-1}AU)X' \,.$$

Em particular, se $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é a matriz do endomorfismo $L \in \mathrm{End}(\mathbb{R}^n)$ relativamente à base $\mathcal{B}$, então a matriz de L relativamente à base $\mathcal{B}'$ é

$$\boxed{A' = U^{-1}AU}$$

e, de consequência,

$$\boxed{A = UA'U^{-1}\,.}$$

Matrizes A e A' relacionadas pelas identidades acima são ditas *semelhantes*. Representam o mesmo endomorfismo em bases possivelmente diferentes.

ex: Verifique que se $A' = U^{-1}AU$ então $A = UA'U^{-1}$.

ex: Use $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ para mostrar que $\mathrm{tr}(U^{-1}AU) = \mathrm{tr}\,A$, ou seja, o traço de uma matriz quadrada apena depende da transformação linear definida pela matriz, e não da base usada.

ex: Determine a matriz de $L(x,y) = (3x, 2y)$ relativamente à base $\mathbf{b}_1 = (1,1)$ e $\mathbf{b}_2 = (1,-1)$.

ex: Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a reflexão na reta $y = x$. Determine a matriz de T relativamente à base canónica e relativamente à base $(1, 1)$ e $(1, -1)$.

ex: Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a reflexão na reta $y = 3x$. Determine a matriz de T relativamente à base canónica.

ex: Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a projecção ortogonal sobre a reta $x + y = 0$. Determine a matriz de T relativamente à base canónica.

ex: Determine a matriz que representa o operador derivação $D : \mathrm{Pol}_{\leq 3}(\mathbb{R}) \to \mathrm{Pol}_{\leq 2}(\mathbb{R})$, definido por $(Df)(t) := f'(t)$, relativamente às bases ordenadas $(1, t, t^2, t^3)$ e $(1, t, t^2)$. Determine umas bases de $\mathrm{Pol}_{\leq 3}(\mathbb{R})$ e $\mathrm{Pol}_{\leq 3}(\mathbb{R})$ tal que o operador $D : \mathbf{V} \to \mathbf{W}$ seja diagonal.

11 Volumes e determinantes

ref: [Ap69] *Vol 2*, **3.1-17** ; [La97] *Ch. VII*

7 dez 2018

Volumes de paralelepípedos e n-formas alternadas. Uma *n-forma (linear)* no espaço vetorial $\mathbb{R}^n$ é uma função escalar que envia n vetores ordenados $\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_n$ de $\mathbb{R}^n$ num escalar $F(\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_n)$, que é multilinear, ou seja, homogénea e aditiva em cada variável. Homogénea significa que

$$F(\dots, \lambda\mathbf{v}, \dots) = \lambda F(\dots, \mathbf{v}, \dots) \tag{11.1}$$

e aditiva significa que

$$F(\dots, \mathbf{v}+\mathbf{w}, \dots) = F(\dots, \mathbf{v}, \dots) + F(\dots, \mathbf{w}, \dots) \tag{11.2}$$

Uma n-forma F é *alternada* (ou *anti-simétrica*) se é nula quando duas varáveis são iguais (ou, pela homogeneidade, proporcionais), i.e.

$$F(\dots, \mathbf{v}, \dots, \mathbf{v}, \dots) = 0 \tag{11.3}$$

e portanto se muda de sinal ao trocar duas variáveis,

$$F(\dots, \mathbf{v}, \dots, \mathbf{w}, \dots) = -F(\dots, \mathbf{w}, \dots, \mathbf{v}, \dots).$$

Fixada uma base de $\mathbb{R}^n$, por exemplo a base canónica $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$, uma n-forma F é determinada pelas suas "coordenadas"

$$f_{ijk\dots} := F(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k, \dots)$$

pois, pela linearidade em cada variável,

$$F(\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots) = \sum_{i,j,k,\dots} x_i y_j z_k \dots F(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k, \dots)$$

Se a n-forma F é alternada, então as coordenadas com (pelo menos) dois índices repetidos são nulas, i.e. $f_{\dots i \dots i \dots} = 0$. Em particular, apenas podem ser diferentes de zero as coordenadas cujos índices são permutações do conjunto $\{1, 2, \dots, n\}$. As coordenadas são também anti-simétricas, i.e. $f_{\dots i \dots j \dots} = -f_{\dots j \dots i \dots}$. Mas toda permutação $\sigma : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$ pode ser obtida da permutação trivial $\sigma(k) = k$ usando trocas repetidas. Se $f_{12\dots n} = \lambda$, então as outras coordenadas não nulas são $\pm\lambda$, ou seja, $f_{\sigma(1)\sigma(2)\dots\sigma(n)} = \pi(\sigma)\lambda$, onde $\pi(\sigma) := (-1)^k$ é a "paridade" da permutação σ (k é o número de trocas que transforma σ na permutação trivial). Portanto, uma n-forma alternada é univocamente determinada pelo seu valor $\lambda = f_{12\dots n}$ sobre os vetores ordenados da base canónica. O espaço linear $\mathcal{A}^n(\mathbb{R}^n)$ das n-formas alternadas em $\mathbb{R}^n$ tem dimensão igual a 1. Em particular,

Teorema 11.1. *Existe uma única n-forma alternada D em $\mathbb{R}^n$ normalizada de maneira tal que $D(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = 1$.*

As coordenadas da "n-forma canónica" D, os números

$$\varepsilon_{ijk\dots} := D(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k, \dots),$$

são $\varepsilon_{ijk\dots} = 0$ se dois índices são iguais, e $\varepsilon_{ijk\dots} = (-1)^{\pi(\sigma)}$ se $(i, j, k, \dots) = \sigma(1, 2, 3, \dots)$ onde σ é uma permutação de $\{1, 2, 3, \dots, n\}$. O símbolo $\varepsilon_{ijk\dots}$ é chamado *símbolo de Levi-Civita*. O valor da forma canónica D sobre os vetores $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$ é

$$D(\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots) = \sum_{i,j,k,\dots} x_i y_j z_k \dots \varepsilon_{ijk\dots}$$

Uma n-forma alternada genérica F em $\mathbb{R}^n$ é proporcional a D, ou seja, é igual a $F = \lambda D$, se $\lambda = F(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ é o seu valor nos vetores da base canónica, ou seja,

$$F(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = F(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)\, D(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$$

O *paralelepípedo* de lados (ou "gerado" pelos vetores) $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ no espaço euclidiano $\mathbb{R}^n$ é o conjunto

$$P = \{t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_n\mathbf{v}_n \ \text{ com } \ t_1, t_2, \dots, t_n \in [0,1]\} .$$

É possível provar que o seu *volume (n-dimensional)* é igual ao valor absoluto do valor da n-forma canónica sobre os lados $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, ou seja,

$$\boxed{\mathrm{Vol}(P) = |D(\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_n)|}$$

Ou seja, a forma canónica D calcula um "volume orientado" dos paralelepípedos, um volume com sinal positivo ou negativo dependendo da maneira em que os lados são ordenados (a prova consiste em mostrar que um volume orientado satisfaz as propriedades que definem a forma canónica, ou seja, que é linear em cada variável/lado, é nulo quando dois lados são iguais ou proporcionais, e é normalizado de maneira tal que o hiper-cubo de lados unitário tem volume igual a um).

ex: Mostre que, se F é uma forma bi-linear, então $F(\mathbf{x}, \mathbf{x}) = 0$ implica $F(\mathbf{x}, \mathbf{y}) + F(\mathbf{y}, \mathbf{x}) = 0$ (calcule $F(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) \dots$)

ex: Verifique que a única 2-forma alternada no plano $\mathbb{R}^2$ normalizada de maneira tal que $D(\mathbf{e}_1, \mathbf{e}_2) = 1$ é

$$D(a\mathbf{e}_1 + b\mathbf{e}_2, c\mathbf{e}_1 + d\mathbf{e}_2) = ad - bc .$$

Determinante. O teorema de unicidade 11.1 implica que existe uma única função *determinante* $\mathrm{Det} : \mathrm{Mat}_{n\times n}(\mathbb{R}) \to \mathbb{R}$

$$A \mapsto \mathrm{Det}A$$

que é uma forma multilinear alternada nas colunas da matriz, e normalizada de maneira tal que $\mathrm{Det}I = 1$. De fato, se $A_1 = (a_{11}, a_{21}, \dots, a_{n1})^\top$, $A_2 = (a_{12}, a_{22}, \dots, a_{n2})^\top$, ..., $A_n = (a_{1n}, a_{2n}, \dots, a_{nn})^\top$ são as colunas de $A = (a_{ij})$, então esta função é

$$\boxed{\mathrm{Det}A := D(A_1, A_2, \dots, A_n)}$$

Outra notação também utilizada é $\mathrm{Det}A = |A|$. Uma fórmula para o determinante é

$$\mathrm{Det}A = \sum_{\sigma\in\mathrm{Per}_n} \pi(\sigma)\, a_{1\sigma(1)}\, a_{2\sigma(2)} \cdots a_{n\sigma(n)} ,$$

onde $\pi(\sigma)$ é a paridade da permutação σ. Usando o símbolo de Levi-Civita, o determinante pode ser definido pela fórmula

$$\mathrm{Det}A = \sum_{ijk\dots} \varepsilon_{ijk\dots}\, a_{1i}\, a_{2j}\, a_{3k} \ \dots .$$

É imediato verificar que $\mathrm{Det}A^\top$ é uma n-forma linear alternada nas linhas $A^1, A^2, \dots, A^n$ da matriz A que vale 1 se as linhas são a base canónica de $\mathbb{R}^n$, ou seja, se $A^\top = I$. Pelo teorema de unicidade 11.1,

$$\mathrm{Det}A^\top = \mathrm{Det}A .$$

Se as colunas (ou as linhas) de uma matriz quadrada A são vetores linearmente dependentes, então uma é combinação linear das outras. A linearidade e a (11.3) então implicam que $\mathrm{Det}A = 0$.

O determinante de ordem n, sendo um volume orientado n-dimensional, é homogéneo de grau n. Ou seja, se $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ e $\lambda \in \mathbb{R}$, então

$$\mathrm{Det}\,(\lambda A) = \lambda^n \mathrm{Det}A$$

Cálculo de determinantes. É possível calcular determinantes usando as propriedades (11.1), (11.2), (11.3) e a normalização $D(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = 1$. As fórmulas ficam logo compridas, pois o número das permutações de n elementos é $n!$, e o fatorial cresce muito rapidamente.

O determinante de uma matriz 2×2 é

$$\boxed{\mathrm{Det}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}\,a_{22} - a_{12}\,a_{21}}$$

O determinante de uma matriz 3×3 é

$$\boxed{\mathrm{Det}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}\,\mathrm{Det}\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12}\,\mathrm{Det}\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13}\,\mathrm{Det}\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}}$$

O determinante de uma matriz diagonal é

$$\boxed{\mathrm{Det}\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = \lambda_1\lambda_2\cdots\lambda_n}$$

(quando os λ_k são positivos, este é o volume de um paralelepípedo de lados dois a dois ortogonais de comprimentos λ_k's).

ex: Calcule o determinante das seguintes matrizes

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \qquad \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

ex: Mostre que

$$\mathrm{Det}\begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix} = (b-a)(c-a)(c-b)$$

ex: Calcule o determinante de $2A$ e $-A$ sabendo que A é uma matriz 5×5 com determinante $\mathrm{Det}A = -3$.

ex: Verifique que uma equação cartesiana da reta que passa pelos pontos (a, b) e (c, d) de $\mathbb{R}^2$ é

$$\mathrm{Det}\begin{pmatrix} x-a & y-b \\ c-a & d-b \end{pmatrix} = 0 \qquad \text{ou} \qquad \mathrm{Det}\begin{pmatrix} x & y & 1 \\ a & b & 1 \\ c & d & 1 \end{pmatrix} = 0$$

ex: [Ap69] **3.6**.

Menores, complemento algébrico e fórmula de Laplace. É possíel determinar fórmulas recursivas que permitem calcular determinantes de ordem n a custa de determinantes de ordem $n-1$. Seja $A = (a_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ uma matriz quadrada de ordem $n \geq 2$. O *menor* ij de A é a matriz $A_{ij} \in \mathrm{Mat}_{(n-1)\times(n-1)}(\mathbb{R})$ obtida da matriz A suprimindo a linha i e a coluna j. O *complemento algébrico* do elemento a_{ij} de A é o número

$$\mathrm{Cal}\,(a_{ij}) := (-1)^{i+j}\mathrm{Det}A_{ij}\,.$$

A *matriz dos complementos algébricos* (ou *dos co-fatores*) de A é a matriz $\mathrm{Cal}\,A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ cujo elemento ij é $\mathrm{Cal}\,(a_{ij})$, ou seja

$$\mathrm{Cal}\,A := (\mathrm{Cal}\,(a_{ij}))\,.$$

O desenvolvimento do determinante da matriz A em função dos elementos da sua i-ésima linha é (*fórmula de Laplace*)

$$\boxed{\mathrm{Det}A = \sum_{j=1}^{n} a_{ij}\,\mathrm{Cal}(a_{ij}) = \sum_{j=1}^{n} a_{ij}\,(-1)^{i+j}\mathrm{Det}A_{ij}}$$

Ao trocar colunas ou ao considerar a matriz transposta, é possível escrever fórmulas análogas que desenvolvem o determinante em função de todas as linhas ou as colunas de uma matriz quadrada.

ex: Calcule a matriz dos complementos algébricos das matrizes

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 0 \\ -7 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

ex: Calcule o determinante das matrizes

$$\begin{pmatrix} 1 & 0 & -1 & 3 \\ -5 & 0 & 2 & -1 \\ 4 & 0 & 1 & 1 \\ 3 & -2 & -1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 4 & -2 \end{pmatrix}$$

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

Determinante e produtos. Seja $L : \mathbb{R}^n \to \mathbb{R}^n$ a transformação linear definida pela matriz quadrada $A = (a_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, que envia o vetor coluna X no vetor coluna $Y = AX$. A função $V_1, V_2, \dots, V_n \mapsto D(AV_1, AV_2, \dots, AV_n)$ é uma n-forma alternada, e portanto, pelo teorema de unicidade 11.1, proporcional à forma canónica D, ou seja,

$$D(AV_1, AV_2, \dots, AV_n) = D(AE_1, AE_2, \dots, AE_n) \cdot D(V_1, V_2, \dots V_n)$$

onde E_i são os vetores coluna da base canónica. As colunas da matriz A são os vetores coluna $A_i = AE_i$. Portanto $D(AE_1, AE_2, \dots, AE_n) = D(A_1, A_2, \dots, A_n) = \mathrm{Det}A$. Se $B \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ denota a matriz cujas colunas são os vetores V_i, então os vetores AV_i são as colunas da matriz AB. Temos portanto

Teorema 11.2. *Se* $A, B \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, *então*

$$\boxed{\mathrm{Det}(AB) = (\mathrm{Det}A)\,(\mathrm{Det}B)}$$

Em particular, se A *é invertível então* $\mathrm{Det}A \neq 0$ *e*

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

Se $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ e $B \in \mathrm{Mat}_{m\times m}(\mathbb{R})$, então matriz "diagonal por blocos"

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in \mathrm{Mat}_{(n+m)\times(n+m)}(\mathbb{R})$$

define a transformação linear $(X, Y) \mapsto (AX, BY)$ de $\mathbb{R}^n \times \mathbb{R}^m \simeq \mathbb{R}^{n+m}$. O seu determinante é

$$\boxed{\mathrm{Det}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = (\mathrm{Det}A)(\mathrm{Det}B)}$$

ex: Verdadeiro ou falso? Dê uma demonstração ou um contra-exemplo.

$$\mathrm{Det}(A+B) = \mathrm{Det}A + \mathrm{Det}B \qquad\qquad \mathrm{Det}\left((A+B)^2\right) = \left(\mathrm{Det}(A+B)\right)^2$$

ex: Observe que

$$\mathrm{Det}\left(A^n\right) = \left(\mathrm{Det}A\right)^n .$$

Deduza que o determinante de uma mariz nilpotente é nulo. O que pode dizer do determinante de uma projeção? E de uma reflexão?

ex: Uma matriz quadrada A é dita *ortogonal* se $A^\top A = A\,A^\top = I$, ou seja, se é invertível e a sua inversa é $A^\top$. Mostre que o determinate de uma matriz ortogonal é ± 1.

ex: [Ap69] **3.11**.

Determinante de um endomorfismo e volumes. Seja $A = (a_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ a matriz quadrada que define a transformação linear $L : \mathbb{R}^n \to \mathbb{R}^n$, relatvamente à base canónica (ou seja, $y_i = \sum_j a_{ij}\,x_j$). Se $F = \lambda E$ é uma n-forma alternada genérica, então

$$F(L\mathbf{v}_1, L\mathbf{v}_2, \dots, L\mathbf{v}_n) = (\mathrm{Det}A)\cdot F(\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_n)$$

Em particular, o determinante da matriz A depende apenas da transformação linear L, e não da base usada para calcular a matriz! De fato, uma mudança de coordenadas envia a matriz A na matriz $A' = U^{-1}AU$, e $\mathrm{Det}A' = \mathrm{Det}A$. Faz portanto sentido definir o "determinante" de um endomorfismo $L : \mathbb{R}^n \to \mathbb{R}^n$ como

$$\boxed{\mathrm{Det}L := \mathrm{Det}A}$$

onde A é qualquer matriz que representa L numa base arbitrária de $\mathbb{R}^n$.

Se

$$Q := [0,1]^n = \{t_1E_1 + t_2E_2 + \cdots + t_nE_n \in \mathbb{R}^n \ \text{ com } \ 0 \le t_k \le 1\}$$

denota o hiper-cubo unitário, ou seja, o paralelepípedo de lados $E_1, E_2, \dots, E_n$, então a imagem $L(Q)$ é o paralelepípedo de lados $L(E_1) = A_1$, $L(E_2) = A_2$, $\dots L(E_n) = A_n$, que são as colunas da matriz A que define L na base canónica. O determinante do endomorfismo L é portanto o quociente

$$\mathrm{Det}L = \pm\frac{\mathrm{Vol}(L(Q))}{\mathrm{Vol}(Q)} ,$$

o sinal sendo positivo ou negativo dependendo se L preserva ou não a orientação. Em geral, se $R \subset \mathbb{R}^n$ é uma região suficientemente regular, e portanto o seu volume pode ser aproximado con precisão arbitrária usando somas de volumes de hiper-cubos, então $\mathrm{Det}L$ é igual a $\pm$ o quociente $\mathrm{Vol}(L(R))/\mathrm{Vol}(R)$.

ex: Calcule o determinate das transformações lineares

$$T(x,y) = (2x, 3y) \qquad T(x,y) = (x, -y) \qquad T(x,y) = (y,x)$$

$$T(x,y,z) = (3z, 2y, x) \qquad T(x,y,z) = (x, x+y, x+y+z) \qquad T(x,y,z) = (y,x,0)$$

Método de Gauß-Jordan para calcular determinantes. O determinante de uma matriz triangular superior (ou triangular inferior) é igual ao produto dos termos diagonais,

$$\boxed{\mathrm{Det}\begin{pmatrix} \lambda_1 & a_{12} & \dots & a_{1n} \\ 0 & \lambda_2 & a_{23} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = \lambda_1\,\lambda_2\,\cdots\,\lambda_n}$$

De consequência, é possível calcular um determinante transformando uma matriz genérica A numa matriz triangular superior pelo método de Gauss-Jordan. Isto acontece porque as operações permitidas têm efeitos simples no determinante: trocar duas linhas muda o sinal do determinante; somar a uma linha um múltiplo de uma outra linha não muda o determinante; e multiplicar uma linha por um escalar $\lambda \neq 0$ transforma $\mathrm{Det} A$ em $\lambda \, \mathrm{Det} A$.

As operações elementares sobre as linhas de uma matriz não alteram a sua caraterística. Em particular, se a caraterística de uma matriz quadrada A de ordem n é igual a $\mathrm{rank} A = n$, ou seja, se as colunas de A são vetores linearmente independentes, então a matriz é equivalente a uma matriz triangular superior com pivots λ_k não nulos. O seu determinante é portanto diferente de zero. Vice-versa, é evidente que o determinante de uma matriz cujas colunas são linearmnete dependentes é nulo (porque uma coluna é combinação linear das outras). De consequência,

Teorema 11.3. *As colunas ou as linhas de uma matriz quadrada A são linearmente independentes sse* $\mathrm{Det} A \neq 0$.

ex: Use o método de eliminação de Gauss-Jordan para calcular o determinante das matrizes

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \qquad \begin{pmatrix} 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{pmatrix}$$

Wronskiano, independência e identidade de Abel. Sejam $x(t)$ e $y(t)$ duas funções diferenciáveis definidas num intervalo da reta real. O *(determinante) Wronskiano* das funções x e y é a função

$$W(t) = \mathrm{Det} \begin{pmatrix} x(t) & y(t) \\ \dot{x}(t) & \dot{y}(t) \end{pmatrix}$$

($\dot{x} = dx/dt$ e $\dot{y} = dy/dt$). Se $x(t)$ e $y(t)$ são linearmemte dependentes então $W(t) = 0$ para todo o t. Portanto, se existe um ponto t onde $W(t) \neq 0$ então as funções x e y são linearmente independentes. A derivada de $W(t)$ é

$$\dot{W}(t) = \mathrm{Det} \begin{pmatrix} x(t) & y(t) \\ \ddot{x}(t) & \ddot{y}(t) \end{pmatrix}$$

(a notação é $\ddot{x} = d^2x/dt^2$ e $\ddot{y} = d^2y/dt^2$). De consequência, se $x(t)$ e $y(t)$ são duas soluções da equação diferencial homogénea de segunda ordem $\ddot{z} + p(t)\,\dot{z} + q(t)\,z = 0$, então

$$\dot{W}(t) = -p(t)W(t) \qquad \text{e portanto} \qquad W(t) = W(t_0)e^{-\int_{t_0}^{t} p(s)\,ds}\,.$$

Em particular, para verificar se $x(t)$ e $y(t)$ são independentes num intervalo, é suficiente calcular $W(t)$ num ponto arbitrário deste intervalo.

Regra de Cramer. Seja $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ uma matriz quadrada com $\mathrm{Det} A \neq 0$. Então as suas colunas $A_1, A_2, \dots, A_n$ geram o espaço $\mathbb{R}^n$, ou seja, a transformação linear $X \mapsto AX$ é invertível. Para todo vetor coluna $B \in \mathbb{R}^n$ existe portanto uma única solução do sistema linear

$$AX = B\,.$$

As coordenadas desta solução são os únicos coeficientes x_k's tais que

$$x_1A_1 + x_2A_2 + \cdots + x_nA_n = B\,.$$

Se substituimos à k-ésima coluna A_k da matriz A o vetor B e calculamos o seu determinante, acontece que

$$\begin{aligned} D(A_1, \dots, B, \dots, A_n) &= D(A_1, \dots, x_1A_1 + x_2A_2 + \cdots + x_nA_n, \dots, A_n) \\ &= x_k\, D(A_1, \dots, A_k, \dots, A_n) \end{aligned}$$

pela multilinearidade e a anti-simeria de D. De consequência,

Teorema 11.4 (regra de Cramer)**.** *Seja $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ uma matriz quadrada com $\mathrm{Det}A \neq 0$. As coordenadas x_k's da única solução do sistema linear $AX = B$ são os quocientes*

$$\boxed{x_k = \frac{\mathrm{Det}\begin{pmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ a_{21} & \dots & b_2 & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{pmatrix}}{\mathrm{Det}\begin{pmatrix} a_{11} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & \dots & a_{2k} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nk} & \dots & a_{nn} \end{pmatrix}}}$$

ex: Resolva os seguines sistema utilizando a regra de Cramer

$$\begin{cases} 3x + 2y + z & = 1 \\ 5x + 3y + 3z & = 2 \\ -x + y + z & = -1 \end{cases} \qquad \begin{cases} 3x + 2y + z & = 1 \\ 2x - 6y + 4z & = 3 \\ x + y + z & = -2 \end{cases}$$

$$\begin{cases} 2x + y & = -6 \\ -x + 2y + 4z & = 1 \\ -x + z & = 3 \end{cases} \qquad \begin{cases} 3x + y + z & = 0 \\ 2x - y + 3z & = 1 \\ x + y + z & = 1 \end{cases}$$

Determinante de Vandermonde. Dados n números reais ou complexos $z_1, z_2, \dots, z_n$, a *matriz de Vandermonde* é a matriz $n \times n$ cujas linhas (ou colunas) sã as progressões geométricas (até o grau $n-1$) dos z_k's, ou seja,

$$V := \begin{pmatrix} 1 & z_1 & z_1^2 & \dots & z_1^{n-1} \\ 1 & z_2 & z_2^2 & \dots & z_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & z_n & z_n^2 & \dots & z_n^{n-1} \end{pmatrix}$$

O determinante de Vandermonde é o produto

$$\mathrm{Det}V = \prod_{i<j} (z_j - z_i)$$

Determinante e matrizes invertíveis. Um matriz quadrada é invertível sse o seu determinante não é nulo. A regra de Cramer 11.4 sugere uma maneira de calcular a inversa.

Teorema 11.5. *Uma matriz quadrada A é invertível sse $\mathrm{Det}A \neq 0$, e a sua inversa é dada por*

$$\boxed{A^{-1} = \frac{1}{\mathrm{Det}A} (\mathrm{Cal}A)^\top}$$

Demonstração. Se A é invertível, a sua matriz inversa $A^{-1} = (x_{ij})$ satisfaz

$$A\, A^{-1} = I\,.$$

As colunas da matriz identidade I são os vetores $E_1, E_2, \dots, E_n$ da base canónica. A identidade acima então diz que as colunas $X_1, X_2, \dots, X_n$ da matriz inversa A^{-1} são as soluções dos sistemas lineares $AX_k = E_k$. Pela regra de Cramer 11.4, as coordenadas x_{ik}de X_k, com $i = 1, 2, \dots, n$, são obtidas ao dividir por $\mathrm{Det}A$ os determinantes das matrizes obtidas ao substituir à i-ésima coluna da matriz A o vetor E_k da base canónica. É imediato verificar que x_{ik} é então o elemento ki da matriz $\mathrm{Cal}A$ dos complementos algébricos de A □

ex: Calcule a inversa das seguintes matrizes

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 3 \\ 7 & 5 \end{pmatrix}$$

ex: Determine os valores de λ para os quais $\lambda I - A$ é singular, quando

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

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & -2 \\ 2 & -2 & 0 \end{pmatrix} \qquad A = \begin{pmatrix} 11 & -2 & 8 \\ 19 & -3 & 14 \\ -8 & 2 & -5 \end{pmatrix}$$

ex: [Ap69] **3.17**.

12 Valores e vetores próprios

ref: [Ap69] *Vol 2*, **4.1-10** ; [La97] *Ch. VIII*, **1-2**

14 dez 2018

Subespaços invariantes. Seja $L : \mathbf{D} \to \mathbf{V}$ um operador linear definido num subespaço $\mathbf{D} \subset \mathbf{V}$ do espaço linear $\mathbf{V}$, real ou complexo. Um subespaço linear $\mathbf{W} \subset \mathbf{D}$ é *invariante* (ou *estável*) se $L(\mathbf{W}) \subset \mathbf{W}$, ou seja, se $\mathbf{v} \in \mathbf{W}$ implica $L\mathbf{v} \in \mathbf{W}$. Subespaços invariantes triviais são o subespaço nulo $\{\mathbf{0}\}$, e o próprio $\mathbf{V}$ quando $\mathbf{D} = \mathbf{V}$.

ex: Mostre que o núcleo $\mathrm{Ker}(L)$ e a imagem $\mathrm{Im}(L)$ (se $\mathbf{D} = \mathbf{V}$) são subespaços invariantes.

ex: Determine os subespaços invariantes da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada ponto (x, y) no seu simétrico em relação à reta $y = x$.

ex: Determine os subespaços invariantes da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada ponto (x, y) na sua projeção ortogonal sobre a reta $y = x$.

ex: Determine os subespaços invariantes do "cisalhamento" (em inglês, *shear*) horizontal de fator $\mu \neq 0$, a transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (x + \mu y, y)$.

Valores e vetores próprios. Seja $L : \mathbf{D} \to \mathbf{V}$ um operador linear definido num subespaço $\mathbf{D} \subset \mathbf{V}$ do espaço linear $\mathbf{V}$, real ou complexo. Um *vetor próprio/autovetor* (em inglês *eigenvector*, do alemão *eigen* = próprio) de L é um vetor não nulo $\mathbf{v} \in \mathbf{D}$ que gera um subespaço invariante $\mathbb{R}\mathbf{v}$ (ou $\mathbb{C}\mathbf{v}$) de dimensão um para L, ou seja, tal que

$$\boxed{L\mathbf{v} = \lambda\mathbf{v}}$$

onde $\lambda \in \mathbb{R}$ (ou $\lambda \in \mathbb{C}$) é um escalar, chamado *valor próprio/autovalor* (em inglês *eigenvalue*) do operador L (associado ao vetor próprio $\mathbf{v}$).

Teorema 12.1. *Se $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ são vetores próprios de $L : \mathbf{D} \to \mathbf{V}$ e se os correspondentes valores próprios $\lambda_1, \lambda_2, \dots, \lambda_k$ são dois a dois distintos (i.e. $\lambda_i \neq \lambda_j$ se $i \neq j$), entaõ os vetores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ são linearmente independentes.*

Demonstração. A prova é por indução. O caso $k = 1$ é trivial. Seja $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{k+1}\mathbf{v}_{k+1} = \mathbf{0}$. Aplicando o operador L ou multiplicando por λ_{k+1}, e depois calculando a diferença, obtemos

$$c_1(\lambda_1 - \lambda_{k+1})\mathbf{v}_1 + c_2(\lambda_2 - \lambda_{k+1})\mathbf{v}_2 + \cdots + c_k(\lambda_k - \lambda_{k+1})\mathbf{v}_k = \mathbf{0}$$

Pela hipótese indutiva todos os coeficientes são nulos, e portanto, sendo os valores próprios distintos, $c_1 = c_2 = \cdots = c_k = 0$. Então também $c_{k+1} = 0$ (pois $\mathbf{v}_{k+1} \neq \mathbf{0}$). □

Se λ é um valor próprio de $L : \mathbf{V} \to \mathbf{V}$, o conjunto

$$\mathbf{V}_\lambda := \{\mathbf{v} \in \mathbf{V} \text{ t.q. } L\mathbf{v} = \lambda\mathbf{v}\} = \mathrm{Ker}(\lambda - L)$$

($\lambda - L$ denota o operador $\mathbf{v} \mapsto \lambda\mathbf{v} - L(\mathbf{v})$) é um subespaço invariante, diferente do espaço trivial $\{\mathbf{0}\}$, dito *subespaço próprio* associado ao valor próprio λ. A restrição do operador linear L a cada espaço próprio $\mathbf{V}_\lambda$ é uma homotetia $\mathbf{v} \mapsto \lambda\mathbf{v}$.

Em particular, se existe uma decomposição de $\mathbf{V}$ como soma direta finita $\mathbf{V} = \oplus_k \mathbf{V}_k$ de espaços próprios $\mathbf{V}_k$ associados aos valores próprios distintos λ_k de L (ou seja, todo $\mathbf{v} \in \mathbf{V}$ é uma sobreposição única $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n$ de vetores $\mathbf{v}_k \in \mathbf{V}_k$ tais que $L\mathbf{v}_k = \lambda_k\mathbf{v}_k$), então o operador é uma soma direta $L = \oplus_k \lambda_k$ de homotetias. Em outras palavras, se $P_k : \mathbf{V} \to \mathbf{V}_k$ denota a projeção sobre $\mathbf{V}_k$ (definida por $P_k(\mathbf{v}) := \mathbf{v}_k$), então $L = \sum_k \lambda_k P_k$. Tais operadores são chamados "diagonalizáveis". De fato, se $\mathbf{V}$ tem dimensão finita, são representados por matrizes diagonais numas bases formadas por vetores próprios.

ex: Todo vetor $\mathbf{v} \neq \mathbf{0}$ é um vetor própro da trasformação identidade $\mathbf{1_V v} := \mathbf{v}$, com valor próprio $\lambda = 1$. Todo vetor $\mathbf{v} \neq \mathbf{0}$ é um vetor própro da trasformação nula $0_\mathbf{V}\mathbf{v} := \mathbf{0}$, com valor próprio $\lambda = 0$. Em geral, todo vetor $\mathbf{v} \neq 0$ é um vetor próprio de uma homotetia $L\mathbf{v} = \lambda\mathbf{v}$, de valor próprio λ.

ex: Determine valores e vetores próprios da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada ponto (x, y) no seu simétrico em relação à reta $y = 2x$.

ex: Determine valores e vetores próprios da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada ponto (x, y) na sua projeção ortogonal sobre a reta $3y = x$.

ex: Uma rotação $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$, definida no espaço vetorial real $\mathbb{R}^2$ por

$$(x, y) \mapsto (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta) \ ,$$

não admite vetores próprios se o ângulo θ não é um múltiplo inteiro de π. No entanto, a mesma rotação pode ser pensada como a transformação $z \mapsto e^{i\theta}z$ definida no espaço vetorial complexo $\mathbb{C} \approx \mathbb{R}^2$, onde $z = x + iy \approx (x, y)$. Neste caso, todo vetor $z \neq 0$ é um vetor próprio, de valor próprio $e^{i\theta}$.

ex: Determine os valores e os vetores próprios das transformações

$$L(x, y) = (x, 0) \qquad L(x, y) = (x/2, 3y) \qquad L(x, y) = (-y, x)$$

$$L(x, y) = (x, x + y) \qquad L(x, y) = (x + \lambda y, y) \qquad L(x, y) = (x + \alpha y, y)$$

$$L(x, y, z) = (0, y, -z) \qquad L(x, y, z) = (y, z, x)$$

ex: Se λ é um valor próprio de $L : \mathbf{V} \to \mathbf{V}$ e $k \in \mathbb{N}$, então λ^k é um valor próprio de L^k.

No entanto, as potências L^k de um operador podem ter mais valores próprios que o próprio L. Por exemplo, uma rotação $R_{\pi/2}$ de um ângulo $\pi/2$ no plano não tem valores próprios, mas o seu quadrádo $R_{\pi/2}R_{\pi/2} = R_\pi$ admite um valor próprio, -1 (cuja raiz quadrada não é um número real!).

ex: Em particular, os valores próprios de um operador nilpotente (um operador tal que alguma potência L^n é o operador nulo) não podem ser diferentes de zero.

ex: Se $L : \mathbf{V} \to \mathbf{V}$ é um automorfismo (i.e. uma transformação linear invertível) e $\mathbf{v} \in \mathbf{V}$ é um vetor próprio de L, então $\mathbf{v}$ é um vetor próprio de L^{-1}

ex: [Ap69] **4.4**.

Resolvente e espetro. Se λ é um valor próprio do operador linear $L : \mathbf{V} \to \mathbf{V}$, então o operador

$$L_\lambda := \lambda - L$$

não é invertível (pois um autovetor $\mathbf{v}$ associado a λ está no seu núcleo). O *espetro* do operador L é o conjunto

$$\sigma(L) := \{\lambda \in \mathbb{C} \text{ t.q. } \lambda - L \text{ não é invertível } \} \subset \mathbb{C} \, ,$$

ou seja, o conjunto dos valores complexos de $z \in \mathbb{C}$ fora dos quais existe o operador *resolvente*

$$R_z := (z - L)^{-1}$$

O número $\rho(L) := \sup_{z\in\sigma(L)} |z|$ é dito *raio espetral* do operador L. Em particular, se λ é um valor próprio de L, então o seu módulo é limitado por $|\lambda| \leq \rho(L)$.

Quando $\mathbf{V}$ tem dimensão finita, o espetro $\sigma(L)$ coincide com o conjunto (finito) dos valores próprios, pois se L_λ não é invertível então o seu núcleo $\ker(L_\lambda)$ não é vazio, e um vetor não nulo do núcleo de L_λ é por definição um vetor próprio de L.

Em geral, em dimensão infinita, o espetro pode conter números que não são valores próprios.

ex: O operador *deslocamento* (em ingês, *shift*) $S : (x_1, x_2, x_3, \dots) \mapsto (x_2, x_3, x_4, \dots)$, definido no espaço $\mathbb{R}^{\mathbb{N}}$ das sucessões reais, não é invertível, portanto $0 \in \sigma(S)$. Determine o espaço próprio $\mathrm{Ker}(S)$.

Operadores derivação e multiplicação. O operador *derivação* envia uma função derivável $f(x)$ na função

$$(\partial f)(x) := f'(x) \, .$$

Pode ser pensado como un endomorfismo de $\mathcal{C}^\infty(\mathbb{R})$, o espaço linear das funções infinitamente diferenciáveis definidas em $\mathbb{R}$ com valores reais ou complexos. O produto $P := -i\hbar\,\partial$, onde $\hbar \simeq 1.054\cdots \times 10^{-34}$ J · s é a constante de Planck reduzida, é o *operador momento* na "representação de Schrödinger" da mecânica quântica. O operador *multiplicação* envia $f(x)$ em

$$(Xf)(x) := x\, f(x) \, .$$

Observe que ∂ e X não comutam. De fato, $\partial X - X\partial$ é o operador identidade.

Todo $\lambda \in \mathbb{C}$ é um valor próprio do operador derivação, e o espaço próprio associado ao valor próprio λ é gerado pela função exponencial $e^{\lambda x}$. Se $\lambda_1, \dots, \lambda_n$ são distintos, então as funções $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$ são linearmente independentes.

ex: O subespaço $\mathrm{Pol}(\mathbb{R}) \subset \mathcal{C}^\infty(\mathbb{R})$ dos polinómios é um subespaço invariante de ∂. O subespaço $\mathrm{Pol}_{\leq n}(\mathbb{R}) \subset \mathrm{Pol}(\mathbb{R})$ dos polinómios de grau $\leq n$ é um subespaço invariante do operador derivação? E do operador multiplicação?

ex: Mostre que os vetores próprios do operador $L = X\partial$, definido em $\mathrm{Pol}(\mathbb{R})$, são os monómios $f(x) = x^n$.

ex: Fixada uma "frequência" λ, real ou complexa, o espaço dos *quase-polinómios.* $p(x)e^{\lambda x}$, com $p \in \mathrm{Pol}(\mathbb{R})$, é um subespaço invariante para qualquer operador diferencial com coeficientes constantes $L = a_n\partial^n + \cdots + a_1\partial + a_0$. Esta observação justifica o método dos coeficientes indeterminados para encontrar uma solução particular de uma EDO linear $Lf = g$ quando a "força" $g(x)$ é um quase-polinómio.

Operador primitivação. O operador *primitivação* envia uma função integrável $f(x)$ numa das suas primitivas $F(x)$, por exemplo a função

$$(Pf)(x) := \int_0^x f(t)\, dt$$

Pode ser pensado como um endomorfismo do espaço linear real $\mathcal{C}^\infty(\mathbb{R})$ das fuções infinitamente diferenciáveis.

ex: Mostre que o operador $P : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$ não tem valores próprios (derive a identidade $Pf = \lambda f$ para obter uma equação diferencial para o suposto vetor próprio $f(x)$, e observe que a mesma identidade também implica uma condição inicial $f(0)\ \dots$).

ex: Mostre que ∂P é o operador identidade. Calcule o comutador $\partial P - P\partial$.

Operadores diferenciais, translações e ondas planas. Sejam $\Omega \subset \mathbb{R}^n$ um aberto e $\mathbf{V} = \mathcal{C}^\infty(\Omega)$ o espaço vetorial complexo das funções $f : \Omega \to \mathbb{C}$ infinitamente diferenciáveis. Dado um multi-índice $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n$, de grau $|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_n$, o operador diferencial $\partial^\alpha : \mathbf{V} \to \mathbf{V}$ é definido por

$$(\partial^\alpha f)(\mathbf{x}) := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}(\mathbf{x}) \, .$$

As ondas planas $e_{\boldsymbol{\xi}}(\mathbf{x}) := e^{i\boldsymbol{\xi}\cdot\mathbf{x}}$, com $\boldsymbol{\xi} \in (\mathbb{R}^n)^* \approx \mathbb{R}^n$, são funções próprias dos operadores diferenciais ∂^α, com $\alpha \in \mathbb{N}^n$, com valores próprios $(i\boldsymbol{\xi})^\alpha := (i\xi_1)^{\alpha_1}(i\xi_2)^{\alpha_2}\dots(i\xi_n)^{\alpha_n}$, ou seja,

$$\partial^\alpha e_{\boldsymbol{\xi}} = (i\boldsymbol{\xi})^\alpha e_{\boldsymbol{\xi}} \,.$$

O operador de *translação* $T_{\mathbf{a}}$, com $\mathbf{a} \in \mathbb{R}^n$, é definido por

$$(T_{\mathbf{a}} f)(\mathbf{x}) := f(\mathbf{x}+\mathbf{a}) \,.$$

As ondas planas $e_{\boldsymbol{\xi}}(\mathbf{x}) = e^{i\boldsymbol{\xi}\cdot\mathbf{x}}$ são também funções próprias dos operadores de translação com valores próprios $\lambda_{\mathbf{a}}(\boldsymbol{\xi}) = e^{i\boldsymbol{\xi}\cdot\mathbf{a}}$, ou seja,

$$T_{\mathbf{a}} e_{\boldsymbol{\xi}} = e^{i\boldsymbol{\xi}\cdot\mathbf{a}} e_{\boldsymbol{\xi}} \,.$$

O operador de *modulação* $M_{\boldsymbol{\xi}}$, com $\boldsymbol{\xi} \in (\mathbb{R}^n)^* \approx \mathbb{R}^n$, é definido por

$$(M_{\boldsymbol{\xi}} f)(\mathbf{x}) := e^{i\boldsymbol{\xi}\cdot\mathbf{x}} f(\mathbf{x}) \,.$$

É imediato verificar que $T_{\mathbf{a}} M_{\boldsymbol{\xi}} = e^{i\boldsymbol{\xi}\cdot\mathbf{a}} M_{\boldsymbol{\xi}} T_{\mathbf{a}}$. Os operadores translação e modulação geram o *grupo de Heisenberg*.

Polinómio caraterístico. Se λ é um valor próprio do operador linear $L : \mathbf{V} \to \mathbf{V}$, então o operador $L_\lambda := \lambda - L$ não é invertível, e os vetores próprios com valor próprio λ são os vetores não triviais do núcleo de L_λ. Se $\mathbf{V} \approx \mathbb{R}^n$ ou $\mathbb{C}^n$ tem dimensão finita, então, fixada uma base, o operador L é $X \mapsto AX$, onde $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ ou $\mathrm{Mat}_{n\times n}(\mathbb{C})$ é uma matriz quadrada (que depende da base escolhida). O escalar $\lambda \in \mathbb{R}$ ou $\mathbb{C}$ é um valor próprio da matriz quadrada A (ou do operador L) sse a matriz $\lambda I - A$ é singular, ou seja, sse

$$\mathrm{Det}\,(\lambda I - A) = 0 \,.$$

O *polinómio caraterístico* da matriz quadrada $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é o polinómio

$$\boxed{P_A(z) := \mathrm{Det}\,(zI - A)}$$

De consequência,

Teorema 12.2. *O escalar λ é um valor próprio da matrriz A (ou do operador definido pela matriz A) sse é uma raiz do polinómio caraterístico, i.e. se $P_A(\lambda) = 0$.*

O espaço próprio associado ao valor próprio λ é $\mathbf{V}_\lambda = \mathrm{Ker}(\lambda - L)$. Em particular, os vetores próprios associados ao valor próprio λ são as soluções não triviais do sistema homogéneo $(\lambda I - A)V = 0$.

ex: Mostre que A e $A^\top$ têm o mesmo polinómio caraterístico.

ex: Verifique que o polinómio caraterístico da matriz 2×2

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

é $P_A(z) = z^2 - (a+d)z + (ad - bc)$. Verifique que $P_A(A) = 0$, ou seja, que

$$A^2 - (\mathrm{tr}A)\,A + (\mathrm{Det}A)\,I = 0$$

ex: Os valores próprios de uma matriz nilpotente (uma matriz A tal que alguma potência $A^n = 0$) não podem ser diferentes de zero.

ex: A matriz quadrada A é unipotente (ou seja, $A - I$ é nilpotente) sse o seu polinómio caraterístico é uma potência de $z - 1$, e portanto os seus valores próprio são todos iguais a 1.

ex: Determine valores e vetores próprios dos endomorfismos definidos pela seguintes matrizes

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 1/3 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

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

$$A = \begin{pmatrix} 2 & 5 & -1 \\ 0 & -3 & 1 \\ 0 & 0 & 7 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 5 & -1 \\ 0 & -2 & 1 \\ -4 & 0 & 3 \end{pmatrix}$$

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix} \qquad A = \begin{pmatrix} 7 & 5 & 1 \\ 0 & -2 & 1 \\ 20 & 0 & 3 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

ex: [Ap69] **4.10**.

Matrizes semelhantes e diagonalização. As matrizes quadradas $A, B \in \text{Mat}_{n\times n}(\mathbb{R})$ são ditas *semelhantes* se existe uma matriz invertível $U \in \text{Mat}_{n\times n}(\mathbb{R})$ tal que

$$B = U^{-1}AU\,,$$

ou seja, se representam a mesma transformação linear $L : \mathbb{R}^n \to \mathbb{R}^n$ em bases que podem ser diferentes. Se A e B são semelhantes então $\text{Det}A = \text{Det}B$. Matrizes semelhantes têm o mesmo polinómio caraterístico, pois

$$\begin{aligned}\text{Det}\left(zI - U^{-1}AU\right) &= \text{Det}\left(zU^{-1}IU - U^{-1}AU\right)\\ &= \text{Det}\left(U^{-1}(zI-A)U\right) = \text{Det}\left(zI - A\right)\end{aligned}$$

e portanto os mesmos valores próprios.

A matriz quadrada $A \in \text{Mat}_{n\times n}(\mathbb{R})$ é *diagonalizável* se é semelhante a uma matriz diagonal. Se (o operador linear $L : \mathbb{R}^n \to \mathbb{R}^n$ definido na base canónica pel)a matriz quadrada $A \in \text{Mat}_{n\times n}(\mathbb{R})$ admite n vetores próprios linearmente independentes $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, com valores próprios $\lambda_1, \lambda_2, \ldots, \lambda_n$, respetivamente (não necessariamente distintos), e se $U \in \text{Mat}_{n\times n}(\mathbb{R})$ é a matriz (invertível, pela independência dos $\mathbf{v}_k$'s) cujas colunas são os vetores $\mathbf{v}_k$, então $U^{-1}AU$ é a matriz diagonal

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Se $\lambda_1, \lambda_2, \ldots, \lambda_n$ são as raizes (em geral, complexas e distintas) do polinómio caraterístico da matriz quadrada $A \in \text{Mat}_{n\times n}(\mathbb{C})$, então

$$\begin{aligned}P_A(t) &= (z-\lambda_1)(z-\lambda_2)\cdots(z-\lambda_n)\\ &= z^n - (\lambda_1+\lambda_2+\cdots+\lambda_n)\,z^{n-1} + \cdots + (-1)^n(\lambda_1\lambda_2\cdots\lambda_n)\end{aligned}$$

Mas $P_A(0) = \text{Det}A$, e portanto

$$\boxed{\text{Det}A = \lambda_1\lambda_2\cdots\lambda_n}$$

Também acontece que

$$\boxed{\text{tr}A = \lambda_1+\lambda_2+\cdots+\lambda_n}$$

(pois $\text{tr}(AB) = \text{tr}(BA)$, e portanto $\text{tr}(\Lambda) = \text{tr}(U^{-1}AU) = \text{tr}(A)$)

ex: Mostre que se A e B são semelhantes então $\text{Det}A = \text{Det}B$.

ex: As matrizes

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{e} \qquad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

são semelhantes?

ex: Diagonalize as seguintes matrizes, ou mostre que não é possível.

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 1 & 3 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Gato de Arnold. Considere a transformação $L : \mathbb{R}^2 \to \mathbb{R}^2$ definida pela matriz

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

ou seja, $L(x, y) = (2x + y, x + y)$. O polinómio caraterístico é $P_A(z) = z^2 - 3z - 5$, e portanto os valores próprios são $\lambda_\pm = (3 \pm \sqrt{5})/2$. Vetores próprios correspondentes, que satisfazem $L(\mathbf{v}_\pm) = \lambda_\pm \mathbf{v}_\pm$, são (por exemplo)

$$\mathbf{v}_+ = (\varphi, 1) \qquad \text{e} \qquad \mathbf{v}_- = (1, -\varphi)$$

onde $\varphi = (1 + \sqrt{5})/2 \simeq 1.6180339887\ldots$ é a "razão" dos gregos. A matriz que representa L na base $(\mathbf{v}_+, \mathbf{v}_-)$ é a matriz diagonal $A' = \operatorname{diag}(\lambda_+, \lambda_-)$. Observem que $\lambda_+ > 1$ e $0 < \lambda_- < 1$, e portanto L estica os vetores da reta $\mathbb{R}\mathbf{v}_+$ e contrae os vetores da reta $\mathbb{R}\mathbf{v}_-$. Observem também que $\operatorname{Det} A = \lambda_+ \lambda_- = 1$. Em particular, L preserva as áreas. A matriz A é invertível, e a sua inversa A^{-1} tem entradas inteira (pois $\operatorname{Det} A = 1$). Em particular, L e L^{-1} preservam o subgrupo $\mathbb{Z}^2 \subset \mathbb{R}^2$, ou seja, $L(\mathbb{Z}^2) = \mathbb{Z}^2$. Isto implica que L define uma transformação invertível $T : \mathbb{T}^2 \to \mathbb{T}^2$ do "toro" $\mathbb{T}^2 := \mathbb{R}^2/\mathbb{Z}^2$, definida por

$$(x, y) + \mathbb{Z}^2 \mapsto (2x + y, x + y) + \mathbb{Z}^2$$

A "dinâmica" da transfomação T, ou seja, o comportamento das trajetórias $\mathbf{v} \mapsto T(\mathbf{v}) \mapsto T(T(\mathbf{v})) \mapsto \ldots$, é particularmente interessante, e T é o arquétipo de uma classe importante de transformações, ditas "hiperbólicas".

Cayley-Hamilton theorem. If $p \in \mathbb{C}[z]$ is a polynomial in the indeterminate z, for example $p(z) = a_k z^k + \cdots + a_1 z + a_0$, and A an $n \times n$ matrix, real or complex, one defines the matrix $p(A)$ as

$$p(A) := a_k A^k + \cdots + a_1 A + a_0 I\,.$$

For example, one may consider the characteristic polynomial $P_A(z)$ of a square matrix A, and try to compute $P_A(A)$.

If $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a diagonal matrix, a simple computation (using the fact that its powers are also diagonal with entries which are powers of the λ_i's) shows that $P_\Lambda(\Lambda) = 0$. A diagonalizable matrix as $D = U^{-1}\Lambda U$ has its powers of the form $D^k = U^{-1}\Lambda^k U$, and consequently $P_D(D) = U^{-1}P_\Lambda(\Lambda)U = 0$, since $P_D = P_\Lambda$.

The set of diagonalizable complex matrices is dense in the space $\operatorname{Mat}_{n\times n}(\mathbb{C}) \approx \mathbb{C}^{n^2}$. Indeed, it is the set of those matrices such that the characteristic polynomial $P_A(z)$ has n distinct complex roots. In particular, for any $n \times n$ matrix A, real or complex, we may find a sequence $(D_m)_{m\in\mathbb{N}}$ of diagonalizable complex matrices such that $D_m \to A$ (w.r.t. some compatible norm). By continuity we get

Teorema 12.3 (Cayley-Hamilton)**.** *Any square matrix A satisfies its characteristic equation, i.e.*

$$\boxed{P_A(A) = 0}$$

ex: If A is an invertible $n \times n$ matrix, then we may multiply the identity $P_A(A) = 0$ by A^{-1}, and obtain a formula for the inverse matrix A^{-1} as a function of the powers A^0, A, A^2, $\dots$, A^{n-1}. For example, show that the inverse of an invertible 2×2 matrix A is

$$A^{-1} = \frac{1}{\mathrm{Det} A} \left((\mathrm{tr} A) I - A \right) .$$

Referências

[Ap69] T.M. Apostol, *Calculus*, John Wiley & Sons, 1969 [*Cálculo*, Editora Reverté, 1999].

[Ar85] V.I. Arnold, *Equações diferenciais ordinárias*, MIR, 1985.

[Ar89] V.I. Arnold, *Metodi geometrici della teoria delle equazioni differenziali ordinarie*, Editori Riuniti - MIR, 1989.

[Ax15] S. Axler, *Linear Algebra Done Right*, Third edition, Springer, 2015.

[Ba77] F. Banino, *Geometria per fisici*, Feltrinelli, 1977.

[Bo89] N. Bourbaki, *Elements of Mathematics, Algebra I*, Springer, 1989.

[BR98] T.S. Blyth and E.F. Robertson, *Basic Linear Algebra*, McGraw Hill, 1998.

[Ef17] J. Efferon, *Linear Algebra*, `http://joshua.smcvt.edu/linearalgebra`, 2017.

[FIS03] S.H. Friedberg, A.J. Insel and L.E. Spence, *Linear Algebra*, Prentice Hall, 2003.

[Go96] R. Godement, *Cours d'algébre* (Troisième édition mise à jour), Hermann Éditeurs, 1996.

[Ha58] P.R. Halmos, *Finite dimensional vetor spaces*, Van Nostrand, 1958.

[KKR62] C. Kittel, W.D. Knight and M.A. Ruderman, *Berkeley Physics*, McGraw-Hill, 1962.

[La87] S. Lang, *Linear Algebra*, Third Edition, UTM Springer, 1987.

[La97] S. Lang, *Introduction to Linear Algebra*, Second Edition, UTM Springer, 1997.

[LL78] L.D. Landau e E.M. Lifshitz, *Mecânica*, MIR, 1978.

[Me00] C.D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, 2000.

[MB99] S. MacLane and G. Birkhoff, *Algebra (Third Edition)*, AMS Chelsea Publishing, 1999.

[RHB06] K.F. Riley, M.P. Hobson and S.J. Bence, *Mathematical Methods for Physics and Engineering*, Cambridge University Press, 2006.

[Se89] E. Sernesi, *Geometria 1*, Bollati Boringhieri, 1989.

[Sh77] E.G. Shilov, *Linear algebra*, Dover, 1977.

[St98] G. Strang, *Linear Algebra and its Applications*, Hartcourt Brace Jonovich Publishers, 1998.

[St09] G. Strang, *Introduction to Linear Algebra*, fourth edition, Wellesley-Cambridge Press and SIAM 2009.
`http://math.mit.edu/linearalgebra/` , MIT Linear Algebra Lectures

[Wa91] B.L. van der Waerden, *Algebra*, Springer, 1991 [*Moderne Algebra*, 1930-1931].

[We52] H. Weyl, *Space Time Matter*, Dover, 1952 [*Raum Zeit Materie*, 1921].