www.im.ufrj.br/cvga

CÁLCULO VETORIAL
&
GEOMETRIA ANALÍTICA

livro 2: o espaço & outros espaços

Felipe Acker

Instituto de Matemática
Universidade Federal do Rio de Janeiro

outubro de 2020

Este trabalho foi contemplado com auxílio financeiro, no âmbito do edital de *Apoio à produção de material didático para atividades de ensino e/ou pesquisa*, 2014, da

FAPERJ

Fundação Carlos Chagas Filho de Amparo
à Pesquisa do Estado do Rio de Janeiro

Sumário

Prefácio

Este **livro 2** do **Cálculo Vetorial e Geometria Analítica** foi sendo construído aos pedaços e acabou aproveitando a viagem pela Geometria Analítica a três dimensões para apresentar, também, uma introdução à Álgebra Linear. Os alunos e as alunas do curso de graduação em Matemática Aplicada da UFRJ têm sido submetidos, ao longo dos anos, às sucessivas edições. Pela primeira vez, em meio à pandemia que nos assola, este texto está servindo de suporte para um curso lecionado de forma não presencial, contando agora também com estudantes do recém iniciado curso de Engenharia Matemática.

Subvertendo um pouco a ordem tradicionalmente aceita, apresento, assim que me pareceu possível, três situações em que a Geometria (plana e espacial) oferece um suporte intuitivo e um charme decisivos para o tratamento de questões em que o número de variáveis é naturalmente bem maior do que três. Da mesma forma, preferi colocar logo, claramente, o desafio de definir a dimensão de um espaço vetorial. Considerando que a proposta é um minicurso de 4 semanas, para alunos de primeiro período, é provável que alguns capítulos, como o que trabalha de forma detalhada o determinante (após uma primeira abordagem simplificada), acabem ficando de fora das aulas. Mas não é má ideia tê-los ali, ao alcance dos olhos de um ou outro leitor mais curioso. Os vídeos das aulas podem ser acessados a partir da página **sites.google.com/matematica.ufrj.br/acker/**, que tem links para as diversas versões do curso.

Na edição de 2018 decidi transferir para o Livro 3 os capítulos referentes a perspectiva e compactificações do plano e do espaço, fazendo com que este Livro 2 se tornasse, praticamente, um texto introdutório de Álgebra Linear. Nesta edição de 2020 resolvi dar ao capítulo de matrizes de Markov um merecido destaque, enfatizando os aspectos geométricos da situação. Um novo capítulo, introduzindo os polinômios de Lagrange e os de Bernstein, foi acrescentado (curiosamente, eu o achei, pronto, na pasta em que estão os arquivos do Livro 2, no computador lá de casa; não tenho a menor lembrança de tê-lo escrito, mas o estilo, de fato, é o meu).

Agradeço mais uma vez ao colega Dinamérico Pombo Jr., pela revisão do texto (mas assumo os erros que introduzi, alterando posteriormente o que já parecia fechado), e a Bernardo da Costa, Monique Carmona, Orestes Piermatei Filho, Ricardo Rosa, Waldecir Bianchini (que criou os applets) e Umberto Hryniewicz, que contribuíram para que o texto viesse à luz. A tradução para arquivos digitais da maioria das minhas figuras a lapiseira Caran d'Ache foi feita por João Paulo Pinto Siqueira. João Paulo foi também meu assistente de direção na maior parte dos vídeos.

Felipe Acker
Santa Teresa, outubro de 2020

Capítulo 1

Vetores

Adotaremos neste texto um ponto de vista algo ”leviano”, com relação à Geometria Sintética. Isto significa afirmar várias coisas sem demonstração, tomando-as como óbvias. Acreditamos que é razoável proceder dessa maneira: o leitor já deve ter alguma experiência prévia com a Geometria, de forma que, provavelmente, não se sentirá ofendido.[1]

1.1 Segmentos orientados e vetores

Como no caso do plano, adotaremos a seguinte definição informal de vetor no espaço: um **vetor** é uma flechinha que pode ser transladada para qualquer ponto de espaço. Para esclarecer um pouco melhor o que isso significa, podemos convencionar:

Definição:

(i) Uma flechinha com pé em A e ponta em B é representada pelo par ordenado (A, B);

(ii) Duas flechinhas (A, B) e (C, D) representam o mesmo vetor se os segmentos AB e CD são paralelos e têm o mesmo comprimento e, além disso, são também paralelos e têm o mesmo comprimento os segmentos AC e BD (segmentos degenerados em um ponto, tipo (A, A), são paralelos a qualquer segmento; dois segmentos sobre a mesma reta são, sempre, paralelos).

Os exercícios a seguir dão sustentação à definição de vetor e às definições das operações com vetores. Se você nunca demonstrou rigorosamente os resultados básicos da Geometria Euclidiana, podem ser muito difíceis e trabalhosos. A ideia não é que você pare agora e vá rever toda a Geometria,mas vale a pena dar uma olhada e, pelo menos, conferir se acredita na veracidade das afirmações que os exercícios contêm. Assim, uma atitude razoável é trocar, nos enunciados, o comando *mostre* por *convença-se*. Em qualquer circunstância, fazer as figuras correspondentes a cada um dos exercícios pode ser quase tão útil quanto fazer as demonstrações.

Exercício 1.1 *Suponha que os segmentos AB e CD são paralelos e têm o mesmo comprimento e, além disso, são também paralelos e têm o mesmo comprimento os segmentos AC e BD. Suponha, também, que*

[1] A verdade, honestamente, é mais dura: demonstrar todos os fatos geométricos que aceitaremos como óbvios exigiria, certamente, um curso inteiro de Geometria Euclidiana

*os segmentos CD e EF são paralelos e têm o mesmo comprimento e, além disso, são também paralelos e têm o mesmo comprimento os segmentos CE e DF. Mostre que os segmentos AB e EF são paralelos e têm o mesmo comprimento e, além disso, são também paralelos e têm o mesmo comprimento os segmentos AE e BF. Conclua que, no conjunto dos pares ordenados de pontos do espaço, representar o mesmo vetor é uma **relação de equivalência**, isto é:*

(i) (A, B) e (A, B) sempre representam o mesmo vetor, quaisquer que sejam os pontos A e B;

(ii) se (A, B) e $C, D)$ representam o mesmo vetor, então (C, D) e (A, B) representam o mesmo vetor;

(iii) se (A, B) e $C, D)$ representam o mesmo vetor e (C, D) e (E, F) representam o mesmo vetor, então (A, B) e $E, F)$ representam o mesmo vetor.

Exercício 1.2 *Conclua, do exercício anterior (mesmo que não o tenha feito), que os pares ordenados de pontos do espaço estão divididos em **classes de equivalência**, dadas da seguinte forma: a classe de equivalência do par (A, B) é o conjunto de todos os pares (C, D) tais que (A, B) e (C, D) representam o mesmo vetor. Um **vetor** é, precisamente, uma tal classe de equivalência. A classe de equivalência de (A, B) é notada por $\overrightarrow{AB}$ e chamada de **vetor** $\overrightarrow{AB}$.*

Exercício 1.3 *Sejam $\overrightarrow{AB}$ um vetor e P um ponto. Mostre que existe um único ponto Q tal que $\overrightarrow{PQ}=\overrightarrow{AB}$. Mostre que existe um único ponto Q' tal que $\overrightarrow{Q'P}=\overrightarrow{AB}$.*

Exercício 1.4 *Sejam A e B pontos distintos e t um número real não negativo. Mostre que existe um único ponto C, na semirreta que começa em A e passa por B, tal que o comprimento do segmento AC é t vezes o comprimento do segmento AB.*

Exercício 1.5 *Mostre que $\overrightarrow{PP}=\overrightarrow{QQ}$, quaisquer que sejam os pontos P e Q.*

Definição: *Se P é um ponto qualquer de E, o vetor $\overrightarrow{PP}$ é chamado de **vetor nulo** e é notado por $\vec{0}$.*

Exercício 1.6 *Mostre que, se $\overrightarrow{AB}=\overrightarrow{CD}$, então os segmentos AB e CD têm o mesmo comprimento.*

Definição: *A **norma** do vetor $\overrightarrow{AB}$, notada por $|\overrightarrow{AB}|$, é o comprimento do segmento AB.*

Exercício 1.7 *Mostre que, se $\overrightarrow{AB}=\overrightarrow{A'B'}$ e $\overrightarrow{CD}=\overrightarrow{C'D'}$, então o ângulo entre as retas AB e CD é igual ao ângulo entre as retas $A'B'$ e $C'D'$.*

Definição: *O **ângulo** entre os vetores $\overrightarrow{AB}$ e $\overrightarrow{CD}$ é o (menor) ângulo entre as retas AB e CD.*

Exercício 1.8 *Mostre que, se $\overrightarrow{AB}=\overrightarrow{CD}$, então $\overrightarrow{BA}=\overrightarrow{DC}$.*

Definição: Sejam $\overrightarrow{AB}$ e $\overrightarrow{CD}$ dois vetores. A **soma** $\overrightarrow{AB} + \overrightarrow{CD}$ é o vetor $\overrightarrow{AE}$, sendo o ponto E tal que $\overrightarrow{BE}=\overrightarrow{CD}$.

Exercício 1.9 *Mostre que a definição acima faz sentido, isto é, depende apenas dos vetores e não das flechinhas escolhidas para representá-los.*

Definição: Sejam $\overrightarrow{AB}$ um vetor e t um número real. O **produto** de t por $\overrightarrow{AB}$ é o vetor $t\ \overrightarrow{AB}=\overrightarrow{AC}$, sendo C o ponto definido da seguinte forma:

(i) $C = A$, se $A = B$;
(ii) Se $A \neq B$ e $t \geq 0$, então C é o ponto da semirreta começando em A e passando por B, tal que o comprimento do segmento AC é t vezes o comprimento do segmento AB.
(iii) Se $A \neq B$ e $t < 0$, então C é o ponto tal que $\overrightarrow{CA}=\overrightarrow{AC'}$, sendo C' o ponto da semirreta começando em A e passando por B, tal que o comprimento do segmento AC' é $|t|$ vezes o comprimento do segmento AB.

Exercício 1.10 *Mostre que a definição acima faz sentido, isto é, depende apenas dos vetores e não das flechinhas escolhidas para representá-los.*

Exercício 1.11 *Mostre que o vetor nulo, $\vec{0}$, é o elemento neutro para a adição, isto é: $\vec{u} + \vec{0} = \vec{0} + \vec{u} = \vec{u}$, para qualquer $\vec{u}$ em V.*

Exercício 1.12 *Mostre que, para qualquer vetor $\overrightarrow{AB}$, vale $\overrightarrow{AB} + \overrightarrow{BA}= \vec{0}$. Chamamos $\overrightarrow{BA}$ de $-\overrightarrow{AB}$.*

Exercício 1.13 *Mostre que $-\overrightarrow{AB}= (-1)\ \overrightarrow{AB}$*

Exercício 1.14 *Mostre que valem as propriedades a seguir, quaisquer que sejam os vetores $\vec{u}$, $\vec{v}$ e $\vec{w}$, e quaisquer que sejam os escalares (reais) s e t:*[2]

$$
\begin{aligned}
&(1)\ \vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}\\
&(2)\ \vec{u} + \vec{v} = \vec{v} + \vec{u}\\
&(3)\ \vec{u} + \vec{0} = \vec{u}\\
&(4)\ \vec{u} + (-\vec{u}) = \vec{0}\\
&(5)\ t(\vec{u} + \vec{v}) = t\vec{u} + t\vec{v}\\
&(6)\ (s + t)\vec{u} = s\vec{u} + t\vec{u}\\
&(7)\ s(t\vec{u}) = (st)\vec{u}\\
&(8)\ 1\vec{u} = \vec{u}
\end{aligned}
$$

Definição: Chamaremos de E o **espaço** (conjunto de pontos) e de V o **conjunto de vetores do espaço**.

Observação: Fixemos um ponto O de E, que chamaremos de **origem**. Pelo que acabamos de ver, fixar O estabelece uma bijeção entre E e V, dada por

[2] O vetor $\vec{0}$ é o definido por $\overrightarrow{PP}$, qualquer que seja o ponto P; se $\vec{u}$ é o vetor dado por $\overrightarrow{AB}$, $-\vec{u}$ é o vetor dado por $\overrightarrow{BA}$

$$\begin{array}{lcl} E: & \longrightarrow & V \\ P & \longrightarrow & \overrightarrow{OP} \end{array}$$

Os físicos costumam chamar $\overrightarrow{OP}$ de **vetor posição** de P.

Observação: Note que o espaço, E, está sendo considerado por nós uma entidade preexistente, com uma série de propriedades (que chamamos de **geométricas**). A partir daí construímos os vetores, por meio de uma relação de equivalência entre pares ordenados de pontos do espaço, que, graças a Deus, podemos agora esquecer e voltar a pensar como flechinhas. De qualquer forma, embora E e V sejam entidades bem diferentes, é maravilhoso que, por meio da simples fixação de um ponto de E (a origem), possamos facilmente identificá-los.

1.2 Coordenadas

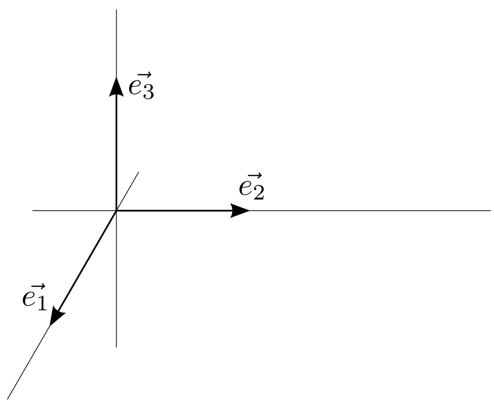


Figura 1.1: a base canônica

Assim como fazemos no plano, vamos considerar a **base canônica** para V, constituída pelos vetores $\vec{e_1}$, $\vec{e_2}$ e $\vec{e_3}$. Usando termos pouco precisos, fica entendido que $\vec{e_3}$ é vertical, com a ponta para cima, e que $\vec{e_1}$ e $\vec{e_2}$ são horizontais. O mais importante é que os três são de comprimento 1 e dois a dois ortogonais. Para cada vetor $\vec{v}$ do espaço, existe um único terno ordenado (x_1, x_2, x_3) de números reais tal que

$$\vec{v} = x_1\vec{e_1} + x_2\vec{e_2} + x_3\vec{e_3}.$$

Exercício 1.15 *Prove esta última afirmação (ou, pelo menos, pense nela).*

Os números $x_1,\ x_2$ e x_3 (ou o terno ordenado (x_1, x_2, x_3)) são ditos as **coordenadas** de $\vec{v}$ na base canônica; dizemos, também, que $\vec{v}$ é dado, na base canônica, por $(x_1, x_2.x_3)$.

Observação: Fixemos uma base canônica para V, que chamaremos de β, constituída pelos vetores $\vec{e_1}$, $\vec{e_2}$ e $\vec{e_3}$. Pelo que acabamos de ver, fixar β estabelece uma bijeção entre V e $\mathbb{R}^3$, dada por

$$\begin{array}{rcl} \mathbb{R}^3 : & \longrightarrow & V \\ (x_1, x_2, x_3) & \longrightarrow & x_1\vec{e_1} + x_2\vec{e_2} + x_3\vec{e_3} \end{array}$$

Essa bijeção é extremamente poderosa, pois preserva as operações: se ao vetor $\vec{u}$ correspondem as coordenadas (x_1, x_2, x_3) e ao vetor $\vec{v}$ correspondem as coordenadas (y_1, y_2, y_3), e se t é um número real, então ao vetor $\vec{u} + t\vec{v}$ correspondem as coordenadas $(x_1, x_2, x_3) + t(y_1, y_2, y_3) = (x_1 + ty_1, x_2 + ty_2, x_3 + ty_3)$. Dito de outra forma, a bijeção que a cada vetor associa suas coordenadas na base canônica (vistas como um terno ordenado) preserva as operações de adição e de multiplicação por escalar.

Exercício 1.16 *Observe que, como a fixação de uma origem, O, determina uma bijeção entre E e V, e a fixação de uma base, β, determina uma bijeção entre V e $\mathbb{R}^3$, segue que a fixação, simultânea, de uma origem, O, e de uma base, β, determina uma bijeção entre E e $\mathbb{R}^3$.*

Exercício 1.17 *Note que a bijeção, de $\mathbb{R}^3$ para E, funciona assim: dado o terno (x_1, x_2, x_3), tomamos o vetor $\vec{v} = x_1\vec{e_1} + x_2\vec{e_2} + x_3\vec{e_3}$; em seguida, tomamos o ponto P tal que $\overrightarrow{OP} = \vec{v}$.*

Definição: Um **sistema de coordenadas canônico** para o espaço E consiste em uma origem O e uma base canônica β (de maneira um pouco pedante, um sistema de coordenadas canônico é um par ordenado (O, β), sendo O um ponto de E e β uma base canônica para V).

1.3 Vetores e combinações lineares

Note que, dado um vetor não nulo $\vec{v}$ do espaço, os **múltiplos** de $\vec{v}$ (vetores $\vec{u}$ da forma $\vec{u} = t\vec{v}$ para algum escalar t), se marcados a partir de um mesmo ponto de origem O, estarão todos sobre uma mesma reta.

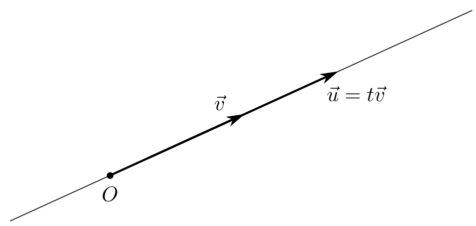


Figura 1.2:

Da mesma forma, se fixarmos dois vetores $\vec{u}$ e $\vec{v}$ tais que nenhum dos dois é múltiplo do outro (neste caso, $\vec{u}$ e $\vec{v}$ são ditos **linearmente independentes**), os vetores $\vec{w}$ da forma $\vec{w} = s\vec{u} + t\vec{v}$, com os escalares s e t percorrendo os números reais, estarão, se marcados a partir de um mesmo ponto de origem O, todos sobre um mesmo plano (isto está um pouco impreciso; veja uma versão mais precisa, como exercício, logo abaixo).

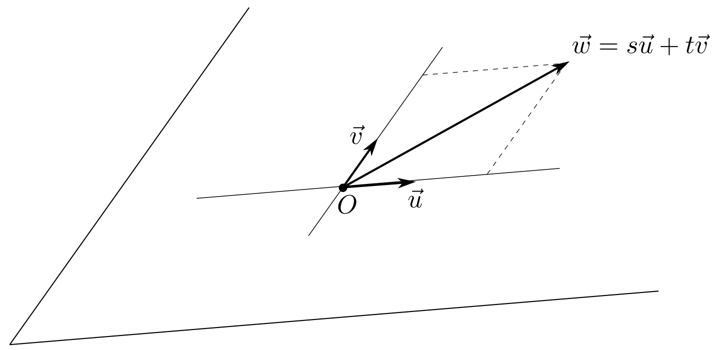


Figura 1.3:

Definição: Dados os vetores $\vec{u}$ e $\vec{v}$, um vetor $\vec{w}$ é dito uma **combinação linear** de $\vec{u}$ e $\vec{v}$ se existem escalares s e t tais que $\vec{w} = s\vec{u} + t\vec{v}$. De maneira mais geral, dados os vetores $\vec{v}_1, \ldots, \vec{v}_n$, um vetor $\vec{u}$ é dito combinação linear de $\vec{v}_1, \ldots, \vec{v}_n$ se existem escalares $t_1, \ldots, t_n$ tais que $\vec{u} = t_1\vec{v}_1 + \ldots + t_n\vec{v}_n$.

Exercício 1.18 *Note que, se fixamos um ponto O e dois vetores, $\vec{u}$ e $\vec{v}$, linearmente independentes, então os pontos P tais que $\overrightarrow{OP}$ é combinação linear de $\vec{u}$ e $\vec{v}$ estão todos sobre um mesmo plano passando por O.*

1.4 Retas e planos

A exemplo do que fizemos no plano, temos uma operação "bastarda", **somando vetores $\vec{v}$ a pontos** P. Neste caso, $P + \vec{v}$ é um novo ponto, Q, definido por: $P + \vec{v} = Q$ se $\overrightarrow{PQ} = \vec{v}$. Às vezes dizemos que o ponto Q é obtido *aplicando* o vetor $\vec{v}$ ao ponto P.

Observe que essa operação também é associativa: para qualquer ponto P e quaisquer vetores $\vec{u}$ e $\vec{v}$, vale $(P + \vec{u}) + \vec{v} = P + (\vec{u} + \vec{v})$.

Uma boa vantagem de lidar com retas de forma paramétrica é que não há diferença entre retas no plano e retas no espaço: dá-se um ponto P (de partida), um vetor não nulo $\vec{v}$ (velocidade) e consideram-se os pontos da forma

$$Q = p + t\vec{v}, \ t \in \mathbb{R}.$$

No caso do espaço, se P é dado por (x_0, y_0, z_0) e $\vec{v}$ por (a, b, c), teremos três equações, ditas **equações paramétricas da reta**:

$$\begin{cases} x = x_0 + at \\ y = y_0 + bt \\ z = z_0 + ct \end{cases}$$

Exercício 1.19 *Note que, ao contrário do que acontece no plano (em que as equações paramétricas são apenas duas), não podemos, eliminando t no sistema, deduzir uma única equação cartesiana para a reta.*

Para a representação paramétrica dos planos, trabalhamos com um ponto de base P e dois vetores $\vec{u}$ e $\vec{v}$; para que tenhamos, de fato, um plano, é preciso que $\vec{u}$ e $\vec{v}$ não sejam *um múltiplo do outro*. As combinações lineares de $\vec{u}$ e $\vec{v}$ nos fornecem um plano passando pela origem. Deslocá-lo para passar por P é como colocar a origem em P e tomar os pontos da forma

$$Q = P + s\vec{u} + t\vec{v},\ s,\ t \in \mathbb{R}.$$

1.5 O mistério da Santíssima Trindade

Pelo que acabamos de ver, a fixação de uma base para os vetores no espaço estabelece uma poderosa bijeção entre V e $\mathbb{R}^3$.[3]. Ao mesmo tempo, a fixação de uma origem, O, estabelece uma bijeção entre o espaço *puro* (E) e o conjunto dos vetores (V), dada por $P \mapsto \overrightarrow{OP}$. Assim, a fixação simultânea de uma base para V e de uma origem em E (que chamamos de um **sistema de coordenadas**) nos fornece duas bijeções, que, compostas, produzem uma bijeção entre pontos e ternos ordenados. Desta forma, fixado um sistema de coordenadas, passamos a identificar três conjuntos distintos, E, V e $\mathbb{R}^3$, enxergando **ponto, vetor e terno ordenado** como se fossem um única entidade, uma espécie de **Santíssima Trindade**.

$$\begin{tikzcd} & V \arrow[dr] & \\ E \arrow[ur] & & \mathbb{R}^3 \arrow[ll] \end{tikzcd}$$

Como trabalharemos sempre com coordenadas, é importante ter claro que, ao terno $(0, 0, 0)$ correspondem, simultaneamente, a origem O e o vetor nulo, assim como ao terno (x_1, x_2, x_3) correspondem, ao mesmo tempo, um ponto P e o vetor $\vec{OP}$.

[3] *Poderosa* no seguinte sentido: a bijeção preserva as operações (se $\vec{u}$ e $\vec{v}$ são associados aos ternos (x_1, x_2, x_3) e (y_1, y_2, y_3) e t é um escalar, então $\vec{u}+\vec{v}$ é associado a $(x_1, x_2, x_3) + (y_1, y_2, y_3)$ e $t\vec{u}$ é associado a $t(x_1, x_2, x_3)$. O termo erudito para uma bijeção preservando as operações é **isomorfismo**

Compreender essa identidade entre pontos, vetores e ternos ordenados é crucial, já que, no meio das contas, somente a intuição geométrica pode nos guiar, na hora de decidir se o terno (x_1, x_2, x_3) representa um ponto ou um vetor. Mais ainda: daqui para a frente, admitiremos que está fixada uma origem O em E e, com frequência, pensaremos vetores como se fossem pontos (estando, ao vetor $\vec{v}$, identificado o ponto P tal que $\vec{v} = \overrightarrow{OP}$). Isto nos permitirá trabalhar diretamente em V e nos referirmos ao *ponto* $\vec{v}$ (que está em V), como se fosse o próprio P (que está em E).

Exercício 1.20 *Observe que, dados um ponto P e um vetor $\vec{v}$, a reta r que passa por P e tem a direção de $\vec{v}$ é dada pelos pontos $Q = P + t\vec{v}$, com t variando em $\mathbb{R}$.*

Exercício 1.21 *Suponha fixado um sistema de coordenadas, no qual o ponto P é dado por (a_1, a_2, a_3) e o vetor $\vec{v}$ por (v_1, v_2, v_3). Traduza a observação do exercício anterior em coordenadas. Mostre que o ponto $Q = (x_1, x_2, x_3)$ está em r se, e somente se, existe um real t tal que*

$$\begin{cases} x_1 = a_1 + tv_1 \\ x_2 = a_2 + tv_2 \\ x_3 = a_3 + tv_3 \end{cases} .$$

*As três equações acima costumam ser chamadas de **equações paramétricas** de r.*

Capítulo 2

O produto escalar

2.1 Definição

A exemplo do que fizemos no plano, vamos definir o **produto escalar** de dois vetores $\vec{u}$ e $\vec{v}$ do espaço por

$$< \vec{u}, \vec{v} >= \vec{u} \cdot \vec{v} = |\vec{u}||\vec{v}| \cos \theta,$$

sendo θ o ângulo entre $\vec{u}$ e $\vec{v}$. Daí concluímos (geometricamente), que o produto escalar tem as propriedades:

$$\begin{array}{l}
(i) < \vec{u}, \vec{v} >=< \vec{v}, \vec{u} > \ \ \forall \vec{u}, \vec{v} \in V \\
(ii) < t\vec{u}, \vec{v} >= t < \vec{u}, \vec{v} > \ \ \forall \vec{u}, \vec{v} \in V, \ \forall t \in \mathbb{R} \\
(iii) < \vec{u}, \vec{v}_1 + \vec{v}_2 >=< \vec{u}, \vec{v}_1 > + < \vec{u}, \vec{v}_2 > \ \ \forall \vec{u}, \vec{v}_1, \vec{v}_2 \in V \\
(iv) < e_i, e_j >= 0, \ i \neq j, \ < e_i, e_j >= 1, \ i = j
\end{array}$$

Quando $< \vec{u}, \vec{v} >= 0$, os vetores $\vec{u}$ e $\vec{v}$ são ditos **ortogonais**.

Segue imediatamente das propriedades que, se os vetores $\vec{u}$ e $\vec{v}$ são dados, na base canônica, por $\vec{u} = u_1\vec{e_1} + u_2\vec{e_2} + u_3\vec{e_3}$ e $\vec{v} = v_1\vec{e_1} + v_2\vec{e_2} + v_3\vec{e_3}$, então

$$< \vec{u}, \vec{v} >= u_1v_1 + u_2v_2 + u_3v_3.$$

Exercício 2.1 *Desenvolva* $< \vec{u}, \vec{v} >=< u_1\vec{e_1} + u_2\vec{e_2} + u_3\vec{e_3}, v_1\vec{e_1} + v_2\vec{e_2} + v_3\vec{e_3} >$, *usando as propriedades (i), (ii), (iii) e (iv), e prove a fórmula acima.*

Daí decorrem algumas coisas importantes. Da definição segue que tanto distâncias como ângulos podem ser expressos por meio do produto escalar:

(i) a norma do vetor $\vec{v}$ é dada por $|\vec{v}| =< \vec{v}, \vec{v} >^{1/2}$;
(ii) o ângulo θ entre os vetores não nulos $\vec{u}$ e $\vec{v}$ é dado por

$$\cos \theta = \frac{< \vec{u}, \vec{v} >}{|\vec{u}||\vec{v}|}.$$

Mas, como acabamos de ver, o produto escalar pode ser calculado diretamente a partir das coordenadas dos vetores envolvidos. As duas fórmulas acima permitem, pois, que medidas de distâncias e de ângulos sejam feitas sem instrumentos, apenas com cálculos sobre as coordenadas. Note que, se no plano ainda é concebível marcar pontos e vetores a partir das coordenadas e efetuar as medidas diretamente com réguas e transferidores, no espaço a coisa seria quase impossível.

Exercício 2.2 *Sejam, no espaço, o triângulo Δ formado pelos pontos A, B e C dados, em coordenadas, por $A = (a_1, a_2, a_3)$, $B = (b_1, b_2, b_3)$ e $C = (c_1, c_2, c_3)$. Use o Teorema de Pitágoras para calcular as medidas dos lados de Δ. Em seguida, use a lei dos cossenos (ou o próprio Teorema de Pitágoras) para calcular o cosseno do ângulo $\hat{A}$. Faça a mesma coisa usando a fórmula do produto escalar e confira.*

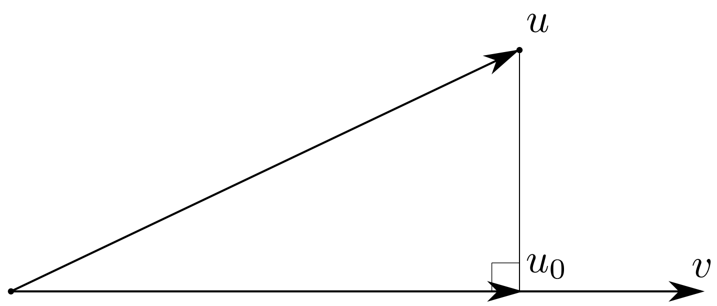


Figura 2.1:

Se $\vec{u}$ e $\vec{v}$ são dois vetores em V, com $|\vec{v}| = 1$, podemos projetar $\vec{u}$ na direção de $\vec{v}$, obtendo

$$\vec{u}_o = \langle \vec{u}, \vec{v} \rangle \, \vec{v}$$

($\vec{u}_o$ é dito a **projeção** de $\vec{u}$ na direção de $\vec{v}$).
Se $\vec{v}$ não é unitário, mas é não nulo, ainda podemos projetar $\vec{u}$ na direção de $\vec{v}$, fazendo

$$\vec{u}_o = \left\langle \vec{u}, \frac{1}{|\vec{v}|}\vec{v} \right\rangle \frac{1}{|\vec{v}|}\vec{v} = \frac{1}{|\vec{v}|^2} \langle \vec{u}, \vec{v} \rangle \, \vec{v} = \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}\vec{v}.$$

Se nossas contas fazem sentido, o triângulo de vértices $\mathbf{0}$, $\vec{u}_o$ e $\vec{u}$ tem um ângulo reto em $\vec{u}_o$.

Exercício 2.3 *Mostre que $\langle \vec{u} - \vec{u}_o, \vec{u}_o \rangle = 0$.*

Proposição: Se representamos um vetor $\vec{v}$ na base canônica, $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3$, então

$$v_1 = < \vec{v}, \vec{e}_1 >, \; v_2 = < \vec{v}, \vec{e}_2 >, \; v_3 = < \vec{v}, \vec{e}_3 > .$$

Demonstração: Podemos *enxergar* isso, ou simplesmente fazer o produto escalar dos dois lados da igualdade $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3$ por cada um dos vetores $\vec{e}_1$, $\vec{e}_2$ e $\vec{e}_3$. Por exemplo:

$$\begin{aligned} <\vec{v},\vec{e_2}> &= <v_1\vec{e_1}+v_2\vec{e_2}+v_3\vec{e_3},\vec{e_2}>= \\ &= <v_1\vec{e_1},\vec{e_2}>+<v_2\vec{e_2},\vec{e_2}>+<v_3\vec{e_3},\vec{e_2}>= \\ &= 0+v_2+0=v_2. \end{aligned}$$

■

2.2 Equação do plano

O produto escalar pode ser usado, de forma bastante simples, para se produzirem equações de planos (a exemplo do que se faz, em Geometria Analítica Plana, para a obtenção da equação da reta). Suponhamos que o plano α passa pelo ponto $P_0 = (x_0, y_0, z_0)$ e é normal ao vetor $\vec{v} = (a, b, c)$. Se buscamos determinar os pontos $P = (x, y, z)$ de α, basta observar que P estará no plano se e somente se o vetor $\overrightarrow{P_0P}$ for normal a $\vec{v}$, ou seja, $<\overrightarrow{P_0P}, \vec{v}>= 0$. Escrevendo em coordenadas, $P = (x, y, z)$ estará em α se e só se

$$0 =< (x - x_0, y - y_0, z - z_0), (a, b, c) >= ax + by + cz - (ax_0 + by_0 + cz_0).$$

Isso significa que, fazendo $d = -(ax_0 + by_0 + cz_0)$, o plano α é dado por

$$\alpha = \left\{(x, y, z) \in \mathbb{R}^3 \mid ax + by + cz + d = 0\right\}.$$

Exercício 2.4 *Faça, na fórmula que acabamos de obter, $\vec{v} = (a, b, c) = (1, 1, 0)$ e $P_0 = (x_0, y_0, z_0) = (0, 0, 0)$. Por que a equação obtida não é a de uma reta passando pela origem?*

Exercício 2.5 *Pense na seguinte forma alternativa de deduzir a equação do plano α passando por $P_0 = (x_0, y_0, z_0)$ e normal ao vetor $\vec{v} = (a, b, c)$: todos os pontos de α se projetam no mesmo ponto da reta que passa pela origem e tem a direção de $\vec{v}$. Consequentemente, $(x, y, z) \in \alpha \Leftrightarrow (x, y, z) \cdot (a, b, c) = (x_0, y_0, z_0) \cdot (a, b, c)$.*

figura

2.3 Segmentos e conjuntos convexos

Dados os pontos A e B, o **segmento de reta** AB (diz-se, também, apenas o **segmento** AB) é o conjunto

$$AB = \left\{A + t\,\overrightarrow{AB},\ t \in [0, 1]\right\}.$$

Exercício 2.6 *Convença-se de que $AB = BA$.*

Exercício 2.7 *Suponha fixada uma origem, de forma que A e B sejam dados, respectivamente, pelos vetores $\vec{u}$ e $\vec{v}$. Mostre que os pontos do segmento AB são dados pelos vetores da forma*

$$s\vec{u} + t\vec{v},\ s,t \in [0,1],\ s+t = 1,$$

ou, o que dá no mesmo,

$$t\vec{u} + (1-t)\vec{v},\ t \in [0,1].$$

Exercício 2.8 *Suponha fixado um sistema de coordenadas, de forma que $A = (x_1, y_1, z_1)$ e $B = (x_2, y_2, z_2)$. Mostre que*

$$(i) AB = \{s(x_1, y_1, z_1) + t(x_2, y_2, z_2),\ s,t \in [0,1],\ s+t = 1\}\,;$$

$$(ii) AB = \{t(x_1, y_1, z_1) + (1-t)(x_2, y_2, z_2),\ t \in [0,1]\}\,;$$

Pode ser útil distinguir **segmentos fechados, abertos** e **semiabertos**, com as notações naturais:

$$[A,B] = \left\{A + t\ \overrightarrow{AB},\ t \in [0,1]\right\};$$

$$[A,B[= \left\{A + t\ \overrightarrow{AB},\ t \in [0,1[\right\};$$

$$]A,B] = \left\{A + t\ \overrightarrow{AB},\ t \in]0,1]\right\};$$

$$]A,B[= \left\{A + t\ \overrightarrow{AB},\ t \in]0,1[\right\}.$$

Definição: Um subconjunto K do espaço é dito **convexo** se o segmento AB está contido em K sempre que A e B estejam em K.

Exemplo 1: Se $\vec{n}$ é um vetor não nulo e c é um número real, o **semiespaço**

$$K = \{\vec{v} \mid \langle \vec{v}, \vec{n}\rangle \leq c\}$$

é convexo. De fato, se $\vec{u}$ e $\vec{v}$ satisfazem $\langle \vec{u}, \vec{n}\rangle \leq c$ e $\langle \vec{v}, \vec{n}\rangle \leq c$, é certo que, sendo s e t não negativos e tais que $s + t = 1$, valem $\langle s\vec{u}, \vec{n}\rangle \leq sc$ e $\langle t\vec{v}, \vec{n}\rangle \leq tc$, de modo que, somando, temos $\langle s\vec{u} + t\vec{v}, \vec{n}, \rangle \leq c$. Note que, em coordenadas, isto significa dizer que, dados quatro reais, α, β, γ e c, com α, β e γ não simultaneamente nulos, o conjunto

$$K = \left\{(x,y,z) \in \mathbb{R}^3 \mid \alpha x + \beta y + \gamma z \leq c\right\}$$

é convexo. Note que K continua convexo se substituirmos $\leq$ por $<$, por $>$ ou por $\geq$.

Exemplo 2: A interseção de uma coleção qualquer de conjuntos convexos é um conjunto convexo. Assim, a interseção de uma coleção de semiespaços é, sempre, um convexo. Em particular, se K é interseção de uma coleção finita de semiespaços, isto é, se existem vetores não nulos, $\vec{u}_1, \ldots, \vec{u}_k$ e escalares $c_1, \ldots, c_k$ tais que

$$K = \{\vec{v} \mid \langle \vec{v}, \vec{u}_i \rangle \leq c_i,\ i = 1, \ldots, k\},$$

então K é convexo. Se K é interseção de uma coleção finita de semiespaços e é, além disso, **limitado** (isto é, se, qualquer que seja o vetor $\vec{u}$, o conjunto $\{\langle \vec{v}, \vec{u} \rangle,\ \vec{v} \in K\}$ é limitado), então K é dito um **poliedro convexo**.

Definição: Se K é convexo e P é um ponto de K, P é dito **ponto interno** de K se, para todo $\vec{v}$ não nulo, existem A e B na reta $\{P + t\vec{v},\ t \in \mathbb{R}\}$, distintos de P, tais que P pertence ao segmento AB; P é dito um **vértice** de K se não existem A e B em K, distintos de P, tais que P está no segmento AB.[1]

Exercício 2.9 *Considere, para cada vetor* ***unitário*** $\vec{u}$ *do espaço (isto é, tal que* $\langle \vec{u}, \vec{u} \rangle = 1$*), o semiespaço*

$$K_{\vec{u}} = \{\vec{v} \mid \langle \vec{v}, \vec{u} \rangle \leq 1\}.$$

Quem é o convexo K *interseção de todos os* $K_{\vec{u}}$*,*

$$K = \bigcap \{K_{\vec{u}}, |\vec{u}| = 1\}?$$

Quem são os vértices de K*?*

Exercício 2.10 *Sejam* $P_1 = (x_1, y_1, z_1)$*,* $P_2 = (x_2, y_2, z_2)$ *e* $P_3 = (x_3, y_3, z_3)$ *três pontos do espaço (poderiam, também, ser três pontos do plano). Mostre que o triângulo (cheio) de vértices* P_1, P_2 *e* P_3 *é*

$$T = \{P = p_1P_1 + p_2P_2 + p_3P_3 \mid p_1 \geq 0, p_2 \geq 0, p_3 \geq 0,\ p_1 + p_2 + p_3 = 1\}.$$

[1] O termo vértice faz mais sentido, é claro, no caso de ser K um poliedro; o que definimos como vértice costuma ser chamado, também de **ponto extremal** de K, mas o termo vértice tem mais apelo; o alcance mais geral que lhe dá esta definição tem lá suas razões de ser (ver Teorema de Krein-Milman)

Capítulo 3

Três problemas exemplares

Os amigos Xavier, Yakecan e Zumbi, também conhecidos com X, Y e Z, serão personagens das três pequenas histórias a seguir.

3.1 Amendoim torradinho

X, Y e Z vendem amendoim torrado em sociedade. Certo dia, X vai vender amendoim em Madureira e Y, no baixo Gávea (Madureira e Gávea são bairros do Rio de Janeiro). Z, que gerencia a sociedade, ficou, na véspera, pensando quanto amendoim cada um dos dois deveria levar para seu ponto de venda. Chamando de x a quantidade que caberá a X, e y a que caberá a Y, as limitações são as seguintes:

(i) $x + y \leq Q$, sendo Q a quantidade total de amendoim em estoque (em unidade de massa, convencionemos);
(ii) $x \leq Q_1$, sendo Q_1 o máximo de amendoim que X consegue carregar (em unidade de massa);
(iii) $y \leq Q_2$, sendo Q_2 o máximo de amendoim que Y consegue carregar (em unidade de massa);
(iv) $p_1 x + p_2 y \leq P$, sendo p_1 a quantidade de papel por unidade de massa necessária para embalar amendoim para venda em Madureira e p_2 a necessária para venda na Gávea (por razões de mercado, as embalagens usadas nos dois bairros são diferentes; P é a quantidade de papel para embalagem em estoque);
(v) $t_1 x + t_2 y \leq T$, sendo t_1 o tempo que Zumbi leva para embalar cada unidade de massa para venda em Madureira e t_2 o tempo que leva para embalar cada unidade de massa para venda na Gávea (Zumbi embala tudo sozinho e só dispõe de um tempo T para concluir esse serviço).

Pela experiência do grupo, ambos os vendedores venderão, em qualquer circunstância, toda a mercadoria que levarem. O objetivo é determinar x e y não negativos, satisfazendo as desigualdades (i), (ii), (iii), (iv) e (v), de tal maneira que o lucro seja máximo. Como o lucro por unidade vendida em Madureira é L_1, e o lucro por unidade vendida na Gávea é L_2, Zumbi tem que determinar x e y tais que

$$L(x, y) = L_1 x + L_2 y$$

seja máximo.[1]

Exercício 3.1 *Desenhe a região R do plano definida pelas condições (i), (ii), (iii), (iv) e (v), além de* $x \geq 0$ *e* $y \geq 0$. *Faça hipóteses razoáveis, tipo* $Q_1 < Q$ *e* $Q_2 < Q$. *Para dar mais graça, suponha* $p_1 Q < P$ *e* $p_2 Q_2 < P < p_2 Q$. *Mostre que a região R, com essas hipóteses, é um polígono convexo. Mostre que mesmo sem essas hipóteses adicionais R é um polígono convexo.*

Exercício 3.2 *Observe que a função* $L(x, y)$ *é constante em cada reta perpendicular ao vetor* (L_1, L_2). *Observe que L aumenta se caminhamos na direção e no sentido indicados pelo vetor* (L_1, L_2).

Exercício 3.3 *Dá para notar que o valor máximo de* $L(x, y)$, *para* (x, y) *em R, é atingido em um vértice de R?*

Exercício 3.4 *Dependendo de* Q, Q_1 Q_2, L_1, L_2, p_1 *e* p_2, *não seria mais lucrativo mandar Xavier para a Gávea e Yakecan para Madureira? Yakecan talez não concorde, pois, com seu jeito de índio, faz enorme sucesso entre as meninas do Baixo Gávea,*

Zumbi pensou um bocado, desenhou a região R e viu que era convexa e limitada. Entendeu o seguinte: cada desigualdade do tipo

$$ax + by \leq c,$$

com a, b e c reais fixos, pode ser reescrita, usando o produto escalar, como

$$\langle (x, y), (a, b) \rangle \leq c,$$

de forma que o conjunto solução,

$$E = \left\{ (x, y) \in \mathbb{R}^2 \mid \langle (x, y), (a, b) \rangle \leq c \right\},$$

corresponde a um semiplano. Como semiplanos são convexos e a interseção de convexos é um convexo, qualquer região do plano definida por desigualdades do tipo $ax + by \leq c$, como as que apareceram em seu problema, é convexa. Dependendo das circunstâncias a tal região pode até ser vazia; com sorte, porém, além de não vazia, será limitada. Contou suas descobertas a Xavier e Yakecan. Xavier disse: ora, é um **polígono convexo**!

Zumbi generalizou: depois de amanhã, quando eu também sair para vender, podemos interpretar as quantidades x, y e z como um ponto do espaço, dado, em coordenadas, por (x, y, z). As desigualdades, então, definem semiespaços; o conjunto de desigualdades define um conjunto K, que é a interseção dos semispaços correspondentes às desigualdades. Os elementos de K são os ternos ordenados (x, y, z) que correspondem às quantidades admissíveis, x, y e z (isto é, cada terno (x, y, z) corresponde a uma possível escolha de quantidades de amendoim a serem atribuídas, respectivamente, a X, a Y e a Z). Como cada semiespaço é um convexo, K é, também, convexo. E, se for limitado, é um **poliedro**, concluiu Xavier.

[1] A diferença de lucratividade é determinada não só pelas diferenças de custo das embalagens, mas também por diferenças nos preços praticados nos dois bairros, assim como outros fatores, como taxas de extorsão a serem pagas às máfias de cada local, suborno a policiais, etc.)

Nesse caso, o lucro pode ser uma função $L(x, y, z)$. Se Xavier vende x em Madureira, com lucro L_1 por unidade, Yakecan, na Gávea lucra L_2 por unidade e vende y, e eu, no Centro, vendo z e consigo $L_3 z$, nosso lucro total será

$$L(x, y, z) = L_1 x + L_2 y + L_3 z = \langle (x, y, z), (L_1, L_2, L_3) \rangle .$$

E se chutarmos, para começar, quantidades x, y e z? sugeriu Xavier. Isso, disse Zumbi! Vamos chamá-las de x_o, y_o e z_o. Se, partindo do ponto (x_o, y_o, z_o), procuramos aumentar o lucro, tem duas coisas claras: andando sobre o plano que passa por (x_o, y_o, z_o) e é normal a (L_1, L_2, L_3), o valor de L não muda; mas se formos para onde aponta o vetor (L_1, L_2, L_3), então L aumenta!

Exercício 3.5 *Volte ao caso do plano. Suponha que a região R é um polígono convexo e chute um ponto de partida, (x_o, y_o), no interior de R. Trace, passando por (x_o, y_o), uma reta, r, perpendicular a (L_1, L_2). Agora faça r ir andando, sempre perpendicular a (L_1, L_2) (e, portanto, paralelamente à posição inicial), no sentido para o qual (L_1, L_2) aponta, até o momento em que r deixa de tocar a região R. Conclua que o valor máximo de $L(x, y)$ é obtido em um vértice de R (note que, mesmo se na posição limite, a reta r contém um lado de R e, portanto, todos os pontos desse lado são de máximo, pode-se dizer que o máximo é atingido em um de seus vértices).*

Faça a mesma coisa em $\mathbb{R}^3$: suponha que K é um poliedro convexo e que queremos maximizar $L(x, y, z) = L_1 x + L_2 y + L_3 z$ em K. Chute um ponto de partida, (x_o, y_o, z_o), trace por ele um plano, α, perpendicular a (L_1, L_2, L_3); mova α, paralelamente à posição inicial, no sentido em que L aumenta, até a posição limite em que α deixa de tocar K. Conclua que o valor máximo de L é atingido em um vértice de K.

No dia seguinte, antes de saírem para o trabalho, Yakecan falou: esta noite tive um sonho. Sonhei que nossa sociedade tinha 623 vendedores. As quantidades de amendoim que cada um devia levar eram 623 números. No começo, estavam, todos os números, em fila, separados por vírgulas, pareciam formar uma enorme cobra. Depois, a cobra começou a se enroscar, se encaracolar...fiquei com medo e saí flutuando, enquanto os números iam se enroscando cada vez mais. Eu fui flutuando, me afastando, e, aos poucos, aquela maçaroca de números, de longe, foi virando apenas um pontinho de um espaço enorme...

Zumbi ficou matutando. Xavier anotou o número, 623, e, sem que os outros vissem, furtivamente, pegou todo o dinheiro que tinham guardado. Chegando a Madureira, jogou tudo no bicho.

3.2 Medindo o papel

X, Y e Z decidem criar uma embalagem padrão, a ser feita enrolando um retângulo de papel. Depois de alguns experimentos, chegam a um retângulo que parece ser o ideal. Para manter o padrão, resolvem medi-lo. Xavier mede, mas, por segurança, Yakecan e Zumbi conferem. O problema é que cada um obtém uma medida diferente para o comprimento. Yakecan acha que a diferença é mínima, irrelevante, mas Xavier insiste em ter um número exato. Yakecan sugere, então, que se adote a média aritmética: somam-se os três números, x, y e z, e divide-se o resultado por três. Zumbi acha que tanto faz, mas tem a curiosidade de saber qual seria a decisão correta, por algum critério razoável. Pensa um pouco e apresenta a ideia a seguir.

Se não houvesse erro, não teríamos três medidas diferentes, x, y e z, mas três medidas iguais; m, m e m. Se pensarmos a três medidas como um vetor,

$$\vec{v} = (x, y, z),$$

deveríamos ter, idealmente, $\vec{v}$ sobre a reta

$$r = \{(t, t, t),\ t \in \mathbb{R}\} = \{t(1, 1, 1),\ t \in \mathbb{R}\}.$$

Não sabemos qual é o ponto certo em r, $\vec{m} = (m, m, m)$; tudo que temos é um ponto aproximado, $\vec{v} = (x, y, z)$. A melhor escolha, então, propõe Zumbi, é o ponto $\vec{v}_o = (\bar{m}, \bar{m}, \bar{m})$ de r mais próximo de $\vec{v}$. Ora, conclui, tal ponto é a projeção de $\vec{v}$ sobre r. Para calculá-la, basta, como já vimos na página 10, tomar o vetor $\vec{w} = (1, 1, 1)$, que dá a direção de r, e fazer

$$\vec{v}_o = \frac{\langle \vec{v}, \vec{w} \rangle}{\langle \vec{w}, \vec{w} \rangle} \vec{w} = \frac{x + y + z}{3} (1, 1, 1),$$

de modo que a melhor escolha, na sua opinião, é

$$\bar{m} = \frac{x + y + z}{3}.$$

Eu não disse?, triunfa Yakecan. Xavier, cético, observa: é, mas o seu método não funciona se, em vez de 3 medidas, tivéssemos 4, 5, ou um número n, qualquer, maior. Zumbi não diz nada; deita na rede de Yakecan, fecha os olhos...e fica matutando. Depois de um tempo levanta, pega papel e lápis, e se põe a escrever e rabiscar; depois deita de novo na rede...fica umas horas nesse deita-levanta-pensa-escreve-rabisca. Por fim, chama os amigos e tasca a seguinte história.

Suponhamos que estamos diante de 10 medidas de um mesmo objeto, todas igualmente confiáveis. Podemos representá-las por um ponto $\mathbf{x} = (x_1, \ldots, x_{10})$ em $\mathbb{R}^{10}$. No entanto, o *resultado esperado*, ao se medir 10 vezes o mesmo objeto, seria $\bar{\mathbf{x}} = (\bar{x}, \ldots, \bar{x})$, com 10 medidas iguais. Nosso problema é, pois, a partir da n-upla $\mathbf{x}$, obter uma nova n-upla, $\bar{\mathbf{x}}$, com todas as coordenadas iguais e que, de alguma forma, esteja o mais próxima possível de $\mathbf{x}$. Ora, se fossem apenas 3 medidas, poderíamos pensar $\mathbf{x}$ como um ponto do espaço; $\bar{\mathbf{x}}$ deveria, então, ser o ponto da reta

$$r = \{(t, t, t), t \in \mathbb{R}\} = \{t(1, 1, 1),\ t \in \mathbb{R}\}$$

mais próximo de $\mathbf{x}$. A solução, geométrica, é projetar $\mathbf{x}$ sobre r. Como r passa pela origem, isso equivale a pensar $\mathbf{x}$ como um vetor e projetá-lo sobre o vetor $\vec{v} = (1, 1, 1)$. A projeção, como já vimos, é dada por

$$\bar{\mathbf{x}} = \frac{\langle \mathbf{x}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}\vec{v} = \frac{1}{3}\langle \mathbf{x}, (1,1,1)\rangle (1,1,1),$$

o que nos dá

$$\bar{\mathbf{x}} = (\bar{x}, \bar{x}, \bar{x}),\ \bar{x} = \frac{x_1 + x_2 + x_3}{3}.$$

Agora pensem um elemento de $\mathbb{R}^{10}$ como um vetor, com as operações de adição e multiplicação por escalar definidas como em $\mathbb{R}^3$. Vamos definir, em $\mathbb{R}^{10}$, o **produto escalar** $< \mathbf{x}, \mathbf{y} >$ por

$$< \mathbf{x}, \mathbf{y} >= x_1 y_1 + \ldots + x_{10} y_{10}.$$

Esse produto escalar, $< \mathbf{x}, \mathbf{y} >= x_1 y_1 + \ldots + x_{10} y_{10}$, tem as boas propriedades do produto escalar:

$$\begin{aligned}
&(i) < \mathbf{u}, \mathbf{v} >=< \mathbf{v}, \mathbf{u} > \ \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^{10} \\
&(ii) < t\mathbf{u}, \mathbf{v} >= t < \mathbf{u}, \mathbf{v} > \ \forall \mathbf{u}, \mathbf{v} \in V,\ \forall t \in \mathbb{R} \\
&(iii) < \mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2 >=< \mathbf{u}, \mathbf{v}_1 > + < \mathbf{u}, \mathbf{v}_2 > \ \forall \mathbf{u}, \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^{10} \\
&(iv) < \mathbf{e_i}, \mathbf{e_j} >= 0,\ i \neq j,\ < \mathbf{e_i}, \mathbf{e_j} >= 1,\ i = j,
\end{aligned}$$

sendo $\mathbf{e_i} = (0, \ldots, 1, \ldots, 0)$, com o 1 na i-ésima coordenada (eu fiz as contas e vi que funciona, mas, depois, cada um de vocês deveria conferir, por conta própria). Podemos, então, definir, como em $\mathbb{R}^3$, a **distância** entre dois pontos, **u** e **v**, por

$$|\mathbf{u} - \mathbf{v}| = \sqrt{< \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} >}.$$

A ideia, agora, é projetar o ponto $\mathbf{x} = (x_1, \ldots, x_{10})$ sobre a reta r, definida em $\mathbb{R}^{10}$ por

$$r = \{(t, \ldots, t), t \in \mathbb{R}\} = \{t(1, \ldots, 1),\ t \in \mathbb{R}\}.$$

Eu tive essa ideia pensando que, em $\mathbb{R}^3$, o ponto de r mais próximo de **x** é a projeção de **x** em r. Eu descobri, conta Zumbi, excitadíssimo, uma versão mais geral do **Teorema de Pitágoras**: se

$$< \mathbf{u}, \mathbf{v} >= 0,$$

então podemos formar uma espécie de triângulo retângulo em $\mathbb{R}^{10}$, de catetos $|\mathbf{u}|$ & $|\mathbf{v}|$ e hipotenuza $|\mathbf{u} + \mathbf{v}|$. Mas $|\mathbf{u} + \mathbf{v}|^2 =< \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} >$, o que dá

$$|\mathbf{u} + \mathbf{v}|^2 =< \mathbf{u}, \mathbf{u} > + < \mathbf{v}, \mathbf{u} > + < \mathbf{u}, \mathbf{v} > + < \mathbf{v}, \mathbf{v} >= |\mathbf{u}|^2 + 0 + 0 + |\mathbf{v}|^2 \ !$$

Se imitarmos o caso de $\mathbb{R}^3$, devemos fazer $\vec{v} = (1, 1, \ldots, 1)$. A **projeção**, como em $\mathbb{R}^3$, será dada por

$$\bar{\mathbf{x}} = \frac{\langle \mathbf{x}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}\vec{v} = \frac{1}{10}\langle \mathbf{x}, (1,1,\ldots,1) \rangle (1,1,\ldots,1),$$

ou seja,

$$\bar{\mathbf{x}} = (\bar{x}, \bar{x}, \ldots, \bar{x}),\ \bar{x} = \frac{x_1 + x_2 + \cdots + x_{10}}{10}.$$

Não precisam acreditar nos meus devaneios geométricos. Vou mostrar que a projeção assim obtida é o ponto $\bar{\mathbf{x}}$ de r mais próximo de $\mathbf{x}$. Acho bom vocês fazerem um desenho, pensando no caso $I\!R^3$ (serve, também, $I\!R^2$). Suponhamos que $\mathbf{y}$ seja um outro ponto de r; vou mostrar que a distância de $\mathbf{y}$ a $\mathbf{x}$ é maior que a de $\bar{\mathbf{x}}$ a $\mathbf{x}$. O importante, aqui, é que, como $\bar{\mathbf{x}}$ e $\mathbf{y}$ estão ambos em r, temos

$$< \mathbf{x} - \bar{\mathbf{x}}, \mathbf{y} - \bar{\mathbf{x}} >= 0.$$

Isso quer dizer que o triângulo $\mathbf{y}\bar{\mathbf{x}}\mathbf{x}$ é retângulo em $\bar{\mathbf{x}}$! Logo, pelo Teorema de Pitágoras-Zumbi, temos

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

concluiu Xavier.

Matou a cobra e mostrou o pau! exclamou Yakecan.

Exercício 3.6 *Mostre que é possível chegar ao mesmo resultado sem recorrer a toda a geometria. Pulando a definição do produto escalar, defina, em* $I\!R^{10}$, *a* ***norma*** *de* $\mathbf{x}$ *por* $|\mathbf{x}| = \sqrt{x_1^2 + \ldots + x_{10}^2}$. *Mostre que* $t = \bar{x}$ *é o ponto de mínimo de* $f(t) = |(x_1, \ldots, x_{10}) - (t, \ldots, t)|$.

Exercício 3.7 *Mas qual seria a mente doentia que, sem o suporte da poderosa analogia geométrica introduzida pelo produto escalar, defenderia a norma acima em detrimento de outras tão mais razoáveis, como* $|\mathbf{x}| = |x_1| + \ldots + |x_{10}|$? *Observe que ao termo* ***norma*** *corresponde a ideia de tamanho do vetor.*

3.3 Com quem está o anel?

Depois de dois meses de trabalho, sem um diazinho de folga, Xavier, Yakecan e Zumbi decidiram passar um domingo com a família de Yakecan (Yakecan tem 6 irmãos e irmãs, e 19 sobrinhos e sobrinhas). Xavier passou o dia todo brincando com a criançada. Na volta, no trem, falou:

Fiquei vendo as crianças brincando de passar o anel. Aos poucos percebi certos padrões: a Kauane, por exemplo, passa mais o anel para a Jaci do que para a Juraci, o Bira passa mais para a Nina do que para o Raoni...fiquei pensando se, anotando tudo por muito tempo, não seria possível ter uma tabela com as probabilidades de que cada criança passasse o anel para cada uma das outras, e se isso serviria para alguma coisa, tipo saber qual seria a probabilidade, a longo prazo, de que o anel ficasse com tal ou qual criança.

Yakecan pega um papel, lápis e rabisca:

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

Que porra é essa, Yakecan?

Ué, Zumbi, a_{ij} é a probabilidade de que a criança j, de posse do anel, passe para a criança i. Pensei usar as iniciais de cada uma, mas estava dando confusão, preferi numerar (no nosso caso, n=19). Por conta disso, a soma dos números em cada coluna vai ter que ser, sempre, igual a 1. Ou seja, para cada j,

$$a_{1j} + a_{2j} + \ldots + a_{nj} = 1.$$

Note, prossegue Yakecan, *que as regras permitem que quem está com o anel o guarde consigo.*

Quer dizer, então, que a coluna j,

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix},$$

nos mostra as probabilidades de que, em uma rodada em que o anel começa com a criança j, ele termine com a própria ou com cada uma das demais?

Exercício 3.8 *Suponha que $n = 3$ e que a matriz de Yakecan seja:*

$$\begin{pmatrix} 0,1 & 0,6 & 0,4 \\ 0,4 & 0,2 & 0,5 \\ 0,5 & 0,2 & 0,1 \end{pmatrix}.$$

Suponha que o anel comece com a criança de número 2. Pelo acima exposto, a probabilidade de que, ao final de uma rodada, o anel esteja com a criança número 1 é 0,6; com a criança número 2, é 0,2 e, com a número 3, é 0,2. Suponhamos, agora que, começando com a criança número 2, joguemos duas vezes. Quais são as probabilidades de que o anel esteja, ao final da segunda rodada, com cada uma das crianças?

Exercício 3.9 *Se essa história parece complicada, pense assim: se o anel começasse, por 10 rodadas, com a criança 2, a boa aposta seria que: em duas rodadas ela guardaria consigo o anel, em seis passaria para a criança 1 e nas outras duas passaria para a criança 3. Como, em cada um desses casos, a criança que recebeu o anel é quem vai passar na rodada seguinte, pode ser interessante trabalharmos com números um pouco maiores: se o anel começasse, por 100 rodadas, com a criança 2, a boa aposta seria que: em 20 rodadas ela guardaria consigo o anel, em 60 passaria para a criança 1 e nas outras 20 passaria para a criança 3. Agora vamos usar todas as informações da matriz de Yakecan. A criança 1 receberia o anel: 12 vezes, correspondendo às 20 rodadas em que a criança 2 ainda tinha o anel no começo da segunda rodada; 8 vezes, correspondendo aos 20 casos em que o anel, no começo da segunda rodada rodada, estaria com a criança 3; e ainda 6 vezes, correspondendo aos 60 casos em que ela mesma teria o anel no início da* *segunda rodada. Ou seja, se fizéssemos 100 vezes o experimento de começar a brincadeira com o anel na mão da criança 2, a boa aposta seria: em 26 dos casos o anel estaria, após duas roadas, com a criança 1, o que dá uma probabilidade de* $0,26$*. Faça o mesmo raciocínio para as crianças 2 e 3.*

Zumbi passou três dias matutando. Na quarta-feira, à noite, veio com a seguinte conversa:

*Andei pensando, pesquisando umas coisas...vamos fazer de conta que o número de crianças é 3, pra poder fazer uns desenhos. A ideia é ficar lá olhando, deixar rolar um monte de rodadas. Vamos chamar a matriz das probabilidades de matriz de Yakecan-**Markov**:*

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Agora imagina que nós não sabemos ao certo com quem começa o anel, só temos um chute sobre as probabilidades: p_1 *é a probabilidade de que comece com a criança 1,* p_2 *a de que comece com a 2,* p_3 *a de que comece com a 3. Notem que, necessariamente,* $p_1 + p_2 + p_3 = 1$*. Assim, podemos considerar uma chute sobre as probabilidades (que vou chamar de um **estado do sistema**) como um ponto no espaço. Mais especificamente, um **estado** vai ser um ponto do plano de equação* $p_1 + p_2 + p_3 = 1$*. Como as probabilidaddes não podem ser negativas, não é difícil descobrir que os estados vivem no triângulo (equilátero) de vértices* $(1,0,0)$, $(0,1,0)$ *e* $(0,0,1)$*. Nosso problema é: dado um estado inicial, como este evolui ao longo da brincadeira?*

Observem o seguinte: nosso estado pode ser visto, também, como um vetor.

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix},$$

Esse vetor evoluirá, após cada rodada, para

$$\begin{pmatrix} a_{11}p_1 + a_{12}p_2 + a_{13}p_3 \\ a_{21}p_1 + a_{22}p_2 + a_{23}p_3 \\ a_{31}p_1 + a_{32}p_2 + a_{33}p_3 \end{pmatrix} = p_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + p_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + p_3 \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix},$$

que é, por definição, o produto

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

da matriz de Yakecan-Markov pelo estado atual. Notem que, após a primeira rodada, os estados, que viviam no triângulo de vértices

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

passarão a viver no triângulo de vértices

$$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}.$$

Caceta, o triângulo encolhe! grita Xavier. *E esse novo triângulo, depois da segunda rodada, deve evoluir para outro. Se for sempre encolhendo, a cada rodada, talvez exista uma situação limite!*

Exercício 3.10 *Entenda o que Xavier está tentando dizer.*

Capítulo 4

O que é uma base?

A base canônica, $\{\vec{e_1},\ \vec{e_2},\ \vec{e_3}\}$ nos permite estabelecer uma bijeção entre o conjunto dos vetores do espaço e o conjunto $\mathbb{R}^3$ dos ternos ordenados de números reais. Como no caso do plano, essa bijeção preserva as operações: se os vetores $\vec{u}$ e $\vec{v}$ são representados, respectivamente, pelos ternos (x_1, x_2, x_3) e (y_1, y_2, y_3), então o vetor $\vec{u} + \vec{v}$ é representado por $(x_1 + y_1, x_2 + y_2, x_3 + y_3)$; se t é um escalar, $t\vec{u}$ é representado por (tx_1, tx_2, tx_3).

Cabem, naturalmente, as perguntas: sob que condições outros três vetores (mesmo que não formem uma base canônica), $\vec{v}_1$, $\vec{v}_2$ e $\vec{v}_3$, podem desempenhar o mesmo papel? o número 3 é mandatório, ou poderíamos conseguir a mesma coisa com 2, ou quem sabe 4 vetores? poderíamos, assim, estabelecer outras bijeções, entre o conjunto dos vetores do espaço e $\mathbb{R}^2$, ou $\mathbb{R}^4$, preservando as operações?

Como de hábito, vamos designar o conjunto dos vetores do espaço por V. Mas tentemos tratar a questão de forma abstrata, *esquecendo* a natureza de V (e o número cabalístico 3, mesmo que, depois, voltemos a ele). Mais precisamente, pensemos V apenas como um conjunto (de vetores), no qual estão definidas duas operações: dados dois vetores $\vec{u}$ e $\vec{v}$ de V, está definida sua **soma**, $\vec{u} + \vec{v}$, que é um vetor de V; da mesma forma, dados um vetor $\vec{v}$ de V e um número real t, está definido o **produto** de t por $\vec{v}$, $t\vec{v}$, que é um vetor de V. As operações assim definidas satisfazem as propriedades:

$$
\begin{array}{l}
(1)\ \vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w} \\
(2)\ \vec{u} + \vec{v} = \vec{v} + \vec{u} \\
(3)\ \vec{u} + \vec{0} = \vec{u} \\
(4)\ \vec{u} + (-\vec{u}) = \vec{0} \\
(5)\ t(\vec{u} + \vec{v}) = t\vec{u} + t\vec{v} \\
(6)\ (s + t)\vec{u} = s\vec{u} + t\vec{u} \\
(7)\ s(t\vec{u}) = (st)\vec{u} \\
(8)\ 1\vec{u} = \vec{u}
\end{array}
$$

(isto é: as propriedades (1) a (8) valem para quaisquer vetores $\vec{u}$, $\vec{v}$ e $\vec{w}$, e para quaisquer números s e t).

Exercício 4.1 *Observe que existem outros possíveis espaços, com as oito propriedades acima, além do que temos considerado: por exemplo, os elementos de* $\mathbb{R}^n$*, se somados e multiplicados por* *escalares da maneira habitual* ($(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$; $t(x_1, \ldots, x_n) = (tx_1, \ldots, tx_n)$)*, podem ser pensados como vetores, já que as oito propriedades são satisfeitas, e isso mesmo que* $n > 3$.

Definição: Um subconjunto ordenado $\beta = \{\vec{v}_1, \ldots, \vec{v}_n\}$ do espaço vetorial V é dito uma **base** de V se, e somente se, para todo vetor $\vec{v}$ de V, existe uma única n-upla $(x_1, \ldots, x_n)$ de $\mathbb{R}^n$ tal que

$$\vec{v} = x_1\vec{v}_1 + \ldots + x_n\vec{v}_n.$$

Os números $x_1, \ldots, x_n$ são chamados **coordenadas** do vetor $\vec{v}$ na base β.

Observe que são duas as condições para que $\beta = \{\vec{v}_1, \ldots, \vec{v}_n\}$ seja uma base de V. A primeira é que β deve **gerar** V (isto é, todo vetor de V deve ser combinação linear de $\vec{v}_1, \ldots, \vec{v}_n$).[1] A segunda é que não podem existir duas n-uplas distintas, $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$, que correspondam a um mesmo vetor, isto é: se

$$x_1\vec{v}_1 + \ldots + x_n\vec{v}_n = y_1\vec{v}_1 + \ldots + y_n\vec{v}_n,$$

então, necessariamente, $x_1 = y_1, \ldots, x_n = y_n$.

Exercício 4.2 *Mostre que nenhum dos vetores de uma base pode ser combinação linear dos demais (perderíamos a unicidade da representação).*

Exercício 4.3 *Entenda que uma base é uma espécie de chave para identificarmos cada vetor de* V *a uma n-upla de* $\mathbb{R}^n$*. Note que a exigência de ter* $\{\vec{v}_1, \ldots, \vec{v}_n\}$ *ordenado é necessária para garantir a unicidade da n-upla* $(x_1, \ldots, x_n)$.[2]

Definição: Os vetores $\vec{v}_1, \ldots, \vec{v}_n$ são ditos **linearmente independentes** se nenhum deles é combinação linear dos demais (diz-se, também, que o conjunto $\{\vec{v}_1, \ldots, \vec{v}_n\}$ é **linearmente independente**). Se, ao contrário, existe um que é combinação linear dos demais, então $\vec{v}_1, \ldots, \vec{v}_n$ são ditos **linearmente dependentes**.

Proposição: Fixada a base $\{\vec{v}_1, \ldots, \vec{v}_n\}$, a bijeção $\vec{v} \leftrightarrow (x_1, \ldots, x_n)$ preserva as operações de adição e de multiplicação por escalar. Isto é, se $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$ são as n-uplas associadas, respectivamente, aos vetores $\vec{u}$ e $\vec{v}$, e t é um escalar qualquer, então a n-upla $(x_1 + ty_1, \ldots, x_n + ty_n) = (x_1, \ldots, x_n) + t(y_1, \ldots, y_n)$ é a associada ao vetor $\vec{u} + t\vec{v}$.

Demonstração: Basta usar as propriedades das operações com vetores e com n-uplas. De fato, se $\vec{u} = x_1\vec{v}_1 + \ldots + x_n\vec{v}_n$, $\vec{v} = y_1\vec{v}_1 + \ldots + y_n\vec{v}_n$ e t é um escalar, então, usando muitas vezes as associatividades, as distributividades e a comutatividade, obtemos

$$\vec{u} + t\vec{v} = (x_1\vec{v}_1 + \ldots + x_n\vec{v}_n) + t(y_1\vec{v}_1 + \ldots + y_n\vec{v}_n) = (x_1 + ty_1)\vec{v}_1 + \ldots + (x_n + ty_n)\vec{v}_n.$$

[1] Neste caso, diz-se também que $\vec{v}_1, \ldots, \vec{v}_n$ **geram** V

[2] Seria mais rigoroso considerar, em vez de conjuntos ordenados de vetores $\{\vec{v}_1, \ldots, \vec{v}_n\}$, n-uplas $(\vec{v}_1, \ldots, \vec{v}_n)$ em V^n, o que é quase a mesma coisa (n-uplas admitem a possibilidade de $\vec{v}_i = \vec{v}_j$, com $i \neq j$). Estamos evitando fazê-lo por conta do risco de confundir o leitor...o que, paradoxalmente, pode gerar alguma confusão

■

É *visível* que nosso espaço V, dos vetores flechinhas, tem infinitas bases (compostas por quaisquer três vetores *não coplanares*). É natural, pois, que afirmemos e tentemos provar que V é **tridimensional**, isto é: que toda base de V tem exatamente três elementos. Seremos, porém, um pouco mais ambiciosos. Vamos tratar a questão, nos próximos capítulos, em um contexto geral: chamaremos de **espaço vetorial** qualquer conjunto em que estejam definidas operações de adição e de multiplicação por escalar, satisfazendo as oito propriedades acima listadas; em tais espaços, a definição de base, que acabamos de introduzir, também faz sentido. A questão passa a ser: se um determinado espaço vetorial, V, tem uma base formada por n vetores, será verdade que toda base de V tem exatamente n vetores? O custo de tal generalização é relativamente baixo, mas os benefícios são altamente compensadores.

Capítulo 5

Espaços vetoriais

5.1 O $\mathbb{R}^n$

O advento da Geometria Analítica, que pode ser datado de 1637, ano da publicação do *La Géométrie*, de Descartes, teve enorme impacto na Matemática. Questões geométricas, a partir daí, passaram a ser tratadas por meio das poderosas ferramentas algébricas que se desenvolviam então e, principalmente, com os métodos analíticos aportados pelo Cálculo Infinitesimal, que veio logo em seguida. Assim, pode-se dizer que teve por consequência uma crescente algebrização da Geometria.

$$\begin{array}{c}
\text{PROBLEMA GEOMÉTRICO} \\
\downarrow \\
\text{TRADUÇÃO ALGÉBRICA} \\
\downarrow \\
\text{SOLUÇÃO ALGÉBRICA} \\
\downarrow \\
\text{INTERPRETAÇÃO GEOMÉTRICA} \\
\downarrow \\
\text{SOLUÇÃO GEOMÉTRICA}
\end{array}$$

As afirmações do parágrafo anterior têm, porém, uma contrapartida um pouco menos óbvia e muito mais impactante: a algebrização da Geometria abre, imediatamente, caminho para a geometrização da Álgebra (e da Análise). Esquematicamente podemos colocar as coisas nos seguintes termos:

PROBLEMA ALGÉBRICO
$\downarrow$
INTERPRETAÇÃO GEOMÉTRICA
$\downarrow$
INSIGHT GEOMÉTRICO
$\downarrow$
MANIPULAÇÃO ALGÉBRICA
$\downarrow$
SOLUÇÃO ALGÉBRICA

Um exemplo elementar em que essa inversão se dá é a análise de gráficos de funções de uma variável: ao representarmos os pares ordenados $(x, f(x))$ por pontos no plano, obtemos um objeto geométrico (o gráfico de f, que é uma curva).[1] O problema de achar um ponto de máximo de f é interpretado geometricamente como a busca de um cume, que corresponde a uma tangente horizontal. Daí para a conclusão de que devemos achar um zero da derivada é um pulo.

Observação: Este é um bom exemplo para notarmos que a passagem da situação algébrica para uma contrapartida geométrica pode ir muito além de uma simples interpretação: o que fazemos, neste caso, é *investir* um problema "algébrico" de um significado geométrico que é, na verdade, não uma interpretação, mas uma verdadeira *fantasia geométrica*, que, poderíamos dizer, está mais próxima da arte e da poesia do que propriamente da técnica.

Um segundo passo é estender a intuição geométrica para *dimensões mais altas*. A solução de um problema envolvendo n variáveis pode não ser, essencialmente, diferente, se $n = 2$ ou se $n = 10$. Do ponto de vista geométrico, porém, no caso $n = 2$ podemos interpretar o par de variáveis (x_1, x_2) como um ponto do plano, e usar a intuição geométrica como ponto de partida para o tratamento do problema. Possivelmente, as mesmas manipulações algébricas que funcionam com $n = 2$, obtidas por um insight geométrico, servirão também para o caso $n = 10$.

Exemplo: Suponhamos que estamos diante de 10 medidas de um mesmo objeto, todas igualmente confiáveis. Podemos representá-las por um ponto $\mathbf{x} = (x_1, \ldots, x_{10})$ em $\mathbb{R}^{10}$. No entanto, o *resultado esperado*, ao se medir 10 vezes o mesmo objeto, seria $\bar{\mathbf{x}} = (\bar{x}, \ldots, \bar{x})$, com 10 medidas iguais. Nosso problema é, pois, a partir da n-upla $\mathbf{x}$, obter uma nova n-upla, $\bar{\mathbf{x}}$, com todas as coordenadas iguais e que, de alguma forma, esteja o mais próxima possível de $\mathbf{x}$. Ora, se fossem apenas 3 medidas, poderíamos pensar $\mathbf{x}$ como um ponto do espaço; $\bar{\mathbf{x}}$ deveria, então, ser o ponto da reta

[1] e isto mesmo em situações em que x e $f(x)$ podem não ter qualquer conotação geométrica: pense em x representando o tempo e $f(x)$ uma temperatura no instante x

$$r = \{(t,t,t), t \in \mathbb{R}\} = \{t(1,1,1),\ t \in \mathbb{R}\}$$

mais próximo de $\mathbf{x}$. A solução, geométrica, é projetar $\mathbf{x}$ sobre r. Como r passa pela origem, isso equivale a pensar $\mathbf{x}$ como um vetor e projetá-lo sobre o vetor $\vec{v} = (1,1,1)$. A projeção, como já vimos na página 10, é dada por

$$\bar{\mathbf{x}} = \frac{\langle \mathbf{x}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}\vec{v} = \frac{1}{3}\langle \mathbf{x}, (1,1,1) \rangle (1,1,1),$$

o que nos dá

$$\bar{\mathbf{x}} = (\bar{x}, \bar{x}, \bar{x}),\ \bar{x} = \frac{x_1 + x_2 + x_3}{3}.$$

Exercício 5.1 *Pense um elemento de $\mathbb{R}^{10}$ como um vetor, com as operações de adição e multiplicação por escalar definidas como em $\mathbb{R}^3$. Defina, em $\mathbb{R}^{10}$, o **produto escalar** $< \mathbf{x}, \mathbf{y} >$ por*

$$< \mathbf{x}, \mathbf{y} >= x_1 y_1 + \ldots + x_{10} y_{10}.$$

Mostre que o produto escalar $< \mathbf{x}, \mathbf{y} >= x_1 y_1 + \ldots + x_{10} y_{10}$ tem as boas propriedades do produto escalar:

$$\begin{array}{l}
(i) < \mathbf{u}, \mathbf{v} >=< \mathbf{v}, \mathbf{u} > \ \forall \mathbf{u}, \mathbf{v} \in V \\
(ii) < t\mathbf{u}, \mathbf{v} >= t < \mathbf{u}, \mathbf{v} > \ \forall \mathbf{u}, \mathbf{v} \in V,\ \forall t \in \mathbb{R} \\
(iii) < \mathbf{u}, \mathbf{v_1} + \mathbf{v_2} >=< \mathbf{u}, \mathbf{v_1} > + < \mathbf{u}, \mathbf{v_2} > \ \forall \mathbf{u}, \mathbf{v_1}, \mathbf{v_2} \in V \\
(iv) < \mathbf{e_i}, \mathbf{e_j} >= 0,\ i \neq j,\ < \mathbf{e_i}, \mathbf{e_j} >= 1,\ i = j,
\end{array}$$

sendo $\mathbf{e_i} = (0, \ldots, 1, \ldots, 0)$, com o 1 na i-ésima coordenada.

Exercício 5.2 *Projete o ponto $\mathbf{x} = (x_1, \ldots, x_{10})$ sobre a reta r, definida em $\mathbb{R}^{10}$ por*

$$r = \{(t, \ldots, t), t \in \mathbb{R}\} = \{t(1, \ldots, 1),\ t \in \mathbb{R}\}.$$

Mostre que a projeção assim obtida é o ponto $\bar{\mathbf{x}}$ de r que minimiza a distância (dada por $|\mathbf{x} - \bar{\mathbf{x}}| = \sqrt{< \mathbf{x} - \bar{\mathbf{x}}, \mathbf{x} - \bar{\mathbf{x}} >}$) a $\mathbf{x}$. Se tudo deu certo, você acaba de legitimar a média aritmética.

No exemplo acima, demos vários passos. Começamos com um problema envolvendo 10 variáveis (poderiam ser 15, 1000, ou um inteiro positivo qualquer). Pensamos no mesmo problema, mas com o número de variáveis reduzido para 3. Nesse caso, traduzimos nosso problema em um problema de geometria. Resolvemos o problema geometricamente. A solução geométrica nos indicou um critério que legitima a escolha de $\bar{x}$ como a melhor aproximação para a medida do nosso objeto, a partir das medidas x_1, x_2 e x_3. O próximo passo foi investir $\mathbb{R}^{10}$ de um caráter geométrico que, algebricamente, é análogo ao que temos em $\mathbb{R}^3$. Isto nos permitiu estabelecer um critério por meio do qual a escolha de $\bar{\mathbf{x}} = (\bar{x}, \ldots, \bar{x})$, com $\bar{x}$ dado pela média aritmética de $x_1, \ldots, x_{10}$, pode ser considerada a mais legítima. Mais ainda, a obtenção de $\bar{\mathbf{x}}$ se deu por meio dos cálculos sugeridos pelo problema análogo com apenas 3 variáveis!

Exercício 5.3 *Mostre que é possível chegar ao mesmo resultado sem recorrer a toda a geometria. Pulando a definição do produto escalar, defina, em $\mathbb{R}^{10}$, a **norma** de $\mathbf{x}$ por $|\mathbf{x}| = \sqrt{x_1^2 + \ldots + x_{10}^2}$. Mostre que $t = \bar{x}$ é o ponto de mínimo de $f(t) = |(x_1, \ldots, x_{10}) - (t, \ldots, t)|$.*

Exercício 5.4 *Mas qual seria a mente doentia que, sem o suporte da poderosa analogia geométrica introduzida pelo produto escalar, defenderia a norma acima em detrimento de outras tão mais razoáveis, como* $|\mathbf{x}| = |x_1| + \ldots + |x_{10}|$ *(observe que ao termo* ***norma*** *corresponde a ideia de tamanho do vetor)?*

Definição: O **espaço vetorial** $\mathbb{R}^n$ é definido pelo conjunto

$$\mathbb{R}^n = \{\mathbf{x} = (x_1, \ldots, x_n),\ x_1, \ldots, x_n \in \mathbb{R}\},$$

e as operações de adição e de multiplicação por escalar, dadas, respectivamente, para $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $t \in \mathbb{R}$, por:

$$(i)\ \mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n);$$
$$(ii)\ t\mathbf{x} = (tx_1, \ldots, tx_n).$$

Exercício 5.5 *Mostre que as operações acima definidas satisfazem às propriedades:*

$$\begin{aligned}
&(1)\ \mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}\\
&(2)\ \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}\\
&(3) \exists \mathbf{0} \in\ V \mid \mathbf{u} + \mathbf{0} = \mathbf{u}\\
&(4)\ \mathbf{u} + (-\mathbf{u}) = \mathbf{0}\\
&(5)\ t(\mathbf{u} + \mathbf{v}) = t\mathbf{u} + t\mathbf{v}\\
&(6)\ (s + t)\mathbf{u} = s\mathbf{u} + t\mathbf{u}\\
&(7)\ s(t\mathbf{u}) = (st)\mathbf{u}\\
&(8)\ 1\mathbf{u} = \mathbf{u}
\end{aligned}$$

(isto é: as propriedades (1) a (8) valem para quaisquer vetores $\mathbf{u}$, $\mathbf{v}$ *e* $\mathbf{w}$ *em* $\mathbb{R}^n$ *e para quaisquer números s e t).*

Definição: O **produto escalar canônico** em $\mathbb{R}^n$ é definido por

$$< \mathbf{x}, \mathbf{y} >= x_1 y_1 + \ldots + x_n y_n.$$

A **base canônica** de $\mathbb{R}^n$ é dada pelos vetores $\mathbf{e_1}, \ldots, \mathbf{e_n}$,

$$\mathbf{e_1} = (1, 0, \ldots, 0), \ldots, \mathbf{e_n} = (0, \ldots, 0, 1).$$

Exercício 5.6 *Mostre que o produto escalar* $< \mathbf{x}, \mathbf{y} >= x_1 y_1 + \ldots + x_n y_n$ *tem as boas propriedades do produto escalar:*

$$\begin{aligned}
&(i) < \mathbf{u}, \mathbf{v} >=< \mathbf{v}, \mathbf{u} > \ \forall \mathbf{u}, \mathbf{v} \in V\\
&(ii) < t\mathbf{u}, \mathbf{v} >= t < \mathbf{u}, \mathbf{v} > \ \forall \mathbf{u}, \mathbf{v} \in V,\ \forall t \in \mathbb{R}\\
&(iii) < \mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2 >=< \mathbf{u}, \mathbf{v}_1 > + < \mathbf{u}, \mathbf{v}_2 > \ \forall \mathbf{u}, \mathbf{v}_1, \mathbf{v}_2 \in V\\
&(iv) < \mathbf{e_i}, \mathbf{e_j} >= 0,\ i \neq j,\ < \mathbf{e_i}, \mathbf{e_j} >= 1,\ i = j,
\end{aligned}$$

Exercício 5.7 *Suponha que os vetores não nulos* $\mathbf{v}_1, \ldots, \mathbf{v}_k$ *de* $\mathbb{R}^n$ *sejam, dois a dois,* ***ortogonais****, isto é:* $< \mathbf{v}_i, \mathbf{v}_j >= 0$*, se* $i \neq j$ *e que*

$$\mathbf{v} = x_1 \mathbf{v}_1 + \ldots + x_n \mathbf{v}_n.$$

Mostre, sem usar coordenadas, que

$$x_i = \frac{< \mathbf{v}, \mathbf{v}_i >}{< \mathbf{v}_i, \mathbf{v}_i >}.$$

5.2 Outros espaços

Passar de $\mathbb{R}^2$ e $\mathbb{R}^3$ a $\mathbb{R}^n$ é um belo salto. Mas há outros *espaços*, envolvidos em situações bastante relevantes, que, também, merecem um pouco da nossa atenção. Comecemos com alguns exemplos. Em uns o espaço envolvido é um $\mathbb{R}^n$; em outros, algo maior.

Exemplo 1: Suponhamos que n corpos se movem no espaço, sujeitos apenas à atração gravitacional mútua. Podemos tomar cada corpo em separado, mas é claro que sua trajetória futura (em $\mathbb{R}^3$) dependerá, não só de suas posição e velocidade atuais, mas também das dos demais $n-1$ corpos. Assim, pode ser interessante pensar que o que temos é um elemento $\mathbf{x}(t)$ de $\mathbb{R}^{3n}$, que se move em função do tempo t (cada 3 coordenadas de $\mathbf{x}(t)$ correspondendo à posição de um dos corpos).

Exemplo 2: Quando vemos uma animação no monitor de um computador, o que varia com o tempo são certos parâmetros que determinam a representação da imagem no monitor. Tipicamente, um monitor pode ser pensado como feito de um monte de pontinhos (chamados **pixels**) arranjados em forma de matriz. Se m é o número de pixels em cada vertical e n o número na horizontal, temos uma matriz $m \times n$. Para definir a imagem, temos que (grosso modo), escolhidas 3 cores primárias (vermelho, verde e azul é uma das escolhas usuais), dizer, para cada pixel, quanto de cada uma vai entrar na definição da *cor* daquele pixel (no caso de imagens em preto e branco, precisaríamos de apenas um número por pixel). Assim, cada imagem é definida por $(m \times n) \times 3$ números. Podemos, considerando que os números que definem o quanto de cada cor entra em cada pixel são reais (na prática, usamos apenas uma gama finita de cada cor, algo como 2^8 ou 2^{16}), dizer que cada imagem é um vetor $\mathbf{x}(t)$ de $\mathbb{R}^{(m\times n)\times 3}$ (na prática, é claro, o tempo t também é picotado em uma quantidade finita de instantes, algo como 24, 30 ou 60 imagens por segundo). Nossa imagem, repetimos, evolui, ao longo do tempo, em $\mathbb{R}^{(m\times n)\times 3}$.

Exemplo 3: Nada nos impede de pensar no vídeo ideal. No lugar de $m \times n$ pixels, a tela poderia ser um retângulo, de dimensões a e b, com *infinitos pixels*: um para cada ponto do retângulo. Para simplificar (ou por razões artísticas), pensemos em preto e branco. Representada em coordenadas, nossa tela pode ser o retângulo $R = [0, a] \times [0, b]$. Cada imagem exigirá um número real para cada ponto do retângulo (0, digamos, para preto; 1 para branco, com valores entre 0 e 1 para os diversos tons de cinza). Assim, nossa imagem será, a cada instante, uma função $F : R \to \mathbb{R}$. O espaço em que evolui nossa imagem não é mais $\mathbb{R}^{(m\times n)\times 3}$, mas o *espaço* V das funções do retângulo R em $\mathbb{R}$.

Exemplo 4: Um exemplo um pouco mais *físico* talvez ajude. Imaginemos que o que se quer é descrever a evolução no tempo da distribuição de temperatura em uma barra de ferro (que, para simplificar, podemos supor unidimensional). Sendo L o comprimento da barra, a distribuição de temperatura, em cada instante t, será dada por uma função $f_t : [0, L] \to \mathbb{R}$ (se um ponto da barra é representado por x, sua temperatura no instante t será $f_t(x)$). Assim, nossa distribuição de temperaturas evolui no *espaço* V das funções de $[0, L]$ em $\mathbb{R}$. Se, no lugar de uma barra, tivéssemos um sólido, representado por um subconjunto B de $\mathbb{R}^3$, nossos possíveis estados estariam no *espaço* V das funções de B em $\mathbb{R}$. Questões como esta aparecem no livro *Théorie analytique de la propagation de la chaleur*, 1822, de Joseph Fourier. Nesse trabalho Fourier lança mão, num *espaço* de funções de um intervalo $[a,b]$ em $\mathbb{R}$, do primeiro **produto escalar** da história:[2] se f e g são funções, definidas no intervalo $[a,b]$ e a valores em $\mathbb{R}$, podemos definir $< f,g >$ por

$$< f,g >= \int_a^b f(x)g(x)dx.$$

Exemplo 5: O problema clássico da **braquistócrona** foi proposto em 1696 por João Bernoulli e chamou a atenção dos maiores matemáticos da época (Newton, Leibniz, Tiago Bernoulli, entre outros). Trata-se do seguinte: dados dois pontos, A e B, do espaço, encontrar uma curva γ ligando A a B (supostamente num plano vertical passando por A e B), tal que uma partícula que sobre ela deslize sem atrito e apenas sob a ação da gravidade, percorra γ no menor tempo possível. Supondo que a altura de A é superior à de B, podemos representar, em $\mathbb{R}^2$, A por $(0,0)$ e B por $(a,-b)$, $b>0$. Dando como óbvio que a curva em questão pode ser representada pelo gráfico de uma função $f:[0,a]\to\mathbb{R}$, pelo menos com derivada primeira contínua, podemos dizer que o que buscamos está no *espaço* V das funções $f:[0,a]\to\mathbb{R}$, tais que f tem derivada primeira contínua, com $f(0)=0$ e $f(a)=-b$.

Exercício 5.8 *Mostre que, nos exemplos 3 e 4, a soma de duas funções do espaço V está em V; mas não no exemplo 5. Mesma coisa para funções do tipo cf, com c real fixo e f em V (isto é: se f é uma função em V e c é um número real fixo, então a função cf, definida por $(cf)(x)=c(f(x))$, está em V, nos exemplos 3 e 4, mas não no exemplo 5).*

Exercício 5.9 *Mostre que toda função f de V, no exemplo 5, pode ser escrita como $f=f_0+h$, com f_0 sendo uma função qualquer de V, fixa, e $h:[0,a]\to\mathbb{R}$, com derivada primeira contínua e $h(0)=0=h(a)$ (f_0 pode ser, por exemplo, o segmento de reta ligando A a B).*

Exercício 5.10 *Seja V_0 o espaço das funções $h:[0,a]\to\mathbb{R}$, com derivada primeira contínua e tais que $h(0)=0=h(a)$. Mostre que a soma de duas funções de V_0 está em V_0; o mesmo para ch, se c é um número real fixo e h é uma função de V_0*

Exercício 5.11 *Seja X um conjunto qualquer. Seja V o espaço das funções $\mathbf{u}$ de X em $\mathbb{R}$. Defina, em V, as operações de **adição** e de **multiplicação por escalar** da seguinte forma: se $\mathbf{u}$ e $\mathbf{v}$ são duas funções de X em $\mathbb{R}$, a função $\mathbf{u}+\mathbf{v}$, de X em $\mathbb{R}$ é definida por $(\mathbf{u}+\mathbf{v})(x)=\mathbf{u}(x)+\mathbf{v}(x)$; se c é um número real fixo e $\mathbf{u}$ é uma função de X em $\mathbb{R}$, a função $c\mathbf{u}$, de X em $\mathbb{R}$, é definida por $(c\mathbf{u})(x)=c\mathbf{u}(x)$. Mostre que valem as seguintes e famosas propriedades:*

$$
\begin{array}{l}
(1)\ \mathbf{u}+(\mathbf{v}+\mathbf{w})=(\mathbf{u}+\mathbf{v})+\mathbf{w}\\
(2)\ \mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}\\
(3)\exists\mathbf{0}\in\ V\mid\mathbf{u}+\mathbf{0}=\mathbf{u}\\
(4)\ \mathbf{u}+(-\mathbf{u})=\mathbf{0}\\
(5)\ t(\mathbf{u}+\mathbf{v})=t\mathbf{u}+t\mathbf{v}\\
(6)\ (s+t)\mathbf{u}=s\mathbf{u}+t\mathbf{u}\\
(7)\ s(t\mathbf{u})=(st)\mathbf{u}\\
(8)\ 1\mathbf{u}=\mathbf{u}
\end{array}
$$

[2]Seria um pouco exagerado dizer que Fourier *inventou* o produto escalar, antes de Hamilton; mas é fato que, no trabalho de Fourier está a semente da ideia de ortogonalidade, em um contexto nada geométrico, que vai inspirar a construção do conceito abstrato que é hoje conhecido como **produto escalar** (ou **produto interno**)

(isto é: as propriedades (1) a (8) valem para quaisquer vetores **u**, **v** *e* **w**, *e para quaisquer números s e t).*

Exercício 5.12 *Note que, fazendo $X = \mathbb{N}$ no exercício anterior, temos o espaço das sequências infinitas de números reais, que é uma espécie de $\mathbb{R}^n$ com n infinito.*

Exemplo 6: Na Mecânica Quântica, o *estado* de uma partícula é dado por uma função de onda, que podemos supor definida em um subconjunto X do espaço. A novidade, nesse caso, é que a função de onda, φ, toma valores no conjunto dos números complexos. Mas isto não chega a ser um problema sério: funções a valores em $\mathbb{C}$ se somam, como as funções a valores em $\mathbb{R}$, e se multiplicam por escalares (complexos!). É fácil ver que, com estas operações, o *espaço* V das funções de X em $\mathbb{C}$ também goza das oito famosas propriedades acima.

Exemplo 7: Já que falamos em $\mathbb{C}$, podemos conceber, pairando em torno de cada $\mathbb{R}^n$, um espaço $\mathbb{C}^n$ de n-uplas de números complexos, que se somam da maneira óbvia e se multiplicam, também da maneira óbvia, por escalares complexos.

Para quem sabe o que é um **corpo**, podemos considerar, já que admitimos escalares complexos, situações em que os escalares sejam elementos de um corpo outro que $\mathbb{R}$ ou $\mathbb{C}$. Particularmente, na Teoria de Galois, é bastante natural a introdução de *espaços* com as mesmas oito propriedades acima, com a diferença que o conjunto dos escalares está (usualmente) em um corpo situado entre os racionais ($\mathbb{Q}$) e os complexos ($\mathbb{C}$).

Exemplo 8: Dado um corpo **K**, podemos considerar, exatamente como fizemos para $\mathbb{R}^n$ e $\mathbb{C}^n$, o *espaço* $\mathbf{K}^n$ das n-uplas de elementos de **K**, com as operações de sempre.

Exemplo 9: Um mesmo espaço pode aparecer disfarçado de várias maneiras. O conjunto dos polinômios a coeficientes reais e grau inferior a n é o $\mathbb{R}^n$ disfarçado. Seu irmão maior (embora mais novo), o dos polinômios a coeficientes complexos e grau inferior a n, é o $\mathbb{C}^n$ disfarçado. Da mesma forma, o conjunto das matrizes $m \times n$ (a coeficientes reais ou complexos) é o $\mathbb{R}^{m\times n}$, ou o $\mathbb{C}^{m\times n}$, em formação de parada militar.

Exemplo 10: Dentro da mesma linha de raciocínio, o conjunto de todos os polinômios a coeficientes reais é, disfarçado, o espaço das sequências $(x_n)_{n\in\mathbb{N}}$ de números reais tais que x_n é não nulo apenas para um número finito de índices n (esse número variando de sequência para sequência). O mesmo para o conjunto dos polinômios a coeficientes complexos e sequências $(z_n)_{n\in\mathbb{N}}$ de números complexos tais que z_n é não nulo apenas para um número finito de índices n.

5.3 Espaços vetoriais

A construção do conceito de **espaço vetorial** culmina, no século XX, com a definição geral que vai englobar o espaço dos vetores flechinhas, os $\mathbb{R}^n$ e todos os espaços mencionados na seção anterior. Mais precisamente, pensemos cada um dos *espaços* apresentados acima como um conjunto **V** de vetores. Dados dois vetores **u** e **v** de **V**, está definida sua **soma**, $\mathbf{u} + \mathbf{v}$, que é um vetor de **V**. Da mesma forma, dados um vetor

$\mathbf{v}$ de $\mathbf{V}$ e um escalar t, está definido o **produto** de t por $\mathbf{v}$, $t\mathbf{v}$, que é um vetor de $\mathbf{V}$. As operações assim definidas satisfazem às oito famosas propriedades.

Definição: Um **espaço vetorial** (real) é uma estrutura que compreende um conjunto, $\mathbf{V}$, e duas operações:

$$\begin{array}{rccc} + \ : & \mathbf{V}\times\mathbf{V} & \longrightarrow & \mathbf{V} \\ & (\mathbf{u},\mathbf{v}) & \longmapsto & \mathbf{u}+\mathbf{v} \end{array}$$

e

$$\begin{array}{rccc} : & \mathbb{R}\times\mathbf{V} & \longrightarrow & \mathbf{V}, \\ & (t,\mathbf{v}) & \longmapsto & t\mathbf{v} \end{array}$$

satisfazendo às oito famosas propriedades abaixo:

$$\begin{array}{l}
(1)\ \mathbf{u}+(\mathbf{v}+\mathbf{w}) = (\mathbf{u}+\mathbf{v})+\mathbf{w} \\
(2)\ \mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u} \\
(3)\ \exists\, \mathbf{0} \in\ V \mid \mathbf{u}+\mathbf{0} = \mathbf{u} \\
(4)\ \forall \mathbf{u} \in \mathbf{V}\ \exists -\mathbf{u} \in \mathbf{V} \mid \mathbf{u}+(-\mathbf{u}) = \mathbf{0} \\
(5)\ t(\mathbf{u}+\mathbf{v}) = t\mathbf{u}+t\mathbf{v} \\
(6)\ (s+t)\mathbf{u} = s\mathbf{u}+t\mathbf{u} \\
(7)\ s(t\mathbf{u}) = (st)\mathbf{u} \\
(8)\ 1\mathbf{u} = \mathbf{u}
\end{array}$$

(isto é: as propriedades (1) a (8) valem para quaisquer vetores $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{w}$ em $\mathbf{V}$, e para quaisquer números s e t).

Observação: O vetor nulo $\mathbf{0}$ mencionado na propriedade (3) é único. De fato, se $\square$ é outro vetor com a mesma propriedade, então

$$\square = \square + \mathbf{0} = \mathbf{0} + \square = \mathbf{0}.$$

Da mesma forma, o simétrico do vetor $\mathbf{u}$, $-\mathbf{u}$, também é único: se $\mathbf{u}+\mathbf{v} = \mathbf{0}$, então

$$\begin{array}{rcl}
\mathbf{v} & = & \mathbf{v}+\mathbf{0} = \mathbf{v}+(\mathbf{u}+(-\mathbf{u})) = (\mathbf{v}+\mathbf{u})+(-\mathbf{u})) = \\
 & = & (\mathbf{u}+\mathbf{v})+(-\mathbf{u}) = \mathbf{0}+(-\mathbf{u}) = -\mathbf{u}+\mathbf{0} = -\mathbf{u}.
\end{array}$$

Exercício 5.13 *Tente provar que* $\mathbf{u}+\mathbf{u} = 2\mathbf{u}$ *sem usar a propriedade* (8).

Observação: O leitor que não se assustou com a ideia de escalares complexos, pode incluir na definição a alternativa de que os escalares estejam não em $\mathbb{R}$, mas em $\mathbb{C}$. O leitor com familiaridade com o conceito de corpo pode considerar, na definição, que, junto com o conjunto $\mathbf{V}$ dos vetores, é dado também o corpo K dos escalares, e que a operação de multiplicação por escalares está definida não em $\mathbb{R}\times\mathbf{V}$, mas em $K\times\mathbf{V}$ (o que inclui, claro, as possibilidades $K = \mathbb{R}$ e $K = \mathbb{C}$). Salvo menção explícita em contrário, suporemos sempre que nossos espaços vetoriais são reais.

Um **subespaço vetorial** do espaço vetorial $\mathbf{V}$ é um subconjunto de $\mathbf{V}$ que, com as operações que "herda" de $\mathbf{V}$, é um espaço vetorial. A definição abaixo é mais operacional.

Definição: Um **subespaço vetorial** de $\mathbf{V}$ é um suconjunto W de $\mathbf{V}$, não vazio, tal que:

$$\begin{array}{ll} (i) & \mathbf{w} \in W,\ t \in \mathbb{R} \Longrightarrow t\mathbf{w} \in W \\ (ii) & \mathbf{w}_1, \mathbf{w}_2 \ \in \ W \Longrightarrow \mathbf{w}_1 + \mathbf{w}_2 \ \in \ W \end{array}$$

Exercício 5.14 *Mostre que todo plano passando pela origem é um subespaço vetorial de* $\mathbb{R}^3$.

Definição: O vetor $\mathbf{v}$ é dito uma **combinação linear** dos vetores $\mathbf{v}_1, \ldots, \mathbf{v}_n$ se existem escalares $t_1, \ldots, t_n$ tais que $\mathbf{v} = t_1\mathbf{v}_1 + \ldots + t_n\mathbf{v}_n$.

Definição: Dado um subconjunto X do espaço vetorial $\mathbf{V}$, o **subespaço gerado** por X é o conjunto de todas as combinações lineares de vetores de X (isto é: todos os vetores da forma $t_1\mathbf{v}_1 + \ldots + t_n\mathbf{v}_n$, sendo os t_i escalares e os $\mathbf{v}_i$ elementos de X). O espaço gerado pelo conjunto vazio é, por convenção, $\{\mathbf{0}\}$.

Exercício 5.15 *Mostre que a definição acima é equivalente a: o* ***subespaço gerado*** *por X é a interseção de todos os subespaços de* $\mathbf{V}$ *que contêm X.*

Exercício 5.16 *Se preferir, adote, como definição de subespaço gerado, a do exercício acima. A partir daí, faz sentido escolher a definição seguinte, mais elegante: sejam* $\mathbf{V}$ *um espaço vetorial e X um subconjunto de* $\mathbf{V}$*; o elemento* $\mathbf{v}$ *de* $\mathbf{V}$ *é* ***combinação linear*** *dos elementos de X se* $\mathbf{v}$ *pertence ao espaço gerado por X. Observe que, usando a definição de espaço gerado do exercício anterior, é natural concluir que* $\mathbf{0}$ *é o único vetor obtido como combinação linear dos elementos do conjunto vazio.*

Capítulo 6

Bases e dimensão

6.1 Bases

Definição: Um subconjunto ordenado $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ do espaço vetorial $\mathbf{V}$ é dito uma **base** de $\mathbf{V}$ se, e somente se, para todo vetor $\mathbf{v}$ de V, existe uma única n-upla $(x_1, \ldots, x_n)$ de $\mathbb{R}^n$ tal que

$$\mathbf{v} = x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n.$$

Os números $x_1, \ldots, x_n$ são chamados **coordenadas** do vetor $\mathbf{v}$.

Exercício 6.1 *Entenda que uma base é uma espécie de chave para identificarmos cada vetor de* $\mathbf{V}$ *a uma n-upla de* $\mathbb{R}^n$*. Note que a exigência de ter* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *ordenado é necessária para garantir a unicidade da n-upla* $(x_1, \ldots, x_n)$*.*[1]

Proposição: Fixada a base $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, a bijeção $\mathbf{v} \leftrightarrow (x_1, \ldots, x_n)$ preserva as operações de adição e de multiplicação por escalar. Isto é, se $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$ são as n-uplas associadas, respectivamente, aos vetores $\mathbf{u}$ e $\mathbf{v}$, e t é um escalar qualquer, então a n-upla $(x_1 + ty_1, \ldots, x_n + ty_n) = (x_1, \ldots, x_n) + t(y_1, \ldots, y_n)$ é a associada ao vetor $\mathbf{u} + t\mathbf{v}$.

Demonstração: Basta usar as propriedades das operações com vetores e com n-uplas. De fato, se $\mathbf{u} = x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n$, $\mathbf{v} = y_1\mathbf{v}_1 + \ldots + y_n\mathbf{v}_n$ e t é um escalar, então, usando muitas vezes as associatividades, as distributividades e a comutatividade, obtemos

$$\mathbf{u} + t\mathbf{v} = (x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n) + t(y_1\mathbf{v}_1 + \ldots + y_n\mathbf{v}_n) = (x_1 + ty_1)\mathbf{v}_1 + \ldots + (x_n + ty_n)\mathbf{v}_n.$$

$\blacksquare$

Exercício 6.2 *Estamos, implicitamente, excluindo a possibilidade de ser vazio o conjunto* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$*. Se o leitor tem tempo para a questão, pense nela: podemos conceber em abstrato o conceito de espaço*

[1] Seria mais rigoroso considerar, em vez de conjuntos ordenados de vetores $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, n-uplas $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ em $\mathbf{V}^n$, o que é quase a mesma coisa (n-uplas admitem a possibilidade de $\mathbf{v}_i = \mathbf{v}_j$, com $i \neq j$). Estamos evitando fazê-lo por conta do risco de confundir o leitor...o que, paradoxalmente, pode gerar alguma confusão

vetorial (um conjunto com operações de adição e de multiplicação por escalar, com as propriedades (1) a (8) apresentadas na seção anterior); neste caso, faz sentido pensar em um espaço reduzido ao vetor nulo, que teria como base o conjunto vazio.

Observe que são duas as condições para que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ seja uma base de $\mathbf{V}$. A primeira é que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ deve **gerar** $\mathbf{V}$ (isto é, todo vetor de $\mathbf{V}$ deve ser combinação linear de $\mathbf{v}_1, \ldots, \mathbf{v}_n$).[2] A segunda é que não podem existir duas n-uplas distintas, $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$, que correspondam a um mesmo vetor. Examinemos essa segunda condição.

Proposição: Sejam $\mathbf{v}_1, \ldots, \mathbf{v}_n$ vetores de $\mathbf{V}$. São equivalentes:

(i)existem duas n-uplas distintas $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$ tais que

$$x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n = y_1\mathbf{v}_1 + \ldots + y_n\mathbf{v}_n;$$

(ii)algum dos vetores $\mathbf{v}_1, \ldots, \mathbf{v}_n$ é combinação linear dos demais;
(iii)existe uma n-upla $(x_1, \ldots, x_n) \neq (0, \ldots, 0)$ tal que

$$\vec{0} = x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n.$$

Demonstração: Provaremos $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)$.

$(i) \Rightarrow (ii)$: Fixemos um i tal que $x_i \neq y_i$ (certamente existe, já que $(x_1, \ldots, x_n) \neq (y_1, \ldots, y_n)$). Como

$$x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n = y_1\mathbf{v}_1 + \ldots + y_n\mathbf{v}_n,$$

podemos concluir que

$$(y_i - x_i)\mathbf{v}_i = (x_1 - y_1)\mathbf{v}_1 + \ldots + (x_{i-1} - y_{i-1})\mathbf{v}_{i-1} + (x_{i+1} - y_{i+1})\mathbf{v}_{i+1} + \ldots + (x_n - y_n)\mathbf{v}_n$$

e, portanto, multiplicando dois lados por $(y_i - x_i)^{-1}$, obtemos $\mathbf{v}_i$ como combinação linear dos demais:

$$\mathbf{v}_i = \frac{(x_1 - y_1)}{(y_i - x_i)}\mathbf{v}_1 + \ldots + \frac{(x_{i-1} - y_{i-1})}{(y_i - x_i)}\mathbf{v}_{i-1} + \frac{(x_{i+1} - y_{i+1})}{(y_i - x_i)}\mathbf{v}_{i+1} + \ldots + \frac{(x_n - y_n)}{(y_i - x_i)}\mathbf{v}_n.$$

$(ii) \Rightarrow (iii)$: Se um certo $\mathbf{v}_i$ é combinação linear dos demais, podemos escrever, para certos $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$,

$$\mathbf{v}_i = (x_1\mathbf{v}_1 + \ldots + x_{i-1}\mathbf{v}_{i-1} + x_{i+1}\mathbf{v}_{i+1} + \ldots + x_n\mathbf{v}_n.$$

Logo,temos a n-upla $x_1, \ldots, x_{i-1}, -1, x_{i+1}, \ldots, x_n \neq (0, \ldots, 0)$, com

$$\vec{0} = x_1\mathbf{v}_1 + \ldots + x_{i-1}\mathbf{v}_{i-1} + (-1)\mathbf{v}_i + x_{i+1}\mathbf{v}_{i+1} + \ldots + x_n\mathbf{v}_n.$$

$(iii) \Rightarrow (i)$: Basta notar que, valendo (iii), temos $(0, \ldots, 0) \neq (x_1, \ldots, x_n)$ e

[2] Neste caso, diz-se também que $\mathbf{v}_1, \ldots, \mathbf{v}_n$ **geram V**

$$0\mathbf{v}_1 + \ldots + 0\mathbf{v}_n = x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n.$$

■

Definição: Os vetores $\mathbf{v}_1, \ldots, \mathbf{v}_n$ são ditos **linearmente independentes** se nenhum deles é combinação linear dos demais. Neste caso, diz-se também que o conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ é **linearmente independente**. No caso contrário, os vetores são ditos **linearmente dependentes** (ou que o conjunto é **linearmente dependente**.

Exercício 6.3 *Mostre que $\mathbf{v}_1, \ldots, \mathbf{v}_n$ são linearmente independentes se, e somente se, $t_1\mathbf{v}_1, \ldots, t_n\mathbf{v}_n = \vec{0}$ apenas no caso em que $t_1 = \ldots = t_n = 0$.*

Corolário: Para que um subconjunto ordenado $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ do espaço vetorial $\mathbf{V}$ seja uma base de $\mathbf{V}$, é necessário e suficiente que as duas condições abaixo sejam simultaneamente satisfeitas:

(i)$\mathbf{v}_1, \ldots, \mathbf{v}_n$ geram $\mathbf{V}$;
(ii)$\mathbf{v}_1, \ldots, \mathbf{v}_n$ são linearmente independentes.[3]

Observação: Pela definição que adotamos no início desta seção, há espaços vetoriais que não têm base (o espaço $\mathbb{R}[x]$ dos polinômios a coeficientes reais, por exemplo, não pode ser gerado por um conjunto finito de polinômios). Mesmo se não vamos trabalhar com bases infinitas, não é difícil dar uma definição mais geral. Diremos que um subconjunto α do espaço $\mathbf{V}$ **gera** $\mathbf{V}$ (ou, equivalentemente, que V é o **espaço gerado** por α) se todo elemento de $\mathbf{V}$ é combinação linear de um número finito de elementos de α (um bom exercício é mostrar que esta definição é equivalente à do exercício da página 37). Da mesma forma, diremos que α é **linearmente independente** (ou que os elementos de α são **linearmente independentes** - diz- se também que são **LI**) se nenhum elemento de α pertencer ao espaço gerado pelos demais (note que não exigimos que α seja finito).

Exercício 6.4 *Gaste um tempinho refletindo sobre as definições acima.*

Começamos esta seção com a definição de base de um espaço vetorial. Concluímos com uma definição equivalente (mas que, se adotarmos as definições alternativas, acima apresentadas, de espaço gerado e de independência linear, pode ser mais abrangente).

Definição: Um subconjunto α de um espaço vetorial $\mathbf{V}$ é dito uma **base** de $\mathbf{V}$ se satisfaz as duas condições:

$$(i)\alpha \text{ gera } \mathbf{V};$$
$$(ii)\alpha \text{ é } LI.$$

Exercício 6.5 *Encontre uma base para o espaço $\mathbb{R}[x]$ dos polinômios a coeficientes reais.*

Questão: É verdade que todo espaço vetorial tem base?

---

[3]Este corolário pode ser usado como definição de base; note que, neste caso, podemos "esquecer" a exigência de ordenação dos $\mathbf{v}_1, \ldots, \mathbf{v}_n$

6.2 Dimensão

É razoavelmente fácil compreender que, se os vetores $\mathbf{v}_1, \ldots, \mathbf{v}_m$ geram o espaço $\mathbf{V}$, podemos extrair de $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ um subconjunto $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ (ordenado) que é uma base de $\mathbf{V}$. Com efeito, se $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ não é linearmente independente, basta ir "jogando fora", passo a passo, os vetores supérfluos, até ficar com um subconjunto linearmente independente que ainda gera $\mathbf{V}$. Uma questão análoga é: e se o conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ é linearmente independente mas não gera $\mathbf{V}$, podemos acrescentar-lhe mais alguns vetores, de forma a obtermos uma base de $\mathbf{V}$?

Lema Fundamental: Suponha que os subconjuntos $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ e $\beta = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ do espaço $\mathbf{V}$ são tais que:

(i)α gera $\mathbf{V}$;
(ii)β é linearmente independente.

Então:

(i)$m \geq n$;
(ii)é possível substituir, em α, n dos $\mathbf{v}_i$ por $\mathbf{u}_1, \ldots, \mathbf{u}_n$, de forma que o novo conjunto assim obtido ainda gere $\mathbf{V}$.

Demonstração: Vamos construir $n+1$ subconjuntos de $\mathbf{V}$, que designaremos por $\alpha_0, \ldots, \alpha_n$, todos com m vetores, de forma que:

- $\alpha_0 = \alpha$
- α_{i+1} é obtido de α_i, se $i < n$, pela substituição de um elemento de α por um elemento de β, de forma que $\beta \subset \alpha_n$

Comecemos por α_1. Como α gera $\mathbf{V}$, podemos encontrar escalares $x_1, \ldots, x_m$ tais que

$$x_1\mathbf{v}_1 + \ldots + x_m\mathbf{v}_m = \mathbf{u}_1.$$

Note também que, como β é linearmente independente, $\mathbf{u}_1$ não é nulo. Logo, podemos escolher um j tal que $x_j \neq 0$; multiplicando dos dois lados por x_j^{-1} e isolando $\mathbf{v}_j$ do lado esquerdo, obtemos

$$\mathbf{v}_j = \frac{-x_1}{x_j}\mathbf{v}_1 + \ldots + \frac{1}{x_j}\mathbf{u}_1 + \ldots + \frac{-x_m}{x_j}\mathbf{v}_m.$$

Concluímos, pois, que $\mathbf{v}_j$ é combinação linear de $\mathbf{v}_1, \ldots, \mathbf{u}_1, \ldots, \mathbf{v}_m$. Assim, se α_1 é o conjunto obtido pela exclusão de $\mathbf{v}_j$ de α e sua substitutição por $\mathbf{u}_1$, temos que α_1 gera $\mathbf{V}$ (já que, com os vetores de α_1 conseguimos gerar $\mathbf{v}_j$ e, a partir daí, temos todos os originalmente em α, que gera $\mathbf{V}$).

Note que podemos escrever

$$\alpha_1 = \{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_m\},$$

com $\boldsymbol{\varepsilon}_1 = \mathbf{u}_1$ e $\{\boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_m\} \subset \alpha$. Suponhamos agora que, para um certo i menor do que m e n (ainda não provamos que $n < m$), tenhamos já definido o conjunto $\alpha_i = \{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_m\}$, tal que:

- $\boldsymbol{\varepsilon_1} = \mathbf{u}_1, \ldots, \boldsymbol{\varepsilon_i} = \mathbf{u}_i$
- $\{\boldsymbol{\varepsilon_{i+1}}, \ldots, \boldsymbol{\varepsilon_m}\} \subset \alpha$
- α_i gera $\mathbf{V}$

Podemos então encontrar escalares $x_1, \ldots, x_m$ tais que

$$x_1\boldsymbol{\varepsilon_1} + \ldots + x_m\boldsymbol{\varepsilon_m} = \mathbf{u}_{i+1},$$

ou seja,

$$x_1\mathbf{u}_1 + \ldots + x_i\mathbf{u}_i + x_{i+1}\boldsymbol{\varepsilon_{i+1}} + \ldots + x_m\boldsymbol{\varepsilon_m} = \mathbf{u}_{i+1}.$$

É hora de observar que, como β é linearmente independente, podemos tomar um $j > i$ tal que $x_j \neq 0$ (caso contrário, $\mathbf{u}_{i+1}$ seria combinação linear de $\mathbf{u}_1, \ldots, \mathbf{u}_i$). Isto significa que, procedendo como no caso de α_1 e definindo α_{i+1} pela exclusão de $\boldsymbol{\varepsilon_j}$ e sua substitutição por $\mathbf{u}_{i+1}$, α_{i+1} terá as mesmas propriedades de α_i acima listadas (devemos, é claro, renomear os vetores de α_{i+1} de forma que $\mathbf{u}_{i+1}$ passe a se chamar $\boldsymbol{\varepsilon_{i+1}}$).

O que acima fizemos mostra que podemos definir os conjuntos α_i, com as três propriedades acima, para todo i de 1 até o menor entre m e n. Isto prova:

(i)$m \geq n$, pois, caso contrário, $\mathbf{u}_n$ seria combinação linear de $\mathbf{u}_1, \ldots, \mathbf{u}_m$
(ii)α_n gera $\mathbf{V}$ e $\alpha_n = \{\mathbf{u}_1, \ldots, \mathbf{u}_n, \boldsymbol{\varepsilon_{n+1}}, \ldots, \boldsymbol{\varepsilon_m}\}$, com $\{\boldsymbol{\varepsilon_{n+1}}, \ldots, \boldsymbol{\varepsilon_m}\} \subset \alpha$ ■

Suponhamos agora que nosso espaço, $\mathbf{V}$, tenha uma base, α, formada por n elementos. Como α gera $\mathbf{V}$, todo conjunto linearmente independente em $\mathbf{V}$ tem, no máximo, n elementos; como α é linearmente independente, todo conjunto que gera $\mathbf{V}$ tem, no mínimo, n elementos. Daí segue um teorema crucial.

Teorema: Se o espaço vetorial $\mathbf{V}$ tem uma base com n elementos, então toda base de $\mathbf{V}$ tem, exatamente, n elementos. Além disso, se $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ é um subconjunto de $\mathbf{V}$ com n elementos, então são equivalentes:

(i)$\mathbf{v}_1, \ldots, \mathbf{v}_n$ formam uma base de $\mathbf{V}$;
(ii)$\mathbf{v}_1, \ldots, \mathbf{v}_n$ geram $\mathbf{V}$;
(iii)$\mathbf{v}_1, \ldots, \mathbf{v}_n$ são linearmente independentes.

Demonstração: A primeira afirmação é o que acabamos de concluir. Além disso, temos, obviamente, (i)$\Leftrightarrow$(ii) e (iii) (juntos). Basta, pois, provar que (ii)$\Leftrightarrow$(iii).

(ii)$\Rightarrow$(iii): Observe que, se fosse possível ter $\mathbf{V}$ gerado por β sem que β fosse linearmente independente, então poderíamos descartar pelo menos um elemento de β (que seja combinação linear dos demais), obtendo um conjunto de $n-1$ elementos que ainda geraria $\mathbf{V}$, o que, como acabamos de observar, não é possível. (ii)$\Rightarrow$(iii): Se, por outro lado, sabemos que β é linearmente independente, podemos, tomando uma base α de $\mathbf{V}$ (que, já sabemos, tem n elementos e gera $\mathbf{V}$), concluir diretamente, do Lema Fundamental, que β gera $\mathbf{V}$. ■

Diante do acima exposto, concluímos que existe um número bem determinado, n, tal que toda base do espaço $\mathbf{V}$ tem n elementos. Dizemos, pois, que a **dimensão** de $\mathbf{V}$ é n.

Capítulo 7

Produto escalar, de novo

Nos dois capítulos anteriores, com o conceito de espaço vetorial, abrimos caminho para *raciocinar geometricamente* em situações absolutamente não geométricas. Um espaço vetorial é, em muitos aspectos, parecido com o espaço físico em que nos movemos e com sua primeira versão abstrata, o espaço da Geometria clássica. No entanto, os axiomas de espaço vetorial não contemplam a possibilidade de se efetuarem medidas de distâncias ou de ângulos.

Quando trabalhamos com vetores flechinhas, no espaço tridimensional *herdado* da Geometria Euclidiana, essas medidas são feitas via produto escalar. Como já vimos, o produto escalar de vetores flechinhas é definido geometricamente por

$$\langle \vec{u}, \vec{v} \rangle = |\vec{u}||\vec{v}| \cos\theta,$$

sendo θ o ângulo entre $\vec{u}$ e $\vec{v}$. A definição puramente algébrica,

$$\langle \vec{u}, \vec{v} \rangle = u_1 v_1 + u_2 v_2 + u_3 v_3,$$

porém, se estende facilmente ao espaço $\mathbb{R}^n$: se $\mathbf{x} = (x_1, \ldots, x_n)$ e $\mathbf{y} = (y_1, \ldots, y_n)$, então

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \ldots + x_n y_n.$$

A exemplo do que aconteceu com o conceito de espaço vetorial, que evoluiu das flechinhas para espaços bem menos geométricos, podemos estabelecer uma definição geral de **produto escalar**, baseada apenas em algumas propriedades fundamentais.

Definição: Dado um espaço vetorial real $\mathbf{V}$, um **produto escalar** (também dito **produto interno**) em $\mathbf{V}$ é uma aplicação

$$\begin{array}{rccc} \langle , \rangle : & \mathbf{V} \times \mathbf{V} & \longrightarrow & \mathbb{R} \\ & (\mathbf{u}, \mathbf{v}) & \longmapsto & \langle \mathbf{u}, \mathbf{v} \rangle \end{array},$$

com as seguites propriedades:

$$\begin{array}{ll} (i) & \langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle \ \ \forall \mathbf{u}, \mathbf{v} \in \mathbf{V} \\ (ii) & \langle t\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v} \rangle = t \langle \mathbf{u}_1, \mathbf{v} \rangle + \langle \mathbf{u}_2, \mathbf{v} \rangle \ \ \forall \mathbf{u}_1, \mathbf{u}_2, \mathbf{v} \in \mathbf{V}, \ \forall t \in \mathbb{R} \\ (iii) & \langle \mathbf{u}, \mathbf{u} \rangle \geq 0 \ \forall \mathbf{u} \in \mathbf{V} \\ (iv) & \langle \mathbf{u}, \mathbf{u} \rangle = 0 \Longrightarrow \mathbf{u} = \mathbf{0} \end{array}$$

Exercício 7.1 *Prove, a partir da definição, as seguintes propriedades:*

$$\begin{array}{ll} (i) & \langle \mathbf{u}_1 + \mathbf{u}_2, \mathbf{v} \rangle = \langle \mathbf{u}_1, \mathbf{v} \rangle + \langle \mathbf{u}_2, \mathbf{v} \rangle \ \forall \mathbf{u}_1, \mathbf{u}_2, \mathbf{v} \in \mathbf{V} \\ (ii) & \langle \mathbf{0}, \mathbf{v} \rangle = 0 \ \forall \mathbf{v} \in \mathbf{V} \\ (iii) & \langle t\mathbf{u}, \mathbf{v} \rangle = t \langle \mathbf{u}, \mathbf{v} \rangle \ \forall \mathbf{u} \in \mathbf{V}, \ \forall t \in \mathbb{R} \end{array}$$

Solução de (ii): $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{0} + \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{0}, \mathbf{v} \rangle + \langle \mathbf{0}, \mathbf{v} \rangle$, etc.

Os dois exemplos a seguir são básicos.

Exemplo 1: Em $\mathbb{R}^n$, como já vimos, podemos definir, se $\mathbf{x} = (x_1, \ldots, x_n)$ e $\mathbf{y} = (y_1, \ldots, y_n)$,

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \ldots + x_n y_n.$$

Exemplo 2: Seja E o espaço das funções contínuas, definidas no intervalo $[a, b]$, a valores em $\mathbb{R}$. Se f e g são elementos de E, definimos

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx.$$

Exercício 7.2 *Mostre que os produtos escalares dos dois exemplos satisfazem, de fato, as quatro propriedades da definição de produto escalar.*

Exemplo 3: Nosso próximo exemplo está fortemente relacionado aos dois anteriores (embora a relação com o segundo não seja evidente). O espaço das sequências de números reais, como já vimos, é uma espécie de $\mathbb{R}^n$, com n *infinito*. Seria natural definirmos o produto escalar da sequência $(x_n)_{n \in \mathbb{N}}$ pela sequência $(y_n)_{n \in \mathbb{N}}$, fazendo

$$\langle (x_n), (y_n) \rangle = \sum_{n=1}^{\infty} x_n y_n.$$

O problema é que a soma infinita à direita nem sempre converge (exercício: ache um exemplo). Podemos apresentar duas saídas:

a) Façamos $\mathbf{V}$ ser o espaço das sequências $(x_n)_{n \in \mathbb{N}}$ tais que $x_n \neq 0$ apenas para um número finito de enes (esse número varia de sequência para sequência);

b)Tomemos como $\mathbf{V}$ o espaço $l^2(\mathbb{R})$ das sequências $(x_n)_{n \in \mathbb{N}}$ tais que

$$\sum_{n=1}^{\infty} x_n^2 < \infty.$$

Exercício 7.3 *Mostre que o exemplo 3a é, de fato, um espaço com produto interno.*

Exercício 7.4 *Mostre que o exemplo 3b é, de fato, um espaço com produto interno. A grande dificuldade é provar que $l^2(\mathbb{R})$ é um espaço vetorial, já que não é evidente que*

$$\sum_{n=1}^{\infty} x_n^2 < \infty, \sum_{n=1}^{\infty} y_n^2 < \infty \Longrightarrow \sum_{n=1}^{\infty} (x_n + y_n)^2 < \infty.$$

Se, depois de dois dias tentando, não conseguir, a resposta virá nas próximas páginas...

Exceto em caso de menção explícita em contrário, todos nossos espaços vetoriais são reais. Mas é importante ressaltar que, no caso em que o corpo dos escalares é $\mathbb{C}$, também podemos falar em produto escalar. A principal observação preliminar é que, enquanto, para números reais, é verdade que $|x| = \sqrt{xx}$, o que vale nos complexos é $|z| = \sqrt{z\bar{z}}$ (lembre que, se $z = x + iy$, então $\bar{z} = x - iy$). Assim, é natural definir, em $\mathbb{C}^n$,

$$\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n) \rangle = z_1\bar{w}_1 + \cdots + z_n\bar{w}_n.$$

Exercício 7.5 *Note que, identificando, da maneira natural, $\mathbb{C}^n$ a $\mathbb{R}^{2n}$, o produto escalar que acabamos de definir é o produto escalar usual de $\mathbb{R}^{2n}$.*

Definição: Dado um espaço vetorial complexo **V**, um **produto escalar** (também dito **produto interno**) em **V** é uma aplicação

$$\begin{array}{rccc} \langle , \rangle : & \mathbf{V} \times \mathbf{V} & \longrightarrow & \mathbb{C} \ , \\ & (\mathbf{u}, \mathbf{v}) & \longmapsto & \langle \mathbf{u}, \mathbf{v} \rangle \end{array}$$

com as seguites propriedades:

$$\begin{array}{rl} (i) & \langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle} \ \forall \mathbf{u}, \mathbf{v} \in \mathbf{V} \\ (ii) & \langle t\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v} \rangle = t \langle \mathbf{u}_1, \mathbf{v} \rangle + \langle \mathbf{u}_2, \mathbf{v} \rangle \ \ \forall \mathbf{u}_1, \mathbf{u}_2, \mathbf{v} \in \mathbf{V}, \ \forall t \in \mathbb{C} \\ (iii) & \langle \mathbf{u}, \mathbf{u} \rangle \geq 0 \ \forall \mathbf{u} \in \mathbf{V} \\ (iv) & \langle \mathbf{u}, \mathbf{u} \rangle = 0 \Longrightarrow \mathbf{u} = \mathbf{0} \end{array}$$

Exercício 7.6 *Prove, a partir da definição, as seguintes propriedades:*

$$\begin{array}{rl} (i) & \langle \mathbf{u}_1 + \mathbf{u}_2, \mathbf{v} \rangle = \langle \mathbf{u}_1, \mathbf{v} \rangle + \langle \mathbf{u}_2, \mathbf{v} \rangle \ \ \forall \mathbf{u}_1, \mathbf{u}_2, \mathbf{v} \in \mathbf{V} \\ (ii) & \langle \mathbf{0}, \mathbf{v} \rangle = 0 \ \forall \mathbf{v} \in \mathbf{V} \\ (iii) & \langle t\mathbf{u}, \mathbf{v} \rangle = t \langle \mathbf{u}, \mathbf{v} \rangle \ \ \forall \mathbf{u} \in \mathbf{V}, \ \forall t \in \mathbb{C} \\ (iv) & \langle \mathbf{u}, t\mathbf{v} \rangle = \bar{t} \langle \mathbf{u}, \mathbf{v} \rangle \ \ \forall \mathbf{u} \in \mathbf{V}, \ \forall t \in \mathbb{C} \end{array}$$

Exercício 7.7 *Note que $\langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle + 2Re \langle \mathbf{u}, \mathbf{v} \rangle$.*

Já que nossos espaços com produto interno serão sempre supostos reais, é um bom exercício verificar, nas demonstrações que virão, quais valem também no caso complexo.

7.1 Distâncias e ângulos

Seja E um espaço vetorial real com produto interno. De acordo com nossa experiência no plano e no espaço geométricos, são naturais as definições seguintes.

Definição: Se $\mathbf{u} \in E$, a **norma** de $\mathbf{u}$ é definida por

$$|\mathbf{u}| = \sqrt{\langle \mathbf{u}.\mathbf{u} \rangle}.$$

Se $|\mathbf{u}| = 1$, $\mathbf{u}$ é dito **unitário**.

Exercício 7.8 *Seja* $\mathbf{v}$ *um vetor não nulo. Defina* $\mathbf{u}$ *por* $\mathbf{u} = \frac{1}{|\mathbf{v}|}\mathbf{v}$. *Mostre que* $|\mathbf{u}| = 1$.

Se $\mathbf{u}$ e $\mathbf{v}$ são dois pontos de E, sua **distância** é dada por $|\mathbf{u} - \mathbf{v}|$.[1]

Definição: Dados os vetores não nulos $\mathbf{u}$ e $\mathbf{v}$ em E, o **ângulo** θ entre $\mathbf{u}$ e $\mathbf{v}$ é (o menor ângulo positivo) definido por

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{|\mathbf{u}||\mathbf{v}|}.$$

Nossas definições, mesmo que diretamente inspiradas pela Geometria Euclidiana, têm satisfações a nos dar. Por exemplo: será que o número que acabamos de chamar de cosseno está, de fato, no intervalo $[-1, 1]$? Precisamos de algumas garantias de que as fantasias geométricas com que pretendemos vestir nossos espaços munidos de produto interno, embora fantasias, tenham alguma verossimilhança.

Curiosamente, se considerarmos que um ângulo reto entre $\mathbf{u}$ e $\mathbf{v}$ é equivalente a $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, o **Teorema de Pitágoras** é um resultado básico.

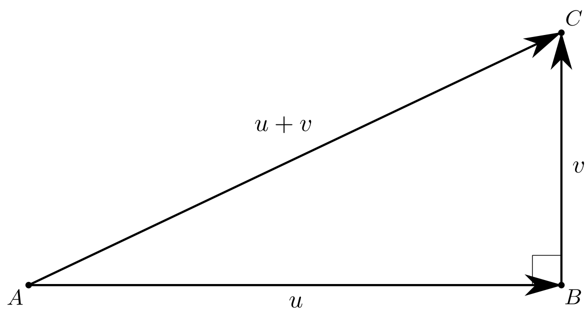


Figura 7.1:

[1]O leitor deve observar que, a exemplo do que fazemos em situações geométricas, costumamos dizer *a norma do* **vetor** *fulano de tal*, mas, quando se trata de distâncias, o usual é *a distância entre os* **pontos** *tal e tal*

Exercício 7.9 *Sejam A, B e C três pontos de E. Considere o triângulo ABC. Faça $\mathbf{u} = B - A$ e $\mathbf{v} = C - B$. Note que $C - A = \mathbf{u} + \mathbf{v}$ e que se o ângulo em B é reto, então $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.*

Teorema de Pitágoras: Se $\mathbf{u}$ e $\mathbf{v}$ são vetores de E tais que $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, então

$$|\mathbf{u} + \mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2.$$

Demonstração: $|\mathbf{u} + \mathbf{v}|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle + 2 \langle \mathbf{u}, \mathbf{v} \rangle = |\mathbf{u}|^2 + \mathbf{v}|^2 + 0.$ ■

Nossa definição de cosseno inspira a introdução da **projeção** de um vetor na direção de outro.

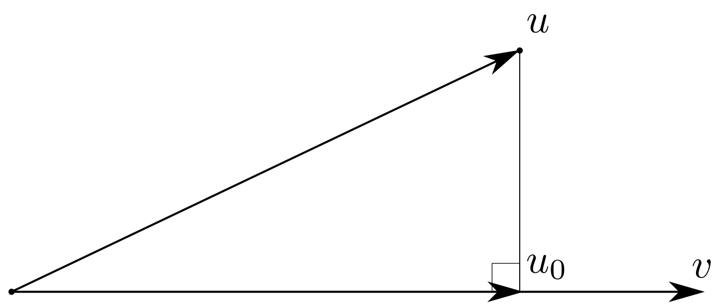


Figura 7.2:

Se $\mathbf{u}$ e $\mathbf{v}$ são dois vetores em E, com $|\mathbf{v}| = 1$, podemos projetar $\mathbf{u}$ na direção de $\mathbf{v}$, obtendo

$$\mathbf{u}_o = \langle \mathbf{u}, \mathbf{v} \rangle \mathbf{v}.$$

Se $\mathbf{v}$ não é unitário, mas é não nulo, ainda podemos projetar $\mathbf{u}$ na direção de $\mathbf{v}$, fazendo

$$\mathbf{u}_o = \left\langle \mathbf{u}, \frac{1}{|\mathbf{v}|}\mathbf{v} \right\rangle \frac{1}{|\mathbf{v}|}\mathbf{v} = \frac{1}{|\mathbf{v}|^2} \langle \mathbf{u}, \mathbf{v} \rangle \mathbf{v} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}\mathbf{v}.$$

Se nossas fantasias geométricas fazem sentido, o triângulo de vértices $\mathbf{0}$, $\mathbf{u}_o$ e $\mathbf{u}$ tem um ângulo reto em $\mathbf{u}_o$.

Exercício 7.10 *Mostre que $\langle \mathbf{u} - \mathbf{u}_o, \mathbf{u}_o \rangle = 0$.*

Uma consequência óbvia do Teorema de Pitágoras é o seguinte fato: a hipotenuza é maior que os catetos. No caso das projeções, isso se traduz por $|\mathbf{u}| \geq |\mathbf{u}_o|$ e recebe um nome pomposo.

Desigualdade de Cauchy-Schwarz-Buniacóvski:

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq |\mathbf{u}||\mathbf{v}| \ \forall \mathbf{u}, \mathbf{v} \in E.$$

Demonstração: Se um dos dois vetores é nulo, vale claramente a igualdade. Podemos, então, supor que são ambos não nulos e definir

$$\mathbf{u}_o = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}.$$

Temos, então,

$$\langle \mathbf{u} - \mathbf{u}_o, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{u}_o, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{v}, \mathbf{v} \rangle = 0.$$

Daí decorre $\langle \mathbf{u} - \mathbf{u}_o, \mathbf{u}_o \rangle = 0$, o que, por Pitágoras, nos dá $|\mathbf{u}|^2 = |\mathbf{u}_o|^2 + |\mathbf{u} - \mathbf{u}_o|^2 \geq |\mathbf{u}_o|^2$, ou seja, $|\mathbf{u}| \geq |\mathbf{u}_o|$. Isso significa

$$|\mathbf{u}| \geq \left| \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v} \right| = \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|}{|\mathbf{v}|^2} |\mathbf{v}| = \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|}{|\mathbf{v}|},$$

que é nossa desigualdade. ■

Observação: A desigualdade de Cauchy-Schwarz-Buniacóvski poderia ser, mais simplesmente, chamada de **desigualdade do cosseno**. De fato, CSB expressa o fato de que aquilo que chamamos de **cosseno** do ângulo entre $\mathbf{u}$ e $\mathbf{v}$,

$$\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{|\mathbf{u}||\mathbf{v}|},$$

é um número entre -1 e 1.

Exercício 7.11 *Sejam $\mathbf{u}$ e $\mathbf{v}$ vetores fixados. Use o fato de que, para todo real t, vale $|\mathbf{u} - t\mathbf{v}|^2 \geq 0$, expresse $|\mathbf{u} - t\mathbf{v}|^2$ como um trinômio do segundo grau em t e dê uma outra demonstração, de inspiração puramente algébrica, para a desigualdade de Cauchy-Schwarz-Buniacóvski.*

Exercício 7.12 *Note que a igualdade, em **CSB**, só pode ocorrer se $\mathbf{u} = \mathbf{0}$, $\mathbf{v} = \mathbf{0}$ ou $\mathbf{u} - \mathbf{u}_o = \mathbf{0}$, ou seja, se $\mathbf{u}$ e $\mathbf{v}$ forem linearmente dependentes.*

A desigualdade **CSB** tem como consequência a **desigualdade triangular**: $|\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|$.

Corolário (Desigualdade Triangular): Sejam $\mathbf{u}$ e $\mathbf{v}$ dois vetores. Então

$$|\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|.$$

Supondo $\mathbf{v}$ não nulo, a igualdade $|\mathbf{u} + \mathbf{v}| = |\mathbf{u}| + |\mathbf{v}|$ ocorre se, e somente se, existe um real estritamente positivo, t, tal que $\mathbf{u} = t\mathbf{v}$.

Demonstração: A desigualdade decorre imediatamente de CSB:

$$|\mathbf{u} + \mathbf{v}|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = |\mathbf{u}|^2 + |\mathbf{v}|^2 + 2 \langle \mathbf{u}, \mathbf{v} \rangle \leq |\mathbf{u}|^2 + |\mathbf{v}|^2 + 2|\mathbf{u}||\mathbf{v}| = (|\mathbf{u}| + |\mathbf{v}|)^2.$$

Por outro lado, temos, se vale a igualdade,

$$|\mathbf{u}|^2 + |\mathbf{v}|^2 + 2 \langle \mathbf{u}, \mathbf{v} \rangle = |\mathbf{u} + \mathbf{v}|^2 = (|\mathbf{u}| + |\mathbf{v}|)^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2 + 2|\mathbf{u}||\mathbf{v}|.$$

Daí decorre, portanto, a igualdade em CSB, o que mostra que $\mathbf{u}$ e $\mathbf{v}$ são linearmente dependentes. Como $\mathbf{v}$ é não nulo, devemos ter $\mathbf{u} = t\mathbf{v}$ para algum t. Como $\langle \mathbf{u}, \mathbf{v} \rangle = |\mathbf{u}||\mathbf{v}| > 0$, t deve ser real positivo. ■

Observação: A desigualdade triangular exprime o fato de que *o menor caminho entre dois pontos é o segmento de reta que os une*.

Exercício 7.13 *Sejam $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$ em $\mathbb{R}^n$. Mostre que*

$$\sum_{j=1}^{n} |x_j y_j| \leq \sqrt{\left(\sum_{j=1}^{n} x_j^2\right)\left(\sum_{j=1}^{n} y_j^2\right)}$$

Exercício 7.14 *Sejam $(x_n)_{n\in\mathbb{N}}$ e $(y_n)_{n\in\mathbb{N}}$ tais que*

$$\sum_{n=1}^{\infty} x_n^2 < \infty, \; \sum_{n=1}^{\infty} y_n^2 < \infty.$$

Passando ao limite a desigualdade do exercício anterior, mostre que

$$\sum_{j=1}^{\infty} |x_j y_j| \leq \sqrt{\left(\sum_{j=1}^{\infty} x_j^2\right)\left(\sum_{j=1}^{\infty} y_j^2\right)}.$$

Conclua que $l^2(\mathbb{R})$ é, de fato, um espaço vetorial.

Exercício 7.15 *Sejam f e g funções contínuas definidas no intervalo $[a, b]$ e a valores em $\mathbb{R}$. Mostre que*

$$\int_a^b |f(x)g(x)|dx \leq \sqrt{\int_a^b f(x)^2 dx \int_a^b g(x)^2 dx}.$$

Capítulo 8

Bases ortogonais

8.1 Bases ortogonais com vetores flechinhas

Como já observamos, se representamos um vetor $\vec{v}$ na base canônica, $\vec{v} = v_1\vec{e_1} + v_2\vec{e_2} + v_3\vec{e_3}$, então

$$v_1 =< \vec{v}, \vec{e_1} >,\ v_2 =< \vec{v}, \vec{e_2} >,\ v_3 =< \vec{v}, \vec{e_3} > .$$

O mesmo truque não funciona se estivermos trabalhando com uma base cujos vetores não sejam ortogonais. Mas funciona razoavelmente bem, se forem.

Proposição: Suponha que os vetores $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$ são não nulos e dois a dois ortogonais. Então:

(i) $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$ formam uma base para V;
(ii) se $\vec{v} = v_1\vec{\varepsilon}_1 + v_2\vec{\varepsilon}_2 + v_3\vec{\varepsilon}_3$, então, para $i = 1, 2, 3$, vale a fórmula

$$v_i = \frac{< \vec{v}, \vec{\varepsilon}_i >}{< \vec{\varepsilon}_i, \vec{\varepsilon}_i >}.$$

Demonstração:

(i) Como estamos com 3 vetores, basta provar que são linearmente independente. Se não fossem, poderíamos escrever um deles, que chamaremos de $\vec{\varepsilon}_i$ como combinação linear dos outros dois: $\vec{\varepsilon}_i = s\vec{\varepsilon}_j + t\vec{\varepsilon}_k$. Como $\vec{\varepsilon}_i$ é não nulo, sabemos que $< \vec{\varepsilon}_i, \vec{\varepsilon}_i > \neq 0$. Mas, então,

$$0 \neq < \vec{\varepsilon}_i, \vec{\varepsilon}_i > = < s\vec{\varepsilon}_j + t\vec{\varepsilon}_k, \vec{\varepsilon}_i > = s < \vec{\varepsilon}_j, \vec{\varepsilon}_i > + t < \vec{\varepsilon}_k, \vec{\varepsilon}_i > = 0.$$

(ii) Se $\vec{v} = v_1\vec{\varepsilon}_1 + v_2\vec{\varepsilon}_2 + v_3\vec{\varepsilon}_3$, então, para $i = 1, 2, 3$, temos

$$\begin{aligned} < \vec{v}, \vec{\varepsilon}_i > &= < v_1\vec{\varepsilon}_1 + v_2\vec{\varepsilon}_2 + v_3\vec{\varepsilon}_3, \vec{\varepsilon}_i > = \\ &= < v_1\vec{\varepsilon}_1, \vec{\varepsilon}_i > + < v_2\vec{\varepsilon}_2, \vec{\varepsilon}_i > + < v_3\vec{\varepsilon}_3, \vec{\varepsilon}_i > = v_i < \vec{\varepsilon}_i, \vec{\varepsilon}_i > . \end{aligned}$$

Como $\vec{\varepsilon}_i$ é não nulo, podemos dividir dos dois lados por $< \vec{\varepsilon}_i, \vec{\varepsilon}_i >$ e o resultado segue. ■

Definição: Uma base composta por vetores dois a dois ortogonais é dita uma **base ortogonal**. Um vetor de norma igual a 1 é dito **unitário**. Uma base ortogonal composta por vetores unitários é dita uma **base ortonormal**.

8.2 Construindo bases ortonormais

Consideremos a seguinte questão: se α é o plano que passa pela origem e é dado pelos vetores $\vec{u}$ e $\vec{v}$, como construir uma base ortonormal para α? Mais claramente, sendo

$$\alpha = \left\{ s\vec{u} + t\vec{v},\ (s,t) \in \mathbb{R}^2 \right\},$$

queremos $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$ em α, unitários e ortogonais.

A solução é simples: começamos fazendo

$$\vec{\varepsilon}_1 = \frac{1}{|\vec{u}|}\vec{u};$$

em seguida, fazemos

$$\vec{v}_1 = \vec{v} - < \vec{v}, \vec{\varepsilon}_1 > \vec{\varepsilon}_1;$$

finalmente, tomamos

$$\vec{\varepsilon}_2 = \frac{1}{|\vec{v}_1|}\vec{v}_1.$$

Exercício 8.1 *Mostre que de fato, $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$ formam base ortonormal para α. Faça uma figura, explicitando $\vec{u}$, $\vec{v}$, $\vec{\varepsilon}_1$, $\vec{v}_1$, $\vec{v} - \vec{v}_1$ e $\vec{\varepsilon}_2$.*

Construída uma base ortonormal para α, é fácil, agora, projetar ortogonalmente um ponto P, dado pelo vetor $\vec{w}$, em α: se Q é a projeção, dada pelo vetor $\vec{w}_0$, e $\vec{w}_1 = \overrightarrow{QP}$, então $\vec{w} = \vec{w}_0 + \vec{w}_1$ e $\vec{w}_1$ e $\vec{\varepsilon}$ são ortogonais, qualquer que seja $\vec{\varepsilon}$ em α.

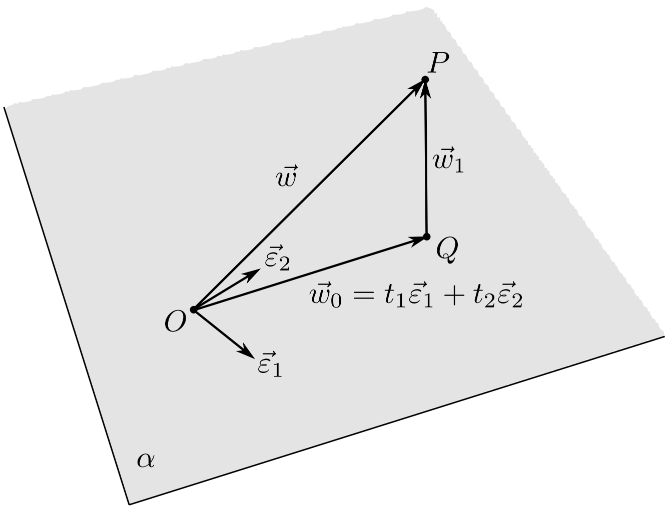


Figura 8.1:

Escrevendo $\vec{w}_0 = t_1\vec{\varepsilon}_1 + t_2\vec{\varepsilon}_2$, com t_1 e t_2 a determinar, temos:

$$t_1 =< \vec{w}_0, \vec{\varepsilon}_1 >, \ t_2 =< \vec{w}_0, \vec{\varepsilon}_2 >$$

Mas

$$< \vec{w}, \vec{\varepsilon} >=< \vec{w}_0 + \vec{w}_1, \vec{\varepsilon} >=< \vec{w}_0, \vec{\varepsilon} > + < \vec{w}_1, \vec{\varepsilon} >=< \vec{w}_0, \vec{\varepsilon} > \ \forall \vec{\varepsilon} \in \alpha.$$

Daí segue $t_1 =< \vec{w}, \vec{\varepsilon}_1 >, t_2 =< \vec{w}, \vec{\varepsilon}_2 >$.

Exercício 8.2 *Note que, se $\vec{w} \notin \alpha$, obtemos, fazendo*

$$\vec{\varepsilon}_3 = \frac{1}{|\vec{w}_1|}\vec{w}_1,$$

uma base ortonormal, $\{\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3\}$, para o espaço V, com dois vetores em α.

Observação: O processo que acabamos de ver (chamado **processo de ortogonalização de Gram-Schmidt**), constrói uma base ortonormal para o subespaço α a partir de uma base dada. Se pensarmos os vetores originalmente fornecidos, $\vec{u}, \vec{v}$ e $\vec{w}$, como uma base para V, o processo nos fornece uma base ortonormal para V,sem fazer qualquer referência à base canônica. Pode parecer pouco interessante no momento... Num contexto mais geral, porém, esta é uma ideia simples mas fundamental.

8.3 Projeções e complemento ortogonais

Seja W um subespaço de V e $\vec{v}$ um vetor de V. A **projeção ortogonal** de $\vec{v}$ em W é o elemento $\vec{v}_0$ de W tal que

$$|\vec{v} - \vec{v}_0| \leq |\vec{v} - \vec{w}| \ \forall \ \vec{w} \in \ W.$$

Proposição: A projeção ortogonal $\vec{v}_0$ de $\vec{v}$ em W existe, é única e satisfaz

$$< \vec{v} - \vec{v}_0, \vec{w} >= 0 \ \forall \ \vec{w} \in \ W.$$

Demonstração: Se $W = \left\{\vec{0}\right\}$, é só tomar $\vec{v} = \vec{0}$; se $W = V$, tomamos $\vec{v}_0 = \vec{v}$. Se W é gerado por $\vec{\varepsilon}_1$, com $|\vec{\varepsilon}_1| = 1$, tome $\vec{v}_0 =< \vec{v}, \vec{\varepsilon}_1 > \vec{\varepsilon}_1$; se W é gerado por $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$, com $|\vec{\varepsilon}_1| = |\vec{\varepsilon}_2| = 1$ e $< \vec{\varepsilon}_1, \vec{\varepsilon}_2 >= 0$, tome $\vec{v}_0 =< \vec{v}, \vec{\varepsilon}_1 > \vec{\varepsilon}_1 + < \vec{v}, \vec{\varepsilon}_2 > \vec{\varepsilon}_2$. Confira. ■

Exercício 8.3 *Seja X um subconjunto de V e seja*

$$X^{\perp} = \{\vec{v} \in \ V \mid \ < \vec{v}, \vec{u} >= 0 \ \forall \ \vec{u} \in \ X\}.$$

Mostre que $X^{\perp}$ é um subespaço vetorial.

Suponhamos que W seja um subespaço vetorial de V. Seja

$$W^{\perp} = \{\vec{v} \in \ V \mid \ < \vec{v}, \vec{w} >= 0 \ \forall \ \vec{w} \in \ W\}.$$

$W^{\perp}$ é chamado de **complemento ortogonal** de W.

Exercício 8.4 *Mostre que* $W \cap W^{\perp} = \left\{\vec{0}\right\}$.

Proposição: Todo elemento de V se escreve, de maneira única, como soma de um elemento de W e um de $W^{\perp}$. Consequentemente, $dim\ W + dim\ W^{\perp} = dim V$.

Demonstração: Seja $\vec{v}$ um elemento de V. Sejam $\vec{v}_0$ a projeção de $\vec{v}$ em W e $\vec{v}_1 = \vec{v} - \vec{v}_0$. Então é fácil ver que $\vec{v}_1 \in W^{\perp}$ e $\vec{v} = \vec{v}_0 + \vec{v}_1$. Se pudéssemos escrever, também, $\vec{v} = \vec{w}_0 + \vec{w}_1$, com $\vec{w}_0 \in W$ e $\vec{w}_1 \in W^{\perp}$, teríamos $\vec{0} = (\vec{v}_0 - \vec{w}_0) + (\vec{v}_1 - \vec{w}_1)$, o que daria $(\vec{v}_0 - \vec{w}_0) = -(\vec{v}_1 - \vec{w}_1)$, com $(\vec{v}_0 - \vec{w}_0) \in W$ e $-(\vec{v}_1 - \vec{w}_1) \in W^{\perp}$, o que implica em $\vec{v}_0 = \vec{w}_0$ e $\vec{v}_1 = \vec{w}_1$. Provada a primeira parte, basta tomar uma base para W e uma para $W^{\perp}$. A união das duas é uma base de V, o que prova a segunda parte. ■

Exercício 8.5 *Mostre que, para qualquer subespaço W de V, tem-se* $(W^{\perp})^{\perp} = W$.

Exercício 8.6 *Seja* $P = (x_1, x_2, x_3)$ *um ponto em* $\mathbb{R}^3$. *Determine a projeção ortogonal,* P_o, *de P sobre o subespaço E gerado por* $\vec{v} = (1,1,1)$. *Mostre que* P_o *é o ponto E mais próximo de P, isto é:* $|P - P_o| \leq |P - Q|\ \forall Q \in E$.

8.4 Bases ortonormais e projeções

Tomemos, como ponto de partida das considerações a seguir, a seguinte questão: massas muito grandes de dados podem ser difíceis de armazenar, transmitir e manipular. Como exemplo extremo, consideremos uma fotografia ideal (em preto e branco), vista como uma função do retângulo $R = [0,a] \times [0,b]$ em $\mathbb{R}$. Uma tal fotografia *vive*, em princípio, em um espaço de dimensão infinita. Uma versão discreta implicaria, por exemplo, em dividir R em pequenos retângulos, de lados a/n e b/m (uma resolução *alta* significando tomar m e n *grandes*). Podemos, pois, supor que nossas fotografias (discretas) são dadas por arquivos que *vivem* em um espaço de dimensão $N = m \times n$ grande. Em determinadas situações, porém, seria preferível ter, dessas fotografias, versões em baixa resolução (de dimensão M, digamos, com $M < N$).

Uma abordagem radicalmente simples da questão acima consistiria em fazer $N = 3$ e $M = 2$. Assim, nossos arquivos originais estariam em $\mathbb{R}^3$, mas suas versões em baixa resolução estariam em um espaço de dimensão 2 (note que estamos evitando dizer que estariam em $\mathbb{R}^2$, já veremos por que). Uma solução simples para a compactação seria jogar fora uma das coordenadas (a segunda, por exemplo). Assim, do arquivo (x_1, x_2, x_3), guardaríamos apenas (x_1, x_3)). Esta não é, provavelmente, a melhor ideia, a menos que a informação contida em x_2 possa, de alguma forma, ser (total ou parcialmente) recuperada a partir de x_1 e x_3.

Tentemos algo um pouco menos chutado. Examinemos a questão de um ponto de vista geométrico. Vamos admitir duas coisas:

1. é razoável *medir* a diferença entre o arquivo $\mathbf{x} = (x_1, x_2, x_3)$ e o arquivo $\mathbf{y} = (y_1, y_2, y_3)$ por $\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$ (que, como sabemos, é a **distância** oriunda do produto escalar canônico de $\mathbb{R}^3$);

2. nossos arquivos (ou, pelo menos, os que nos interessam) não estão totalmente dispersos pelo espaço, mas apresentam um certo padrão, que nos permite dizer que se situam não muito longe de um certo plano, α (que, para simplificar, suporemos passar pela origem).

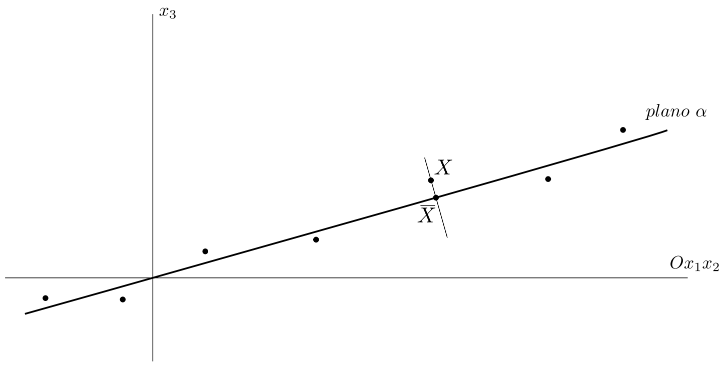


Figura 8.2:

Nas condições acima, podemos adotar o seguinte procedimento:

1. aproximamos cada arquivo (ponto) $\mathbf{x} = (x_1, x_2, x_3)$ por um arquivo (ponto) $\bar{\mathbf{x}}$ em α, de forma que $\bar{\mathbf{x}}$ seja a projeção ortogonal de $\mathbf{x}$ sobre α (o que, pelo que sabemos, garante que $\bar{\mathbf{x}}$ seja o elemento de α mais *próximo* de $\mathbf{x}$);

2. escolhendo uma base $\{\boldsymbol{\varepsilon_1}, \varepsilon_2\}$ de α, representamos $\bar{\mathbf{x}}$ por suas coordenadas $(\bar{x}_1, \bar{x}_2)$ na base $\{\boldsymbol{\varepsilon_1}, \varepsilon_2\}$

3. guardamos (ou utilizamos) o arquivo $(\bar{x}_1, \bar{x}_2)$, que nos permite, a qualquer momento (estando de posse de $\{\boldsymbol{\varepsilon_1}, \varepsilon_2\}$), recuperar $\bar{\mathbf{x}} = \bar{x}_1\boldsymbol{\varepsilon_1} + \bar{x}_2\varepsilon_2$.

A partir daí, temos duas dificuldades a enfrentar:

1. como encontrar a projeção $\bar{\mathbf{x}}$ de $\mathbf{x}$ sobre α?

2. como representar $\bar{\mathbf{x}}$ como combinação linear de $\boldsymbol{\varepsilon_1}$ e $\boldsymbol{\varepsilon}_2$?

Exercício 8.7 *Como você enfrentaria essas questões?*

Podemos resumir nosso problema da seguinte forma: dados o vetor $\mathbf{x}$ em $\mathbb{R}^3$ e a base $\{\boldsymbol{\varepsilon_1}, \varepsilon_2\}$ de α, queremos encontrar o par ordenado $(\bar{x}_1, \bar{x}_2) \in \mathbb{R}^2$ tal que

$$\langle \mathbf{x} - (\bar{x}_1\boldsymbol{\varepsilon_1} + \bar{x}_2\varepsilon_2), v\rangle = 0 \; \forall \; v \in \alpha.$$

Na realidade, como $\{\boldsymbol{\varepsilon_1}, \varepsilon_2\}$ é base de α, é (necessário e) suficiente que

$$\langle \mathbf{x} - (\bar{x}_1\boldsymbol{\varepsilon_1} + \bar{x}_2\varepsilon_2), \boldsymbol{\varepsilon_1}\rangle = 0$$

e

$$\langle \mathbf{x} - (\bar{x}_1\boldsymbol{\varepsilon_1} + \bar{x}_2\varepsilon_2), \boldsymbol{\varepsilon}_2\rangle = 0.$$

Exercício 8.8 *Observe que isso significa que devemos resolver um sistema de duas equações e duas incógnitas. Note que, voltando ao caso geral em que estávamos em um espaço de dimensão N e queríamos reduzir o tamanho de nosso arquivo de N para M, poderíamos imitar o caso $N = 3,\ M = 2$; chegaríamos, então, a um sistema de M equações e M incógnitas.*

Suponhamos agora que os vetores $\boldsymbol{\varepsilon_1}$ e $\boldsymbol{\varepsilon_2}$ sejam ortogonais. O sistema já vem resolvido! Temos, imediatamente,

$$\bar{x}_1 = \frac{\langle \mathbf{x}, \boldsymbol{\varepsilon_1} \rangle}{\langle \boldsymbol{\varepsilon_1}, \boldsymbol{\varepsilon_1} \rangle},$$

$$\bar{x}_2 = \frac{\langle \mathbf{x}, \boldsymbol{\varepsilon_2} \rangle}{\langle \boldsymbol{\varepsilon_2}, \boldsymbol{\varepsilon_2} \rangle}.$$

Exercício 8.9 *Suponha que $\mathbf{V}$ é um espaço com produto interno e que $\mathbf{V}_o$ é um subespaço de $\mathbf{V}$, com dim $V = M$. Suponha, ainda, que $\{\boldsymbol{\varepsilon_1}, \ldots, \boldsymbol{\varepsilon_M}\}$ é base de $\mathbf{V}_o$ e que*

$$\langle \boldsymbol{\varepsilon_i}, \boldsymbol{\varepsilon_j} \rangle = 0\ \forall\ i, j = 1, \ldots, M,\ i \neq j.$$

Mostre que, se $\mathbf{v} = x_1\boldsymbol{\varepsilon_1} + \cdots + x_M\boldsymbol{\varepsilon_M}$, então

$$x_i = \frac{\langle \mathbf{v}, \boldsymbol{\varepsilon_i} \rangle}{\langle \boldsymbol{\varepsilon_i}, \boldsymbol{\varepsilon_i} \rangle}\ \forall i = 1, \ldots, M.$$

Seja $\mathbf{v}$ um vetor em V. Suponha que

$$\mathbf{v} = \mathbf{v}_o + \mathbf{u},$$

com $\mathbf{v}_o \in \mathbf{V}_o$ e

$$\langle \mathbf{u}, \mathbf{w} \rangle = 0\ \forall \mathbf{w} \in \mathbf{V}.$$

Mostre que, escrevendo $\mathbf{v}_o$ na base $\{\boldsymbol{\varepsilon_1}, \ldots, \boldsymbol{\varepsilon_M}\}$, temos $\mathbf{v}_o = x_1\boldsymbol{\varepsilon_1} + \cdots + x_M\boldsymbol{\varepsilon_M}$, com

$$x_i = \frac{\langle \mathbf{v}, \boldsymbol{\varepsilon_i} \rangle}{\langle \boldsymbol{\varepsilon_i}, \boldsymbol{\varepsilon_i} \rangle}\ \forall i = 1, \ldots, M.$$

A proposição a seguir é a versão geral das ideias geométricas que acabamos de discutir.

Proposição 1: Sejam $\mathbf{V}$ um espaço vetorial com produto interno, $\mathbf{V}_o$ um subespaço de $\mathbf{V}$, de dimensão M, e $\boldsymbol{\varepsilon_1}, \ldots, \boldsymbol{\varepsilon_M}$ vetores em $\mathbf{V}_o$ tais que

$$\langle \boldsymbol{\varepsilon_i}, \boldsymbol{\varepsilon_j} \rangle = 0\ \forall\ i, j = 1, \ldots, M,\ i \neq j.$$

Então:

1. $\boldsymbol{\varepsilon_1}, \ldots, \boldsymbol{\varepsilon_M}$ são linearmente independentes (e, como dim $\mathbf{V}_o = M$, formam base de $\mathbf{V}_o$).

2. Para todo $\mathbf{v}$ em $\mathbf{V}$, o vetor $\bar{\mathbf{v}} = x_1\boldsymbol{\varepsilon_1} + \cdots + x_M\boldsymbol{\varepsilon_M}$, com

$$x_i = \frac{\langle \mathbf{v}, \boldsymbol{\varepsilon_i} \rangle}{\langle \boldsymbol{\varepsilon_i}, \boldsymbol{\varepsilon_i} \rangle}\ \forall i = 1, \ldots, M,$$

é tal que

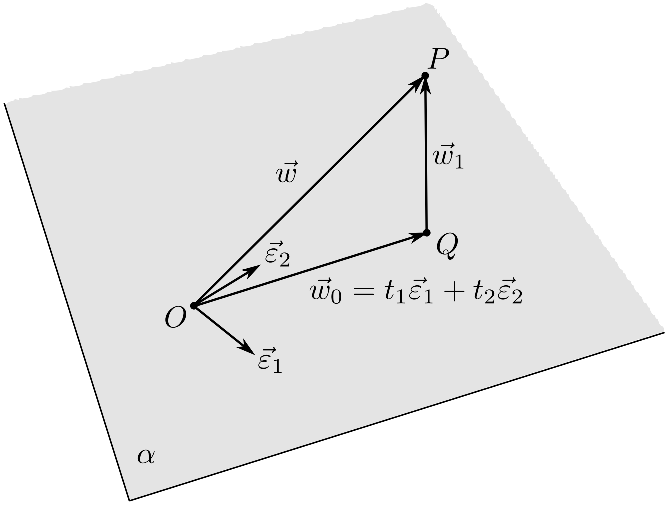


Figura 8.3:

(a) $\langle \mathbf{v} - \bar{\mathbf{v}}, \mathbf{w} \rangle = 0 \; \forall \mathbf{w} \in \mathbf{V}_o$;

(b) $|\mathbf{v} - \bar{\mathbf{v}}| \leq |\mathbf{v} - \mathbf{w}| \; \forall \mathbf{w} \in \mathbf{V}_o$;

(c) se $\mathbf{v}_o$ pertence a $\mathbf{V}_o$ e $|\mathbf{v} - \mathbf{v}_o| \leq |\mathbf{v} - \mathbf{w}| \; \forall \mathbf{w} \in \mathbf{V}_o$, então $\mathbf{v}_o = \bar{\mathbf{v}}$.

Demonstração:

1. Se $0 = x_1\boldsymbol{\varepsilon_1} + \cdots + x_M\boldsymbol{\varepsilon}_M$, então, multiplicando escalarmente por $\boldsymbol{\varepsilon_i}$ dos dois lados, temos:

$$0 = \langle x_1\boldsymbol{\varepsilon_1} + \cdots + x_M\boldsymbol{\varepsilon}_M, \boldsymbol{\varepsilon_i} \rangle = x_1 \langle \boldsymbol{\varepsilon_1}, \boldsymbol{\varepsilon_i} \rangle + \cdots + x_M \langle \boldsymbol{\varepsilon}_M, \boldsymbol{\varepsilon_i} \rangle = x_i \langle \boldsymbol{\varepsilon_i}, \boldsymbol{\varepsilon_i} \rangle .$$

Como $\boldsymbol{\varepsilon_i} \neq \mathbf{0}$, temos $\langle \boldsymbol{\varepsilon_i}, \boldsymbol{\varepsilon_i} \rangle \neq 0$, o que nos dá $x_i = 0$ e demonstra a independência linear de $\boldsymbol{\varepsilon_1}, \cdots, \boldsymbol{\varepsilon}_M$.

2. Sejam $\mathbf{v}$ um elemento de $\mathbf{V}$ e $\bar{\mathbf{v}}$ definido como acima. Seja $\mathbf{u} = \mathbf{v} - \bar{\mathbf{v}}$. Então:

(a) Para cada $i = 1, \ldots, M$, temos
$\langle \mathbf{u}, \boldsymbol{\varepsilon_i} \rangle = \langle \mathbf{v}, \boldsymbol{\varepsilon_i} \rangle - \langle x_1\boldsymbol{\varepsilon_1} + \cdots + x_M\boldsymbol{\varepsilon}_M, \boldsymbol{\varepsilon_i} \rangle = \langle \mathbf{v}, \boldsymbol{\varepsilon_i} \rangle - x_i \langle \boldsymbol{\varepsilon_i}, \boldsymbol{\varepsilon_i} \rangle =$
$= \langle \mathbf{v}, \boldsymbol{\varepsilon_i} \rangle - \frac{\langle \mathbf{v}, \boldsymbol{\varepsilon}_i \rangle}{\langle \boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_i \rangle} \langle \boldsymbol{\varepsilon_i}, \boldsymbol{\varepsilon_i} \rangle = 0.$
Como $\boldsymbol{\varepsilon_1}, \cdots, \boldsymbol{\varepsilon}_M$ formam base de $\mathbf{V}_o$, isso mostra que $\langle \mathbf{v} - \bar{\mathbf{v}}, \mathbf{w} \rangle = 0 \; \forall \mathbf{w} \in \mathbf{V}_o$

(b) Se $\mathbf{w}$ está em $\mathbf{V}_o$, então $\mathbf{w} - \bar{\mathbf{v}}$ está em $\mathbf{V}_o$ e é, portanto, ortogonal a $\mathbf{u}$. Segue do Teorema de Pitágoras que $|\mathbf{v} - \mathbf{w}|^2 = |\mathbf{u}|^2 + |\mathbf{w} - \bar{\mathbf{v}}|^2$, o que prova que $|\mathbf{v} - \mathbf{w}| \geq |\mathbf{u}|$

(c) A demonstração do item anterior prova, também, que se $\mathbf{v}_0$ está em $\mathbf{V}_o$ e $\mathbf{v}_o \neq \bar{\mathbf{v}}$, então $|\mathbf{v} - \mathbf{v}_o| > |\mathbf{u}|$.

■

Definição: Sejam $\mathbf{V}$ um espaço vetorial com produto interno, $\mathbf{v}$ um elemento de $\mathbf{V}$ e $\mathbf{V}_o$ um subespaço vetorial de $\mathbf{V}$. O vetor $\bar{\mathbf{v}}$ de $\mathbf{V}_o$ é chamado de **projeção ortogonal** de $\mathbf{v}$ sobre $\mathbf{V}_o$ se

$$\langle \mathbf{v} - \bar{\mathbf{v}}, \mathbf{w} \rangle = 0 \; \forall \; \mathbf{w} \in \mathbf{V}_o.$$

Proposição 2: A projeção ortogonal é única.

Demonstração: A ideia geométrica é que, se $\mathbf{v}_1$ e $\mathbf{v}_2$ são projeções ortogonais de $\mathbf{v}$ sobre $\mathbf{V}_o$, então o triângulo $\mathbf{v}_1\mathbf{v}\mathbf{v}_2$ tem dois ângulos retos: já que $\mathbf{w} = \mathbf{v}_1 - \mathbf{v}_2$ está em $\mathbf{V}_o$, teremos

$$\langle \mathbf{v} - \mathbf{v}_1, \mathbf{w} \rangle = 0 = \langle \mathbf{v} - \mathbf{v}_2, \mathbf{w} \rangle,$$

Daí vem

$$0 = \langle (\mathbf{v} - \mathbf{v}_2) - (\mathbf{v} - \mathbf{v}_1), \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{w} \rangle,$$

o que dá $|\mathbf{v}_1 - \mathbf{v}_2|^2 = 0$ e prova que $\mathbf{v}_1 = \mathbf{v}_2$. ■

Proposição 3: Sejam $\mathbf{V}$ um espaço vetorial com produto interno, $\mathbf{v}$ um elemento de $\mathbf{V}$ e $\mathbf{V}_o$ um subespaço vetorial de $\mathbf{V}$. O vetor $\bar{\mathbf{v}}$ de $\mathbf{V}_o$ é a projeção ortogonal de $\mathbf{v}$ sobre $\mathbf{V}_o$ se, e somente se,

$$|\mathbf{v} - \bar{\mathbf{v}}| \leq |\mathbf{v} - \mathbf{w}| \; \forall \; \mathbf{w} \in \mathbf{V}_o.$$

Demonstração: Se $\bar{\mathbf{v}}$ é a projeção ortogonal de $\mathbf{v}$ sobre $\mathbf{V}_o$ e $\mathbf{w} \in \mathbf{V}_o$, temos $\langle \mathbf{v} - \bar{\mathbf{v}}, \bar{\mathbf{v}} - \mathbf{w} \rangle = 0$, o que, pelo Teorema de Pitágoras, nos dá $|\mathbf{v} - \mathbf{w}|^2 = |\mathbf{v} - \bar{\mathbf{v}}|^2 + |\bar{\mathbf{v}} - \mathbf{w}|^2$ e mostra que $|\mathbf{v} - \bar{\mathbf{v}}| \leq |\mathbf{v} - \mathbf{w}|$.

Reciprocamente, seja $\bar{\mathbf{v}}$ em $\mathbf{V}_o$ tal que $|\mathbf{v} - \bar{\mathbf{v}}| \leq |\mathbf{v} - \mathbf{w}| \; \forall \; \mathbf{w} \in \mathbf{V}_o$. Imaginemos que, para um certo $\mathbf{u}$ em $\mathbf{V}_o$, tivéssemos $\langle \mathbf{v} - \bar{\mathbf{v}}, \mathbf{u} \rangle \neq 0$. Como poderíamos trocar $\mathbf{u}$ por $-\mathbf{u}$ e tomar o unitário de $\mathbf{u}$ sem sair de $\mathbf{V}_o$, não haveria mal algum em supor $|\mathbf{u}| = 1$ e $\langle \mathbf{v} - \bar{\mathbf{v}}, \mathbf{u} \rangle > 0$. Projetando $\mathbf{v} - \bar{\mathbf{v}}$ sobre $\mathbf{u}$ e fazendo

$$\mathbf{v}_o = \bar{\mathbf{v}} + \langle \mathbf{v} - \bar{\mathbf{v}}, \mathbf{u} \rangle \, \mathbf{u},$$

teríamos, pelo Teorema de Pitágoras,

$$|\mathbf{v} - \bar{\mathbf{v}}|^2 = |\mathbf{v} - \mathbf{v}_o|^2 + |\langle \mathbf{v} - \bar{\mathbf{v}}, \mathbf{u} \rangle \, \mathbf{u}|^2 = |\mathbf{v} - \mathbf{v}_o|^2 + (\langle \mathbf{v} - \bar{\mathbf{v}}, \mathbf{u} \rangle)^2 > |\mathbf{v} - \mathbf{v}_o|^2,$$

o que entraria em choque com nossa hipótese. ■

Definição: Uma base $\{\boldsymbol{\varepsilon}_1, \ldots, \varepsilon_n\}$ do espaço vetorial $\mathbf{V}$ com produto interno é dita uma **base ortogonal** de V se

$$\langle \varepsilon_i, \varepsilon_j \rangle = 0 \; \forall \; i, j = 1, \ldots, n, \; i \neq j.$$

Se, além disso, tivermos $\langle \boldsymbol{\varepsilon_i}, \boldsymbol{\varepsilon_i} \rangle = 1 \; \forall \; i = 1, \ldots, n$, então a base é dita **ortonormal**.

Exercício 8.10 *Veja se entendeu mesmo. Suponha que $\{\boldsymbol{\varepsilon}_1, \ldots, \varepsilon_n\}$ é uma base ortonormal do espaço* $\mathbf{V}$. *Seja* $\mathbf{v}$ *um elemento de* $\mathbf{V}$. *Mostre que* $\mathbf{v} = x_1\boldsymbol{\varepsilon}_1 + \cdots + x_n\boldsymbol{\varepsilon}_n$, *com*

$$x_i = \langle \mathbf{v}, \boldsymbol{\varepsilon_i} \rangle \; \forall \; i = 1, \ldots, n.$$

Exercício 8.11 *Por via das dúvidas, veja se está claro: se* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *é base ortogonal de* $\mathbf{V}$, *então* $\{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n\}$, *com*

$$\boldsymbol{\varepsilon}_i = \frac{1}{|\mathbf{v}_i|}\mathbf{v}_i, \ i = 1, \ldots, n,$$

é base ortonormal de $\mathbf{V}$.

Exercício 8.12 *Um subconjunto K do espaço vetorial* $\mathbf{V}$ *é dito* ***convexo*** *se, para quaisquer* $\mathbf{u}$ *e* $\mathbf{v}$ *em K, o segmento* $[\mathbf{u}, \mathbf{v}] = \{t\mathbf{v} + (1-t)\mathbf{u}, \ t \in [0,1]\}$ *está contido em K (note que todo subespaço vetorial de* $\mathbf{V}$ *é convexo). Se* $\mathbf{V}$ *tem produto interno, a* ***projeção*** *sobre K do vetor* $\mathbf{v}$ *é o elemento* $p_K(\vec{v})$ *de K dado por*

$$|\mathbf{v} - p_k(\mathbf{v})| \leq |\mathbf{v} - \mathbf{w}| \ \forall \mathbf{w} \in K.$$

A existência da projeção $p_K(\mathbf{v})$, *para qualquer* $\mathbf{v}$ *em* $\mathbf{V}$, *não é garantida, a menos que K seja um* ***subconjunto fechado*** *de* $\mathbf{V}$.[1] *Independente disso, prove:*

1. $p_K(\mathbf{v})$, *se existe, é única;*

2. *O elemento* $\mathbf{v}_o$ *de K é a projeção de* $\mathbf{v}$ *sobre K se, e somente se,*

$$\langle \mathbf{v} - \mathbf{v}_o, \mathbf{w} - \mathbf{v}_o \rangle \leq 0 \ \forall \ \mathbf{w} \in K.$$

3. *Se existem* $p_K(\mathbf{u})$ *e* $p_K(\mathbf{v})$, *então* $|p_K(\mathbf{u}) - p_K(\mathbf{v})| \leq |\mathbf{u} - \mathbf{v}|$.

Exercício 8.13 *Seja* $\mathbf{V}$ *o espaço das funções contínuas definidas em* $[0, 2\pi]$ *e a valores em* $\mathbb{R}$, *com o produto interno*

$$\langle f, g \rangle = \int_0^{2\pi} f(x)g(x)dx.$$

Seja, para cada $n = 0, 1, 2, \ldots$, *o elemento* $\boldsymbol{\varepsilon}_n$ *de* $\mathbf{V}$ *definido por*

$$\varepsilon_n = \begin{cases} \cos \frac{n}{2}x, \ n \ par \\ \sin \frac{n+1}{2}x, \ n \ ímpar \end{cases}$$

Mostre que, se $n \neq m$, $\langle \boldsymbol{\varepsilon}_n, \boldsymbol{\varepsilon}_m \rangle = 0$.

8.5 O processo de Gram-Schmidt

A seção anterior destacou o papel fundamental das bases ortonormais em espaços com produto interno. Até agora, porém, não demonstramos a existência de bases ortonormais em um contexto geral. Mais claramente: supondo que $\mathbf{V}$ é um espaço vetorial com produto interno, podemos garantir que $\mathbf{V}$ tem uma base ortonormal? Todo o trabalho que já tivemos com as projeções deve ter deixado pistas de que a resposta, se $\mathbf{V}$ é de dimensão finita, é sim e indicar o processo que nos permite construir, a partir de uma base qualquer de $\mathbf{V}$, uma base ortonormal.

[1] K é dito fechado se seu complementar em $\mathbf{V}$ é aberto. Um subconjunto A de $\mathbf{V}$ é dito **aberto** se $\forall x \in A \ \exists \varepsilon > 0 \mid |y - x| < \varepsilon \Rightarrow y \in A$

Lema de Gram-Schmidt: Todo espaço vetorial de dimensão finita com produto interno tem base ortonormal. Se $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ é base de $\mathbf{V}$, então $\{\boldsymbol{\varepsilon}_\mathbf{1}, \ldots, \boldsymbol{\varepsilon}_n\}$, construída recursivamente por:

1. $\boldsymbol{\varepsilon}_\mathbf{1} = \frac{1}{|\mathbf{v}_1|}\mathbf{v}_1$,
2. $\mathbf{u}_{i+1} = \mathbf{v}_{i+1} - \sum_{j=1}^{i} - \langle \mathbf{v}_{i+1}, \boldsymbol{\varepsilon}_j \rangle \boldsymbol{\varepsilon}_j$,
3. $\boldsymbol{\varepsilon}_{i+1} = \frac{1}{|\mathbf{u}_{i+1}|}\mathbf{u}_{i+1}$,

é base ortonormal de $\mathbf{V}$.

Demonstração: Os $\boldsymbol{\varepsilon}_i$ são unitários por construção. A demonstração de que $\mathbf{u}_{i+1}$ (e, portanto, $\boldsymbol{\varepsilon}_{i+1}$) é ortogonal a $\boldsymbol{\varepsilon}_\mathbf{1}, \ldots, \boldsymbol{\varepsilon}_i$ é uma simples conta. ■

Podemos, agora, juntar as ideias que já desenvolvemos sobre projeções e o Lema de Gram-Schmidt em um teorema.

Teorema da Projeção: Sejam $\mathbf{V}$ um espaço vetorial com produto interno e $\mathbf{V}_o$ um subespaço vetorial de $\mathbf{V}$. Suponhamos que $\mathbf{V}_o$ é de dimensão finita. Então existe uma única aplicação $P : \mathbf{V} \to \mathbf{V}_o$ satisfazendo, para cada $\mathbf{v}$ em $\mathbf{V}$, as duas propriedades equivalentes:

1. $\langle \mathbf{v} - P\mathbf{v}, \mathbf{w} \rangle = 0 \; \forall \mathbf{w} \in \mathbf{V}_o$
2. $|\mathbf{v} - P\mathbf{v}| \leq |\mathbf{v} - \mathbf{w}| \; \forall \mathbf{w} \in \mathbf{V}_o$

A transformação P assim definida é linear; além disso, se $\{\boldsymbol{\varepsilon}_\mathbf{1}, \ldots, \boldsymbol{\varepsilon}_n\}$ é base ortonormal de $\mathbf{V}_o$, então a P é dada, para cada $\mathbf{v}$ em $\mathbf{V}$, por

$$P\mathbf{v} = \sum_{i=1}^{n} \langle \mathbf{v}, \boldsymbol{\varepsilon}_i \rangle \boldsymbol{\varepsilon}_i.$$

Exercício 8.14 *Mostre que, para quaisquer* $\mathbf{v}_1$ *e* $\mathbf{v}_2$ *em* $\mathbf{V}$, *vale* $|P\mathbf{v}_1 - P\mathbf{v}_2| \leq |\mathbf{v}_1 - \mathbf{v}_2|$, *com igualdade se, e somente se,* $\mathbf{v}_1 - \mathbf{v}_2 \in \mathbf{V}_o$.

Capítulo 9

O determinante

Este capítulo é apenas uma primeira apresentação dos determinantes: retornaremos ao assunto, com mais detalhes, no capítulo 14. Na verdade, esta é uma segunda apresentação do assunto. Uma leitura do capítulo sobre determinantes de nosso livro de Geometria Analítica plana é uma boa introdução, e facilitará bastante a vida do leitor. Vamos começar reproduzindo parte do material lá contido.

9.1 Áreas

Até agora somos capazes de medir distâncias e de determinar ângulos através de coordenadas. Vejamos agora como lidar com o cálculo de áreas. Em princípio, se sabemos calcular os comprimentos de dois vetores e o seno (que podemos obter do cosseno) do ângulo entre eles, temos certeza de poder chegar à área de qualquer paralelogramo.

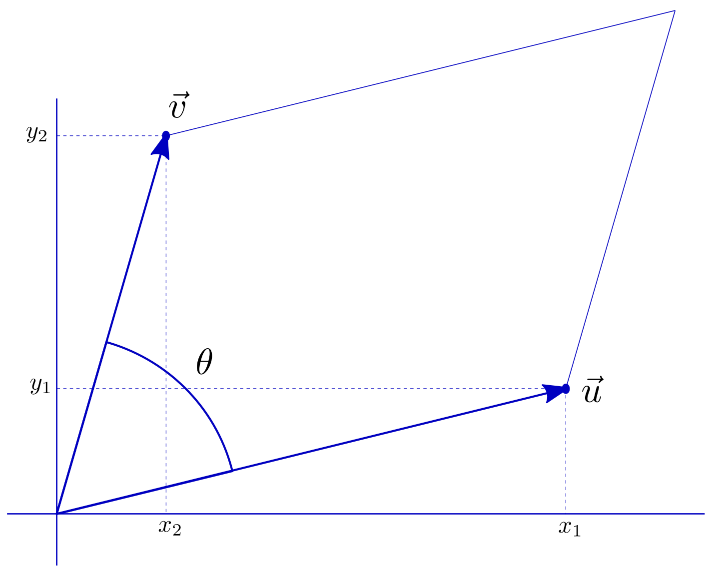


Figura 9.1:

Vamos, porém, partir para uma abordagem direta: tentaremos associar a cada par de vetores, $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$, a área do paralelogramo por eles formado, expressa diretamente em função de x_1, y_1, x_2, y_2. Veremos, depois de algumas peripécias, que tal área é dada pelo valor absoluto do **determinante**

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1.$$

9.2 Orientação

Comecemos definindo a **orientação** de um par de vetores. Sejam $\vec{u}_1, \vec{u}_2$ dois vetores não paralelos e não nulos. Diremos que o par $\vec{u}_1, \vec{u}_2$ tem orientação positiva se o seno do ângulo θ entre $\vec{u}_1$ e $\vec{u}_2$, medido de $\vec{u}_1$ para $\vec{u}_2$ no sentido trigonométrico, é positivo (ou, o que é equivalente, se, "para girarmos o ponteiro $\vec{u}_1$ para o ponteiro $\vec{u}_2$ pelo menor ângulo, andamos no sentido trigonométrico").

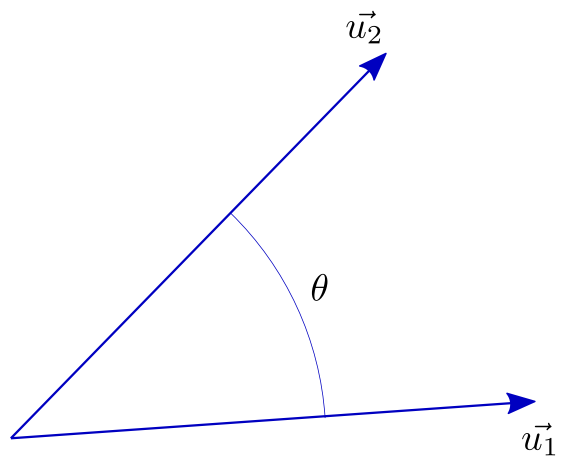


Figura 9.2:

Note que a orientação depende da ordem em que tomamos os vetores, e que se a orientação de $\vec{u}_1, \vec{u}_2$ é positiva, então a de $\vec{u}_2, \vec{u}_1$ é negativa. Assim, quando falarmos "a orientação de $\vec{u}, \vec{v}$", estará sempre implícito que se trata de um par ordenado. Diremos que dois pares de vetores $\vec{u}_1, \vec{u}_2$ e $\vec{v}_1, \vec{v}_2$ têm **a mesma orientação** se as respectivas orientações são simultaneamente positivas ou simultaneamente negativas. Assim, por exemplo, o par $\vec{u}, \vec{v}$ tem orientação positiva se e só se tem a mesma orientação que o par formado pela base canônica, $\vec{e}_1, \vec{e}_2$.

Exercício 9.1 *Verifique que se $\vec{u}, \vec{v}$ tem orientação positiva e t é um número real não nulo, então $t\vec{u}, \vec{v}$ e $\vec{u}, t\vec{v}$ têm orientação positiva se $t > 0$ e negativa se $t < 0$.*

Exercício 9.2 *Considere o vetor $\vec{u} = (x, y)$ identificado com o ponto $P = (x, y)$. Considere a reta OP, coloque-se sobre a origem e olhe para P. Verifique que o par $\vec{u}, \vec{v}$ tem orientação positiva se e só se o ponto correspondente a $\vec{v}$ está à sua esquerda.*

Exercício 9.3 *Mostre que $\vec{u}, \vec{v}$ e $\vec{u}, \vec{v} + t\vec{u}$ têm a mesma orientação, qualquer que seja t.*

Exercício 9.4 *Suponha que $\vec{u}, \vec{v}$ tem orientação positiva. Gire $\vec{u}$ de um ângulo reto no sentido trigonométrico, obtendo o vetor $\vec{u}^{\perp}$. Mostre que o produto escalar*

$$\left\langle \vec{u}^{\perp}, \vec{v} \right\rangle$$

é positivo.

Exercício 9.5 *Sejam $\vec{u} = (a_{11}, a_{21})$ e $\vec{v} = (a_{12}, a_{22})$. Use a observação do exercício anterior para mostrar que $\vec{u}, \vec{v}$ tem orientação positiva se, e somente se,*

$$a_{11}a_{22} - a_{21}a_{12} > 0.$$

9.3 Áreas com sinal

Vamos agora definir uma função d, que a cada par (ordenado) de vetores $\vec{u}, \vec{v}$ associa a área do paralelogramo por eles formado.

Fica entendido que se $\vec{u}$ e $\vec{v}$ são paralelos (o que inclui a possibilidade de um dos dois ser nulo, ou ambos), então $d(\vec{u}, \vec{v}) = 0$. Incluiremos na definição de d, porém, uma novidade, que a distingue do que comumente chamamos área: $d(\vec{u}, \vec{v})$ será a área do paralelogramo formado por $\vec{u}$ e $\vec{v}$, mas com um sinal, positivo, se o par $\vec{u}, \vec{v}$ tiver orientação positiva, e negativo, se a orientação de $\vec{u}, \vec{v}$ for negativa. É claro que o leitor não é obrigado a aceitar áreas negativas assim à toa, e daremos boas razões algébricas para a ousadia.

A primeira razão algébrica é a seguinte: se t é positivo, a área do paralelogramo formado por $t\vec{u}$ e $\vec{v}$ é t vezes a do paralelogramo formado por $\vec{u}$ e $\vec{v}$, o que nos leva a conjecturar que

$$d(t\vec{u}, \vec{v}) = td(\vec{u}, \vec{v}).$$

Mas na verdade isso não pode valer para t negativo, a menos que admitamos valores negativos para d ou que modifiquemos um pouco a fórmula acima. Podemos ainda notar que o problema que surge diz respeito apenas ao sinal. Ora, se d troca de sinal quando trocamos a orientação, então a definição que demos está boa, pois t negativo troca o sinal dos dois lados da igualdade.

O leitor argumentará, talvez, que bastaria escrever $d(t\vec{u}, \vec{v}) = |t|d(\vec{u}, \vec{v})$. Poderíamos contra-argumentar dizendo que trabalhar com $|t|$ é chatíssimo, mas preferimos lançar mão de nossa segunda razão algébrica, que é um verdadeiro canhão.

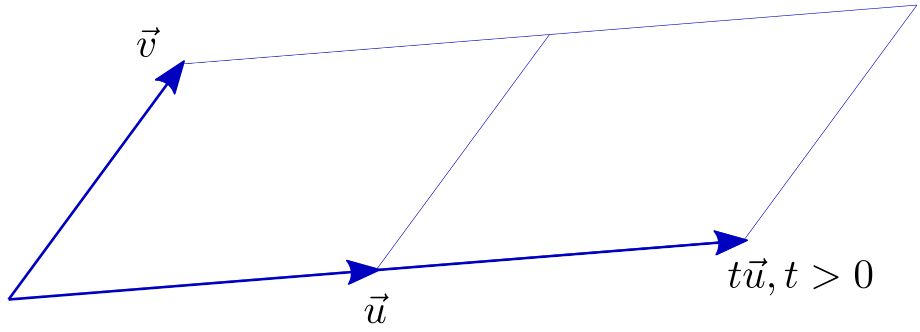


Figura 9.3:

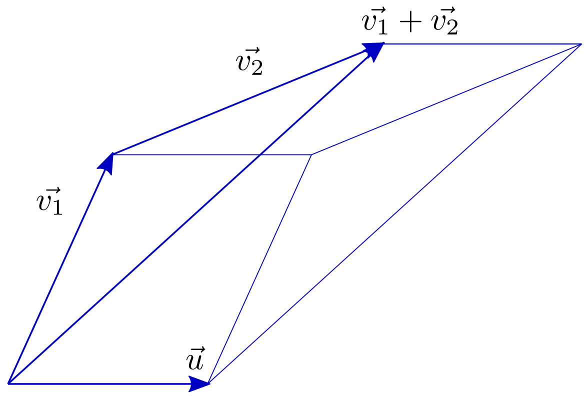


Figura 9.4:

A figura abaixo nos sugere a seguinte propriedade, pensando em termos de áreas:

$$d(\vec{u}, \vec{v}_1 + \vec{v}_2) = d(\vec{u}, \vec{v}_1) + d(\vec{u}, \vec{v}_2).$$

No entanto, a figura seguinte já sugere outra coisa:

$$d(\vec{u}, \vec{v}_1 + \vec{v}_2) = d(\vec{u}, \vec{v}_1) - d(\vec{u}, \vec{v}_2).$$

Pois é...Na primeira figura, podemos observar, os pares $\vec{u}, \vec{v}_1$ e $\vec{u}, \vec{v}_2$ têm a mesma orientação; já na segunda, as orientações são opostas.

Exercício 9.6 *Pegue papel e lápis e desenhe todos os casos que achar necessário até se convencer de que, trabalhando com áreas negativas (isto é, com a definição de d dada acima), vale a propriedade*

$$d(\vec{u}, \vec{v}_1 + \vec{v}_2) = d(\vec{u}, \vec{v}_1) + d(\vec{u}, \vec{v}_2) \quad \forall \vec{u}, \vec{v}_1, \vec{v}_2 \in \mathbb{R}^2.$$

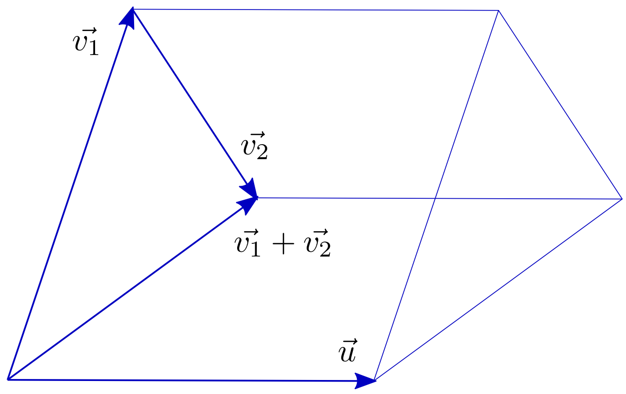


Figura 9.5:

Vamos tratar nossa função d, agora, através de certas propriedades notáveis. Vamos ver que tais propriedades caracterizam d e nos permitem deduzir uma expressão simples para seu cálculo.

d é uma função que a cada par (ordenado) $\vec{u}, \vec{v}$ de vetores do plano associa um número real $d(\vec{u}, \vec{v})$, com as seguintes propriedades:

$$\begin{array}{ll}
(i) d(\vec{u}, \vec{v}) = -d(\vec{v}, \vec{u}) & \forall \vec{u}, \vec{v} \in \mathbb{R}^2; \\
(ii) d(t\vec{u}, \vec{v}) = td(\vec{u}, \vec{v}) & \forall \vec{u}, \vec{v} \in \mathbb{R}^2,\ \forall t \in \mathbb{R}; \\
(iii) d(\vec{u}, \vec{v}_1 + \vec{v}_2) = d(\vec{u}, \vec{v}_1) + d(\vec{u}, \vec{v}_2) & \forall \vec{u}, \vec{v}_1, \vec{v}_2 \in \mathbb{R}^2; \\
(iv) d(\vec{e_1}, \vec{e_2}) = 1. &
\end{array}$$

As propriedades (i), (ii) e (iii) foram discutidas na seção precedente; a propriedade (iv) parece óbvia, mas não teríamos como deduzi-la das demais. Três outras propriedades com as quais contamos podem ser deduzidas de (i), (ii) e (iii):

$$\begin{array}{ll}
(i)' d(\vec{u}, \vec{u}) = 0 & \forall \vec{u}; \\
(ii)' d(\vec{u}, t\vec{v}) = td(\vec{u}, \vec{v}) & \forall \vec{u}, \vec{v} \in \mathbb{R}^2,\ \forall t \in \mathbb{R}; \\
(iii)' d(\vec{u}_1 + \vec{u}_2, \vec{v}) = d(\vec{u}_1, \vec{v}) + d(\vec{u}_2, \vec{v}) & \forall \vec{u}_1, \vec{u}_2, \vec{v} \in \mathbb{R}^2.
\end{array}$$

As demonstrações são simples e puramente algébricas:

(i)′ segue do fato que $d(\vec{u}, \vec{u}) = -d(\vec{u}, \vec{u})$ (por (i));

(ii)′ se deduz notando que, por (i) e (ii), $d(\vec{u}, t\vec{v}) = -d(t\vec{v}, \vec{u}) = -td(\vec{v}, \vec{u}) =$
=-t(-d($\vec{u}$,$\vec{v}$))=td($\vec{u}$,$\vec{v}$).

Exercício 9.7 *Prove (iii)′ usando apenas (i) e (iii).*

Vamos agora, sem mais delongas, proceder ao cálculo de $d(\vec{u}, \vec{v})$, usando as propriedades acima. Sendo $\vec{u} = (x_1, y_1) = x_1\vec{e_1} + y_1\vec{e_2}$, $\vec{v} = (x_2, y_2) = x_2\vec{e_1} + y_2\vec{e_2}$, temos:

$$d(\vec{u},\vec{v}) = d(x_1\vec{e_1} + y_1\vec{e_2}, x_2\vec{e_1} + y_2\vec{e_2}) =$$

$$= x_1x_2d(\vec{e_1},\vec{e_1}) + x_1y_2d(\vec{e_1},\vec{e_2}) + y_1x_2d(\vec{e_2}.\vec{e_1}) + y_1y_2d(\vec{e_2},\vec{e_2}).$$

Agora basta substituir

$$d(\vec{e_1},\vec{e_1}) = 0, d(\vec{e_2},\vec{e_2}) = 0, d(\vec{e_1},\vec{e_2}) = 1, d(\vec{e_2},\vec{e_1}) = -d(\vec{e_1},\vec{e_2}) = -1$$

para obter

$$d(\vec{u},\vec{v}) = x_1y_2 - x_2y_1,$$

ou, usando a notação consagrada,

$$d(\vec{u},\vec{v}) = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}[1].$$

Assim, a **área** (com sinal) do paralelogramo formado por $\vec{u} = (x_1, y_1)$ e $\vec{v} = (x_2, y_2)$ é dada por $x_1y_2 - x_2y_1$. Se fizermos questão da área "mesmo", basta tomarmos o valor absoluto.

9.4 Volumes com sinal

A exemplo do que fizemos no plano, vamos procurar associar, a cada trio de vetores, $\vec{u}$, $\vec{v}$ e $\vec{w}$, um número que corresponda ao **volume** do paralelepípedo p definido a partir de $\vec{u}$, $\vec{v}$ e $\vec{w}$.

Exercício 9.8 *Entenda que o palalelepípedo de que falamos é o subconjunto p de V definido por $p = \{r\vec{u} + s\vec{v} + t\vec{w},\ (r,s,t) \in [0,1] \times [0,1] \times [0,1]\}$.*

Como no caso do plano, é razoável supor que aceitemos volumes negativos, dependendo da ordem em que tomamos os vetores $\vec{u}$, $\vec{v}$ e $\vec{w}$.[2] Assim, vamos, na verdade, buscar uma função det que, a cada terno ordenado $(\vec{u},\vec{v},\vec{w})$ associe um número, $det(\vec{u},\vec{v},\vec{w})$, que chamaremos de **determinante** de $(\vec{u},\vec{v},\vec{w})$, com as propriedades:

$$\begin{array}{ll}
(i) & det(\vec{u},\vec{v},\vec{w}) = 0, \text{ se } \vec{u},\vec{v} \text{ e } \vec{w} \text{ são linearmente dependentes} \\
(ii) & det(\vec{u}_1 + t\vec{u}_2,\vec{v},\vec{w}) = det(\vec{u}_1,\vec{v},\vec{w}) + t\, det(\vec{u}_2,\vec{v},\vec{w})\ \forall t,\vec{u}_1,\vec{u}_2,\vec{v},\vec{w} \\
 & det(\vec{u},\vec{v}_1 + t\vec{v}_2,\vec{w}) = det(\vec{u},\vec{v}_1,\vec{w}) + t\, det(\vec{u},\vec{v}_2,\vec{w})\ \forall t,\vec{u},\vec{v}_1,\vec{v}_2,\vec{w} \\
 & det(\vec{u},\vec{v},\vec{w}_1 + t\vec{w}_2) = det(\vec{u},\vec{v},\vec{w}_1) + t\ (\vec{u},\vec{v},\vec{w}_2)\ \forall t,\vec{u},\vec{v},\vec{w}_1,\vec{w}_2 \\
(iii) & det(\vec{e_1},\vec{e_2},\vec{e_3}) = 1
\end{array}$$

---

[1] $\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1y_2 - x_2y_1$ é chamado **determinante** da matriz $\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}$

[2] *razoável*, aqui, significa que vale a pena, como no caso do plano, abrir mão da exigência de que volumes sejam, sempre, positivos, em troca de boas propriedades algébricas; desta forma, o *verdadeiro* volume será dado pelo valor absoluto do número que vamos definir

Por conta da propriedade (ii), o determinante é uma **aplicação trilinear** (ou **forma trilinear**, neste caso). Por conta da propriedade a ser provada no exercício a seguir, o determinante é uma **forma trilinear alternada**.

Exercício 9.9 *Dedique o tempo que achar necessário a ver se acredita mesmo que o as propriedades acima caracterizam o volume (com sinal) do paralelepípedo p.*[3]

Exercício 9.10 *Mostre que, das propriedades (i) e (ii) acima, decorre que o sinal de det muda, se trocarmos a ordem de dois dos vetores.*

Solução: para a troca, por exemplo, de $\vec{u}$ e $\vec{w}$, começamos observando que $0 = det(\vec{u}+\vec{w},\vec{v},\vec{u}+\vec{w}) = det(\vec{u},\vec{v},\vec{u}) + det(\vec{u},\vec{v},\vec{w}) + det(\vec{w},\vec{v},\vec{u}) + det(\vec{w},\vec{v},\vec{w}) = det(\vec{u},\vec{v},\vec{w}) + det(\vec{w},\vec{v},\vec{u})$. Logo, $det(\vec{u},\vec{v},\vec{w}) = -det(\vec{w},\vec{v},\vec{u})$.

Definição: Uma **forma trilinear alternada** em **V** é uma aplicação $\omega : \mathbf{V}\times\mathbf{V}\times\mathbf{V} \to \mathbb{R}$ tal que:

$$\begin{array}{ll}
(i) & \omega(\vec{u},\vec{v},\vec{w}) = 0, \textit{ se } \vec{u},\vec{v} \textit{ e } \vec{w} \textit{ são linearmente dependentes} \\
(ii) & \omega(\vec{u}_1+t\vec{u}_2,\vec{v},\vec{w}) = \omega(\vec{u}_1,\vec{v},\vec{w}) + t\,\omega(\vec{u}_2,\vec{v},\vec{w})\ \forall t,\vec{u}_1,\vec{u}_2,\vec{v},\vec{w} \\
 & \omega(\vec{u},\vec{v}_1+t\vec{v}_2,\vec{w}) = \omega(\vec{u},\vec{v}_1,\vec{w}) + t\,\omega(\vec{u},\vec{v}_2,\vec{w})\ \forall t,\vec{u},\vec{v}_1,\vec{v}_2,\vec{w} \\
 & \omega(\vec{u},\vec{v},\vec{w}_1+t\vec{w}_2) = \omega(\vec{u},\vec{v},\vec{w}_1) + t\,(\vec{u},\vec{v},\vec{w}_2)\ \forall t,\vec{u},\vec{v},\vec{w}_1,\vec{w}_2
\end{array}$$

9.5 A fórmula

Não é difícil ver que, a exemplo do que ocorre no plano, as propriedades (i), (ii) e (iii) determinam, para cada terno $(\vec{u},\vec{v},\vec{w})$, um único número $det(\vec{u},\vec{v},\vec{w})$, que pode ser calculado em termos das coordenadas de $\vec{u}$, $\vec{v}$ e $\vec{w}$ na base canônica.

Proposição: Suponhamos que os vetores $\vec{u}$, $\vec{v}$ e $\vec{w}$ sejam dados, na base canônica, por

$$\begin{array}{l}
\vec{u} = a_{11}\vec{e}_1 + a_{21}\vec{e}_2 + a_{31}\vec{e}_3, \\
\vec{v} = a_{12}\vec{e}_1 + a_{22}\vec{e}_2 + a_{32}\vec{e}_3, \\
\vec{w} = a_{13}\vec{e}_1 + a_{23}\vec{e}_2 + a_{33}\vec{e}_3.
\end{array}$$

Então, das propriedades (i), (ii) e (iii) decorre:

$$det(\vec{u},\vec{v},\vec{w}) = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

Demonstração: É um simples exercício. Basta escrever

$$det(\vec{u},\vec{v},\vec{w}) = det(a_{11}\vec{e}_1 + a_{21}\vec{e}_2 + a_{31}\vec{e}_3, a_{12}\vec{e}_1 + a_{22}\vec{e}_2 + a_{32}\vec{e}_3, a_{13}\vec{e}_1 + a_{23}\vec{e}_2 + a_{33}\vec{e}_3)$$

[3] Historicamente, o determinante (restrito ao caso 2×2) aparece no Ars Magna, de Cardano, no século XVI, associado à solução de sistemas de equações lineares; o caso geral só será devidamente estudado cerca de duzentos anos depois. Sua interpretação geométrica, em termos de áreas e volumes, vem apenas na segunda metade do século XVIII, com Lagrange. Voltaremos à relação entre determinantes e volumes no final do capítulo sobre o produto vetorial

e fazer as contas. ■

Escólio 1: Segue diretamente da demonstração que, se ω é uma outra forma trilinear alternada em V (isto é, ω associa um número a cada terno ordenado $(\vec{u}, \vec{v}, \vec{w})$ de vetores de V, satisfazendo as propriedades (i) e (ii), mas não necessariamente a propriedade (iii)), então, obrigatoriamente,

$$\omega(\vec{u}, \vec{v}, \vec{w}) = \omega(a_{11}\vec{e_1} + a_{21}\vec{e_2} + a_{31}\vec{e_3}, a_{12}\vec{e_1} + a_{22}\vec{e_2} + a_{32}\vec{e_3}, a_{13}\vec{e_1} + a_{23}\vec{e_2} + a_{33}\vec{e_3}) = (a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31})\omega(\vec{e_1}, \vec{e_2}, \vec{e_3}).$$

Isto merece ser chamado de Proposição.

Proposição 1: Para cada forma trilinear alternada ω em V existe um número α (que só depende de ω e é dado por $\alpha = \omega(\vec{e_1}, \vec{e_2}, \vec{e_3})$) tal que

$$\omega(\vec{u}, \vec{v}, \vec{w}) = \alpha\, det(\vec{u}, \vec{v}, \vec{w}), \ \forall (\vec{u}, \vec{v}, \vec{w}) \in \ V \times V \times V.$$

■

Escólio 2: Podemos inverter o raciocínio. Suponhamos que $\vec{u}$, $\vec{v}$ e $\vec{w}$ sejam linearmente independentes e, portanto, formem uma base de V. Consideremos uma forma trilinear alternada ω. Como acabamos de ver, ω é determinada por $\omega(\vec{e_1}, \vec{e_2}, \vec{e_3})$. Por outro lado, podemos,visto que $\vec{u}$, $\vec{v}$ e $\vec{w}$ formam base para V, escrever $\vec{e_1}$, $\vec{e_2}$ e $\vec{e_3}$ como combinações lineares de $\vec{u}$, $\vec{v}$ e $\vec{w}$:

$$\begin{aligned} \vec{e_1} &= b_{11}\vec{u} + b_{21}\vec{v} + b_{31}\vec{w}, \\ \vec{e_2} &= b_{12}\vec{u} + b_{22}\vec{v} + b_{32}w, \\ \vec{e_3} &= b_{13}\vec{u} + b_{23}\vec{v} + b_{33}w. \end{aligned}$$

Desenvolvendo $\omega(\vec{e_1}, \vec{e_2}, \vec{e_3})$, obtemos, exatamente como na Proposição,

$$\omega(\vec{e_1}, \vec{e_2}, \vec{e_3}) = (b_{11}b_{22}b_{33} - b_{11}b_{23}b_{32} + b_{12}b_{23}b_{31} - b_{12}b_{21}b_{33} + b_{13}b_{21}b_{32} - b_{13}b_{22}b_{31})\omega(\vec{u}, \vec{v}, \vec{w}).$$

Assim, se $\omega(\vec{u}, \vec{v}, \vec{w}) = 0$, ω será identicamente nula. Em particular, fazendo $\omega = det$, podemos concluir que

Proposição 2: $\det(\vec{u}, \vec{v}, \vec{w}) \neq 0$ se, e somente se, $\vec{u}$, $\vec{v}$ e $\vec{w}$ forem linearmente independentes. ■

O número

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

é, usualmente, notado por

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

e chamado de **determinante** da matriz

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

O determinante é, assim, uma **forma trilinear alternada** dos vetores coluna da matriz. Mais precisamente, se representarmos por (a_{ij}) a matriz, por $|a_{ij}|$ seu determinante e por (a_j) cada um dos vetores dados, em coordenadas, por $(a_j) = (a_{1j}, a_{2j}, a_{3j})$, então $|a_{ij}|$ é uma forma trilinear alternada do terno de vetores $((a_1), (a_2), (a_3))$.

Exercício 9.11 *Demonstre essa última afirmação. Isto é: mostre, a partir da fórmula, que o determinante é uma forma trilinear alternada dos vetores coluna da matriz. Mostre que o determinante da **matriz identidade** (aquela que tem $a_{ij} = 1$, se $i = j$, e $a_{ij} = 0$, se $i \neq j$) é 1.*

Exercício 9.12 *Mostre, também a partir da fórmula, que o determinante é uma forma trilinear alternada dos vetores linha da matriz (os vetores linha são os $(\bar{a}_i)$ dados, em coordenadas, por $(\bar{a}_i) = (a_{i1}, a_{i2}, a_{i3})$).*

Exercício 9.13 *Conclua que é nulo o determinante de qualquer matriz em que uma das colunas é combinação linear das outras duas; o mesmo para matrizes em que uma das linhas é combinação linear das outras duas.*

Exercício 9.14 *Observe que uma forma simples de decorar a fórmula é escrever*

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{vmatrix}$$

e observar que os produtos com sinal + são obtidos, seguindo setas $\searrow$, partindo de a_{11}, a_{21} e a_{31}; os com sinal - são obtidos, seguindo setas $\swarrow$, partindo de a_{13}, a_{23} e a_{33}.

Exercício 9.15 *Usando a fórmula do determinante para matrizes 2×2, observe que*

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Exercício 9.16 *De maneira mais geral, mostre que o determinante $|a_{ij}|$ é dado, para qualquer linha i, por*

$$|a_{ij}| = \sum_{j=1}^{3} (-1)^{i+j} a_{ij} |\tilde{a}_{ij}|,$$

*com $|\tilde{a}_{ij}|$ definido como o determinante da matriz 2×2 obtida de (a_{ij}) pela exclusão da linha i e da coluna j. Chamaremos de **cofator** de a_{ij} o número $(-1)^{i+j}|\tilde{a}_{ij}|$. Essa terminologia não é unânime: como o determinante $|\tilde{a}_{ij}|$ já tem nome (é chamado de **determinante menor** de aij), usa-se também chamar de **cofator** de a_{ij} o número $(-1)^{i+j}$.*

Exercício 9.17 *Mostre que, para qualquer coluna j, vale*

$$|a_{ij}| = \sum_{i=1}^{3}(-1)^{i+j}a_{ij}|\tilde{a}_{ij}|.$$

Exercício 9.18 *Sejam A e B as matrizes* 3×3 *dadas por* (a_{ij}) *e* (b_{ij}). *O* ***produto*** *de A e B é a matriz* $C = (c_{ij})$, *, assim definida:*

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix},$$

com $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j}$ *(note que* c_{ij} *é o produto escalar da linha i de A peela coluna j de B).*

Suponha que a matriz (a_{ij}) *tem determinante não nulo. Seja* (b_{ij}) *a matriz definida por* $b_{ij} = (det(a_{ij}))^{-1}(-1)^{i+j}|\tilde{a}_{ji}|$ *(note que, a menos do fator* $(det(a_{ij}))^{-1}$, *trata-se da transposta da matriz dos cofatores). Mostre que* $(a_{ij})(b_{ij}) = (\delta_{ij})$ *(*(δ_{ij}) *é a* ***matriz identidade***, *I, dada por* $\delta_{ij} = 0$, *se* $i \neq j$, *e* $\delta_{ij} = 1$, *se* $i = j$*).*

Mostre que $IA = AI = A$ *para toda matriz* 3×3 *A.*

Exercício 9.19 *Suponha que* $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ *seja solução do sistema*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = y_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = y_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = y_3, \end{cases}$$

ou seja,

$$\begin{aligned} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \\ &= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}. \end{aligned}$$

Observe que, se trocarmos a primeira coluna da matriz $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ *por* $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, *teremos:*

$$det \begin{pmatrix} y_1 & a_{12} & a_{13} \\ y_2 & a_{22} & a_{23} \\ y_3 & a_{32} & a_{33} \end{pmatrix} =$$

$$= det \left(x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} \quad \begin{matrix} a_{12} & a_{13} \\ a_{22} & a_{23} \\ a_{32} & a_{33} \end{matrix} \right) =$$

$$= x_1 det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} +$$

$$+ x_2 det \begin{pmatrix} a_{12} & a_{12} & a_{13} \\ a_{22} & a_{22} & a_{23} \\ a_{32} & a_{32} & a_{33} \end{pmatrix} + x_3 det \begin{pmatrix} a_{13} & a_{12} & a_{13} \\ a_{23} & a_{22} & a_{23} \\ a_{33} & a_{32} & a_{33} \end{pmatrix} =$$

$$= x_1 det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

*Logo, se o determinante da matriz (a_{ij}) (também chamado de **determinante do sistema**) for não nulo, temos:*

$$x_1 = \frac{det \begin{pmatrix} y_1 & a_{12} & a_{13} \\ y_2 & a_{22} & a_{23} \\ y_3 & a_{32} & a_{33} \end{pmatrix}}{det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}}.$$

*Esta é a chamada **fórmula de Cramer** para a solução x_1 do sistema.*

Exercício 9.20 *Obtenha a fórmula de Cramer para x_2 e para x_3.*

Observação: Nosso estudo de determinantes foi apresentado de um ponto de vista absolutamente geométrico. Por uma questão de honestidade histórica, devemos salientar que a origem do conceito é algébrica (**Cardano**, dá a expressão para o caso 2×2 (1545); **Leibniz** trabalha com a ideia (1693) e **Seki Kowa**, na mesma época, obtém resultados semelhantes (casos 3×3 e 4×4, com boas tentativas para o caso $n \times n$); a ideia vai ser detalhadamente desenvolvida por **MacLaurin** (1748) e Cramer(1750); a fórmula dita de Cramer foi obtida, de forma independente e um pouco antes, por MacLaurin). A utilização de determinantes para o cálculo de áreas e volumes é posterior (Lagrange (1773), Cauchy (1812)).

9.6 Orientação

Tratemos, agora, de uma questão que vem sendo adiada desde o começo do capítulo (aproveitando a oportunidade para agradecer a boa vontade do leitor que, pacientemente, aguarda as devidas explicações): o que significa, geometricamente, termos um determinante negativo? Ou, de um outro ponto de vista, o que distingue os ternos ordenados com determinante negativo daqueles para os quais o determinante é positivo.

Como já sabemos, $det(\vec{u},\vec{v},\vec{w}) \neq 0$ se e somente se $\vec{u}$, $\vec{v}$ e $\vec{w}$ são linearmente independentes (e, portanto, constituem base para V). Vamos estabelecer o conceito de orientação.

Definição: Duas bases ordenadas $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ e $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ têm a mesma **orientação** se existem três funções contínuas $\vec{\varepsilon}_1,\vec{\varepsilon}_2,\vec{\varepsilon}_3 : [a,b] \to V$ (com a e b reais, $a < b$) tais que:

$$\begin{array}{rl} (i) & \vec{\varepsilon}_i(a) = \vec{u}_i,\ i = 1,2,3 \\ (ii) & \vec{\varepsilon}_i(b) = \vec{v}_i,\ i = 1,2,3 \\ (iii) & \vec{\varepsilon}_1(t),\vec{\varepsilon}_2(t),\vec{\varepsilon}_3(t) \text{ são linearmente independentes } \forall t \in [a,b] \end{array}$$

Chamaremos de **deformação** entre as bases $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ e $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ um terno de funções contínuas $\vec{\varepsilon}_1,\vec{\varepsilon}_2,\vec{\varepsilon}_3 : [a,b] \to V$ com as três propriedades acima.

Observação: Entenda que duas bases ordenadas têm a mesma orientação se e somente se podemos *deformar* uma na outra de maneira que, durante a deformação, os três vetores permaneçam linearmente independentes o tempo todo. De maneira ainda mais visual, as bases $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ e $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ têm a mesma orientação se o paralelepípedo formado por $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ pode ser deformado no formado por $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ sem, em qualquer momento, degenerar (entendido que a deformação deve levar cada $\vec{u}_i$ no respectivo $\vec{v}_i$).

Observação: Note que ter a mesma orientação é uma relação de equivalência: uma base, ficando parada, se deforma em si mesma; se uma base $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ se deforma na base $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ por meio de $\vec{\varepsilon}_1,\vec{\varepsilon}_2,\vec{\varepsilon}_3 : [a,b] \to V$, podemos usar $\vec{\delta}_1,\vec{\delta}_2,\vec{\delta}_3 : [-b,-a] \to V$, dadas por $\vec{\delta}_i(t) = \vec{\varepsilon}_i(-t)$, para deformar $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ em $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$; finalmente, se a base $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ se deforma na base $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ e a base $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ se deforma na base $(\vec{w}_1,\vec{w}_2,\vec{w}_3)$, então podemos deformar $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ em $(\vec{w}_1,\vec{w}_2,\vec{w}_3)$ *passando por* $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$.

Exercício 9.21 *Explique como podemos deformar* $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ *em* $(\vec{w}_1,\vec{w}_2,\vec{w}_3)$ *passando por* $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$. *Isto é, construa, a partir das funções que deformam* $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ *em* $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ *e das que deformam* $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ *em* $(\vec{w}_1,\vec{w}_2,\vec{w}_3)$, *novas funções, que deformem* $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ *em* $(\vec{w}_1,\vec{w}_2,\vec{w}_3)$.

Exercício 9.22 *Suponha que as bases ordenadas* $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ *e* $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ *têm a mesma orientação. Mostre que as bases ordenadas* $(\vec{u}_1,\vec{u}_2,-\vec{u}_3)$ *e* $(\vec{v}_1,\vec{v}_2,-\vec{v}_3)$ *têm a mesma orientação.*

Proposição: Se duas bases ordenadas $(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ e $(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ têm a mesma orientação, então $det(\vec{u}_1,\vec{u}_2,\vec{u}_3)$ e $det(\vec{v}_1,\vec{v}_2,\vec{v}_3)$ têm o mesmo sinal.

Demonstração: Pensemos $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ e $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ dados em coordenadas, assim como as funções contínuas $\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3 : [a,b] \to V$ tais que:

$$\begin{array}{rl} (i) & \vec{\varepsilon}_i(0) = \vec{u}_i,\ i = 1,2,3 \\ (ii) & \vec{\varepsilon}_i(1) = \vec{v}_i,\ i = 1,2,3 \\ (iii) & \vec{\varepsilon}_1(t), \vec{\varepsilon}_2(t), \vec{\varepsilon}_3(t) \textit{ são linearmente independentes } \forall t \in [a,b]. \end{array}$$

Temos que a continuidade das $\vec{\varepsilon}_i$ se traduz na continuidade das $a_{ij}(t)$, que são as entradas da matriz dada pelas coordenadas dos $\vec{\varepsilon}_i(t)$:

$$\begin{pmatrix} a_{11}(t) & a_{12}(t) & a_{13}(t) \\ a_{21}(t) & a_{22}(t) & a_{23}(t) \\ a_{31}(t) & a_{32}(t) & a_{33}(t) \end{pmatrix}.$$

Mas, então, a função $f : [a,b] \to \mathbb{R}$, dada por

$$f(t) = det(\vec{\varepsilon}_1(t), \vec{\varepsilon}_2(t), \vec{\varepsilon}_3(t)) = a_{11}(t)a_{22}(t)a_{33}(t) - a_{11}(t)a_{23}(t)a_{32}(t) +$$
$$+ a_{12}(t)a_{23}(t)a_{31}(t) - a_{12}(t)a_{21}(t)a_{33}(t) + a_{13}(t)a_{21}(t)a_{32}(t) - a_{13}(t)a_{22}(t)a_{31}(t),$$

é contínua e não se anula (se se anulasse em um certo t, $\vec{\varepsilon}_1(t), \vec{\varepsilon}_2(t), \vec{\varepsilon}_3(t)$ deixariam de ser linearmente independentes). Logo, $f(a) = det(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ e $f(b) = det(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ não podem ter sinais diferentes. ■

Questão: Será verdadeira a recíproca? Isto é, supondo que $det(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ e $det(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ tenham o mesmo sinal, então será que as bases ordenadas $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ e $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ têm necessariamente a mesma orientação? Neste caso, como construir a deformação entre $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ e $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$?

A resposta é sim. Vamos esboçar a construção de uma deformação entre duas bases ordenadas com determinante de mesmo sinal. Vamos proceder em duas etapas: primeiro, mostramos que toda base ortonormal com determinante positivo tem a mesma orientação que a base canônica (e, consequentemente, toda base ortonormal com determinante negativo tem a mesma orientação que a base $(\vec{e}_1, \vec{e}_2, -\vec{e}_3)$); em seguida, mostramos que toda base tem a mesma orientação que uma base ortonormal (esta, construída pelo processo de Gram-Schmidt).

Teorema: Duas bases ordenadas $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ e $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ têm a mesma orientação se, e somente se, $det(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ e $det(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ têm o mesmo sinal.

Demonstração: Já provamos, na proposição acima, que bases que têm a mesma orientação têm determinantes com o mesmo sinal. Para cuidar da outra parte, procederemos, conforme anunciado, em duas etapas. Vamos, na verdade, abusar um pouco da intuição, para não fazer boa parte do trabalho sujo...

Etapa 1: Suponhamos que $(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3)$ seja uma base ortonormal de V, com determinante positivo. Vamos mostrar como deformá-la na base canônica. Vamos fazer duas rotações do espaço, dando como intuitivo que, quando rodamos o espaço todo em torno de um eixo, bases ortonormais continuam, durante todo o tempo da rotação, bases ortonormais (já que rotações preservam a ortogonalidade e as distâncias). A primeira rotação, supondo que $\vec{\varepsilon}_1 \neq \vec{e}_1$, destina-se a levar $\vec{\varepsilon}_1$ até $\vec{e}_1$. Tomamos, como eixo de rotação, a reta que passa pela origem e é perpendicular ao plano formado por $\vec{\varepsilon}_1$ e $\vec{e}_1$; rodamos até que $\vec{\varepsilon}_1$ coincida com $\vec{e}_1$. No fim dessa rotação $\vec{\varepsilon}_1$ coincide com $\vec{e}_1$; $\vec{\varepsilon}_2$ vai parar em algum outro lugar. Se o novo $\vec{\varepsilon}_2$ já coincidir com $\vec{e}_2$, ótimo; caso contrário, temos que o novo $\vec{\varepsilon}_2$ e $\vec{e}_2$ formam um plano perpendicular a $\vec{e}_1 = \vec{\varepsilon}_1$. Fazemos, então uma nova rotação, em torno da reta que passa pela origem e tem a direção de $\vec{e}_1$, até que o novo $\vec{\varepsilon}_2$ coincida com $\vec{e}_2$.

Chegamos, até aqui, à seguinte situação: temos novos $\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3$, com $\vec{\varepsilon}_1 = \vec{e}_1$ e $\vec{\varepsilon}_2 = \vec{e}_2$; e o novo $\vec{\varepsilon}_3$, onde foi parar? Ora, como rotações não alteram ortogonalidade nem distâncias, o novo $\vec{\varepsilon}_3$ é unitário e ortogonal a $\vec{\varepsilon}_1 = \vec{e}_1$ e a $\vec{\varepsilon}_2 = \vec{e}_2$. Ora, só existem dois vetores nestas condições: $\vec{e}_3$ e $-\vec{e}_3$ (provar isto é um bom exercício!). Mas temos uma informação a mais: o sinal do determinante, durante a deformação, não mudou de sinal (já que $(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3)$ se manteve, o tempo todo, uma base ortonormal); logo, o novo $\vec{\varepsilon}_3$ não pode ser $-\vec{e}_3$.

Para concluir a etapa 1; observamos que, caso $(\vec{\delta}_1, \vec{\delta}_2, \vec{\delta}_3)$ seja uma base ortonormal de V, com determinante negativo, então $(\vec{\delta}_1, \vec{\delta}_2 -, \vec{\delta}_3)$ é base ortonormal de V, com determinante positivo. Como acabamos de provar, existe uma deformação, dada por três funções, $\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3 : [a, b] \to V$, começando em $(\vec{\delta}_1, \vec{\delta}_2 -, \vec{\delta}_3)$ e terminando em $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$. Trocando a função $\vec{\varepsilon}_3(t)$ por $-\vec{\varepsilon}_3(t)$, teremos uma deformação de $(\vec{\delta}_1, \vec{\delta}_2, \vec{\delta}_3)$ em $(\vec{e}_1, \vec{e}_2, -\vec{e}_3)$.

Etapa 2: Suponhamos, agora, que $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ é uma base ordenada qualquer. Vamos deformá-la em uma base ortonormal, usando o processo de Gram-Schmidt. Começamos *encolhendo*, ou *esticando* $\vec{u}_1$ até virar $\vec{\varepsilon}_1 = |\vec{u}|^{-1}\vec{u}$ (se $|\vec{u}| > 1$, por exemplo, por meio de $t \longmapsto -t\vec{u}$, $t \in [-1, -|\vec{u}|^{-1}]$). Em seguida, movemos $\vec{u}_2$ até $\vec{v}_2 = \vec{u}_2 - < \vec{u}_2, \vec{\varepsilon}_1 > \vec{\varepsilon}_1$, por meio de $t \longmapsto \vec{u}_2 - t < \vec{u}_2, \vec{\varepsilon}_1 > \vec{\varepsilon}_1$, $t \in [0, 1]$. A seguir, *encolhemos*, ou *esticamos*, conforme o caso, $\vec{v}_2$ até virar $\vec{\varepsilon}_2 = |\vec{v}_2|^{-1}\vec{v}_2$. Agora, tomamos a projeção de $\vec{u}_3$ sobre o plano gerado por $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$, dada por $\vec{u}_0 =< \vec{u}_3, \vec{\varepsilon}_1 > \vec{\varepsilon}_1 + < \vec{u}_3, \vec{\varepsilon}_2 > \vec{\varepsilon}_2$, e movemos $\vec{u}_3$ até $\vec{v}_3 = \vec{u}_3 - \vec{u}_0$, por meio de $t \longmapsto \vec{u}_3 - t\vec{u}_0$, $t \in [0, 1]$. Finalmente, *encolhemos*, ou *esticamos*, conforme o caso, $\vec{v}_3$ até virar $\vec{\varepsilon}_3 = |\vec{v}_3|^{-1}\vec{v}_3$.

Assim, provamos que toda base ordenada com determinante positivo tem a mesma orientação que a base canônica. E toda base ordenada com determinante negativo tem a mesma orientação que a base $(\vec{e}_1, \vec{e}_2, -\vec{e}_3)$. Como já vimos, duas bases que têm a mesma orientação que uma terceira têm a mesma orientação. O teorema está provado. ■

Definição: Uma base ordenada $(\vec{u}, \vec{v}, \vec{w})$ tem **orientação positiva** (negativa) se $det(\vec{u}, \vec{v}, \vec{w}) > 0$ (< 0). Neste caso, diz-se também que $(\vec{u}, \vec{v}, \vec{w})$ é **positivamente orientada** (negativamente orientada).

Capítulo 10

Quatérnions e produto vetorial

10.1 Os quatérnions

Os números complexos começam a entrar na Matemática no século XVI, com a manipulação formal de raízes quadradas de números negativos, dentro do processo de cálculo de raízes de polinômios do 3º grau. É apenas na virada do século XVIII para o XIX, porém, que passam a ser identificados com pontos do plano e interpretados geometricamente. Um dos maiores entusiastas da identificação de $\mathbb{C}$ com $\mathbb{R}^2$ foi o matemático e físico irlandês William Rowan Hamilton (Hamilton publicou sua versão da identificação entre $\mathbb{C}$ e $\mathbb{R}^2$ em 1833). Sua obsessão, naturalmente, já que $\mathbb{R}$ se identifica à reta e $\mathbb{C}$ ao plano, passou a ser estender os números complexos a um conjunto numérico que pudesse ser identificado ao espaço $\mathbb{R}^3$ (o que, na verdade, é impossível).[1] Em 1843, finalmente, durante uma caminhada com sua senhora, Hamilton, subitamente,teve a ideia de passar diretamente para um espaço de dimensão 4.

Hamilton introduziu dois novos números, j e k, réplicas do imaginário i. Os **quatérnions** surgem, então, como combinações lineares dos números 1, i, j e k, ou seja, são números da forma

$$t + xi + yj + zk,$$

com t, x, y e z reais. Hamilton utilizava, originalmente, a letra w no lugar de t. Usar t, porém, não trai a memória do irlandês, que já em 1828, publicara um trabalho em que falava da *indissociável conexão entre espaço e tempo*.

10.2 O produto escalar e o produto vetorial

A adição de quatérnions se faz da maneira mais natural:

[1]Embora o primeiro trabalho com a interpretação geométrica dos números complexos, devido a Caspar Wessel, tenha sido publicado em 1799 e já abordasse a questão de obter algo semelhante para o espaço, a ideia só *pegou* depois que Gauss publicou a sua versão, em 1831. Gauss, na verdade, já tinha a versão geométrica de $\mathbb{C}$ desde 1799 e, também tentara achar entidades numéricas que correspondessem ao espaço. O mesmo haviam feito Argand, em 1806, e Mourey, em 1828. Ao que tudo indica, nenhum sabia dos trabalhos dos demais, nem mesmo Hamilton, que só tomou conhecimento do trabalho de Gauss em 1852. Mas a ideia estava no ar...

$$(t_1 + x_1 i + y_1 j + z_1 k) + (t_2 + x_2 i + y_2 j + z_2 k) =$$

$$= (t_1 + t_2) + (x_1 + x_2)i + (y_1 + y_2)j + (z_1 + z_2)k.$$

Já a multiplicação, associativa e distributiva em relação à adição, fica definida pela fórmula que, entusiasmado, o próprio Hamilton gravou com o canivete em uma pedra da Brougham Bridge no instante seguinte à descoberta:[2]

$$i^2 = j^2 = k^2 = ijk = -1.$$

Exercício 10.1 *Deduza, da fórmula acima, que*

$$ij = k = -ji, \;\; jk = i = -kj, \;\; ki = j = -ik.$$

Assim, a multiplicação de quatérnions não é comutativa (propor números com multiplicação não comutativa, para a época, era uma ousadia).

Os quatérnions tiveram, como subproduto, um papel decisivo na consolidação da ideia de vetor: de fato, o próprio Hamilton introduz os termos **escalar** e **vetor** para designar, respectivamente, a parte real (t) e a parte imaginária ($xi + yj + zk$) de um quatérnion (observe que, para todos os efeitos, os quatérnions i, j e k correspondem aos vetores $\vec{e_1}$, $\vec{e_2}$ e $\vec{e_3}$, que compõem a base canônica do espaço V). Isso tem consequências interessantíssimas. Se multiplicarmos quatérnions sem parte escalar, ou seja, dois vetores, $x_1 i + y_1 j + z_1 k$ e $x_2 i + y_2 j + z_2 k$, obtemos (faça as contas):

$$-(x_1x_2 + y_1y_2 + z_1z_2) + (y_1z_2 - z_1y_2)i + (z_1x_2 - x_1z_2)j + (y_1z_2 - z_1y_2)k.$$

Ou seja, o produto (como quatérnions) de dois vetores é um quatérnion que tem uma **parte escalar**:

$$-(x_1x_2 + y_1y_2 + z_1z_2)$$

e uma **parte vetorial**:

$$(y_1z_2 - z_1y_2)i + (z_1x_2 - x_1z_2)j + (x_1y_2 - y_1x_2)k.$$

A primeira resulta ter, como já sabemos, uma importante interpretação geométrica, pois, a menos do sinal, é o nosso já estudado **produto escalar**.[3] Ocupemo-nos da segunda, que recebe, naturalmente, o nome de **produto vetorial**.

[2]Esta história foi relatada pelo próprio Hamilton, anos depois, em uma carta a seu filho Archibald

[3]Não deixa de ser admirável o fato de que este produto, tão geométrico, tenha nascido de devaneios puramente algébricos. Mais admirável ainda é o fato de que, em se tratando de história do produto escalar, ainda não contamos, nem vamos contar neste texto, da missa a metade

Definição: Dados os vetores $\vec{u} = x_1\vec{e_1} + y_1\vec{e_2} + z_1\vec{e_3}$ e $\vec{v} = x_2\vec{e_1} + y_2\vec{e_2} + z_2\vec{e_3}$, seu **produto vetorial**, designado por $\vec{u} \otimes \vec{v}$ ou por $\vec{u} \times \vec{v}$, é definido por

$$\vec{u} \otimes \vec{v} = (y_1z_2 - z_1y_2)\vec{e_1} + (z_1x_2 - x_1z_2)\vec{e_2} + (x_1y_2 - y_1x_2)\vec{e_3}.$$

Observação: por conta da evidente identificação entre os quatérnions i, j e k e os vetores $\vec{e_1}$, $\vec{e_2}$ e $\vec{e_3}$, ainda hoje é usual designar os vetores da base canônica de V por i, j e k. Também o faremos, ocasionalmente. Neste sentido, é particularmente útil a apresentação do produto vetorial de $\vec{u} = x_1i + y_1j + z_1k$ e $\vec{v} = x_2i + y_2j + z_2k$ como um determinante:

$$\begin{vmatrix} i & j & k \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix}$$

Exercício 10.2 *Verifique que o determinante acima é, de fato, igual ao produto vetorial (observe que a fórmula do determinante pode, de fato, ser aplicada não apenas a matrizes cujas entradas sejam números reais, mas também àquelas em que as entradas sejam complexos, quatérnions, ou mesmo outras coisas que se multipliquem, se somem e se subtraiam).*

Exercício 10.3 *Prove que o produto vetorial tem as seguintes propriedades:*

$$\begin{array}{ll} (i) & \vec{u} \otimes (t\vec{v}_1 + \vec{v}_2) = t(\vec{u} \otimes \vec{v}_1) + (\vec{u} \otimes \vec{v}_2) \ \forall\, t, \vec{u}, \vec{v}_1, \vec{v}_2 \\ (ii) & \vec{u} \otimes \vec{v} = -\vec{v} \otimes \vec{u} \ \forall\, \vec{u}, \vec{v} \end{array}$$

Exercício 10.4 *Conclua, do exercício acima, que $\vec{u} \otimes \vec{u} = \vec{0} \ \forall\, \vec{u}$ e que*

$$(t\vec{u}_1 + \vec{u}_2) \otimes \vec{v} = t(\vec{u}_1 \otimes \vec{v}) + (\vec{u}_2 \otimes \vec{v}) \ \forall\, t, \vec{u}_1, \vec{u}_2, \vec{v}.$$

Exercício 10.5 *Mostre que $\vec{u} \otimes \vec{v} = \vec{0}$ se, e somente se, $\vec{u}$ e $\vec{v}$ são linearmente dependentes.*

Exercício 10.6 *Mostre, com um exemplo, que o produto vetorial não é associativo.*

Exercício 10.7 *Calcule, dados $\vec{u}, \vec{v}$ e $\vec{w}$,*

$$(\vec{u} \otimes \vec{v}) \otimes \vec{w} + (\vec{v} \otimes \vec{w}) \otimes \vec{u} + (\vec{w} \otimes \vec{u}) \otimes \vec{v}.$$

10.3 Interpretação geométrica do produto vetorial

A primeira propriedade geométrica do produto vetorial a ser destacada é: **$\vec{u} \otimes \vec{v}$ é perpendicular a $\vec{u}$ e a $\vec{v}$**, ou seja,

$$< \vec{u}, \vec{u} \otimes \vec{v} >= 0 =< \vec{v}, \vec{u} \otimes \vec{v} > .$$

A demonstração é uma simples conta, mas fica mais sugestiva se observarmos que, se $\vec{u} = x_1 i + y_1 j + z_1 k$, $\vec{v} = x_2 i + y_2 j + z_2 k$ e $\vec{w} = xi + yj + zk$, então

$$< \vec{w}, \vec{u} \otimes \vec{v} >= \begin{vmatrix} x & y & z \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix}.$$

Exercício 10.8 *Conclua, da identidade acima, que*

$$0 \leq < \vec{u} \otimes \vec{v}, \vec{u} \otimes \vec{v} >= det(\vec{u} \otimes \vec{v}, \vec{u}, \vec{v}).$$

Assim, como $\vec{u} \otimes \vec{v} = \vec{0}$ se e somente se $\vec{u}$ e $\vec{v}$ são linearmente dependentes, temos que, se $\vec{u}$ e $\vec{v}$ são linearmente independentes, então $det(\vec{u} \otimes \vec{v}, \vec{u}, \vec{v}) > 0$ e, portanto, $(\vec{u} \otimes \vec{v}, \vec{u}, \vec{v})$ é uma base ordenada para V, com a mesma orientação que a base canônica.

Cuidemos, agora, de uma propriedade menos evidente: **o comprimento (norma) de $\vec{u} \otimes \vec{v}$ é igual à área do paralelogramo formado por $\vec{u}$ e $\vec{v}$**. Ou seja: se θ é o ângulo entre $\vec{u}$ e $\vec{v}$, então

$$|\vec{u} \otimes \vec{v}| = |\vec{u}||\vec{v}| \sin \theta.$$

A demonstração lançará mão de dois resultados geométricos que enunciamos e demonstramos a seguir.

Proposição: Sejam α e β planos que se cortam segundo o ângulo θ. Se F é uma figura em β e $\bar{F}$ sua projeção ortogonal sobre α, então

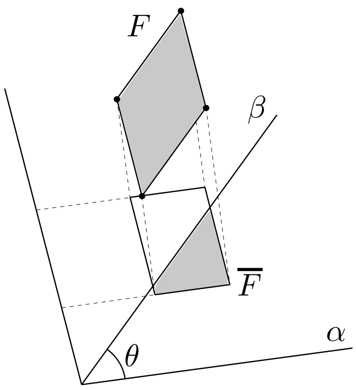
$$\acute{a}rea de \bar{F} = \cos \theta \; \acute{a}rea de F.$$

Demonstração: Podemos desconsiderar os casos $\theta = 0$ e $\theta = \pi/2$, em que o resultado é evidente. Chamemos de r a reta interseção entre α e β.

Para obter a área de F, ladrilhemos o plano β com retângulos de lados paralelos ou perpendiculares a r. Projetando-os ortogonalmente a α, obtemos um ladrilhamento de α com retângulos também de lados paralelos ou perpendiculares a r. Mas se R é um dos ladrilhos de β e $\bar{R}$ é sua projeção em α, o lado de R paralelo a r tem seu comprimento preservado pela projeção, enquanto o lado perpendicular a r, ao ser projetado, tem seu comprimento multiplicado por $\cos \theta$. Logo, a área de $\bar{R}$ é igual à de R multiplicada por $\cos \theta$. Fazendo aproximações por dentro e por fora de F com ladrilhos como acima e observando as correspondentes aproximações de $\bar{F}$, obtemos o resultado por passagem ao limite. ■

O segundo resultado é uma interessante generalização do Teorema de Pitágoras.

Teorema: Considere um sistema de três planos, α_1, α_2 e α_3, dois a dois ortogonais. Dados um quarto plano β e uma figura F em β, sejam F_1, F_2 e F_3 as projeções ortogonais de F sobre, respectivamente, α_1, α_2 e α_3. Então as áreas das figuras, dadas por $|F|$, $|F_1|$, $|F_2|$ e $|F_3|$, satisfazem a relação:



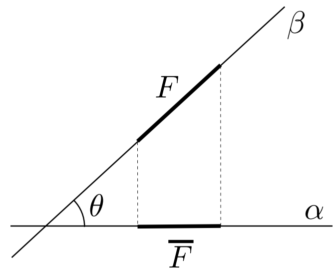


Figura 10.1:

$$|F|^2 = |F_1|^2 + |\ F_2|^2 + |F_3|^2.$$

Demonstração: Podemos supor que os planos α_i são normais aos vetores $\vec{e}_i$, de forma que, se chamarmos de θ_i o ângulo entre α_i e β e de $\vec{n}$ um vetor unitário e normal a β, teremos θ_i igual ao ângulo entre $\vec{e}_i$ e $\vec{n}$ (o ângulo entre dois planos é igual ao ângulo entre seus vetores normais[4]). Assim, se $\vec{n}$ é dado, na base canônica, por $\vec{n} = n_1\vec{e_1} + n_2\vec{e_2} + n_3\vec{e_3}$, temos

$$\cos\theta_i = |< \vec{n}, \vec{e}_i >| = n_i.$$

Mas, então,

$$|F_i|^2 = (|F|\cos\theta_i)^2 = |F|^2 n_i^2.$$

Logo,

$$|F_1|^2 + |\ F_2|^2 + |F_3|^2 = |F|^2(n_1^2 + n_2^2 + n_3^2) = |F|^2,$$

já que $\vec{n}$ é unitário. ■

Consideremos agora os vetores $\vec{u}$ e $\vec{v}$, dados, em coordenadas, por $\vec{u} = (x_1, y_1, z_1)$ e $\vec{v} = (x_2, y_2, z_2)$. Vamos provar que a norma de $\vec{u} \otimes \vec{v}$ é igual à área do paralelogramo F formado por $\vec{u}$ e $\vec{v}$, examinando as projeções de F nos planos coordenados e aplicando o teorema que acabamos de demonstrar. Chamaremos de α_i o plano coordenado normal ao vetor $\vec{e}_i$; $\vec{u}_i$ e $\vec{v}_i$ serão as projeções respectivas de $\vec{u}$ e $\vec{v}$ sobre α_i. Em coordenadas, teremos (faça os desenhos, se precisar):

[4] há, na realidade "dois" ângulos possíveis entre os normais, mas seus cossenos diferem apenas pelo sinal

$$\begin{array}{lll} \vec{u}_1 = (y_1, z_1) & \vec{u}_2 = (z_1, x_1) & \vec{u}_3 = (x_1, y_1) \\ \vec{v}_1 = (y_2, z_2) & \vec{v}_2 = (z_2, x_2) & \vec{v}_3 = (x_2, y_2) \end{array}$$

Chamando de F_i a projeção ortogonal de F sobre α_i, temos que F_i é o paralelogramo formado por $\vec{u}_i$ e $\vec{v}_i$. Ora, pelo que aprendemos sobre determinantes de matrizes 2×2, a área de F_i é o módulo do determinante de $\vec{u}_i$ e $\vec{v}_i$, e esse determinante é a i-ésima coordenada de $\vec{u} \otimes \vec{v}$, ou seja:

$$\vec{u} \otimes \vec{v} = (y_1z_2 - z_1y_2)\vec{e_1} + (z_1x_2 - x_1z_2)\vec{e_2} + (x_1y_2 - y_1x_2)\vec{e_3},$$

e

$$\begin{aligned} |F_1| &= |y_1z_2 - z_1y_2|, \\ |F_1| &= |z_1x_2 - x_1z_2|, \\ |F_1| &= |x_1y_2 - y_1x_2| \end{aligned}$$

Aplicando o teorema, temos imediatamente

$$|F|^2 = |F_1|^2 + |\ F_2|^2 + |F_3|^2 = |\vec{u} \otimes \vec{v}|^2,$$

o que completa a prova

10.4 Determinantes e volume, de novo

No capítulo referente ao determinante, usamos, como ponto de partida, a busca de algo que correspondesse ao **volume** do paralelepídedo p gerado pelos vetores $\vec{u}$, $\vec{v}$ e $\vec{w}$. Naquela ocasião, demos como óbvio que o volume com sinal deveria ser uma forma trilinear alternada em $\vec{u}$, $\vec{v}$ e $\vec{w}$.[5] O leitor, com razão, pode não ter achado tão óbvio assim. Voltemos, pois, ao assunto, agora fortalecidos por nossos conhecimentos sobre o produto vetorial. Como já sabemos que a área do paralelogramo formado por $\vec{v}$ e $\vec{w}$ é a norma de $\vec{v} \otimes \vec{w}$, podemos obter o volume de p multiplicando $|\vec{v} \otimes \vec{w}|$ pela projeção de $\vec{u}$ na direção normal ao plano α gerado por $\vec{v}$ e $\vec{w}$.

Ora, como já vimos, $\vec{v} \otimes \vec{w}$ é normal ao plano α. Basta, pois, tomar

$$< \vec{u}, \vec{v} \otimes \vec{w} >,$$

observando que o sinal será positivo se $\vec{u}$ e $\vec{v} \otimes \vec{w}$ estiverem do mesmo lado do plano α, negativo se estiverem em lados opostos. Ora, é fácil ver que, no primeiro caso, a base $(\vec{u}, \vec{v}, \vec{w})$ tem orientação positiva e, no segundo, tem orientação negativa. Também é claro que o volume de p, dado por $\vec{u} \cdot \vec{v} \otimes \vec{w}$ (que costuma ser chamado de **produto misto** de $\vec{u}$ e $\vec{v} \otimes \vec{w}$), é dado, se $\vec{u} = (x_1, y_1, z_1)$, $\vec{v} = (x_2, y_2, z_2)$ e $\vec{w} = (x_3, y_3, z_3)$, pelo determinante

[5] óbvio, pelo menos, a partir de nossa experiência com o determinante de dois vetores, no plano

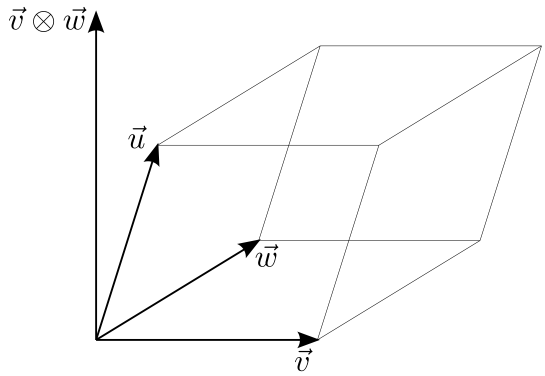


Figura 10.2:

$$\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}.$$

Agora sim, podemos acreditar que

$$volume\ de\ p\ = < \vec{u}, \vec{v} \otimes \vec{w} > =\ det(\vec{u}, \vec{v}, \vec{w}).$$

Exercício 10.9 *Suponha que $\vec{u}$ e $\vec{v} \otimes \vec{w}$ estão do mesmo lado em relação ao plano α. Construa uma deformação entre as bases ordenadas $(\vec{u}, \vec{v}, \vec{w})$ e $(\vec{v} \otimes \vec{w}, \vec{v}, \vec{w})$.*

10.5 O módulo

Dado um quatérnion $q = t + xi + yj + zk$, seu **módulo** é, naturalmente, definido por

$$|q| = \sqrt{t^2 + x^2 + y^2 + z^2}.$$

Um quatérnion de módulo igual a 1 é dito unitário. Chamando $xi + yj + zk$ de $\vec{v}$ e definindo o **conjugado** de $q = t + \vec{v}$ por $q' = t - \vec{v}$, temos

$$|q|^2 = t^2 + |\vec{v}|^2 = qq'.$$

Daí decorre que todo quatérnion q não nulo tem por inverso multiplicativo o quatérnion q^{-1} dado por

$$q^{-1} = \frac{q'}{|q|^2}.$$

Menos evidente é a **desigualdade triangular**, $|q_1 + q_2| \leq |q_1| + |q_2|$, que decorre de ser $\langle q_1, q_2\rangle = t_1t_2 + \langle v_1, v_2\rangle$ um produto escalar e da desigualdade de Cauchy-Schwarz-Buniacóvsqui (aqui, $q_1 = t_1 + v_1$, $q_2 = t_2 + v_2$ e $\langle v_1, v_2\rangle$ é o produto escalar usual em $\mathbb{R}^3$). Mas a propriedade mais emocionante é dada na Proposição a seguir.

Proposição: Dados dois quatérnions, q_1 e q_2, vale

$$|q_1q_2| = |q_1||q_2|.$$

Demonstração: Escrevendo $q_1 = r + \vec{u}$ e $q_2 = s + \vec{v}$, temos:

$$\begin{aligned}
|q_1q_2|^2 &= |rs + r\vec{v} + s\vec{u} - <\vec{u}, \vec{v}> + \vec{u} \otimes \vec{v}|^2 = \\
&= |rs - <\vec{u}, \vec{v}>|^2 + |r\vec{v} + s\vec{u}|^2 + |\vec{u} \otimes \vec{v}|^2 = \\
&= (rs)^2 - 2rs <\vec{u}, \vec{v}> + <r\vec{v} + s\vec{u}, r\vec{v} + s\vec{u}> + <\vec{u}, \vec{v}>^2 + |\vec{u} \otimes \vec{v}|^2 = \\
&= r^2s^2 + r^2|\vec{v}|^2 + s^2|\vec{u}|^2 + |\vec{u}|^2|\vec{v}|^2 = |q_1|^2|q_2|^2.
\end{aligned}$$

■

Exercício 10.10 *Mostre que, dados dois quatérnions, q_1 e q_2, vale $(q_1q_2)' = q_2'q_1'$.*

10.6 Quatérnions e rotações

Consideremos o problema de rodar os pontos do espaço, de um ângulo θ, em torno do eixo passando pela origem e dado pelo vetor $\vec{u}$. Vamos supor, para facilitar as contas, que $\vec{u}$ é unitário. Suporemos, também, que, "visto de $\vec{u}$", o plano α, perpendicular a $\vec{u}$ e passando pela origem, roda no sentido trigonométrico. Chamaremos de R a transformação que associa, a cada vetor $\vec{v}$, seu rodado.

Seja $\vec{v}$ um vetor qualquer. Comecemos decompondo $\vec{v}$ como soma de um vetor $\vec{v}_0$, na direção de $\vec{u}$, e outro, $\vec{w}$, perpendicular a $\vec{u}$, isto é, em α. Usando o produto escalar, temos:

$$\vec{v}_0 = <\vec{v}, \vec{u}> \vec{u}.$$

Exercício 10.11 *Verifique que $R\vec{v} = \vec{v}_0 + R\vec{w}$.*

Agora, vamos usar o produto vetorial de maneira um pouco menos evidente:

$$\vec{u} \otimes \vec{v} = \vec{u} \otimes (\vec{w} + \vec{v}_0) = \vec{u} \otimes \vec{w}.$$

Chamando $\vec{u} \otimes \vec{v}$ de $\vec{w}^\perp$, temos $\vec{w} = \vec{w}^\perp \otimes \vec{u}$ e $R\vec{w} = \cos\theta\vec{w} + \sin\theta\vec{w}^\perp$. Assim,

$$R\vec{v} = <\vec{v}, \vec{u}> \vec{u} + \cos\theta\ (\vec{u} \otimes \vec{v}) \otimes \vec{u} + \sin\theta\ \vec{u} \otimes \vec{v}.$$

Exercício 10.12 *Confira.*

Vejamos, agora, como o mesmo resultado pode ser obtido com quatérnions. Considere os quatérnions

$$q = \cos\frac{\theta}{2} + \sin\frac{\theta}{2}\,\vec{u}, \;\; q' = \cos\frac{\theta}{2} - \sin\frac{\theta}{2}\,\vec{u}.$$

Exercício 10.13 *Entenda isso (pense o vetor $\vec{u}$ como $\vec{u} = ai + bj + zk$). Mostre que $qq' = q'q = 1$.*

Exercício 10.14 *Pense o vetor $\vec{v}$ como um quatérnion ($\vec{v} = xi + yj + zk$). Faça as contas (não é preciso usar coordenadas) e mostre que*

$$R\vec{v} = q\vec{v}q'.$$

Assim, rotações são reduzidas a multiplicações por quatérnions unitários. Não chega a ser grande coisa, mas simplifica um pouco a vida, quando se trata de compor rotações. De fato, se q_1 e q_2 correspondem às rotações R_1 e R_2, dadas, respectivamente, por ângulos θ_1 e θ_2 e eixos $\vec{u}_1$ e $\vec{u}_2$, então a composta $R = R_1R_2$ é dada por

$$R\vec{v} = R_1(R_2\vec{v}) = q_1(q_2\vec{v}q_2')q_1' = q\vec{v}q',$$

com $q = q_1q_2$. Como o produto de quatérnions unitários é um quatérnion unitário, segue um resultado não tão evidente.

Proposição: A composta de duas rotações em torno de eixos passando por um mesmo ponto O é uma rotação em torno de eixo passando por O.

Demonstração: ■

Capítulo 11

Transformações lineares

11.1 Em três dimensões

Ocupemo-nos, agora, das **transformações**, isto é, de processos pelos quais um objeto se torna outro objeto.

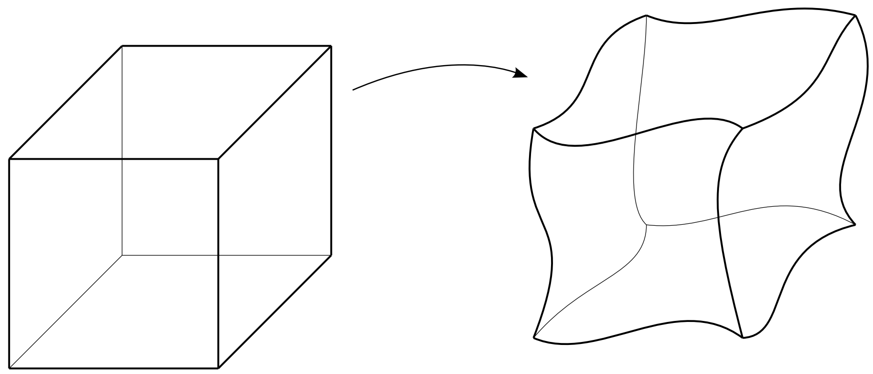


Figura 11.1:

As possibilidades são, porém, amplas demais. Modestamente, cuidemos de um tipo particularmente simples de transformação: aquele em que as combinações lineares são preservadas. Mais rigorosamente, identificando pontos e vetores, vejamos o que conseguimos dizer sobre funções

$$T : V \longrightarrow V$$

tais que

$$T(s\vec{u} + t\vec{v}) = sT\vec{u} + tT\vec{v}.$$

Por enquanto, V é o nosso velho e bom espaço dos vetores flechinhas, que, eventualmente, identificamos a $\mathbb{R}^3$.

Definição: Uma aplicação $T : V \to V$ tal que, para quaisquer vetores $\vec{u}$ e $\vec{v}$, e para qualquer escalar t, vale $T(\vec{u} + t\,\vec{v}) = T\vec{u} + t\,T\vec{v}$ é dita uma **transformação linear**.

Exercício 11.1 *Mostre que $T : V \to V$ é uma transformação linear se, e somente se, valem as duas propriedades:*

$$\begin{array}{cl} (i) & T(t\vec{u}) = tT\vec{u}\;\forall\,\vec{u} \in V,\ t \in \mathbb{R} \\ (ii) & T(\vec{u}+\vec{v}) = T\vec{u} + T\vec{v}\;\forall\,\vec{u},\vec{v} \in V. \end{array}$$

Prove primeiro que $T(\vec{u}+\vec{v}) = T\vec{u} + T\vec{v}$, fazendo $t = 1$; depois prove que $T\vec{0} = \vec{0}$ (veja abaixo); em seguida, para $T(t\vec{u})$, faça $T(t\vec{u}) = T(\vec{0} + t\vec{u}) = T\vec{0} + tT\vec{u}$

Exercício 11.2 *Mostre que, se $T : V \to V$ é uma transformação linear, então $T\vec{0} = \vec{0}$ e $T(-\vec{v}) = -T\vec{v}\;\forall\,\vec{v} \in V$.*

$T\vec{0} = T(\vec{0}+\vec{0}) = T\vec{0} + T\vec{0})$

Exercício 11.3 *Mostre que, se $T : V \to V$ é uma transformação linear, $\vec{u}$ e $\vec{v}$ são dois vetores fixos, r é a reta dada por $r = \{\vec{u} + t\vec{v},\ t \in \mathbb{R}\}$ e $T\vec{v} \neq \vec{0}$, então $T(r)$ é uma reta.*

Exemplo 1: Considere o sistema linear

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = y_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = y_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = y_3 \end{cases}$$

No lugar de pensar em resolver o sistema, pense nele como uma transformação que, a cada terno ordenado (x_1, x_2, x_3), associa o terno ordenado (y_1, y_2, y_3), definido por $y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3$, $y_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3$, $y_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3$. Fixe, agora, em V, uma base (a canônica, por exemplo). Considere, então que nossa transformação, T, leva o vetor $\vec{v}$ no vetor $\vec{w} = T\vec{v}$, da seguinte forma: se $\vec{v} = x_1\vec{e_1} + x_2\vec{e_2} + x_3\vec{e_3}$, então $\vec{w} = y_1\vec{e_1} + y_2\vec{e_2} + y_3\vec{e_3}$, com (y_1, y_2, y_3) definido como acima.

Exercício 11.4 *Mostre que a transformação que acabamos de definir é linear.*

Exercício 11.5 *Mostre que a ideia é boa: se o sistema tiver só duas equações, poderíamos definir uma transformação linear que a cada vetor do espaço associasse um vetor do plano. Da mesma forma, se tivéssemos só duas variáveis (x_1 e x_2) e três equações, poderíamos definir uma transformação linear que a cada vetor do plano associasse um do espaço. mais geral mente, podemos considerar que, por trás de cada sistema linear com m equações e n incógnitas, está uma transformação linear do espaço $\mathbb{R}^n$ no espaço $\mathbb{R}^m$.*

Exemplo 2: Como sabemos, cada vetor $\vec{v}$ de V se escreve, de maneira única, como $\vec{v} = x\vec{e_1} + y\vec{e_2} + z\vec{e_3}$. Escolha, arbitrariamente, três vetores, $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$ em V. Defina, então, $T : V \to V$ por

$$x\vec{e_1} + y\vec{e_2} + z\vec{e_3} = \vec{v} \longmapsto T\vec{v} = x\vec{\varepsilon}_1 + y\vec{\varepsilon}_2 + z\vec{\varepsilon}_3.$$

Exercício 11.6 *Mostre que a transformação T definida acima é linear.*

Exercício 11.7 *Mostre que toda transformação linear $T : V \to V$ cabe no exemplo acima. Solução: considere uma transformação linear $T : V \to V$; defina $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$ por $\vec{\varepsilon}_1 = T\vec{e}_1$, $\vec{\varepsilon}_2 = T\vec{e}_2$, $\vec{\varepsilon}_3 = T\vec{e}_3$; faça as contas.*

Exercício 11.8 *Pense o exemplo 1 em termos do exemplo 2. Expresse os vetores $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$ em termos dos coeficientes do sistema.*

Os exemplo 1 e 2 abarcam tudo, são gerais demais. Sejamos mais específicos.

Exemplo 3: Consideremos uma rotação do espaço em torno do eixo vertical. Convença-se de que trata-se de uma transformação linear. Chamando de θ o ângulo de rotação (no sentido trigonométrico, *olhando de cima*), teremos:

$$T\vec{e}_1 = \cos\theta\vec{e}_1 + \sin\theta\vec{e}_2, \quad T\vec{e}_2 = -\sin\theta\vec{e}_1 + \cos\theta\vec{e}_2, \quad T\vec{e}_3 = \vec{e}_3.$$

Assim, se o vetor $\vec{v}$ é dado, em coordenadas canônicas, por (x, y, z), então

$$\begin{aligned} T\vec{v} &= T(x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3) = xT\vec{e}_1 + yT\vec{e}_2 + zT\vec{e}_3 = \\ &= x(\cos\theta\vec{e}_1 + \sin\theta\vec{e}_2) + y(-\sin\theta\vec{e}_1 + \cos\theta\vec{e}_2) + z\vec{e}_3 = \\ &= (x\cos\theta - y\sin\theta)\vec{e}_1 + (x\sin\theta + y\cos\theta)\vec{e}_2 + z\vec{e}_3. \end{aligned}$$

Exercício 11.9 *Conclua que a imagem do vetor de coordenadas (na base canônica) (x, y, z) é dada, em coordenadas, por*

$$(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta, z).$$

Se você sabe multiplicar matrizes, observe que, se (x', y', z') são as coordenadas da imagem, então

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Exercício 11.10 *Seja T a transformação linear dada pela rotação do espaço em torno do eixo que passa pela origem e tem a direção de um vetor $\vec{v}$ qualquer. Os casos $\vec{v} = \vec{e}_2$ e $\vec{v} = \vec{e}_3$ são, claro, análogos ao exemplo 3. Suponha que $\vec{v}$ seja dado, em coordenadas canônicas, por $\vec{v} = (a, b, c)$. Como obter as coordenadas, na base canônica, da imagem por T de um vetor $\vec{u}$ de coordenadas (x, y, z)?*

Exemplo 4: Considere a seguinte transformação: um vetor horizontal $\vec{\varepsilon}$ é fixado (isto é, ε é combinação linear de $\vec{e}_1$ e $\vec{e}_2$). Cada ponto do espaço anda um pouco na direção dada por $\vec{\varepsilon}$; o quanto anda é proporcional a sua altura. Mais precisamente, fixamos um número α; cada vetor $\vec{v} = x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3$ é, então, levado em $T\vec{v} = \vec{v} + \alpha z\vec{\varepsilon}$. Uma tal transformação é dita um **cisalhamento** (horizontal, na direção do vetor $\vec{\varepsilon}$).

Exercício 11.11 *Mostre que nossa transformação faz com que os planos horizontais se desloquem sobre si mesmos. Isto é: para cada z_0, os pontos do plano horizontal de altura z_0 sofrem, todos, uma mesma translação (qual?).*

Exercício 11.12 *Mostre que cisalhamentos preservam volumes.*

Exercício 11.13 *Se $\vec{\varepsilon} = a\vec{e_1} + b\vec{e_2}$, escreva as coordenadas, na base canônica, de $T\vec{v}$ em função das de $\vec{v}$. Mostre que, de fato, cisalhamentos horizontais preservam volumes.*

Exercício 11.14 *É claro que podemos definir, de forma análoga, cisalhamentos em que $\vec{\varepsilon}$ seja combinação linear de dois outros vetores da base canônica. Faça isso.*

Exercício 11.15 *E, se $\vec{\varepsilon}$ for um vetor qualquer, como definir um cisalhamento na direção de ε?*

Exemplo 5: Seja T a reflexão através do plano horizontal. Em coordenadas, T é dada por: $T(x_1, x_2, x_3) = (x_1, x_2, -x_3)$. O exemplo é bobo, mas serve para colocar a seguinte questão:

Exercício 11.16 *Seja α um plano qualquer, passando pela origem (pode ser dado por seu vetor normal, ou por um par de vetores linearmente independentes de α). Como expressar a reflexão através de α?*

11.2 O trepa-trepa catalão

Vamos agora trabalhar um pouco mais a geometria das transformações lineares. Tomemos como ponto de partida a seguinte ideia: trata-se de transformações T tais que $T(x\vec{e_1} + y\vec{e_2} + z\vec{e_3}) = x\vec{\varepsilon}_1 + y\vec{\varepsilon}_2 + z\vec{\varepsilon}_3$, ou seja, mantemos as coordenadas (x, y, z) de cada ponto, mas, essencialmente, mudamos a base do sistema de coordenadas.[1]

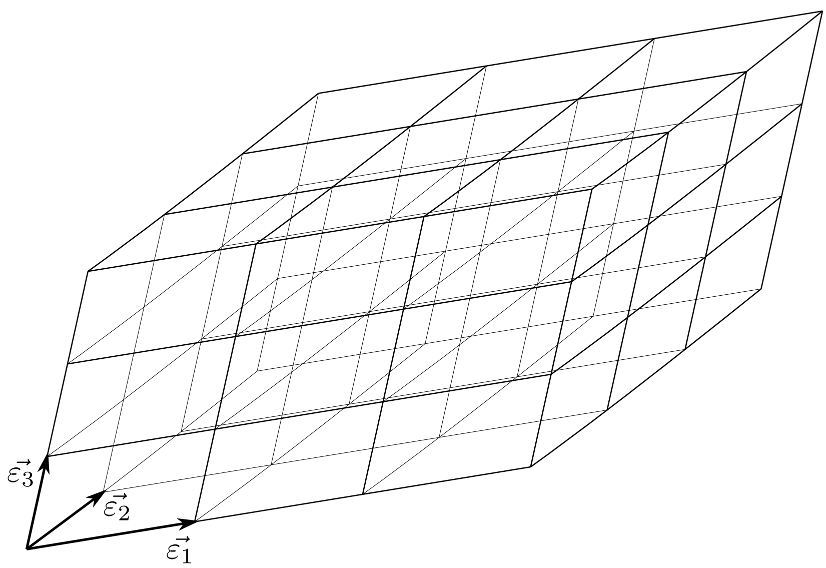


Figura 11.2:

Mais precisamente, consideremos um objeto φ, que consiste em um conjunto de pontos do espaço dados, em coordenadas canônicas, por ternos ordenados (digamos que estes ternos ordenados estão guardados em um arquivo que também chamaremos de φ).

[1]Quando dizemos *essencialmente*, é por que, na verdade, não é preciso, para que T seja linear, que $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$ sejam linearmente independentes. O caso das projeções é um bom exemplo

Isto significa que a cada terno ordenado (x, y, z) corresponde um ponto P cujo vetor posição é

$$x\vec{e_1} + y\vec{e_2} + z\vec{e_3}.$$

Mas o que acontece se resolvermos mudar a base de nosso sistema de coordenadas, trocando $\{\vec{e_1},\vec{e_2},\vec{e_3}\}$ por $\{\vec{\varepsilon}_1,\vec{\varepsilon}_2,\vec{\varepsilon}_3\}$? A resposta é simples: quando formos ao arquivo e trouxermos o terno ordenado (x, y, z), nosso sistema nos dará não mais o ponto P cujo vetor posição é $x\vec{e_1} + y\vec{e_2} + z\vec{e_3}$, mas o ponto P', cujo vetor posição é

$$x\vec{\varepsilon}_1 + y\vec{\varepsilon}_2 + z\vec{\varepsilon}_3.$$

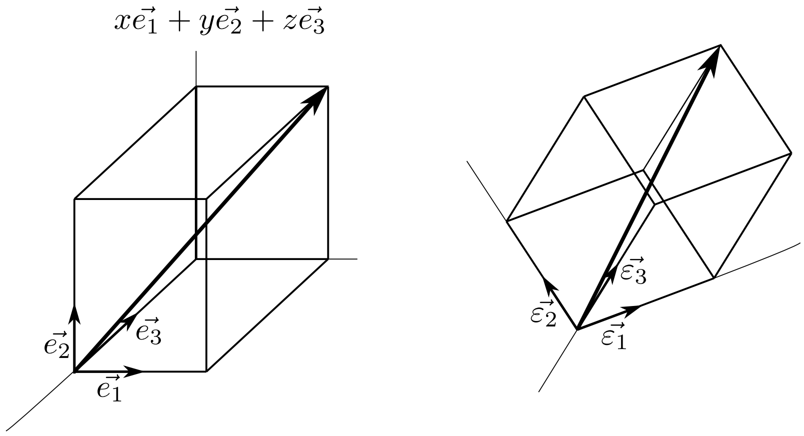


Figura 11.3:

Observação: Estamos, na verdade, fazendo uma opção no que diz respeito *à maneira como vamos representar, em nossas cabeças*, uma transformação linear. O efeito visual da transformação T é claro: transportamos o ponto P do reticulado que tem por base um cubo (definido por meio de $\vec{e_1}$, $\vec{e_2}$ e $\vec{e_3}$) para o ponto P' (de mesmas coordenadas) do reticulado que tem por base um paralelepípedo (definido por $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$). É como se a prefeitura contratasse um arquiteto (provavelmente catalão), para remodelar as pracinhas. O arquiteto, então, altera os brinquedos conhecidos como trepa-trepa, fazendo com que deixem de ser tão certinhos, assumindo, na nova versão, formas mais oblíquas.

Exercício 11.17 *Seja $T : V \to V$ uma transformação linear, com $T\vec{e_1} = \vec{\varepsilon}_1$, $T\vec{e_2} = \vec{\varepsilon}_2$, $T\vec{e_3} = \vec{\varepsilon}_3$. Identifique, no espaço, pontos a vetores. Considere um sólido φ. Dividindo o espaço em cubinhos com lados nas direções dos vetores da base canônica, obtenha uma aproximação para o volume de φ. Considere, agora, a imagem de φ por T, $T(\varphi)$. Aproxime o volume de $T(\varphi)$ usando os paralelepípedos (com lados nas direções de $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$) imagens dos que usou para aproximar o volume de φ (veja a figura correspondente, em dimensão 2).*

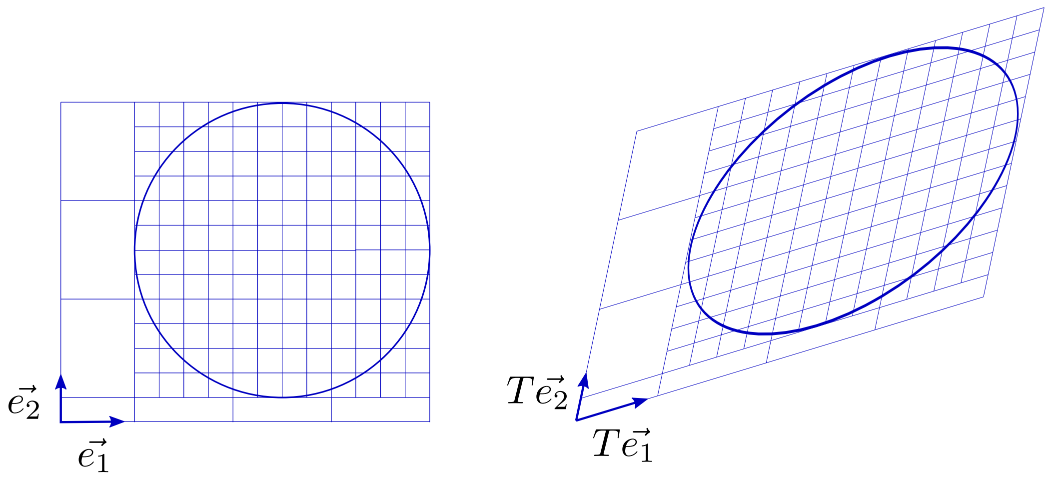


Figura 11.4: relação entre volumes, versão 2d

Note que a razão entre o volume de cada paralelepipedozinho e o correspondente cubinho é igual à que existe entre o volume do paralelepípedo P_ε, formado por $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$, e o volume do paralelepípedo formado por $\vec{e_1}$, $\vec{e_2}$ e $\vec{e_3}$. Mostre que essa razão é igual a $|det(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3)|$. Conclua que a razão entre o volume de $T(\varphi)$ e o de φ é

$$\frac{\textit{volume de } T(\varphi)}{\textit{volume de } \varphi} = |det(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3)|.$$

11.3 Transformações lineares

Definição: Sejam **V** e **W** dois espaços vetoriais. Uma função $T : \mathbf{V} \to \mathbf{W}$ é dita uma **transformação linear** se

$$T(\mathbf{u} + t\mathbf{v}) = T\mathbf{u} + tT\mathbf{v} \ \forall\ \mathbf{u}, \mathbf{v} \in \mathbf{V},\ \forall\ t \in \mathbb{R}.$$

Exercício 11.18 *Mostre que $T : \mathbf{V} \to \mathbf{W}$ é uma transformação linear se, e somente se, valem as duas propriedades:*

$$\begin{array}{ll} (i) & T(t\vec{u}) = tT\vec{u} \ \forall\ \vec{u} \in V,\ t \in \mathbb{R} \\ (ii) & T(\vec{u} + \vec{v}) = T\vec{u} + T\vec{v} \ \forall\ \vec{u}, \vec{v} \in V. \end{array}$$

Exemplo 1: Por trás do sistema linear

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = y_1 \\ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = y_m \end{cases}$$

está a transformação linear $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, dada por

$$A(x_1, \ldots, x_n) = (y_1, \ldots, y_m),$$

com, para cada $i = 1, \ldots, m$,

$$y_i = a_{i1}x_1 + \ldots + a_{in}x_n = \sum_{j=1}^{n} a_{ij}x_j.$$

Exemplo 2: Se $\mathbf{V} = C^1(\mathbb{R}, \mathbb{R})$ e $\mathbf{W} = C^o(\mathbb{R}, \mathbb{R})$ (o que significa que $\mathbf{V}$ é o espaço das funções com derivada primeira contínua e $\mathbf{W}$ é o espaço das funções contínuas), então o operador $D : \mathbf{V} \to \mathbf{W}$ de derivação, definido por $Df = f'$ (isto é, D leva f em sua derivada) é linear.

Exemplo 3: Se $\mathbf{W} = C^o(\mathbb{R}, \mathbb{R})$, o operador de integração de a até b, $I : \mathbf{W} \longrightarrow \mathbb{R}$, dado por

$$If = \int_a^b f(x)dx,$$

é linear.

Exemplo 4: Se $\mathbf{V}$ é o espaço das sequências infinitas de números reais (um elemento de $\mathbf{V}$ é uma sequência $(x_n)_{n \in \mathbb{N}}$) então o operador *shift*, $S : \mathbf{V} \longrightarrow \mathbf{V}$, dado por

$$S(x_1, x_2, x_3, x_4, \ldots) = (0, x_1, x_2, x_3, \ldots),$$

é linear.

Exercício 11.19 *Mostre que a transformação D do exemplo 2 é sobrejetiva mas não injetiva (use o Teorema Fundamental do Cálculo).*

Exercício 11.20 *Mostre que a transformação S do exemplo 4 é injetiva mas não sobrejetiva.*

11.4 Um pouquinho de Álgebra

O exercício a seguir é tão importante que poderia até ser convertido em Proposição e demonstrado no texto. Se você entendeu a história de dimensão, o que é crucial, deve ser fácil.

Exercício 11.21 *Sejam $\mathbf{V}$ um espaço vetorial de dimensão finita e $T : \mathbf{V} \to \mathbf{V}$ linear. Seja $\{\boldsymbol{\varepsilon_1}, \ldots, \boldsymbol{\varepsilon_n}\}$ uma base de $\mathbf{V}$. Observe que a imagem de T, isto é, o conjunto $ImT = \{T\mathbf{v},\ \mathbf{v} \in \mathbf{V}\}$ é o conjunto das combinações lineares de $T\boldsymbol{\varepsilon_1}, \ldots, T\boldsymbol{\varepsilon_n}$. Mostre que T é injetiva se e somente se é sobrejetiva.*

Vamos considerar a operação **produto de transformações lineares**, que nada mais é que a composição de funções: se $T : \mathbf{V}_1 \to \mathbf{V}_2$ e $S : \mathbf{V}_2 \to \mathbf{V}_3$ são transformações lineares, seu produto é a transformação linear $ST : \mathbf{V}_1 \to \mathbf{V}_3$, dada por $(ST)\mathbf{v} = S(T\mathbf{v})$.

Exercício 11.22 *Mostre que ST é, de fato, linear. Observe que, como a composição de funções, o produto de transformações lineares é associativo, isto é: sempre vale $(RS)T = R(ST)$.*

Exercício 11.23 *Mostre que o produto é distributivo em relação à adição:*[2] *se $R, S : \mathbf{V}_1 \to \mathbf{V}_2$ e $T : \mathbf{V}_2 \to \mathbf{V}_3$ são lineares, então $T(R+S) = TR + TS$; da mesma forma, se $T : \mathbf{V}_1 \to \mathbf{V}_2$ e $R, S : \mathbf{V}_2 \to \mathbf{V}_3$ são lineares, então $(R+S)T = RT + ST$.*

Definição: Seja $I : \mathbf{V} \to \mathbf{V}$ a transformação (linear) **identidade**, isto é, $I\mathbf{v} = \mathbf{v}\ \forall\ \mathbf{v} \in \mathbf{V}$. Se $T : \mathbf{V} \to \mathbf{V}$ é linear e existe $T^{-1} : \mathbf{V} \to \mathbf{V}$ tal que $T^{-1}T = I = TT^{-1}$, T é dita **invertível** (e T^{-1} é dita **inversa** de T). Transformações lineares invertíveis são também chamadas de **isomorfismos**.

Exercício 11.24 *Mostre que a inversa, caso exista, é linear e única.*

Exercício 11.25 *Suponha que $\mathbf{V}$ seja de dimensão finita, que $T : \mathbf{V} \to \mathbf{V}$ seja linear e que exista $S : \mathbf{V} \to \mathbf{V}$ tal que $ST = I$ (não se supõe S linear). Mostre que, então, T tem inversa e que $S = T^{-1}$.*

Exercício 11.26 *Suponha que R e S sejam transformações lineares invertíveis. Mostre que RS é invertível e que $(RS)^{-1} = S^{-1}R^{-1}$.*

Exercício 11.27 *Seja $L(\mathbf{V})$ o conjunto das transformações lineares de $\mathbf{V}$ em $\mathbf{V}$. Se A e B pertencem a $L(\mathbf{V})$, define-se o **colchete de Lie** $[A, B]$ por $[A, B] = AB - BA$. Mostre que, para quaisquer A, B e C, vale*

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

11.5 Isometrias

Transformações lineares são os **homomorfismos** de espaços vetoriais (transformações de um espaço em outro que preservam as operações). Se o espaço, além das operações de adição e multiplicação por escalar, é dotado de um produto interno, podemos considerar transformações que preservem não apenas as operações de espaço vetorial, mas também o próprio produto interno. Ao contrário das simples transformações lineares, que não precisam ser injetivas, transformações desse tipo especial não podem deixar de sê-lo.

Proposição: Sejam $\mathbf{V}_1$ e $\mathbf{V}_2$ espaços vetoriais com produto interno e $T : \mathbf{V}_1 \to \mathbf{V}_2$ uma trensformação linear tal que

[2] Se $R, S : \mathbf{V}_1 \to \mathbf{V}_2$ são lineares, sua **soma**, $R + S$ é definida da maneira óbvia: $(R+S)\mathbf{v} = R\mathbf{v} + S\mathbf{v}$

$$\langle T\mathbf{u}, T\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle \ \forall\ \mathbf{u}, \mathbf{v} \in \mathbf{V}_1.$$

Então T é injetiva.

Demonstração: Basta provar que $\mathbf{u} \neq \mathbf{0} \Rightarrow T\mathbf{u} \neq \mathbf{0}$, o que equivale a $|T\mathbf{u}| \neq 0$. Mas isso é claramente verdade, já que $\mathbf{u} \neq \mathbf{0} \Rightarrow 0 \neq \langle \mathbf{u}, \mathbf{u}\rangle = \langle T\mathbf{u}, T\mathbf{u}\rangle \Rightarrow |T\mathbf{u}| \neq 0$. ■

Observemos que preservar produto interno implica em preservar distâncias e ângulos. Na demonstração da proposição anterior, porém, utilizamos apenas a preservação das distâncias.

Proposição: Sejam $\mathbf{V}_1$ e $\mathbf{V}_2$ espaços vetoriais com produto interno e $T : \mathbf{V}_1 \to \mathbf{V}_2$ uma transformação linear tal que $|T\mathbf{u}| = |\mathbf{u}|$ para todo $\mathbf{u}$ em $\mathbf{V}_1$ (ou seja: T preserva a norma). Então T preserva o produto interno, isto é:

$$\langle T\mathbf{u}, T\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle \ \forall\ \mathbf{u}, \mathbf{v} \in \mathbf{V}_1.$$

Demonstração: Da hipótese, temos, para quaisquer $\mathbf{u}$ e $\mathbf{v}$ em $\mathbf{V}_1$, $|T(\mathbf{u}+\mathbf{v})| = |\mathbf{u}+\mathbf{v}|$. Daí vem

$$|\mathbf{u}|^2 + |\mathbf{v}|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle = |\mathbf{u}+\mathbf{v}|^2 = |T\mathbf{u}+T\mathbf{v}|^2 = |T\mathbf{u}|^2 + |T\mathbf{v}|^2 + 2\langle T\mathbf{u}, T\mathbf{v}\rangle.$$

Como $|\mathbf{u}|^2 = |T\mathbf{u}|^2$ e $|\mathbf{v}|^2 = |T\mathbf{v}|^2$, segue $\langle \mathbf{u}, \mathbf{v}\rangle = \langle T\mathbf{u}, T\mathbf{v}\rangle$. ■

Tentemos ir um pouco além: vamos manter a hipótese de preservação das distâncias mas retirar a linearidade.

Exercício 11.28 *Note que, sem a hipótese de linearidade, a preservação das distâncias não é garantida se supusermos apenas a preservação da norma.*

Exercício 11.29 *Sejam $\mathbf{V}_1$ e $\mathbf{V}_2$ espaços vetoriais com produto interno e $f : \mathbf{V}_1 \to \mathbf{V}_2$ uma transformação preservando distâncias, isto é: $|f(\mathbf{u}) - f(\mathbf{v})| = |\mathbf{u}-\mathbf{v}| \ \forall\ \mathbf{u},\ \mathbf{v} \in \mathbf{V}_1$. Mostre, com um exemplo (tipo $f : \mathbb{R}^2 \to \mathbb{R}^2$), que não necessariamente temos f linear.*

Exercício 11.30 *Seja $f : \mathbb{R}^2 \to \mathbb{R}^2$ preservando distâncias e tal que $f(\vec{0}) = \vec{0}$. Use argumentos geométricos para provar que f é linear. Mostre que o mesmo vale para $f : \mathbb{R}^2 \to \mathbb{R}^3$.*

Exercício 11.31 *Refaça a demonstração do exercício anterior usando apenas argumentos algébricos, isto é: $\mathbb{R}^2$ é um espaço vetorial real, com produto interno, de dimensão 2, o mesmo valendo para $\mathbb{R}^3$, com dimensão 3).*

Os exercícios acima nos dão motivos de sobra para acreditar na veracidade da proposição a seguir. Se você não resolveu os exercícios, a proposição pode ser vista como uma solução elegante para o último, sem a limitação a $\mathbb{R}^2$ ou $\mathbb{R}^3$.

Proposição: Sejam $\mathbf{V}_1$ e $\mathbf{V}_2$ espaços vetoriais reais com produto interno e $f : \mathbf{V}_1 \to \mathbf{V}_2$ uma função preservando distâncias (isto é: $|f(\mathbf{u}) - f(\mathbf{v})| = |\mathbf{u}-\mathbf{v}| \ \forall\ \mathbf{u},\ \mathbf{v} \in \mathbf{V}_1$). Suponhamos que, além disso, $f(\mathbf{0}) = \mathbf{0}$. Então f é uma transformação linear.

Demonstração: Sejam $\mathbf{u},\ \mathbf{v}$ em $\mathbf{V}_1$ e t em $\mathbb{R}$. Provemos que $f(\mathbf{u}+t\mathbf{v}) = f(\mathbf{u}) + tf(\mathbf{v})$.

Comecemos observando que, como $f(\mathbf{0}) = \mathbf{0}$, f preserva norma (já que $|f(\mathbf{u})| = |f(\mathbf{u}) - f(\mathbf{0})|$). Daí vem

$$|f(\mathbf{u})|^2 + |f(\mathbf{v}|^2 - 2\langle f(\mathbf{u}), f(\mathbf{v})\rangle = |f(\mathbf{u}) - f(\mathbf{v})|^2 =$$
$$= |\mathbf{u} - \mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2 - 2\langle \mathbf{u}, \mathbf{v}\rangle .$$

Logo, temos $\langle f(\mathbf{u}), f(\mathbf{v})\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$ (entenda isso como *triângulos de lados congruentes têm ângulos correspondentes também congruentes*, o que implica em dizer que a preservação de distâncias nos dá a preservação de ângulos). Daí vem:

$$|f(\mathbf{u} + t\mathbf{v}) - (f(\mathbf{u}) + tf(\mathbf{v}))|^2 =$$
$$= \langle f(\mathbf{u} + t\mathbf{v}) - (f(\mathbf{u}) + tf(\mathbf{v})), f(\mathbf{u} + t\mathbf{v}) - (f(\mathbf{u}) + tf(\mathbf{v}))\rangle =$$
$$|f(\mathbf{u} + t\mathbf{v}) - f(\mathbf{u})|^2 + t^2|f(\mathbf{v})|^2 - 2t\langle f(\mathbf{u} + t\mathbf{v}) - f(\mathbf{u}), f(\mathbf{v})\rangle =$$
$$|(\mathbf{u} + t\mathbf{v}) - \mathbf{u}|^2 + t^2|\mathbf{v}|^2 - 2t\langle (\mathbf{u} + t\mathbf{v}) - \mathbf{u}, \mathbf{v}\rangle = 0,$$

o que prova que $f(\mathbf{u} + t\mathbf{v}) = f(\mathbf{u}) + tf(\mathbf{v})$. ■

Exercício 11.32 *Note que a mesma demonstração se aplica a espaços vetoriais complexos (trocando $\langle , \rangle$ por Re $\langle , \rangle$ quando necessário).*

Transformações preservando distâncias são ditas **isométricas**. Bijeções preservando distâncias são chamadas de **isometrias**.

Exercício 11.33 *Observe que, pelo que acabamos de ver, transformações isométricas entre espaços vetoriais com produto interno são, a menos de translações, lineares (isto é: se $|f(\mathbf{u}) - f(\mathbf{v})| = |\mathbf{u} - \mathbf{v}|\ \forall\ \mathbf{u},\ \mathbf{v} \in \mathbf{V}_1$), então existem $\mathbf{w}_o \in \mathbf{V}_2$ e $T : \mathbf{V}_1 \to \mathbf{V}_2$ linear tais que $f(\mathbf{u}) = \mathbf{w}_o + T\mathbf{u}\ \forall \mathbf{u} \in \mathbf{V}_1$).*

Se T é uma transformação linear de um espaço vetorial com produto interno, V, em si mesmo, a preservação de distâncias é, como acabamos de ver, equivalente à preservação do produto interno.

Exercício 11.34 *Suponha que $\mathbf{V}$ é espaço vetorial de dimensão finita, com produto interno e que $T : \mathbf{V} \to \mathbf{V}$ é linear. Mostre que são equivalentes:*

(i) $|T\mathbf{u}| = |\mathbf{u}|\ \forall \mathbf{u} \in V$;
(ii) $\langle T\mathbf{u}, T\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle\ \forall \mathbf{u}, \mathbf{v} \in \mathbf{V}$;
(iii) T transforma bases ortonormais de $\mathbf{V}$ em bases ortonormais de $\mathbf{V}$;
(iv) T transforma alguma base ortonormal de $\mathbf{V}$ em uma base ortonormal de $\mathbf{V}$.

Definição: Seja $\mathbf{V}$ um espaço vetorial com produto interno. Uma transformação linear bijetiva $T : \mathbf{V} \to \mathbf{V}$ é dita **ortogonal** se $\langle T\mathbf{u}, T\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$ para quaisquer $\mathbf{u}$ e $\mathbf{v}$ em $\mathbf{V}$.

Capítulo 12

A matriz de uma transformação linear

A noção de matriz surge naturalmente a partir de sistemas lineares. O sistema

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

pode ser representado, na forma matricial, por

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

É também natural considerar, associada à **matriz** $m \times n$ [1]

$$(a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},$$

a transformação linear $A : \mathbb{R}^n \to \mathbb{R}^m$, dada por $A\mathbf{x} = \mathbf{y}$, com

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Vamos, porém, inverter a ordem natural das coisas. Partiremos das transformações lineares para associar-lhes matrizes.

12.1 No plano

Para simplificar um pouco, voltemos à tela do computador e ao plano. Recordemos que a representação dos pontos na tela se faz por meio de um sistema de coordenadas

[1] Embora o conceito de matriz seja bastante simples, não custa dar uma definição precisa: uma matriz $m \times n$ (a_{ij}) com entradas no conjunto X é uma aplicação $a : \{1, \ldots, m\} \times \{1, \ldots, n\} \to X$. É usual notar $a(i, j)$ por a_{ij}

fixo. Assim, ao par ordenado (x, y) corresponderá sempre, na tela, o ponto P cujo vetor posição é $x\vec{e_1} + y\vec{e_2}$ ($\{\vec{e_1}, \vec{e_2}\}$ é suposta ser a base do sistema de coordenadas da tela). Consideremos agora dois vetores, $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$, e a tansformação T que leva $x\vec{e_1} + y\vec{e_2}$ em $x\vec{\varepsilon}_1 + y\vec{\varepsilon}_2$. Teoricamente está lindo, abstrato, mas nosso problema prático é: dado um ponto P, como vamos saber qual é o ponto P' em que T transforma P? Digamos assim: se o ponto P é dado por suas coordenadas (x, y), como vamos calcular as coordenadas (x', y') de seu transformado P'?

Exercício 12.1 *Note que o que queremos são as coordenadas de P' no sistema de base $\{\vec{e_1}, \vec{e_2}\}$, que é o único que nosso computador conhece.*

A primeira coisa a saber é, é claro, quem são os vetores $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$. Ora, se nosso computador só conhece o sistema de base $\{\vec{e_1}, \vec{e_2}\}$, $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$ devem ser indicados por suas coordenadas nesse sistema. Digamos que

$$\vec{\varepsilon}_1 = (3, 2), \ \ \vec{\varepsilon}_2 = (1, 4).$$

Exercício 12.2 *Pegue papel quadriculado e desenhe. Marque os vetores $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$.*

Exercício 12.3 *Marque o ponto P' cujo vetor posição é $2\vec{\varepsilon}_1+\vec{\varepsilon}_2$.*

Consideremos o ponto P dado por suas coordenadas, $P = (2, 1)$. Isto significa que o vetor posição de P é $2\vec{e_1}+\vec{e_2}$. Logo, P será transformado no ponto P' cujo vetor posição é $2\vec{\varepsilon}_1+\vec{\varepsilon}_2$.

E agora **atenção para o pulo do gato**: se temos as coordenadas de $\vec{\varepsilon}_1$ e as de $\vec{\varepsilon}_2$ no sistema de base $\{\vec{e_1}, \vec{e_2}\}$ ($\vec{\varepsilon}_1$=(3,2), $\vec{\varepsilon}_2$=(1,4)), então temos as de $2\vec{\varepsilon}_1+\vec{\varepsilon}_2$:

$$2\vec{\varepsilon}_1 + \vec{\varepsilon}_2 = 2(3, 2) + (1, 4) = (6, 4) + (1, 4) = (7, 8).$$

Ora, mas as coordenadas do vetor posição de P' são as coordenadas do ponto P'. Logo, $P' = (7, 8)$.

Exercício 12.4 *Certifique-se de que entendeu tudo.*

Exercício 12.5 *Veja se as coisas ficaram claras. Continuemos no plano. Pegue duas folhas de papel quadriculado e marque a origem O em cada um dos dois. Desenhe uma figura (não muito grande nem muito complicada) em uma das folhas (pode ser, por exemplo, um dos quadradinhos). Escolha dois vetores $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$ ($\vec{\varepsilon}_1$=$(2, 1)$, $\vec{\varepsilon}_2$=$(-1, 3)$, por exemplo) e marque-os (com pé na origem) na outra folha. Construa na segunda folha, por cima do quadriculado original, o reticulado de base $\{\vec{\varepsilon}_1, \vec{\varepsilon}_2\}$. Desenhe nesta última folha, com auxílio do reticulado, a imagem de sua figura pela transformação T que leva $x\vec{e_1} + y\vec{e_2}$ em $x\vec{\varepsilon}_1 + y\vec{\varepsilon}_2$. Um ponto P' da nova figura foi marcado usando o sistema de coordenadas de base $\{\vec{\varepsilon}_1, \vec{\varepsilon}_2\}$. Suponha que suas coordenadas, neste sistema, são (x, y) (note que isto quer dizer que seu vetor posição é $x\vec{\varepsilon}_1 + y\vec{\varepsilon}_2$ e que P' é imagem do ponto P da figura original cujo vetor posição é $x\vec{e_1} + y\vec{e_2}$). Pois bem, o ponto P' tem coordenadas (x', y'), referentes ao papel quadriculado, que são, é claro, diferentes das coordenadas (x, y) de P (note que (x, y) são as coordenadas de P', só que no sistema de base $\{\vec{\varepsilon}_1, \vec{\varepsilon}_2\}$). O que acabamos de mostrar é que*

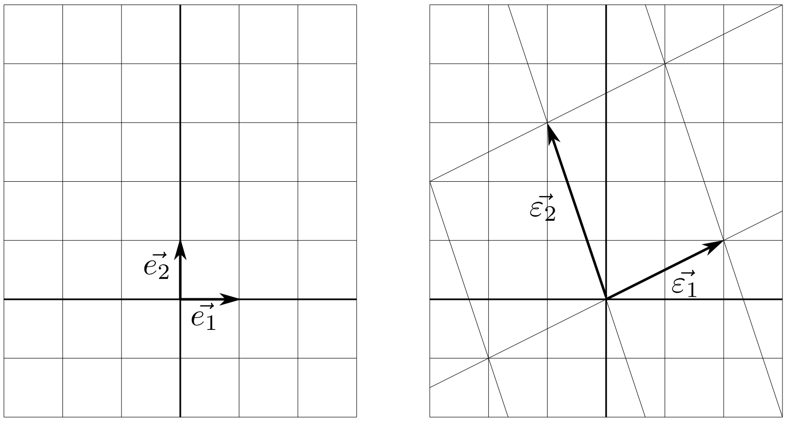


Figura 12.1:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} =: x \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + y \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} =:$$

$$=: \begin{pmatrix} a_{11}x \\ a_{21}x \end{pmatrix} + \begin{pmatrix} a_{12}y \\ a_{22}y \end{pmatrix} =: \begin{pmatrix} a_{11}x + a_{12}y \\ a_{21}x + a_{22}y \end{pmatrix},$$

onde

$$\vec{\varepsilon}_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \ \vec{\varepsilon}_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}.$$

12.2 Matriz de uma transformação linear

Vamos utilizar com frequência, como já fizemos no plano, uma notação alternativa: ternos ordenados em pé. Isto é, podemos representar o terno ordenado (x, y, z) pelo **vetor coluna**

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Usando a nova notação, o terno ordenado P' correspondente a

$$-2(3, 2, 1) + (-3)(1, 4, 7)$$

será indicado por

$$-2\begin{pmatrix}3\\2\\1\end{pmatrix}+(-3)\begin{pmatrix}1\\4\\7\end{pmatrix}=\begin{pmatrix}-9\\-16\\-23\end{pmatrix}.$$

O terno ordenado P' correspondente a

$$\pi(3,2,1)+\sqrt{5}(1,4,7)$$

será indicado por

$$\pi\begin{pmatrix}3\\2\\1\end{pmatrix}+\sqrt{5}\begin{pmatrix}1\\4\\7\end{pmatrix}=\begin{pmatrix}3\pi\\2\pi\\\pi\end{pmatrix}+\begin{pmatrix}\sqrt{5}\\4\sqrt{5}\\7\sqrt{5}\end{pmatrix}=\begin{pmatrix}3\pi+\sqrt{5}\\2\pi+4\sqrt{5}\\\pi+7\sqrt{5}\end{pmatrix}.$$

Façamos $\vec{\varepsilon}_1=(3,2,1)$, $\vec{\varepsilon}_2=(1,4,7)$, $\vec{\varepsilon}_3=(1,1,0)$ e consideremos a transformação T dada por $T(x\vec{e_1}+y\vec{e_2}+z\vec{e_3})=x\vec{\varepsilon}_1+y\vec{\varepsilon}_2+z\vec{\varepsilon}_3$. Em termos práticos, interessa-nos conhecer as coordenadas, na base canônica, do ponto $P'=(x',y',z')$, imagem por T do ponto $P=(x,y,z)$. Ora, não é difícil ver que

$$\begin{pmatrix}x'\\y'\\z'\end{pmatrix}=x\begin{pmatrix}3\\2\\1\end{pmatrix}+y\begin{pmatrix}1\\4\\7\end{pmatrix}+z\begin{pmatrix}1\\1\\0\end{pmatrix}=\begin{pmatrix}3x+y+z\\2x+4y+z\\x+7y\end{pmatrix}.$$

Assim, para indicar a transformação T correspondente aos vetores $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$, usaremos uma matriz tendo $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$ nas colunas, definindo sua multiplicação por uma matriz coluna (que representa um vetor) como abaixo:

$$\begin{pmatrix}3&1&1\\2&4&1\\1&7&0\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=x\begin{pmatrix}3\\2\\1\end{pmatrix}+y\begin{pmatrix}1\\4\\7\end{pmatrix}+z\begin{pmatrix}1\\1\\0\end{pmatrix}=\begin{pmatrix}3x+y+z\\2x+4y+z\\x+7y\end{pmatrix}.$$

Exercício 12.6 *Se o ponto P' é a imagem de P ($P'=T(P)$) pela transformação definida acima, quais seriam as coordenadas (x',y',z') de P' se $P=(-2,-3,1)$? E se $P=(\pi,\sqrt{5},e)$?*

Exercício 12.7 *Quais seriam as coordenadas (x',y',z') de P' se $P=(x,y,z)$ e mudássemos $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$ para $\vec{\varepsilon}_1$=(1,1,0), $\vec{\varepsilon}_2$=(-2,3,0), $\vec{\varepsilon}_3$=(1,1,3)?*

Exercício 12.8 *Quais seriam as coordenadas (x',y',z') de P' se $P=(x,y,z)$ e $\vec{\varepsilon}_1=(a_{11},a_{21},a_{31})$, $\vec{\varepsilon}_2=(a_{12},a_{22},a_{32})$, $\vec{\varepsilon}_3=(a_{13},a_{23},a_{33})$?*

Já é hora de explicitarmos as definições.

Definição: Dados os vetores $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$, seja $T:V\to V$ a transformação linear que associa, a cada vetor $(x,y,z)=x\vec{e_1}+y\vec{e_2}+z\vec{e_3}$, o vetor $(x',y',z')=x\vec{\varepsilon}_1+y\vec{\varepsilon}_2+z\vec{\varepsilon}_3$. Se as coordenadas de $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ e $\vec{\varepsilon}_3$ na base $(\vec{e_1},\vec{e_2},\vec{e_3})$ são dadas por

$$\vec{\varepsilon}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \quad \vec{\varepsilon}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \quad \vec{\varepsilon}_3 = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix},$$

a **matriz da transformação** T é

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Definição: Dados a matriz $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ e o vetor coluna $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$, o **produto**

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

é definido por

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + y \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + z \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11}x \\ a_{21}x \\ a_{31}x \end{pmatrix} + \begin{pmatrix} a_{12}y \\ a_{22}y \\ a_{32}y \end{pmatrix} + \begin{pmatrix} a_{13}z \\ a_{23}z \\ a_{33}z \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{pmatrix}.$$

Observação: Nossas definições foram feitas sob medida para que, dado o ponto $P = (x, y, z)$, as coordenadas (x', y', z') do ponto $P' = T(P)$ sejam dadas por

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Exercício 12.9 *Determine a matriz de cada uma das transformações definidas nos exemplos 1 a 5 da primeira seção deste capítulo.*

Exercício 12.10 *Considere as transformações lineares $A : V \to V$ e $B : V \to V$, dadas, respectivamente, pelas matrizes (a_{ij}) e (b_{ij}),*

$$(a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad (b_{ij}) = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$$

Seja $C : V \to V$, $C = AB$, isto é: C é dada por $C\vec{v} = A(B\vec{v})$. Mostre que a matriz de C é o produto das matrizes de A e B, isto é: a matriz de C, (c_{ij}), é dada pelo produto de (a_{ij}) e (b_{ij}),

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix},$$

com $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j}$.

Definição: A i-ésima **linha** da matriz (a_{ij}) é o terno ordenado (a_{i1}, a_{i2}, a_{i3}) (por analogia, o vetor representado, na base canônica por uma linha de (a_{ij}) é chamado de **vetor linha** de (a_{ij})). Analogamente, a j-ésima **coluna** da matriz (a_{ij}) é o terno ordenado (a_{1j}, a_{2j}, a_{3j}) (e o vetor representado, na base canônica por uma coluna de (a_{ij}) é chamado de **vetor coluna** de (a_{ij})).

Assim, dizemos que o produto das matrizes A e B é a matriz $C = AB$, cujas **entradas** c_{ij} são dadas pelo produto escalar da i-ésima linha de A pela j-ésima coluna de B.

Exercício 12.11 *Seja I a matriz dada por*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Mostre que $IA = A = AI$, qualquer que seja a matriz 3×3 A.

Definição: A matriz I definida acima é chamada **matriz identidade**. Se A é uma matriz 3×3 e A^{-1} é outra matriz 3×3, tal que $A^{-1}A = I = AA^{-1}$, A^{-1} é chamada **matriz inversa** de A.

Exercício 12.12 *Mostre que nem toda matriz tem inversa.*

Exercício 12.13 *Mostre que a inversa, se existe, é única.*

Exercício 12.14 *Sejam A uma matriz 3×3 e B outra matriz 3×3, tal que $BA = I$. Mostre que, então, A tem inversa e que $A^{-1} = B$.*

Exercício 12.15 *A matriz $A = (a_{ij})$ é dita **simétrica** se $a_{ij} = a_{ji}$ para todo par ij. Mostre que o conjunto das matrizes simétricas é um espaço vetorial (de que dimensão?). Analogamente, a matriz $A = (a_{ij})$ é dita **antissimétrica** se $a_{ij} = -a_{ji}$ para todo par ij. Mostre que o conjunto das matrizes antissimétricas é um espaço vetorial (de que dimensão?).*

Exercício 12.16 *Mostre que toda matriz se escreve, de maneira única, como soma de uma simétrica com uma antissimétrica.*

Exercício 12.17 *Seja A matriz antissimétrica 3×3. Mostre que existe um vetor a em $\mathbb{R}^3$ tal que $Av = a \otimes v$ para todo v em $\mathbb{R}^3$.*

Exercício 12.18 *Retorne à seção Um pouquinho de Álgebra, do capítulo 7, e refaça tudo, com matrizes no lugar de transformações lineares.*

12.3 O caso geral

Sejam $\mathbf{V}$ e $\mathbf{W}$ espaços vetoriais de dimensão finita, $T : \mathbf{V} \to \mathbf{W}$ transformação linear. Sejam

$$\alpha = \{\mathbf{e_1}, \ldots, \mathbf{e_n}\}, \; \beta = \{\boldsymbol{\varepsilon_1}, \ldots, \boldsymbol{\varepsilon_m}\}$$

bases de $\mathbf{V}$ e $\mathbf{W}$, respectivamente. Suponhamos que $\mathbf{v} = x_1\mathbf{e_1} + \cdots + x_n\mathbf{e_n}$ seja um vetor qualquer de $\mathbf{V}$. Queremos obter a representação de $T\mathbf{v}$ na base β. Como $T\mathbf{v} = x_1 T\mathbf{e_1} + \cdots + x_n T\mathbf{e_n}$, fica claro que basta conhecer as expressões, na base β, de $T\mathbf{e_1}, \cdots, T\mathbf{e_n}$. Admitindo que $T\mathbf{e_j} = a_{1j}\boldsymbol{\varepsilon_1} + \cdots + a_{mj}\boldsymbol{\varepsilon_m}$, é só fazer as contas.

As coisas podem ficar mais simples se convencionarmos que, para qualquer n-upla $(x_1, \ldots, x_n)$, o vetor $\mathbf{v} = x_1\mathbf{e_1} + \cdots + x_n\mathbf{e_n}$ de $\mathbf{V}$ será representado pelo vetor coluna $(\mathbf{v})^\alpha$,

$$(\mathbf{v})^\alpha = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

enquanto, para cada m-upla $(y_1, \ldots, y_m)$, o vetor $\mathbf{w} = y_1\boldsymbol{\varepsilon_1} + \cdots + y_m\boldsymbol{\varepsilon_m}$ de $\mathbf{W}$ será representado pelo vetor coluna $(\mathbf{w})^\beta$,

$$(\mathbf{w})^\beta = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Teremos, então,

$$(T\mathbf{e_1})^\beta = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \ldots, (T\mathbf{e_j})^\beta = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}, \ldots, (T\mathbf{e_n})^\beta = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix},$$

o que nos dá, se $\mathbf{v} = x_1\mathbf{e_1} + \cdots + x_n\mathbf{e_n}$ e $T\mathbf{v} = y_1\boldsymbol{\varepsilon_1} + \cdots + y_m\boldsymbol{\varepsilon_m}$,

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_j \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix},$$

ou seja,

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1j}x_j + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mj}x_j + \cdots + a_{mn}x_n \end{pmatrix}.$$

Dizemos, então, que a **matriz de** T, referente às bases α para $\mathbf{V}$ e β para $\mathbf{W}$, é $(T)^\beta_\alpha$, dada por

$$(T)_\alpha^\beta = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix},$$

de modo que as **colunas** de $(T)_\alpha^\beta$ são as representações na base β das imagens por T dos vetores da base α. Assim, a representação na base β da imagem de $\mathbf{v}$ por T é dada por $(T\mathbf{v})^\beta = (T)_\alpha^\beta (\mathbf{v})^\alpha$, sendo o produto definido por

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1j}x_j + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mj}x_j + \cdots + a_{mn}x_n \end{pmatrix} =$$

$$= x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_j \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

Exercício 12.19 *Sejam* $\mathbf{U}$, $\mathbf{V}$ *e* $\mathbf{W}$ *espaços vetoriais de dimensão finita nos quais fixamos bases (respectivas)* α, β *e* γ. *Sejam* $T : \mathbf{U} \to \mathbf{V}$ *e* $S : \mathbf{V} \to \mathbf{W}$ *transformações lineares. Suponhamos que* $(S)_\beta^\gamma = (a_{ij})$ *e* $(T)_\alpha^\beta = (b_{kl})$. *Mostre que* $(ST)_\alpha^\gamma = (a_{ij})\,(b_{jl})$, *definida por*

$$\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \vdots & \vdots \\ a_{p1} & \dots & a_{pm} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \vdots & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \vdots & \vdots \\ c_{p1} & \dots & a_{pn} \end{pmatrix},$$

com

$$c_{il} = \sum_{j=1}^{m} a_{ij} b_{jl}.$$

Exercício 12.20 *Seja* $S : \mathbb{C}^n \to \mathbb{C}^n$ *dada por* $S(z_1, z_2, \ldots, z_n) = (z_n, z_1, z_2, \ldots, z_{n-1})$. *Exiba a matriz* $(S)_\alpha^\alpha$, *sendo* α *a base canônica de* $\mathbb{C}^n$.

Exercício 12.21 *Considere a transformação* S *do exercício anterior. Encontre todos os números complexos* λ *para os quais existe* $\mathbf{z} \in \mathbb{C}^n$, $\mathbf{z} \neq (0, \ldots, 0)$ *tal que* $S\mathbf{z} = \lambda\mathbf{z}$. *Escolha, para cada um dos tais* λ*s, um* $\boldsymbol{\varepsilon}$, *unitário, tal que* $S\boldsymbol{\varepsilon} = \lambda\boldsymbol{\varepsilon}$. *Mostre que esses* $\boldsymbol{\varepsilon}$*s formam uma base ortonormal de* $\mathbb{C}^n$. *Encontre a matriz de* S *nessa base.*

Capítulo 13

Matrizes de Markov

Importantes por suas aplicações, mas, acima de tudo, fascinantes por seu caráter geométrico, as **matrizes de Markov** têm suas raízes nos processos estocásticos. Vamos abordar aqui apenas um pequeno resultado referente ao assunto.

13.1 Probabilidades de um processo de transição

Nosso ponto de partida é um sistema que evolui aos saltos, mudando de estado a cada intervalo de tempo Δt (Δt é considerado fixo e não tem importância em nossa análise). Existe apenas um número finito de estados possíveis, que designaremos pelos números

$$1,\ 2, \ldots, n.$$

Vamos também nos referir aos tempos,

$$0,\ \Delta t,\ 2\Delta t,\ 3\Delta t, \ldots,\ k\Delta t, \ldots$$

como "tempo 0", "tempo 1", "tempo 2",..., "tempo k", etc.

A informação sobre a evolução do sistema, do tempo k para o tempo $k+1$, é dada pelas probabilidades de transição:

$$p_{ij},\ 1 \leq i, j \leq n,$$

p_{ij} representando a probabilidade de que o sistema, estando no estado j, no tempo k, salte para o estado i, no tempo $k+1$ (p_{ii} representa, claro, a probabilidade de que o sistema, estando em i no tempo k, permaneça em i no tempo $k+1$).

Exercício 13.1 *Assegure-se de que tem claro que, necessariamente,*

$$p_{1j} + p_{2j} + \cdots + p_{nj} = 1,\ \forall j = 1, \ldots, n.$$

Suponhamos, agora, que tenhamos, em um dado tempo, k, as probabilidades de que o sistema esteja em cada um dos possíveis estados, p_i representando a probabilidade de que se encontre no estado i. Temos, é claro,

$$p_1 + \cdots + p_n = 1.$$

Qual será, a partir daí, a probabilidade de que nosso sistema salte para o estado i, no tempo $k+1$?

Ora, como, em princípio, o sistema pode ter vindo de qualquer um dos n estados possíveis, basta observar que, para cada j, a probabilidade de que estivesse, no tempo k, no estado j e tenha saltado deste para i, é $p_j p_{ij}$, de forma que a probabilidade de que venha para o estado i, no tempo $k+1$, é

$$p_1 p_{i1} + p_2 p_{i2} + \cdots + p_n p_{in}.$$

Assim, se considerarmos o **vetor das probabilidades** no tempo k,

$$\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix},$$

então o vetor das probabilidades, no tempo $k+1$, será dado pelo produto matricial

$$\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}.$$

Chamaremos de **matriz de transição**, ou **matriz de Markov** do sistema a matriz $[P]$, dada por

$$[P] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}.$$

Embora $[P]$ possa variar com k, suporemos que $[P]$ é constante (em particular, isto significa que as probabilidades de transição dependem apenas do estado atual do sistema e não da sua história). Assim, se p é o vetor das probabilidades dos diversos estados no tempo 0, então o vetor das probabilidades no tempo k será

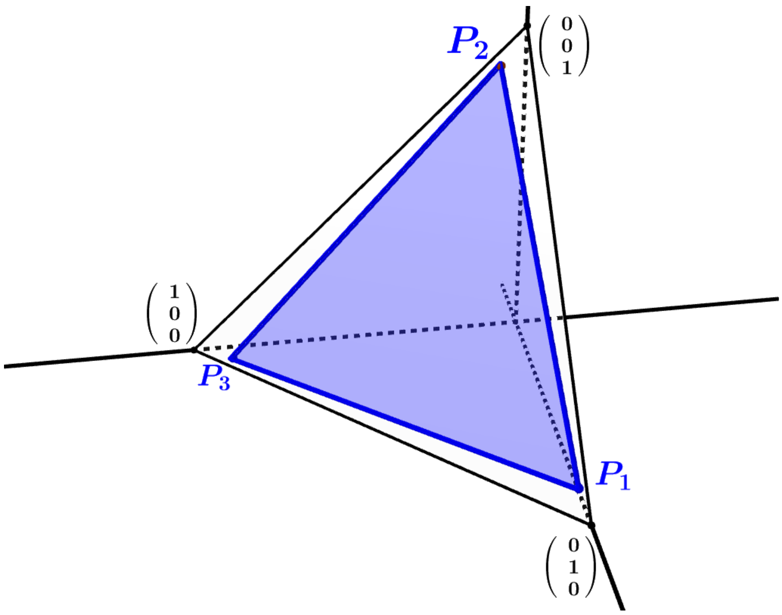
$$[P]^k p.$$

A pergunta que nos colocaremos é, dentro das condições acima, o que podemos dizer sobre a evolução de longo prazo do nosso sistema (faremos, eventualmente, hipóteses adicionais que facilitem a obtenção de alguma resposta).

13.2 Geometria

Embora devamos ter em vista situações em que o número de estados do sistema , n, é grande, vamos fazer uma hipótese radical: $n = 3$. Temos, pois, associada à matriz de transição, uma transformação linear, $P : \mathbb{R}^3 \to \mathbb{R}^3$. Mais ainda: nossos vetores de probabilidades são pontos do plano

$$\alpha = \{p = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} \in \mathbb{R}^3 \mid p_1 + p_2 + p_3 = 1\}.$$



Exercício 13.2 *Note que α é perpendicular ao vetor* $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

E ainda mais: como as probabilidades p_1, p_2 e p_3 são não negativas, nossos vetores de probabilidades estão condenados a viver, eternamente, no triângulo T interseção de α com o primeiro octante, ou seja, se $p \in T$, então $P^k p \in T$, para todo k inteiro positivo.

Exercício 13.3 *Certifique-se de que enxerga que T é o triângulo de vértices em* $\vec{e_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{e_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ *e* $\vec{e_3} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Exercício 13.4 *Mais: note que $T = \{p_1\vec{e_1} + p_2\vec{e_2} + p_3\vec{e_3},\ p_,p_2, p_3 \geq 0,\ p_1 + p_2 + p_3 = 1\}$. Conclua que, como P é linear, $P(T)$ é o triângulo de vértices $P_1 = P\vec{e_1}$, $P_2 = P\vec{e_2}$ e $P_3 = P\vec{e_3}$.*

Agora, que estamos enxergando, algumas coisas ficam óbvias:

1. P é uma transformação linear de $\mathbb{R}^3$ em $\mathbb{R}^3$, com $P(T) \subset T$.

2. Por indução, temos

$$T \supset P(T) \supset P^2(T) \supset \cdots \supset P^k(T) \supset P^{k+1}(T) \supset \cdots .$$

3. O comportamento assintótico do sistema será determinado por informações sobre de que maneira encolhem os triângulos $P^k(T)$. É razoável conjecturar que, sob alguma hipótese adicional, encolham para um ponto $p*$.

Exercício 13.5 *Mostre que, se P permuta os vértices de T, então T não encolhe.*

Há mais uma coisinha quase óbvia para nós, que conhecemos o Teorema de Brouwer (ver o Livro 1):

Teorema de Brouwer: Se $B = \{p \in \mathbb{R}^n \mid p_1^2 + \cdots + p_n^2 \leq 1\}$ e $f : B \to B$ é contínua, então existe $p* \in B$ tal que $f(p*) = p*$.[1]

Teorema de Brouwer para polígonos convexos: Se B é um polígono convexo e $f : B \to B$ é contínua, então existe $p* \in B$ tal que $f(p*) = p*$.

A demonstração da versão do Teorema de Brouwer para polígonos convexos está baseada no seguinte exercício (mais ou menos evidente, mas que pode ser trabalhoso).

Exercício 13.6 *Mostre que existe uma bijeção contínua, $h : B \to T$, com inversa contínua, qualquer que seja o polígono convexo T (tente fazer para o nosso caso, em que T é um triângulo equilátero).*[2]

Demonstração do Brouwer para polígonos convexos: a demonstração só depende da existência de uma função h com as propriedades definidas no exercício; vale, pois, para outros subconjuntos do plano, ou mesmo do espaço. Basta definir, a partir de f e de h, a função $g : B \to B$, dada por $g(x) = h^{-1}(f(h(x)))$. Como g satisfaz as condições do Teorema para o disco, g tem um ponto fixo, $x*$. De $x* = g(x*) = h^{-1}(f(h(x*)))$, segue imediatamente que $f(p*) = p*$, com $p = h(x*)$.*

Assim, da versão estendida do Teorema de Brouwer e dado que transformações lineares são contínuas, temos que existe um ponto $p*$, em T, tal que $P(p*) = p*$. Daí segue, imediatamente, que

$$p* \in P^k(T)\ \forall k = 1, 2, 3, \ldots .$$

Vamos, porém, dispensar o Teorema de Brouwer e introduzir uma condição geométrica, necessária e suficiente, que garantirá o "encolhimento"da sequência $P^n(T)$.

[1]Como estamos em um plano, vamos precisar apenas da versão bidimensional do Teorema

[2]Estamos excluindo polígonos convexos degenerados

13.3 Convergência

O leitor, caso se sinta mais confortável, pode supor que estamos lidando apenas com o caso em que $n = 3$, de forma que tudo vai se passar em um plano e pode ser desenhado. Mas a demonstração, no caso geral, muda apenas pelo raio da maior bola contida em T.

Para facilitar a vida na hora da demonstração, comecemos com uma definição e dois exercícios.

Definição: Se A e B são subconjuntos de $\mathbb{R}^n$ e t é um número real, então o **produto de** A **por** t, tA, e a **soma de** A **e** B, $A + B$ são definidos por

$$tA = \{x \in \mathbb{R}^n \mid \exists (t, a) \in \mathbb{R} \times A,\ x = ta\}$$

$$A + B = \{x \in \mathbb{R}^n \mid \exists (a, b) \in A \times B,\ x = a + b\}$$

Exercício 13.7 *Suponha que $P : \mathbb{R}^n \to \mathbb{R}^n$ é linear, que A e B são subconjuntos de $\mathbb{R}^n$ e que t é um número real. Mostre que $P(A + B) = P(A) + P(B)$ e que $P(tA) = tP(A)$.*

Exercício 13.8 *Sejam x em $\mathbb{R}^n$, t e ε reais positivos. Mostre que, se, de maneira, geral, designamos a **bola de centro** y **e raio** r por $B_r y)$,*

$$B_r(y) = \{z \in \mathbb{R}^n \mid |z - y| < r\},$$

então $B_{t\varepsilon}(x) = x + tB_\varepsilon(0)$.

Lema do Encolhimento: Sejam $P : \mathbb{R}^n \to \mathbb{R}^n$ uma matriz com entradas não negativas,

$$[P] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix},$$

tal que

$$p_{1j} + p_{2j} + \cdots + p_{nj} = 1,\ \forall j = 1, \ldots, n.$$

Seja

$$T = \{p_1\mathbf{e_1} + p_2\mathbf{e_2} + \cdots + p_n\mathbf{e_n} \mid p_1 + p_2 + \cdots + p_n = 1,\ p_1, p_2 \ldots, p_n \geq 0\},$$

ou seja, T é composto pelos vetores

$$\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix},$$

de entradas positivas e tais que

$$p_1 + \cdots + p_n = 1.$$

Nestas condições, são equivalentes:

(i) existe p^* em T tal que $Pp^* = p^*$ e

$$\lim_{k\to\infty} P^k(T) = p^*,$$

ou seja,

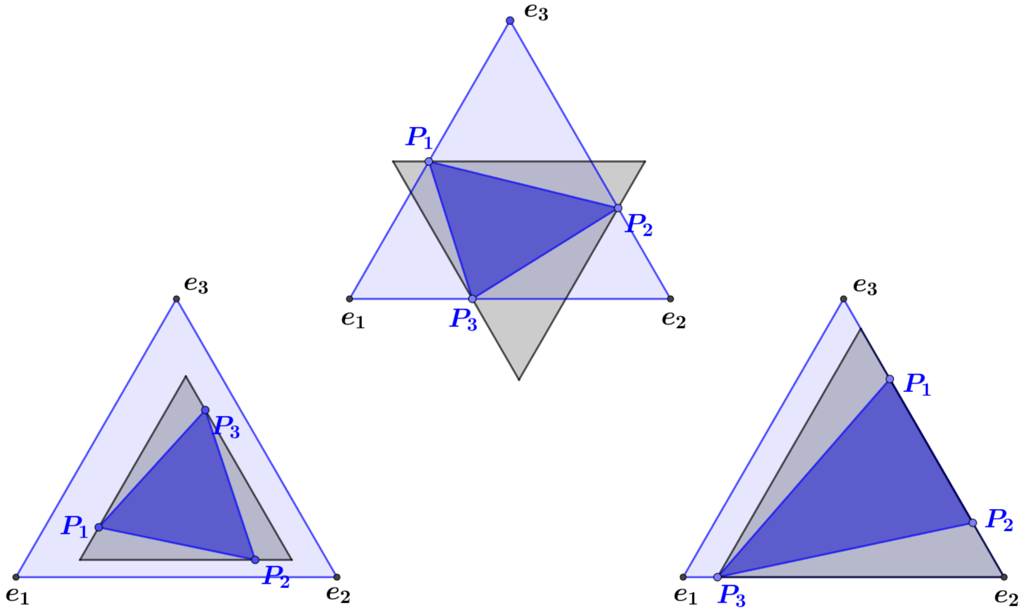
$$\forall \varepsilon > 0\ \exists k_o \in \mathbb{N} \mid k > k_o \ \Rightarrow P^k(T) \subset B_\varepsilon(p^*)$$

sendo $B_\varepsilon(p^*)$ a bola de centro p^* e raio ε;

(ii) para algum escalar α, com $|\alpha| < 1$, se tem

$$P^k(T) \subset b_1 + \alpha T = \{b_1 + \alpha x,\ x \in T\},$$

para um certo inteiro positivo k e um certo b_1 em $\mathbb{R}^n$.



Demonstração: Na figura acima estão ilustrados três possíveis situações em que a condição (ii) é satisfeita, com $k = 1$.

(ii)⇒(i) : Como, para todo natural m, vale $P^{m+1}(T) \subset P^m(T)$, basta provar o resultado para o caso $k = 1$ (faça $P_1 = P^k$; a convergência de $P_1^m(T)$ para p^* implica na de $P^m(T)$) . Suponhamos, pois, que

$$P(T) \subset b_1 + \alpha T = \{b_1 + \alpha x,\ x \in T\},$$

para um certo b_1 em $\mathbb{R}^n$ e um certo escalar α, com $|\alpha| < 1$. A ideia é simples: vamos provar que, para todo natural k, existe b_k tal que

$$P^k(T) \subset b_k + \alpha^k T.$$

O caso $k = 1$ está nas hipóteses; se o resultado vale para k, então

$$P^{k+1}(T) = P(P^k(T)) \subset Pb_k + \alpha^k P(T) \subset (Pb_k + \alpha^k b_1) + \alpha^{k+1} T,$$

o que prova nossa assertiva, com $b_{k+1} = (Pb_k + \alpha^k b_1)$. Desta forma, se $\delta(X)$ representa o diâmetro do conjunto X, temos, para todo natural k,

$$\delta(P^k(T)) \leq |\alpha^k| \delta(T).$$

Em particular, se tomarmos um x_o qualquer em T, a sequência $(P^k x_o)_{k \in \mathbb{N}}$ é de Cauchy e, portanto, convergente. Basta então fazer $p^* = \lim_{k \to \infty} P^k x_o$. A continuidade de P garante $Pp^* = p^*$. Como T é fechado e limitado, cada $P^k(T)$ é fechado, o que garante $p^* \in P^k(T)\ \forall\, k \in \mathbb{N}$. Desta forma, fixado $\varepsilon > 0$, basta tomar k_o em $\mathbb{N}$ tal que $|\alpha^{k_o}|\delta(T) < \varepsilon$. Suponhamos $k > k_o$. Teremos então, se $x \in T$,

$$|P^k x - p^*| = |P^k x - P^k p^*| \leq \delta(P^k(T)) \leq |\alpha^k| \delta(T) < \varepsilon.$$

Para a recíproca, comece com os seguintes exercícios:

Exercício 13.9 *Suponhamos $n \geq 2$. Sejam $a = (n^{-1}, n^{-1}, \ldots, n^{-1}) \in \mathbb{R}^n$ e $\varepsilon = \sqrt{\frac{1}{n(n-1)}}$. Prove que, se $p = (p_1, \ldots, p_n)$ é tal que $p_1 + \cdots + p_n = 1$ e $|p - a| = \sqrt{(p_1 - k^{-1})^2 + \cdots + (p_n - n^{-1})^2} < \varepsilon$, então p está em T.*

(i)⇒(ii) : Sejam $a = (n^{-1}, n^{-1}, \ldots, n^{-1}) \in \mathbb{R}^n$ e $\varepsilon = \sqrt{\frac{1}{n(n-1)}}$. Seja F o subespaço afim de $\mathbb{R}^n$ definido por

$$F = \{p \in \mathbb{R}^n \mid p_1 + \cdots + p_n = 1\}.$$

Pelo exercício acima, a interseção com F da bola de centro a e raio ε, $B_\varepsilon(a) \cap F$, está contida em T. Como tudo se passa em F, isto é,

$$Pp \in F\ \forall\, p \in P,$$

temos que, se $P^k(T) \to p^*$, então, existe k tal que

$$P^k(T) \subset B_{\varepsilon/2}(p^*) \cap F \subset p^* - \frac{1}{2}a + \frac{1}{2}(B_\varepsilon(a) \cap F) \subset p^* - \frac{1}{2}a + \frac{1}{2}T,$$

o que conclui a demonstração.

■

Nos exercícios a seguir, P é matriz de Markov e T é definido por

$$T = \{p_1\mathbf{e_1} + p_2\mathbf{e_2} + \cdots + p_n\mathbf{e_n} \mid p_1 + p_2 + \cdots + p_n = 1,\ p_1, p_2 \ldots, p_n \geq 0\}.$$

Exercício 13.10 *Suponha $n = 3$. Classifique todos os casos em que não existe p^* tal que $P^k(T) \to p^*$. Mostre que, em todos, há um subconjunto I de $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, não vazio, tal que $P(I) = I$.*

Exercício 13.11 *Mostre que, para todo $n \geq 4$ há exemplos de matrizes de MarkovP para as quais não existe p^* tal que $P^k(T) \to p^*$, mas também não há um subconjunto I de $\{\mathbf{e_1}, \mathbf{e_2}, \cdots \mathbf{e_n}\}$, não vazio, tal que $P(I) = I$.*

Exercício 13.12 *A distância do ponto a ao conjunto X é definida por*

$$d(a, X) = inf\{|x - a|, \ x \in X\}.$$

*Mostre que o mínimo de $d(x, \partial T)$, com x em $P^k(T)$, é atingido em um dos vértices de $P^k(T)$ (∂T, **dito bordo de** T, é o conjunto dos p em T com alguma coordenada nula).*

Exercício 13.13 *Mostre que, se todas as entradas de P^k são não nulas, então a condição (ii) do Teorema é satisfeita.*

Teorema: Seja P matriz real $n \times n$, com entradas não negativas e tal que a soma das entradas de cada coluna é 1. Seja T dado por

$$T = \{p_1\mathbf{e_1} + p_2\mathbf{e_2} + \cdots + p_n\mathbf{e_n} \mid p_1 + p_2 + \cdots + p_n = 1, \ p_1, p_2 \ldots, p_n \geq 0\}.$$

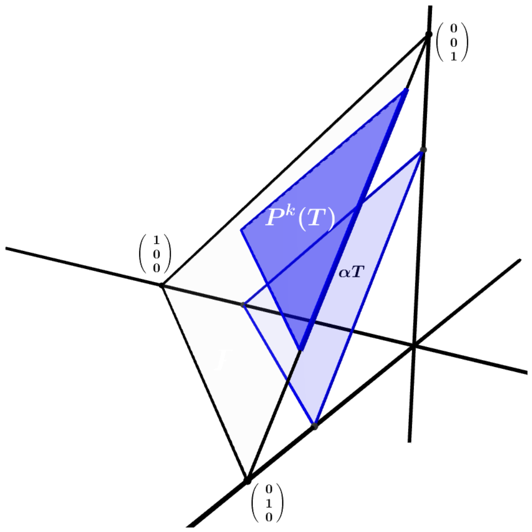
Se, para algum natural k, P^k tem uma linha com todas as entradas não nulas, então $P^m(T)$ converge, quando $m \to \infty$, para um ponto p^* de T. Se além disso, sendo $P^k = [p_{ij}]$, definimos β_i, $i = 1 \ldots, n$, e α por

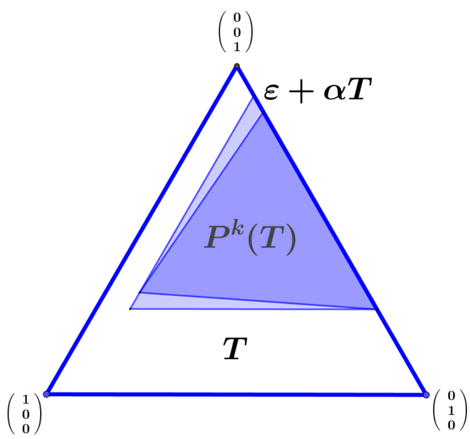
$$\beta_i = min\{p_{ij}, \ j = 1, \ldots, n\},$$

$$\alpha = 1 - \sum_{i=1}^{n} \beta_i,$$

temos, para todo p em T e para todo natural m,

$$|P^{mk}p - p^*| \leq \alpha^m \sqrt{2}.$$





Demonstração: A necessidade é clara, pois p^* não pode ser o vetor nulo. Passemos à suficiência e à estimativa de erro. Podemos desconsiderar o caso $\alpha = 0$, que é trivial. Suporemos, pois, $\alpha > 0$. Sejam. para $i = 1, \ldots, n$, β_i como definidos acima; seja também α como acima. Chamaremos de $\mathbf{e_1}, \ldots, \mathbf{e_n}$ os vetores da base canônica de $\mathbb{R}^n$. Seja, então,

$$\boldsymbol{\varepsilon} = \sum_{i=1}^{n} \beta_i \mathbf{e_i}.$$

A hipótese do Teorema nos diz que, para $p = (p_1, \ldots, p_n)$ em T, vale

$$p \in P^k(T) \Rightarrow p_i \geq \beta_i \; \forall \, i = 1, \ldots, n.$$

Logo, para $p = (p_1, \ldots, p_n)$ em $P^k(T)$, temos

$$p = \sum_{i=1}^{n} p_i \mathbf{e_i} = \boldsymbol{\varepsilon} + \alpha \sum_{i=1}^{n} q_i \mathbf{e_i}, \; com \; q_i = \frac{p_i - \beta_i}{\alpha}, \; i = 1, \ldots, n,$$

o que prova

$$P^k(T) \subset \boldsymbol{\varepsilon} + \alpha T.$$

Como as hipóteses do Teorema garantem que $\alpha < 1$, o Lema de Encolhimento completa a prova. ■

Definição: Uma **face** de dimensão $k - 1$ de T é um subconjunto (não vazio) F de T da forma

$$F = \{p_1 u_1 + \cdots + p_k u_k, \; (p_1, \ldots, p_k) \in [0,1]^k, \; p_1 + \cdots + p_k = 1\},$$

sendo $u_1, \ldots, u_k$ distintos e

$$\{u_1, \ldots, u_k\} \subset \{\mathbf{e_1}, \ldots, \mathbf{e_n}\}.$$

Definição: Uma face F é dita **recorrente** se existe um natural k tal que $P^k(F) \subset F$.

Exercício 13.14 *Mostre que, se existe $p^* \in T$ tal que $P^k(T) \to p^*$, então, a condição a seguir é satisfeita:*

(*) se F e G são faces recorrentes, então $F \cap G \neq \phi$.

Exercício 13.15 *A condição (*) é suficiente para a existência de $p^* \in T$ tal que $P^k(T) \to p^*$?*

Capítulo 14

Mudanças de base

14.1 Um exemplo

Exemplo 1: Rotação. Suponhamos dados um vetor não nulo, $\vec{\varepsilon}$, e um ângulo, θ. Queremos expressar a rotação R, de ângulo θ, em torno da reta que passa pela origem e tem direção $\vec{\varepsilon}$. Supomos que $\vec{\varepsilon}$ seja dado por suas coordenadas na base canônica, (a, b, c). Para simplificar (e não retornar a um caso já tratado), suporemos $a^2 + b^2 + c^2 = 1$ e $\vec{\varepsilon} \neq (0, 0, 1)$.

Nossa primeira observação é: se tivermos dois outros vetores, $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$, unitários, ortogonais a $\vec{\varepsilon}$ e ortogonais entre si (e, se não for pedir muito, tais que a base $(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon})$ tenha orientação positiva), então, escolhendo o *sentido trigonométrico* para a rotação, teremos:

$$R\vec{\varepsilon}_1 = \cos\theta\vec{\varepsilon}_1 + \sin\theta\vec{\varepsilon}_2, \quad R\vec{\varepsilon}_2 = -\sin\theta\vec{\varepsilon}_1 + \cos\theta\vec{\varepsilon}_2, \quad R\vec{\varepsilon} = \vec{\varepsilon}.$$

Exercício 14.1 *Entenda isso. Note que nada muda, se trocarmos a exigência $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$ unitários para $|\vec{\varepsilon}_1| = |\vec{\varepsilon}_2| \neq 0$.*

Exercício 14.2 *Encontre $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$, unitários, ortogonais a $\vec{\varepsilon}$ e ortogonais entre si (e, abusando um pouco, tais que a base $(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon})$ tenha orientação positiva). Exiba as coordenadas de $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$ na base canônica. Sugestão: comece com $\vec{\varepsilon}_1 = (a^2 + b^2)^{-1/2}(-b, a, 0)$; faça $\vec{\varepsilon}_2 = \vec{\varepsilon} \otimes \vec{\varepsilon}_1$.*

Suponhamos, pois, que temos nossos vetores $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$ como acima. Se $\vec{v} = y_1\vec{\varepsilon}_1 + y_2\vec{\varepsilon}_2 + y_3\vec{\varepsilon}$, então

$$R\vec{v} = y_1 R\vec{\varepsilon}_1 + y_2 R\vec{\varepsilon}_2 + y_3 R\vec{\varepsilon} = y_1(\cos\theta\vec{\varepsilon}_1 + \sin\theta\vec{\varepsilon}_2) + y_2(-\sin\theta\vec{\varepsilon}_1 + \cos\theta\vec{\varepsilon}_2) + y_3\vec{\varepsilon},$$

ou seja,

$$R\vec{v} = (y_1\cos\theta - y_2\sin\theta)\vec{\varepsilon}_1 + (y_1\sin\theta + y_2\cos\theta)\vec{\varepsilon}_2 + y_3\vec{\varepsilon}.$$

Podemos, então, concluir que, dadas as coordenadas (y_1, y_2, y_3) de $\vec{v}$ na base $\alpha = (\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon})$, as correspondentes coordenadas (y_1', y_2', y_3') de $R\vec{v}$, na mesma base α, serão dadas por

$$\begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

Exercício 14.3 *Observe que as colunas da matriz* $\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ *são as respectivas coordenadas, na base* α*, de* $R\vec{\varepsilon}_1$*,* $R\vec{\varepsilon}_2$ *e* $R\vec{\varepsilon}$*.*

Vamos chamar a matriz

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

de **matriz da transformação** R **na base** α; vamos notá-la por $(R)_\alpha$. Chamando de β a base canônica, nosso problema, ainda, é obter a matriz de R em relação a β (que chamaremos de $(R)_\beta$).

Exercício 14.4 *Note que o problema estará resolvido se obtivermos dois conversores: o primeiro, que chamaremos de* $(I)^\alpha_\beta$*, converte as coordenadas* (x_1, x_2, x_3) *(do vetor* $\vec{v}$ *na base* β*) nas coordenadas* (y_1, y_2, y_3) *(do vetor* $\vec{v}$ *na base* α*); o segundo, que chamaremos de* $(I)^\beta_\alpha$*, reverte as coordenadas* (y_1', y_2', y_3') *(do vetor* $R\vec{v}$ *na base* α*) nas coordenadas* (x_1', x_2', x_3') *(do vetor* $R\vec{v}$ *na base* β*).*

Já temos as coordenadas, (a, b, c), de $\vec{\varepsilon}$ na base β. Suponhamos que as de $\vec{\varepsilon}_1$ e de $\vec{\varepsilon}_2$ sejam, respectivamente, (a_1, b_1, c_1) e (a_2, b_2, c_2). É fácil, então, obter as coordenadas de $R\vec{v}$ na base canônica, a partir de suas coordenadas na base α:

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = y_1' \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} + y_2' \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} + y_3' \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a \\ b_1 & b_2 & b \\ c_1 & c_2 & c \end{pmatrix} \begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix}.$$

O que acabamos de obter, na verdade, é bem geral: acabamos de criar um método geral para converter as coordenadas de um vetor em uma base qualquer nas coordenadas do mesmo vetor em outra base qualquer. Vamos patenteá-lo, sob forma de proposição.

Proposição: Sejam α e β são duas bases e $\vec{v}$ um vetor de V. Se $(\vec{v})_\alpha$ e $(\vec{v})_\beta$ são, respectivamente, os vetores coluna representando $\vec{v}$ nas bases α e β, então $(\vec{v})_\beta = (I)^\beta_\alpha (\vec{v})_\alpha$, sendo $(I)^\beta_\alpha$ a matriz cujos vetores coluna são os vetores da base α representados na base β.

Voltando ao nosso problema, já sabemos converter coordenadas de vetores na base α para as coordenadas na base canônica. Para converter coordenadas na base canônica para coordenadas na base α, pela Proposição, precisamos das coordenadas dos vetores da base canônica na base α. Teremos, então, a matriz $(I)^{\alpha}_{\beta}$. Não custa nada observar que, para qualquer vetor $\vec{v}$, teremos

$$(\vec{v})_{\alpha} = (I)^{\alpha}_{\beta}(\vec{v})_{\beta} = (I)^{\alpha}_{\beta}((I)^{\beta}_{\alpha}(\vec{v})_{\alpha}) = ((I)^{\alpha}_{\beta}((I)^{\beta}_{\alpha})(\vec{v})_{\alpha},$$

ou seja: $(I)^{\alpha}_{\beta}$ é a inversa de $(I)^{\beta}_{\alpha}$.

Exercício 14.5 *Como você faria para inverter nossa matriz* $\begin{pmatrix} a_1 & a_2 & a \\ b_1 & b_2 & b \\ c_1 & c_2 & c \end{pmatrix}$*?*

Exercício 14.6 *Mostre que é possível obter, diretamente, a matriz da rotação na base canônica.*

Sugestão: dado um vetor $\vec{v}$, comece por projetar $\vec{v}$ sobre o eixo de rotação, obtendo $\vec{v}_{\varepsilon} = \langle \vec{v}, \vec{\varepsilon} \rangle \, \vec{\varepsilon}$; tome $\vec{v}_1 = \vec{v} - \vec{v}_{\varepsilon}$ e $\vec{v}_2 = \vec{\varepsilon} \otimes \vec{v}_1$; faça $R\vec{v} = \vec{v}_{\varepsilon} + \cos\theta \vec{v}_1 + \sin\theta \vec{v}_2$.

14.2 Outros exemplos em $\mathbb{R}^3$

Exemplo 2: Consideremos um cisalhamento perpendicular ao vetor $\vec{\varepsilon}_1$ (que suporemos unitário). Ou seja, planos perpendiculares a $\vec{\varepsilon}_1$ deslizam, sobre si mesmos, na direção de um certo vetor $\vec{\varepsilon}_2$, normal a $\vec{\varepsilon}_1$ (para simplificar, suporemos, também, $\vec{\varepsilon}_2$ unitário). O deslizamento é proporcional à altura do plano (medida na direção de $\vec{\varepsilon}_1$). Fazendo $\vec{\varepsilon}_3 = \vec{\varepsilon}_1 \otimes \vec{\varepsilon}_2$, podemos então definir nosso cisalhamento da seguinte forma:

$$T(y_1\vec{\varepsilon}_1 + y_2\vec{\varepsilon}_2 + y_3\vec{\varepsilon}_3) = y_1\vec{\varepsilon}_1 + y_2\vec{\varepsilon}_2 + y_3\vec{\varepsilon}_3 + ay_1\vec{\varepsilon}_2 = y_1\vec{\varepsilon}_1 + (y_2 + ay_1)\vec{\varepsilon}_2 + y_3\vec{\varepsilon}_3$$

(a é um coeficiente fixo, que indica a intensidade do deslizamento).

Exercício 14.7 *Confira que, na base* $\alpha = (\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3)$*, a matriz da nossa transformação é*

$$\begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Exercício 14.8 *Suponha conhecidas as coordenadas, na base canônica, de* $\vec{\varepsilon}_1$ *e* $\vec{\varepsilon}_2$*. Determine a matriz de* T *na base canônica.*

Exercício 14.9 *Observe que poderíamos também ter definido* T*, diretamente, por* $T\vec{v} = \vec{v} + a \langle \vec{v}, \vec{\varepsilon}_1 \rangle \, \vec{\varepsilon}_2$*.*

Exemplo 3: Nossa transformação, agora, é a **projeção ortogonal sobre um plano** que passa pela origem. O plano, como de hábito, pode ser dado por um vetor normal $\vec{n}$, que suporemos unitário. Tomando, no plano, dois vetores unitários e ortogonais, $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$, teremos, fazendo $\vec{\varepsilon}_3 = \vec{n}$, uma base ortonormal $\alpha = (\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3)$. A matriz da projeção, nessa base, será, então,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Exercício 14.10 *Confira.*

Exercício 14.11 *Suponha dado $\vec{n} = a\vec{e}_1 + b\vec{e}_2 + c\vec{e}_3$, com $a^2 + b^2 + c^2 = 1$. Encontre, no plano α normal a $\vec{n}$ e passando pela origem, dois vetores unitários e ortogonais, $\vec{\varepsilon}_1$ e $\vec{\varepsilon}_2$ (note que as possibilidades são infinitas). Determine a matriz da projeção ortogonal sobre α na base $\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$, $\vec{n}$; a partir daí, obtenha a matriz da projeção na base canônica.*

Exercício 14.12 *Note que o problema pode ser resolvido, diretamente, subtraindo de cada vetor sua componente na direção de $\vec{\varepsilon}_1$.*

Exemplo 4: Numa variante do exemplo anterior, consideremos a **reflexão** através de um plano passando pela origem.

Exercício 14.13 *Usando a mesma base construída no exemplo 3, note que a matriz é*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Exercício 14.14 *Mostre que, a exemplo do caso anterior, neste, também, podemos dar uma solução direta.*

Exercício 14.15 *Resolva os problemas análogos, para projeções sobre e reflexões através de retas passando pela origem.*

14.3 Matrizes ortogonais

A lição, até agora, é: dependendo do problema, pode ser mais simples trabalhar com uma base outra que a canônica; isto pode resultar em obtermos, de forma quase imediata, a matriz da transformação linear na base escolhida. O preço, porém, é que, embora a matriz que converte coordenadas na nova base em coordenadas na base canônica venha de graça, temos que invertê-la, para obter a matriz que converte coordenadas na base canônica em coordenadas na nova base. Inverter matrizes é, em geral, trabalhoso. Mas, em algumas situações (é o caso em todos os exemplos que escolhemos), a nova base não é qualquer base: é, também, ortonormal!

Exercício 14.16 *Suponha que os vetores coluna da matriz (b_{ij}) sejam unitários e, dois a dois, ortogonais. Seja (a_{ij}) a **matriz transposta** de (b_{ij}), isto é, $a_{ij} = b_{ji}\ \forall(i,j)$. Calcule o produto $(a_{ij})(b_{ij})$.*

O leitor pode ter preguiça de fazer o exercício acima, o melhor é incluir no texto a solução... No produto de duas matrizes,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix},$$

procedemos como se multiplicássemos escalarmente cada linha da primeira por cada coluna da segunda. Mas, no nosso caso, as linhas da primeira são as colunas da segunda. Como os vetores coluna da segunda são unitários e, dois a dois, ortogonais, o resultado é a matriz identidade.

Exercício 14.17 *Note que, como sabemos, se as matrizes 3×3 A e B são tais que $AB = I$, então $BA = I$. Mostre que, se os vetores coluna de uma matriz são unitários e, dois a dois, ortogonais, então o mesmo vale para seus vetores linha.*

Definição: A **transposta** da matriz $A = (a_{ij})$ é a matriz $A^T = (b_{ij})$, dada por $b_{ij} = a_{ji}$.

Definição: A matriz 3×3 A é dita **ortogonal** se satisfaz qualquer uma das propriedades equivalentes abaixo:

- Os vetores coluna de A são unitários e, dois a dois, ortogonais
- $A^T A = I$
- $A^{-1} = A^T$
- $AA^T = I$
- Os vetores linha de A são unitários e, dois a dois, ortogonais

O conjunto das matrizes 3×3 ortogonais é chamado de **grupo ortogonal** (3×3) e notado por $O(3)$.

Exercício 14.18 *Mostre que as cinco propriedades acima são, de fato, equivalentes.*

Exercício 14.19 *Mostre que, se A e B são ortogonais, então AB é ortogonal.*

Exercício 14.20 *Seja $T : V \to V$ uma transformação linear que preserva o produto interno, isto é:*

$$\langle T\vec{u}, T\vec{v}\rangle = \langle \vec{u}, \vec{v}\rangle \ \forall \ \vec{u}, \vec{v} \in V.$$

Mostre que a matriz de T na base canônica é ortogonal. Mostre que a matriz de T em qualquer base ortonormal é ortogonal.

14.4 O caso geral

De maneira geral, dados um espaço vetorial $\mathbf{V}$ e duas bases, α e β, de $\mathbf{V}$, a **matriz de mudança de base** de α para β é

$$(I)^{\beta}_{\alpha}.$$

Exercício 14.21 *Veja se entendeu. I é a transformação identidade de $\mathbf{V}$ em $\mathbf{V}$. As colunas de $(I)^{\beta}_{\alpha}$ são dadas pelos vetores de α escritos na base β.*

Obviamente, a matriz de mudança de base de β para α, $(I)^{\alpha}_{\beta}$, é a inversa de $(I)^{\beta}_{\alpha}$.

Um caso particularmente interessante é aquele em que o espaço $\mathbf{V}$ tem produto interno e as duas bases em consideração, α e β, são ortonormais. Nesse caso, se $\mathbf{u}$ e $\mathbf{v}$ são dois vetores em $\mathbf{V}$, representados, em β, por $(\mathbf{u})^{\beta}$ e $(\mathbf{v})^{\beta}$, respectivamente, seu produto escalar é dado, no caso real, por

$$\langle \mathbf{u}, \mathbf{v}\rangle = u_1 v_1 + \ldots + u_n v_n.$$

No caso complexo:

$$\langle \mathbf{u}, \mathbf{v}\rangle = u_1 \bar{v}_1 + \ldots + u_n \bar{v}_n.$$

Exercício 14.22 *Prove isso.*

Assim, os vetores coluna de $(I)^{\beta}_{\alpha}$ são ortogonais (em $\mathbb{R}^n$ ou em $\mathbb{C}^n$, conforme o caso). É, então, imediato, a exemplo do que já discutimos em $\mathbb{R}^3$, que a inversa de $(I)^{\beta}_{\alpha}$ é sua transposta, caso **V** seja espaço vetorial real (ou a transposta conjugada, se **V** é complexo).

Exercício 14.23 *Prove.*

Exercício 14.24 *Seja n um número natural, $n > 1$. Considere u e v em $\mathbb{C}$, distintos e tais que $u^n = v^n = 1$. Faça $\mathbf{u} = (u, u^2, \ldots, u^n)$ e $\mathbf{v} = (v, v^2, \ldots, v^n)$. Mostre que $\langle \mathbf{u}, \mathbf{v} \rangle = 0$*

Notação: A transposta da matriz A é notada por A^T. Caso estejamos trabalhando com escalares complexos, a **transposta conjugada** (cujas entradas são os conjugados das entradas de A^T) é notada por A^* (no caso real, as duas coincidem).

Definição: A matriz $n \times n$ A é dita **ortogonal** se satisfaz qualquer uma das propriedades equivalentes abaixo:

- Os vetores coluna de A são unitários e, dois a dois, ortogonais
- $A^*A = I$
- $A^{-1} = A^*$
- $AA^* = I$
- Os vetores linha de A são unitários e, dois a dois, ortogonais

O conjunto das matrizes $n \times n$ ortogonais é chamado de **grupo ortogonal** ($n \times n$) e notado por $O(n)$.

Exercício 14.25 *Prove que as cinco propriedades da definição são, de fato, equivalentes.*

Exercício 14.26 *Seja n um número natural, $n > 1$. Seja $u = \cos(2\pi/n) + i\sin(2\pi/n)$. Seja $A = [a_{ij}]$ a matriz $n \times n$ dada por*

$$a_{ij} = \frac{u^{ij}}{n}.$$

Mostre que A é ortogonal.

Capítulo 15

Outro exemplo: Polinômios

Nosso propósito, neste capítulo, é apresentar um exemplo, simples, em que diferentes problemas conduzem, naturalmente, à utilização de diferentes bases. Nosso espaço é o dos polinômios de grau inferior a n, que chamaremos de $\mathcal{P}_n$. Vamos trabalhar com polinômios a coeficientes reais, mas não hesitaremos em estender os resultados ao caso complexo.

Começamos observando que $\mathcal{P}_n$ tem uma base natural: $\{1, x, x^2, \ldots, x^{n-1}\}$

15.1 Questão 1: achar polinômio dados seus valores em n pontos

Suponhamos fixados n números distintos, $x_1, \ldots, x_n$; suponhamos, ainda, dados outros n números, que não precisam ser diferentes, $y_1, \ldots, y_n$. Queremos encontrar um polinômio p, de grau inferior a n, tal que $p(x_i) = y_i$, $i = 1, \ldots, n$. Você pode supor que os números dados são reais, mas não faz diferença: o caso complexo, ou mesmo se trabalharmos em outro corpo qualquer, é análogo.

Uma abordagem simples é escrever

$$p(x) = a_{n-1}x^{n-1} + \ldots + a_1 x + a_o,$$

considerar $a_{n-1}, \ldots, a_o$ como incógnitas e, substituindo cada x_i, obter n equações. Matricialmente[1], temos o sistema

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix} \begin{bmatrix} a_o \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Uma outra abordagem, mais esperta, parte da observação seguinte: se multiplicarmos todos os monômios $(x - x_i)$, exceto $(x - x_j)$, teremos um polinômio, de grau $n - 1$,

[1]esta matriz é conhecida como **matriz de Vandermonde**

que se anula em todos os x_i, exceto x_j. Assim, usando o símbolo $\prod$ para **produtório**, podemos definir, para $j = 1, \ldots, n$, os polinômios p_j, dados por

$$p_j(x) = \frac{\prod_{i=1, i\neq j}^{n}(x - x_i)}{\prod_{i=1, i\neq j}^{n}(x_j - x_i)}.$$

É imediato que

$$p_j(x_i) = \begin{cases} 0, \ i \neq j \\ 1, \ i = j \end{cases},$$

de forma que nosso polinômio, p, é dado por

$$p(x) = \sum_{j=1}^{n} y_j p_j(x).$$

15.2 Polinômios de Lagrange, uma nova base para $\mathcal{P}_n$

Resolvido nosso problema, podemos, da solução, extrair algum ensinamento. Visto que escrevemos nosso polinômio como combinação linear dos polinômios p_j (que são n e, todos, de grau $n-1$) é natural perguntar, e imediatamente responder de forma afirmativa, se os p_j não constituiriam uma base para $\mathcal{P}_n$.

Proposição: Dados n números reais (podem ser complexos) distintos, $x_1, \ldots, x_n$, os correspondentes polinômios (ditos de **Lagrange** $\ell_1, \ldots, \ell_n$, dados por

$$\ell_j(x) = \frac{\prod_{i=1, i\neq j}^{n}(x - x_i)}{\prod_{i=1, i\neq j}^{n}(x_j - x_i)},$$

formam uma base para $\mathcal{P}_n$.

Demonstração: Como temos n elementos de $\mathcal{P}_n$, que é de dimensão n, basta provar que são linearmente independentes. Suponhamos, pois, que, para escalares $\alpha_1, \ldots, \alpha_n$, seja identicamente nulo o polinômio

$$p(x) = \alpha_1 \ell_1(x) + \cdots + \alpha_n \ell_n(x).$$

Fazendo, sucessivamente, $x = x_1, \ldots, x = x_n$, obtemos $\alpha_1 = \ldots = \alpha_n = 0$. ■

Corolário: Suponhamos fixados n números distintos, $x_1, \ldots, x_n$; suponhamos, ainda, dados outros n números, que não precisam ser diferentes, $y_1, \ldots, y_n$. Então existe um único polinômio p de grau inferior a n tal que $p(x_i) = y_i$, $i = 1, \ldots, n$.

Demonstração: Acabamos de provar a existência. Se houvesse duas soluções distintas, p_1 e p_2 sua dferença seria um polinômio de grau inferior a n com n raízes, já que se anularia em cada um dos x_i. ■

Exercício 15.1 *Considere, dado o espaço vetorial E, o conjunto E' das transformações lineares de E em $\mathbb{R}$ (ou em $\mathbb{C}$, se for o caso). Mostre que, com as operações naturais, E' é um espaço vetorial. E' é chamado de* ***espaço dual***[2] *de E.*

Exercício 15.2 *Se $v_1, \dots, v_n$ formam base de E, mostre que existe uma única base, $f_1, \dots, f_n$, de E', tal que*

$$f_j(v_i) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}.$$

A base formada pelos f_j é dita ***base dual*** *à formada pelos v_j.*

Exercício 15.3 *Seja E o espaço vetorial das funções de $\mathbb{R}$ em $\mathbb{R}$ (ou de $\mathbb{C}$ em $\mathbb{C}$, se for o caso). Mostre que, para cada x em $\mathbb{R}$ (ou em $\mathbb{C}$), a aplicação de E em $\mathbb{R}$ (ou em $\mathbb{C}$), dada por $f \mapsto f(x)$, é um elemento de E'*

Exercício 15.4 *Note que $\mathcal{P}_n$ é subespaço vetorial do espaço E definido no exercício anterior. Conclua que todo elemento de E', quando restrito a $\mathcal{P}_n$, é elemento do dual de $\mathcal{P}_n$.*

Exercício 15.5 *Sejam $x_1, \dots, x_n$ números reais (ou, eventualmenete, complexos) distintos e $\ell_1, \dots, \ell_n$ os correspondentes polinômios de Lagrange. Mostre que a base dual à formada pelos ℓ_j é a formada pelos $f_j : p \mapsto p(x_j)$.*

15.3 Questão 2: toda curva polinomial é de Bézier?

Dados os pontos $A_o, \dots, A_n$ (usualmente, em $\mathbb{R}^2$; mas podem estar, também, em $\mathbb{R}^3$ ou, a rigor, em qualquer $\mathbb{R}^N$), a correspondente **curva de Bézier**[3] é dada por

$$B(A_o, \dots, A_n)(t) = \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{n-k} A_k,\ t \in [0,1].$$

Olhando para as coordenadas, temos polinômios (de grau, no máximo, n) em t. É natural, então, conjecturarmos que toda curva $c : [0,1] \to \mathbb{R}^N$, dada, em cada coordenada, por um polinômio, seja uma curva de Bézier. Com a experiência adquirida nas seções anteriores, é natural observar que temos, em cada coordenada, uma combinação linear dos polinômios, todos de grau n,

$$B_k(t) = \binom{n}{k} t^k (1-t)^{n-k}.$$

15.4 Polinômios de Bernstein, uma nova base para $\mathcal{P}_n$

Os $n+1$ polinômios $B_{k,n}$, todos de grau n, definidos por

---
[2] eventualmente, de **dual algébrico** de E

[3] Se precisar, dê uma olhada no Livro 1

$$B_{k,n}(t) = \binom{n}{k} t^k (1-t)^{n-k},$$

são ditos os **polinômios de Bernstein** de ordem n.

Proposição: Os polinômios de Bernstein de ordem n formam uma base para $\mathcal{P}_{n+1}$.

Demonstração: Temos $n+1$ elementos de um espaço, $\mathcal{P}_{n+1}$, de dimensão $n+1$. Basta, pois, mostrar que são linearmente independentes. Suponhamos, pois, que, para certos reais $\alpha_o, \ldots, \alpha_n$, seja identicamente nulo o polinômio p dado por

$$p(t) = \sum_{k=0}^{n} \alpha_k \binom{n}{k} t^k (1-t)^{n-k}.$$

Fazendo $t = 0$, obtemos diretamente $\alpha_o = 0$. Em seguida, já com $\alpha_o = 0$ e pondo em evidência o fator comum t, obtemos

$$t \sum_{k=1}^{n} \alpha_k \binom{n}{k} t^{k-1} (1-t)^{n-k},$$

que deve, também, ser identicamente nulo. Ora, isto significa que o somatório deve ser nulo para t diferente de zero. como se trata de um polinômio,

$$\sum_{k=1}^{n} \alpha_k \binom{n}{k} t^{k-1} (1-t)^{n-k}$$

tem que ser identicamente nulo. Fazendo, de novo, $t = 0$, temos $\alpha_1 = 0$. Repetindo indutivamente o raciocínio, obtemos $\alpha_o = \alpha_1 = \cdots = \alpha_n = 0$. ∎

Capítulo 16

O Teorema do Núcleo e da Imagem

16.1 O Teorema do Núcleo e da Imagem em $\mathbb{R}^3$

Nesta seção, **V** representa o espaço dos vetores flechinha, que pode ser identificado a $\mathbb{R}^3$. Se $T : V \to V$ é um isomorfismo, o trepatrepa catalão nos dá uma boa noção de como funciona T. Mas, se o trepatrepa degenera, as coisas ficam menos claras.

Definição: Se $T : V \to V$ é uma transformação linear, a imagem de T é o subespaço vetorial de V definido por $Im(T) = \{T\vec{v},\ \vec{v} \in V\}$.

Exercício 16.1 *Mostre que a imagem de uma transformação linear é, de fato, um subespaço do contradomínio.*

Observação: Se $(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3)$ é uma base qualquer de V, a imagem de T é o conjunto das combinações lineares de $T\vec{\varepsilon}_1$, $T\vec{\varepsilon}_2$ e $T\vec{\varepsilon}_3$. Note que temos, então quatro possibilidades:

- $T\vec{\varepsilon}_1$, $T\vec{\varepsilon}_2$ e $T\vec{\varepsilon}_3$ são linearmente independentes. Neste caso, T é um isomorfismo e $Im(T) = V$. Neste caso, $Im(T)$ é de **dimensão** 3 ($dim Im(T) = 3$).

- Dentre $T\vec{\varepsilon}_1$, $T\vec{\varepsilon}_2$ e $T\vec{\varepsilon}_3$ há dois, mas não três, que são linearmente independentes. Neste caso, $Im(T)$ é o conjunto das combinações lineares de dois vetores e é, portanto, um plano passando pela origem. Neste caso, $Im(T)$ é de **dimensão** 2 ($dim Im(T) = 2$).

- Os três, $T\vec{\varepsilon}_1$, $T\vec{\varepsilon}_2$ e $T\vec{\varepsilon}_3$, estão alinhados (e pelo menos um é não nulo). Então $Im(T)$ é o conjunto dos múltiplos de um vetor; é, portanto, uma reta passando pela origem. Neste caso, $Im(T)$ é de **dimensão** 1 ($dim Im(T) = 1$).

- Caso totalmente degenerado: $Im(T) = \left\{\vec{0}\right\}$. Neste caso, $Im(T)$ é de **dimensão** 0 ($dim Im(T) = 0$).

Consideremos os casos em que a dimensão de $Im(T)$ é menor do que 3. Para mandarmos o V, que é de dimensão 3, em algo de dimensão menor, alguma compressão deve ter sido feita. Um primeiro passo é olharmos para a imagem inversa de um ponto de $Im(T)$, $\vec{w} = T\vec{v}_0$. Como sabemos, se T não é sobrejetiva, não pode ser injetiva. Uma observação simples é: se $T\vec{u} = \vec{0}$, então $T(\vec{v}_0 + \vec{u}) = T\vec{v}_0 + T\vec{u} = T\vec{v}_0 = \vec{w}$.

Proposição: Se $T : V \to V$ é uma transformação linear, seja $N(T) = T^{-1}(\vec{0})$, isto é:

$$N(T) = \left\{ \vec{u} \in V \mid T\vec{u} = \vec{0} \right\}.$$

Se $\vec{v}_0$ é um vetor qualquer tal que $T\vec{v}_0 = \vec{w}$, então a imagem inversa de $\vec{w}$, $T^{-1}(\vec{w}) = \{\vec{v} \in V \mid T\vec{v} = \vec{w}\}$, é dada por

$$T^{-1}(\vec{w}) = \vec{v}_0 + N(T) = \{\vec{v}_0 + \vec{u},\ \vec{u} \in N(T)\}.$$

Demonstração: Se $T\vec{v}_0 = \vec{w}$ e $\vec{v}$ está em $\vec{v}_0 + N(T)$, temos $\vec{v} = \vec{v}_0 + \vec{u}$, para algum $\vec{u}$ em $N(T)$. Mas, então, temos $T\vec{v} = T(\vec{v}_0 + \vec{u}) = T\vec{v}_0 + T\vec{u} = T\vec{v}_0 = \vec{w}$, o que mostra que $\vec{v}$ está em $T^{-1}(\vec{w})$ e prova que $\vec{v}_0 + N(T) \subset T^{-1}(\vec{w})$. Reciprocamente, se $\vec{v}$ está em $T^{-1}(\vec{w})$, temos $T\vec{v} = \vec{w}$. Fazendo $\vec{u} = \vec{v} - \vec{v}_0$, temos $T\vec{u} = T\vec{v} - T\vec{v}_0 = \vec{w} - \vec{w} = \vec{0}$, o que nos dá $\vec{v} = \vec{v}_0 + \vec{u}$, com $\vec{u} \in N(T)$; ou seja, $\vec{v} \in \vec{v}_0 + N(T)$. Isso mostra que $T^{-1}(\vec{w}) \subset \vec{v}_0 + N(T)$, e conclui a prova. ■

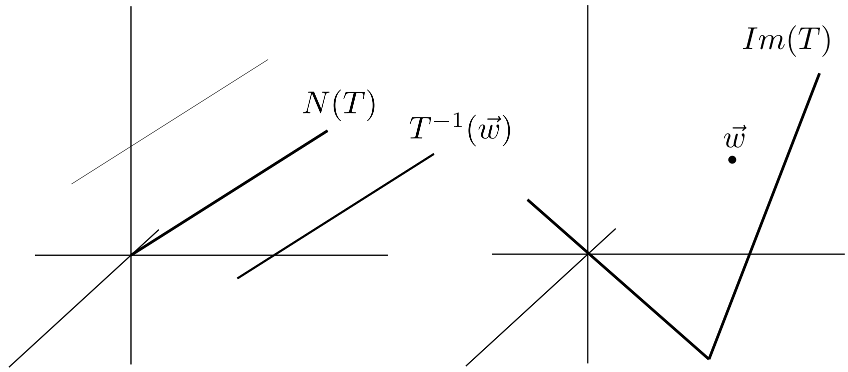


Figura 16.1:

Definição: O conjunto

$$N(T) = \left\{ \vec{u} \in V \mid T\vec{u} = \vec{0} \right\}$$

é chamado de **núcleo** da transformação linear T.

Pelo que acabamos de ver, o espaço V, visto como o domínio da transformação T, é decomposto em subespaços afins paralelos ao núcleo, cada um dos quais é levado em um ponto da imagem. Excluindo os casos em que a transformação é sobrejetiva ou totalmente degenerada, teremos, como veremos a seguir, seja um feixe de retas paralelas, seja um *fatiamento* em planos paralelos.

Para melhor entendermos $N(T)$, consideremos a matriz de T na base canônica. Temos, então, se $(T) = (a_{ij})$ e $\vec{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $T\vec{u}$ é dado por:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + y \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + z \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11}x \\ a_{21}x \\ a_{31}x \end{pmatrix} + \begin{pmatrix} a_{12}y \\ a_{22}y \\ a_{32}y \end{pmatrix} + \begin{pmatrix} a_{13}z \\ a_{23}z \\ a_{33}z \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{pmatrix}.$$

Assim, para que $\vec{u}$ esteja no núcleo de T, é preciso que suas coordenadas satisfaçam às três equações:

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = 0 \\ a_{21}x + a_{22}y + a_{23}z = 0 \\ a_{31}x + a_{32}y + a_{33}z = 0 \end{cases}$$

Exercício 16.2 *Mostre que $N(T)$ é o complemento ortogonal do subespaço gerado pelos vetores linha da matriz de T na base canônica.*

Em princípio, cada uma das três equações define um plano passando pela origem (a menos que os três coeficientes sejam nulos, o que daria $0 = 0$). Como $N(T)$ é a interseção, temos as seguintes possibilidades:

- Os planos se interceptam em um ponto (que, necessariamente é $\vec{0}$, já que todos passam pela origem). Neste caso, a dimensão de $N(T)$ é 0 ($dimN(T) = 0$).
- Os planos se interceptam em uma reta passando pela origem: $N(T)$ é uma reta passando pela origem. Neste caso, a dimensão de $N(T)$ é 1 ($dimN(T) = 1$).
- Os planos são coincidentes: $N(T)$ é um plano passando pela origem. Neste caso, a dimensão de $N(T)$ é 2 ($dimN(T) = 2$).
- Caso totalmente degenerado: todos as entradas da matriz de T são nulas, as três equações são $0 = 0$, $N(T) = V$. Neste caso, a dimensão de $N(T)$ é 3 ($dimN(T) = 3$).

É interessante observar que o primeiro caso corresponde a T ser um isomorfismo, já que $T\vec{e_1}$, $T\vec{e_2}$ e $T\vec{e_3}$ terão que ser linearmente independentes (caso contrário, teríamos $\vec{0} = xT\vec{e_1} + yT\vec{e_2} + zT\vec{e_3}$, para algum terno $(x, y, z) \neq (0, 0, 0)$, o que daria um elemento não nulo, $x\vec{e_1} + y\vec{e_2} + z\vec{e_3}$ em $N(T)$). Por acaso, isto nos dá, como $dim\ V = 3$,

$$dim\ N(T)\ +\ dim\ Im(T)\ =\ dim\ V.$$

Exercício 16.3 *Mostre que, também no caso em que* $dimN(T) = 3$*, temos*

$$dim\ N(T)\ +\ dim\ Im(T)\ =\ dim\ V.$$

Note que é razoável imaginar que a fórmula acima seja geral: para jogar V, que é tridimensional, em um plano, que é bidimensional (ou seja, $dim\ Im(T) = 2$), devemos *perder* uma dimensão, possivelmente jogando retas do domínio em pontos da imagem, o que significa que o núcleo deve ser unidimensional; se, por outro lado, a imagem de T for unidimensional, então devemos *perder* duas dimensões, jogando planos em pontos da imagem, e o núcleo deve ser bidimensional.

Teorema do Núcleo e da Imagem:[1] Se $T : V \to V$ é uma transformação linear, então

$$dim\ N(T)\ +\ dim\ Im(T)\ =\ dim\ V.$$

Demonstração: Os casos em que $dimN(T) = 0$ ou $dimN(T) = 3$ são simples, já fizemos.

Suponhamos, pois, caso 1, que $dimN(T) = 1$. Podemos, então, tomar um vetor não nulo, $\vec{u}$, em $N(T)$. O Lema Fundamental nos garante que podemos tomar dois vetores, $\vec{v}_1$ e $\vec{v}_2$, tais que $(\vec{u}, \vec{v}_1, \vec{v}_2)$ seja base de V (podemos, inclusive, exigir que $\vec{v}_1$ e $\vec{v}_2$ sejam uma base do complemento ortogonal de $N(T)$). Afirmamos que $T\vec{v}_1$ e $T\vec{v}_2$ são linearmente independentes e geram $Im(T)$ (isto é $Im(T)$ é o conjunto das combinações lineares de $T\vec{v}_1$ e $T\vec{v}_2$). De fato, como $T\vec{u}$, $T\vec{v}_1$ e $T\vec{v}_2$ geram $Im(T)$ e $T\vec{u} = \vec{0}$, temos que $T\vec{v}_1$ e $T\vec{v}_2$ geram $Im(T)$. Resta provar que $T\vec{v}_1$ e $T\vec{v}_2$ são linearmente independentes. Se não fossem, um deles, digamos $T\vec{v}_2$, seria múltiplo do outro. Mas, então, supondo $T\vec{v}_2 = tT\vec{v}_1$, teríamos $T(\vec{v}_2 - t\vec{v}_1) = \vec{0}$, o que daria $\vec{v}_2 - t\vec{v}_1 \in N(T)$. Logo, para algum escalar s, teríamos $\vec{v}_2 - t\vec{v}_1 = s\vec{u}$, e $\vec{u}$, $\vec{v}_1$ e $\vec{v}_2$ não seriam linearmente independentes. Portanto, $dim\ Im(T) = 2$.

Suponhamos agora, caso 2, que $dimN(T) = 2$. Podemos, então tomar dois vetores linearmente independentes, $\vec{u}_1$ e $\vec{u}_2$, tais que $N(T)$ seja o plano gerado por $\vec{u}_1$ e $\vec{u}_2$. De novo, usando o Lema Fundamental, podemos completar $\vec{u}_1$ e $\vec{u}_2$ por um vetor $\vec{v}$, que podemos supor um gerador de $N(T)^{\perp}$, de forma que $\vec{u}_1$ e $\vec{u}_2$ e $\vec{v}$ formem base de V. Como no caso anterior, temos que $T\vec{v}$ gera $Im(T)$. Além disso, $T\vec{v}$ não pode ser nulo, ou $\vec{v}$ estaria em $N(T)$ e seria combinação linear de $\vec{u}_1$ e $\vec{u}_2$. Logo, $dimIm(T) = 1$. ■

Escólio: Nossa demonstração contém mais do que o prometido no enunciado. Provamos que a imagem de T tem a mesma dimensão que o complemento ortogonal do núcleo e que T leva os elementos de $N(T)^{\perp}$ bijetivamente nos de $Im(T)$. Suponha que a matriz de T na base canônica seja

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Já observamos que $Im(T)$ é o subespaço de V gerado pelos vetores coluna da matriz. Mas temos também que o núcleo de T é o conjunto dos vetores ortogonais aos vetores

[1] O que aqui estamos enunciando é uma versão simplificada de um teorema mais geral, que veremos logo à frente. Achamos melhor manter o nome, *do Núcleo e da Imagem*, e o título de nobreza, *Teorema*

linha da matriz. Na verdade, temos que o complemento ortogonal de $N(T)$ é o subespaço gerado pelos vetores linha da matriz. Assim, a dimensão do subespaço gerado pelas colunas da matriz é igual à do subespaço gerado por suas linhas.

16.2 O caso geral

Definição: Sejam $\mathbf{V}$ e $\mathbf{W}$ espaços vetoriais (reais ou complexos) e $T : \mathbf{V} \to \mathbf{W}$ uma transformação linear. O **núcleo** de T é o subespaço $N(T)$ de $\mathbf{V}$, definido por

$$N(T) = T^{-1}(\mathbf{0}) = \{\mathbf{v} \in \mathbf{V} \mid T\mathbf{v} = \mathbf{0}\}.$$

A **imagem** de T é o subespaço $Im(T)$ de $\mathbf{V}$, definido por

$$Im(T) = T(\mathbf{V}) = \{T\mathbf{v},\ \mathbf{v} \in \mathbf{V}\}.$$

É imediato que, se, para um certo $\mathbf{u} \in \mathbf{V}$ e um certo $\mathbf{w} \in \mathbf{W}$, se tem $T\mathbf{u} = \mathbf{w}$, então

$$T^{-1}(\mathbf{w}) = \mathbf{u} + N(T) = \{\mathbf{u} + \mathbf{v} \mid \mathbf{v} \in N(T)\}$$

(é claro que $T(\mathbf{u} + \mathbf{v}) = T\mathbf{u} + T\mathbf{v} = \mathbf{w} + \mathbf{0} = \mathbf{w}$, $\forall \mathbf{v} \in N(T)$; reciprocamente, se $\mathbf{u}_1 \in T^{-1}(\mathbf{w})$, então $T(\mathbf{u}_1 - \mathbf{u}) = \mathbf{w} - \mathbf{w} = \mathbf{0}$, o que mostra que $\mathbf{u}_1 - \mathbf{u} \in N(T)$, e $\mathbf{u}_1 = \mathbf{u} + (\mathbf{u}_1 - \mathbf{u})$).

Exercício 16.4 *Mostre que o núcleo é, de fato, subespaço vetorial de* $\mathbf{V}$. *Mostre que a imagem é subespaço de* W.

Consideremos, no caso geral, dois espaços vetoriais (podem ser reais ou complexos, nada muda), $\mathbf{V}$ e $\mathbf{W}$ e uma transformação linear, T, de $\mathbf{V}$ em $\mathbf{W}$. Vamos supor que $\mathbf{V}$ é de dimensão finita (o que assegura que $Im(T) = \{Tv,\ v \in \mathbf{V}\}$, também, é de dimensão finita).

Teorema do Núcleo e da Imagem: Se $T : \mathbf{V} \to \mathbf{W}$ é linear e $\mathbf{V}$ é de dimensão finita, então

$$dim\ N(T)\ +\ dim\ Im(T)\ =\ dim\ \mathbf{V}.$$

Demonstração: Como $N(T)$ é, forçosamente, de dimensão finita, podemos tomar uma base, $\alpha = \{\varepsilon_1, \ldots, \varepsilon_k\}$, de $N(T)$. Podemos ainda, graças a nosso velho e bom Lema Fundamental (que vale, sem alterações, para espaços complexos), complementá-la com $n - k$ vetores, $\varepsilon_{k+1}, \ldots, \varepsilon_n$, de forma que $\beta = \{\varepsilon_1, \ldots, \varepsilon_n\}$ seja base de $\mathbf{V}$. Seja, então, $\mathbf{V}_1$ o espaço gerado por $\varepsilon_{k+1}, \ldots, \varepsilon_n$ (note que $dim\ \mathbf{V}_1 = n - k = dim\ \mathbf{V} - dim\ N(T)$).[2] Consideremos, agora, os vetores $\mathbf{w}_1 = T\varepsilon_{k+1}, \ldots, \mathbf{w}_{n-k} = T\varepsilon_n$ em $Im(T) \subset \mathbf{W}$. Vamos provar que $\mathbf{w}_1, \ldots, \mathbf{w}_{n-k}$ formam base de $Im(T)$, o que encerrará a demonstração.

[2] Neste caso, $\mathbf{V}$ é **soma direta** de $N(T)$ com $\mathbf{V}_1$, isto é, todo elemento de $\mathbf{V}$ se obtém, de forma única, como soma de um de $N(T)$ com um de $\mathbf{V}_1$. Se $\mathbf{V}$ for espaço com produto interno, podemos nos dar ao luxo de escolher $\mathbf{V}_1 = N(T)^{\perp} = \{\mathbf{v} \in V \mid \langle \mathbf{v}, n \rangle = 0\ \forall\, n \in N(T)\}$, isto é, o complemento ortogonal de $N(T)$)

1. Se $t_1\mathbf{w}_1 + \ldots + t_{n-k}\mathbf{w}_{n-k} = \mathbf{0}$, então $T(t_1\varepsilon_{k+1} + \ldots + t_{n-k}\boldsymbol{\varepsilon_n}) = \mathbf{0}$, o que significa que $t_1\varepsilon_{k+1} + \ldots + t_{n-k}\boldsymbol{\varepsilon_n} \in N(T)$. Ora, isto só é possível se $t_1 = \ldots = t_{n-k} = 0$, o que mostra que $\mathbf{w}_1, \ldots, \mathbf{w}_{n-k}$ são linearmente independentes.

2. Se $\mathbf{w} \in Im(T)$, podemos tomar $\mathbf{v}$ em $\mathbf{V}$ tal que $T\mathbf{v} = \mathbf{w}$. Fazendo $\mathbf{v} = t_1\boldsymbol{\varepsilon_1} + \ldots + t_n\boldsymbol{\varepsilon_n}$, temos $\mathbf{w} = T\mathbf{v} = t_1T\boldsymbol{\varepsilon_1} + \ldots + t_kT\boldsymbol{\varepsilon_k} + t_{k+1}\mathbf{w}_1 + \ldots + t_n\mathbf{w}_{n-k} = t_{k+1}\mathbf{w}_1 + \ldots + t_n\mathbf{w}_{n-k}$, o que encerra a demonstração. ■

Escólio: Nossa demonstração contou com o apoio de um subespaço auxiliar $\mathbf{V}_1$ de V, tal que $dim\mathbf{V}_1 + \dim N(T) = dim\mathbf{V}$ e $dim\mathbf{V}_1 = dim\ Im(T)$. Como já observamos, caso $\mathbf{V}$ tenha produto interno, podemos escolher $\mathbf{V}_1 = N(T)^{\perp}$. Assim, quando $\mathbf{V}$ tem produto interno, nosso teorema também diz que $dim\ Im(T) = dim N(T)^{\perp}$.

Suponhamos que $\mathbf{V} = \mathbb{R}^n$, que $\mathbf{W} = \mathbb{R}^m$ e que T seja a transformação linear dada pala matriz $m \times n$ $[a_{ij}]$. Ora, neste caso, a imagem de T é o espaço gerado pelos vetores coluna da matriz e o núcleo de T é o espaço dos vetores que são ortogonais aos vetores linha de $[a_{ij}]$.

Exercício 16.5 *Prove as duas últimas afirmações acima.*

Corolário: Seja $[a_{ij}]$ matriz $m \times n$ a coeficientes reais. Sejam E o subespaço de $\mathbb{R}^n$ gerado pelas linhas de $[a_{ij}]$ e F o subespaço de $\mathbb{R}^m$ gerado pelas colunas de $[a_{ij}]$. Então $dimE = dimF$.

Demonstração: Basta notar que $F = Im(T)$ e que $N(T) = E^{\perp}$, ou seja: $E = N(T)^{\perp}$ e $F = Im(T)$. ■

Exercício 16.6 *Seja $[a_{ij}]$ matriz $m \times n$ a coeficientes complexos. Mostre que o núcleo da transformação linear $T : \mathbb{C}^n \to \mathbb{C}^m$ dada por $[a_{ij}]$ é o espaço ortogonal ao gerado pelas linhas de $[\bar{a}_{ij}]$ (a barra indicando conjugação complexa) Observe que isto equivale a ser $N(T)$ ortogonal à imagem da transformação linear $T^* : \mathbb{C}^m \to \mathbb{C}^n$ dada pela matriz transposta de $[\bar{a}_{ij}]$.*

Capítulo 17

O determinante, de novo

17.1 Determinante de transformação linear em $\mathbb{R}^3$

Talvez você tenha pulado a seção "O trepa-trepa catalão"...Volte a ela (página 90) e ataque o exercício que está no final.

Uma das propriedades das transformações lineares é tratar de maneira uniforme o espaço: a imagem de um cubinho de centro na origem é congruente à do mesmo cubinho transladado para outro ponto qualquer do espaço. Mais precisamente: sejam V espaço vetorial real de dimensão três e $T : V \to V$ linear. Seja X um subconjunto qualquer de V. Suponhamos que transladem todos os pontos de X, usando um vetor $\vec{v}_0$, obtendo o conjunto X_0:

$$X_0 = \{x + \vec{v}_0,\ x \in X\}\,.$$

Como a translação preserva distâncias e ângulos, X_0 é congruente a X. Mas o mesmo acontece com $T(X)$ e $T(X_0)$:

$$\begin{aligned} T(X_0) &= \{Tx_0,\ x_0 \in X_0\} = \{T(x + \vec{v}_0),\ x \in X\} = \\ &= \{Tx + T\vec{v}_0,\ x \in X\} = \{y + T\vec{v}_0,\ y \in\ T(X)\}\,. \end{aligned}$$

Da mesma forma, se X_1 é a imagem de X por uma homotetia de razão r,

$$X_1 = \{rx,\ x \in X\}\,,$$

então $T(X_1)$ é a imagem de $T(X)$ pela mesma homotetia:

$$\begin{aligned} T(X_1) &= \{Tx_1,\ x_1 \in X_1\} = \{T(rx),\ x \in X\} = \\ &= \{r(Tx),\ x \in X\} = \{ry,\ y \in T(X)\}\,. \end{aligned}$$

Exercício 17.1 *Mostre, com um exemplo, que o mesmo não acontece com figuras rodadas: se rodarmos um cubo Q, obtendo o cubo Q_0, então Q e Q_0 são congruentes, mas um cisalhamento T pode nos dar $T(Q)$ e $T(Q_0)$ não congruentes.*

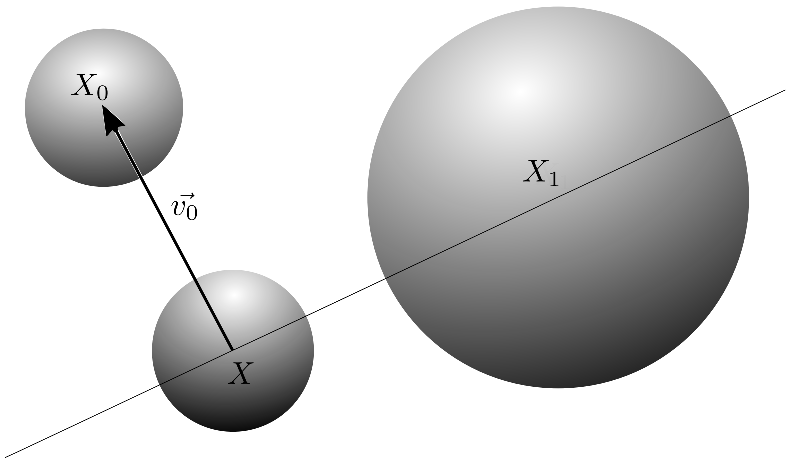


Figura 17.1:

As observações acima devem ser suficientes para que o leitor se convença da veracidade do seguinte resultado:

Teorema: Se $T : V \to V$ é uma transformação linear, então existe um número α tal que, para todo *sólido* $X \subset V$, se tem *volume de* $T(X) = \alpha$ *volume de* X.

Demonstração: Entendemos por *sólido*, sem entrar em demasiados detalhes, qualquer subconjunto X de V cujo volume possa ser calculado por meio de aproximações, *por dentro e por fora*: as aproximações por dentro consistem em colocar, dentro de X, cubinhos cujos interiores não se interceptem; as aproximações por fora consistem em colocar X dentro de uma união de cubinhos cujos interiores não se inteceptam.[1]

Pelo que acabamos de ver, se fizermos, para X, aproximações, por dentro e por fora, usando cubinhos de lados paralelos aos vetores da base canônica, teremos, automaticamente, aproximações para $T(X)$, por dentro e por fora, usando os paralelepípedos imagens dos cubinhos usados para X. Mas pelo que aprendemos sobre determinantes, a razão entre o volume do paralelepípedo imagem e o volume do cubinho original é $|det(T\vec{e_1}, T\vec{e_2}, T\vec{e_3})|$, qualquer que seja o cubinho (veja se acredita mesmo). Tomando supremos e ínfimos (e roubando um pouquinho, ou gastando mais tempo e espaço para cuidar dos detalhes), concluímos que *volume de* $T(X) = \alpha$ *volume de* X, com $\alpha = |det(T\vec{e_1}, T\vec{e_2}, T\vec{e_3})|$. ■

Escólio: Pelo que acabamos de ver, a razão α entre os volumes de $T(X)$ e de X é dada por $\alpha = |det(T\vec{e_1}, T\vec{e_2}, T\vec{e_3})|$. É um pouco desagradável termos esse número atrelado à base canônica. No entanto, é imediato que, se $\vec{u}$, $\vec{v}$ e $\vec{w}$ são linearmente independentes, então $|det(T\vec{u}, T\vec{v}, T\vec{w})| = \alpha|det(\vec{u}, \vec{v}, \vec{w})|$. É claro, também, que o sinal de $det(T\vec{e_1}, T\vec{e_2}, T\vec{e_3})$ *indica se* T *está preservando ou invertendo a orientação do espaço*.

[1] Fica entendido que os cubinhos terão lados paralelos aos vetores da base canônica. Podemos, eventualmente, ter sólidos degenerados, que terão volume nulo; neste caso, as aproximações por dentro são feitas por zero cubinhos. A hipótese é que o ínfimo dos volumes das aproximações por fora é igual ao supremo dos volumes das aproximações por dentro

Lembramos que a transformação linear T é um **isomorfismo** se é bijetiva, o que equivale a dizer que a imagem por T de qualquer base de V é uma base de V.

Proposição: Suponhamos que a transformação linear $T : V \to V$ seja um isomorfismo e que as bases $(\vec{u}_1, \vec{v}_1, \vec{w}_1)$ e $(\vec{u}_2, \vec{v}_2, \vec{w}_2)$ tenham a mesma orientação (isto é, $det(\vec{u}_1, \vec{v}_1, \vec{w}_1)$ e $det(\vec{u}_2, \vec{v}_2, \vec{w}_2)$ tenham o mesmo sinal). Então $(T\vec{u}_1, T\vec{v}_1, T\vec{w}_1)$ e $(T\vec{u}_2, T\vec{v}_2, T\vec{w}_2)$ têm a mesma orientação (ou seja, o sinal de $det(T\vec{u}_1, T\vec{v}_1, T\vec{w}_1)$ e $det(T\vec{u}_2, T\vec{v}_2, T\vec{w}_2)$ é o mesmo).

Demonstração: Se $(\vec{u}_1, \vec{v}_1, \vec{w}_1)$ $(\vec{u}_2, \vec{v}_2, \vec{w}_2)$ têm a mesma orientação, podemos fazer uma deformação $(\vec{\varepsilon}_1(t), \vec{\varepsilon}_2(t), \vec{\varepsilon}_3(t))$, $t \in [a,b]$, começando em $(\vec{u}_1, \vec{v}_1, \vec{w}_1)$ e terminando em $(\vec{u}_2, \vec{v}_2, \vec{w}_2)$, de forma que, para todo t em $[a,b]$, $(\vec{\varepsilon}_1(t), \vec{\varepsilon}_2(t), \vec{\varepsilon}_3(t))$ é uma base. Mas, então, $(T\vec{\varepsilon}_1(t), T\vec{\varepsilon}_2(t), T\vec{\varepsilon}_3(t))$, $t \in [a,b]$ é uma deformação começando em $(T\vec{u}_1, T\vec{v}_1, T\vec{w}_1)$ e terminando em $(T\vec{u}_2, T\vec{v}_2, T\vec{w}_2)$ (note que, como cada $\vec{\varepsilon}_i(t)$ é contínua, as correspondentes $T\vec{\varepsilon}_i(t)$ também são contínuas; além disso, como acabamos de observar, o fato de ser T um isomorfismo garante que $(T\vec{\varepsilon}_1(t), T\vec{\varepsilon}_2(t), T\vec{\varepsilon}_3(t))$ é uma base, para todo t em $[a,b]$). ■

Agora já estamos suficientemente motivados para definir o **determinante de uma transformação linear**.

Definição: Se $T : V \to V$ é uma transformação linear e $(\vec{u}, \vec{v}, \vec{w})$ é uma base qualquer de V, o **determinante** de T é a razão entre $det(T\vec{u}, T\vec{v}, T\vec{w})$ e $det(\vec{u}, \vec{v}, \vec{w})$. Notação: $det(T)$.

Observação: Note que, pela Proposição 2 da página 70, $det(\vec{u}, \vec{v}, \vec{w})$ é não nulo sempre que $\vec{u}, \vec{v}$ e $\vec{w}$ são linearmente independentes.

Proposição: Se S e T são transformações lineares de V em V, então $det(ST) = det(S)det(T)$.

Demonstração: Suponhamos, primeiro, que T seja um isomorfismo.Neste caso, se $(\vec{u}, \vec{v}, \vec{w})$ é uma base qualquer de V, $(T\vec{u}, T\vec{v}, T\vec{w})$ será, também, base de V. Logo,

$$\frac{det(ST\vec{u}, ST\vec{v}, ST\vec{w})}{det(\vec{u}, \vec{v}, \vec{w})} = \frac{det(S(T\vec{u}), S(T\vec{v}), S(T\vec{w}))}{det(T\vec{u}, T\vec{v}, T\vec{w})} \frac{det(T\vec{u}, T\vec{v}, T\vec{w})}{det(\vec{u}, \vec{v}, \vec{w})},$$

e o resultado vale. Caso T não seja isomorfismo, teremos $det(T) = 0$ e, como ST também não será isomorfismo, $det(ST) = 0$. Logo, independente de como seja S, teremos $det(ST) = det(S)det(T)$. ■

Deixamos como exercícios um conjunto de propriedades importantes do determinante.

Exercício 17.2 *Mostre que $det(T) = 0$ se, e somente se, T não é isomorfismo.*

Exercício 17.3 *Mostre que, se T é isomorfismo, então $det(T^{-1}) = (det(T))^{-1}$.*

Exercício 17.4 *Seja $(T)_\beta$ a matriz da transformação T na base canônica. Mostre que $det(T) = det(T)_\beta$.*

Exercício 17.5 *Sejam A e B duas matrizes 3×3. Mostre que $det(AB) = det(A)det(B)$.*

Exercício 17.6 *Seja* $(T)_\alpha$ *a matriz da transformação* T *em uma base qualquer, que designaremos por* α*. Seja* $(I)^\alpha_\beta$ *a matriz cujas colunas são as coordenadas dos vetores da base canônica na base* α*. Mostre que a matriz de* T *na base canônica,* $(T)_\beta$*, é dada por* $(T)_\beta = ((I)^\alpha_\beta)^{-1}(T)_\alpha(I)^\alpha_\beta$*.*

Exercício 17.7 *Seja* $(T)_\alpha$ *a matriz da transformação* T *em uma base qualquer, que designaremos por* α*. Mostre que* $det(T) = det(T)_\alpha$*.*

Está tudo muito bom, mas, pensando bem, nossa definição repousa sobre um Teorema que não é trivial e, principalmente, cuja demonstração deixou pontos nebulosos.[2] A próxima seção tem por objetivo apresentar a mesma definição em termos puramente algébricos, sem recorrer ao Teorema *bonitinho mas suspeito*.

17.2 Formas trilineares alternadas

Em nossa primeira discussão (capítulo 4), o determinante foi apresentado como uma função ω que, a cada terno ordenado $(\vec{u},\vec{v},\vec{w})$, associa um número, $\omega(\vec{u},\vec{v},\vec{w})$, com as propriedades:

$$\begin{array}{ll}
(i) & \omega(\vec{u},\vec{v},\vec{w}) = 0, \textit{ se } \vec{u},\vec{v} \textit{ e } \vec{w} \textit{ são linearmente dependentes} \\
(ii) & \omega(\vec{u}_1 + t\vec{u}_2,\vec{v},\vec{w}) = \omega(\vec{u}_1,\vec{v},\vec{w}) + t\,\omega(\vec{u}_2,\vec{v},\vec{w})\ \forall t,\vec{u}_1,\vec{u}_2,\vec{v},\vec{w} \\
 & \omega(\vec{u},\vec{v}_1 + t\vec{v}_2,\vec{w}) = \omega(\vec{u},\vec{v}_1,\vec{w}) + t\,\omega(\vec{u},\vec{v}_2,\vec{w})\ \forall t,\vec{u},\vec{v}_1,\vec{v}_2,\vec{w} \\
 & \omega(\vec{u},\vec{v},\vec{w}_1 + t\vec{w}_2) = \omega(\vec{u},\vec{v},\vec{w}_1) + \ (\vec{u},\vec{v},\vec{w}_2 t)\ \forall t,\vec{u},\vec{v},\vec{w}_1,\vec{w}_2
\end{array}$$

Naquela ocasião observamos que, das propriedades (i) e (ii), segue que ω troca de sinal, se trocamos de posição dois dos vetores. Assim, chamamos uma função com as propriedades (i) e (ii) de **forma trilinear alternada**. Em seguida, escrevendo os vetores $\vec{u}$, $\vec{v}$ e $\vec{w}$ na base canônica,

$$\begin{aligned}
\vec{u} &= a_{11}\vec{e_1} + a_{21}\vec{e_2} + a_{31}\vec{e_3}, \\
\vec{v} &= a_{12}\vec{e_1} + a_{22}\vec{e_2} + a_{32}\vec{e_3}, \\
\vec{w} &= a_{13}\vec{e_1} + a_{23}\vec{e_2} + a_{33}\vec{e_3},
\end{aligned}$$

e usando as propriedades (i) e (ii), obtivemos

$$\omega(\vec{u},\vec{v},\vec{w}) = \omega(a_{11}\vec{e_1} + a_{21}\vec{e_2} + a_{31}\vec{e_3}, a_{12}\vec{e_1} + a_{22}\vec{e_2} + a_{32}\vec{e_3}, a_{13}\vec{e_1} + a_{23}\vec{e_2} + a_{33}\vec{e_3}) =$$
$$(a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31})\omega(\vec{e_1},\vec{e_2},\vec{e_3}).$$

Observação: Assim, o determinante $det(\vec{u},\vec{v},\vec{w})$ é caracterizado por uma escolha da unidade de volume: escolhemos $det(\vec{e_1},\vec{e_2},\vec{e_3}) = 1$. Daí decorre que toda forma trilinear alternada é apenas um múltiplo do determinante (é como se trocássemos a unidade de medida: todos os volumes seriam multiplicados por um mesmo fator). Mais precisamente,

[2]A definição de volume pede cuidados: se está dada em termos de aproximações por cubinhos, então é preciso provar que o volume de um paralelepípedo é a área da base vezes a altura. Aliás... o que é área?

Proposição: Seja $\omega : V \times V \times V \to V$ uma forma trilinear alternada e seja $\alpha = \omega(\vec{e}_1, \vec{e}_2, \vec{e}_3)$. Então

$$\omega(\vec{u}, \vec{v}, \vec{w}) = \alpha det(\vec{u}, \vec{v}, \vec{w}), \; \forall \; \vec{u}, \vec{v}, \vec{w} \; \in \; V.$$

Corolário: Se $T : V \to V$ é uma transformação linear, então, qualquer que seja a base $(\vec{u}, \vec{v}, \vec{w})$ de V, vale

$$\frac{det(T\vec{u}, T\vec{v}, T\vec{w})}{det(\vec{u}, \vec{v}, \vec{w})} = det(T\vec{e}_1, T\vec{e}_2, T\vec{e}_3).$$

Demonstração: Seja $\omega : V \times V \times V \to V$ a forma trilinear alternada dada por $\omega(\vec{u}, \vec{v}, \vec{w}) = det(T\vec{u}, T\vec{v}, T\vec{w})$ (exercício: verifique que ω é, de fato, trilinear e alternada). Pela Proposição, temos

$$det(T\vec{u}, T\vec{v}, T\vec{w}) = det(T\vec{e}_1, T\vec{e}_2, T\vec{e}_3)\; det(\vec{u}, \vec{v}, \vec{w}), \; \forall \; \vec{u}, \vec{v}, \vec{w} \; \in \; V.$$

Se $(\vec{u}, \vec{v}, \vec{w})$ é base de V, então $det(\vec{u}, \vec{v}, \vec{w}) \neq 0$; logo,

$$\frac{det(T\vec{u}, T\vec{v}, T\vec{w})}{det(\vec{u}, \vec{v}, \vec{w})} = det(T\vec{e}_1, T\vec{e}_2, T\vec{e}_3).$$

■

17.3 Formas de medir volumes

Se queremos levar para dimensões mais altas as mesmas ideias que nos conduziram ao determinante em $\mathbb{R}^3$, devemos começar com

Definição: Seja **V** um espaço vetorial (real ou complexo), de dimensão n. Uma **forma de medir volumes** (n-dimensionais) em **V** é uma aplicação $\omega : \mathbf{V}^n \to \mathbb{R}$ ou $\mathbb{C}$ tal que:

(i) ω é ***n*-linear**, isto é, é linear em cada coordenada[3] e

(ii) $\omega(v_1, v_2, \ldots, v_n) = 0$ sempre que $v_1, v_2, \ldots, v_n$ forem linearmente dependentes.

Exercício 17.8 *Mostre que, na presença da condição (i) acima, a condição (ii) é equivalente a*
(iii)$\omega(v_1, v_2, \ldots, v_n) = 0$ *sempre que existam dois índices distintos, i e j, tais que* $v_i = v_j$.

Exercício 17.9 *Seja* $\omega : \mathbf{V}^n \to \mathbb{R}$ *ou* $\mathbb{C}$ *uma forma de medir volumes. Mostre que* ω *é uma* **forma n-linear alternada**, *isto é:*
(iv)$\omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -\omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$ *para quaisquer vetores* $v_1, \ldots, v_n$ *em* **V**.

---

[3] ou seja: para todos $v_1, \ldots, v_j, u_j, \ldots, v_n$ e para todo λ, vale

$$\omega(v_1, \ldots, v_j + \lambda u_j, \ldots, v_n) = \omega(v_1, \ldots, v_j, \ldots, v_n) + \lambda\omega(v_1, \ldots, u_j, \ldots, v_n)$$

Exercício 17.10 *Suponha que* $\omega : \mathbf{V}^n \to \mathbb{R}$ *ou* $\mathbb{C}$ *é linear em cada coordenada. Mostre que as condições (ii) (da definição), (iii) e (iv) (dos exercícios acima) são equivalentes.*

De maneira geral, se $\omega : \mathbf{V}^p \to \mathbb{R}$ ou $\mathbb{C}$ é linear em cada coordenada e satisfaz $\omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_p) = -\omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_p)$ para quaisquer vetores $v_1, \ldots, v_p$ em $\mathbf{V}$, ω é dita uma **forma p-linear alternada**. Usaremos, para formas de medir volumes, o nome forma n-linear alternada, que é mais usual. O espaço das formas p-lineares alternadas em $\mathbf{V}$, qualquer que seja p em $\mathbb{N}$, é denotado por $\boldsymbol{\mathcal{A}_p(\mathbf{V})}$.

Exercício 17.11 *Mostre que* $\mathcal{A}_p(\mathbf{V})$*, com as operações* $(\omega_1 + \omega_2)(v_1, \ldots, v_p) = \omega_1(v_1, \ldots, v_p) + \omega_2(v_1, \ldots, v_p)$ *&* $(t\omega)(v_1, \ldots, v_p) = t(\omega(v_1, \ldots, v_p))$*, é um espaço vetorial.*

Exercício 17.12 *Mostre que, se* ω *é forma p-linear alternada e* $p > n$*, então* $\omega(v_1, \ldots, v_p) = 0$*, quaiquer que sejam* $v_1, \ldots, v_p$*.*

Exercício 17.13 *Parece razoável que, se dim*$\mathbf{V} = n$*, então o espaço das formas n-lineares alternadas de* $\mathbf{V}$ *tem dimensão 1?*

Exercício 17.14 *Sejam* $\mathbf{V}$ *um espaço vetorial de dimensão n e* ω *uma forma n-linear alternada em* $\mathbf{V}$*. Mostre que são equivalentes:*
a) ω *é identicamente nula;*
b)existe base $v_1, \ldots, v_n$ *de* $\mathbf{V}$ *tal que* $\omega(v_1, \ldots, v_n) = 0$*;*
c)para toda base $v_1, \ldots, v_n$ *de* $\mathbf{V}$ *se tem* $\omega(v_1, \ldots, v_n) = 0$*.*

O exercício acima nos coloca diante de uma questão aparentemente simples mas que merece alguma reflexão: se ω é uma forma n-linear alternada e conhecemos $\omega(v_1, \ldots, v_n)$, então, para qualquer permutação σ dos índices i, está perfeitamente determinado o valor de $\omega(v_{\sigma_1}, \ldots, v_{\sigma_n})$ (dado, é claro por $\pm\omega(v_1, \ldots, v_n)$)? A ideia é que podemos, a partir de $v_1, \ldots, v_n$, ir trocando de posição dois vetores de cada vez, até chegarmos à configuração $v_{\sigma_1}, \ldots, v_{\sigma_n}$; a cada troca, muda o sinal de ω. A questão é: existem infinitas maneiras de partir da configuração inicial e chegar à final; será que todas nos darão o mesmo resultado (isto é, o mesmo sinal)?

É hora de darmos uma arrumada na casa...

17.4 Permutações

Definição: Uma **permutação** do conjunto X é uma bijeção de X em X. O conjunto das permutações de X é notado $S(X)$. Se $X = \{1, 2, \ldots, n\}$ $S(X)$ é trocado por S_n (se σ pertence a S_n, $\sigma(k)$ é notado σ_k). Usamos chamar de **produto** a composta de duas permutações.

Entre os elementos de S_n, destacaremos as **transposições**, que são as que mantêm fixos todos os elementos de $\{1, 2, \ldots, n\}$, exceto dois (a mesma definição, a rigor, vale em qualquer $S(X)$). A transposição que *troca* i e j é designada por (ij). Mais geralmente, dizemos que σ em $S(X)$ é um **ciclo** de ordem k, $\infty > k \geq 2$, se existe um subconjunto A de X, com exatamente k elementos, tal que:
1. $\sigma(x) = x \ \forall\, x \notin A$;

2. $\exists\, x \in A \mid \forall\, y \in A\ \exists\, j \mid y = \sigma^j(x)$.

Chamaremos A de conjunto dos pontos **não fixos** do ciclo σ.

Exercício 17.15 *Entenda que, se σ é um ciclo, o conjunto A dos pontos não fixos é obtido assim: excolhe-se um x qualquer tal que $\sigma(x) \neq x$. Tomam-se os elementos $x, \sigma(x)$, $\sigma^2(x)$, $\sigma^3(x), \ldots$. Como A é finito, depois de um certo número, k, de iterações, volta-se a algum dos elementos anteriores. Mostre que o primeiro que volta é o próprio x. Assim, $A = \{\sigma(x),\ \sigma^2(x),\ \sigma^3(x), \ldots,\ \sigma^k(x) = x\}$.*

Exercício 17.16 *Sejam σ um ciclo e A o conjunto de seus pontos não fixos. Seja x um elemento qualquer de A. Mostre que $\forall\, y \in A\ \exists\, j \mid y = \sigma^j(x)$.*

Exercício 17.17 *Mostre que todo elemento de S_n é produto de **ciclos disjuntos** (dois ciclos são ditos disjuntos se todo elemento de $\{1, 2, \ldots, n\}$ fica fixo por, pelo menos, um dos dois).*[4]

Exercício 17.18 *Mostre que todo ciclo de ordem k é produto de $k-1$ transposições. Mostre que não dá para fazer com menos.*

Exercício 17.19 *Conclua que todo elemento de S_n é produto de transposições.*

Proposição 1: Toda permutação em S_n é produto de transposições.

Demonstração: Trata-se de resolver os exercícios acima.

1. Comecemos provando que toda permutação em S_n é produto de ciclos. Vamos fazer a prova por indução sobre n. O resultado é obviamente válido (na forma de produto de 0 transposições, entendido que o produto de 0 números é 1 e o produto de 0 permutações é a identidade) se $n = 1$. Suponhamos $n \geq 1$ e o resultado válido em S_k, $k < n$. Seja σ um elemento de S_n. Se σ é a identidade, está terminado. Se não, tomemos $x \in \{1, \ldots, n\}$ tal que $\sigma(x) \neq x$. Seja
$$A_1 = \left\{x = \sigma^0(x), \sigma(x), \sigma^2(x), \ldots\right\}.$$
Como A_1 é finito, existe um menor k tal que $\sigma^k(x) = \sigma^l(x)$ para algum l inferior a k. Então, necessariamente, $l = 0$, pois, caso contrário, $\sigma(\sigma^{l-1}(x)) = \sigma(\sigma^{k-1}(x))$, contrariando a injetividade de σ. Logo,
$$A_1 = \left\{\sigma(x), \sigma^2(x), \ldots, \sigma^k(x) = x\right\}.$$
Se $k = n$, terminamos. Se não, excluindo de $\{1, \ldots, n\}$ os elementos de A_1, temos um conjunto B com menos do que n elementos. Mas, então, a restrição de σ a B é uma permutação de B e se escreve, pela hipótese de indução, como produto de ciclos. Agora é só juntar.

2. Resta mostrar que todo ciclo é produto de transposições. Considere o ciclo σ, dado por $\sigma(x), \sigma^2(x), \ldots, \sigma^k(x) = x$. Então
$$\sigma = (\sigma(x)\sigma^k(x)) \ldots (\sigma(x)\sigma^3(x))(\sigma(x)\sigma^2(x)).$$
Confira. ■

[4]O produto de k permutações está definido, para $k > 2$, por conta da associatividade; o de uma só permutação, claro, é a própria; o produto de 0 permutações é, por definição, a permutação identidade

Nossa ideia é que, se ω é alternada, cada transposição troca o sinal de ω. Diante de uma permutação, queremos saber se esta foi produzida por uma número par ou ímpar de transposições. Como esse número não é único, a questão é: se a permutação σ de S_n é produto de uma certo número k_1, de transposições, mas, também, de um outro número, k_2 (de outras transposições), então, necessariamente, $k_1 - k_2$ é par?

Exercício 17.20 *Mostre que se, para um certo p, existir uma permutação que seja produto tanto de um número par como de um número ímpar de transposições, então toda forma p-linear alternada (em qualquer espaço vetorial, real ou complexo, seja de que dimensão for) será identicamente nula.*

Uma observação interessante é a seguinte: um dos efeitos de uma transposição é, sempre, um número ímpar de **inversões**. O número de inversões é obtido assim: dada σ em S_n, contamos quantos são os pares não ordenados $\{i, j\}$ tais que o sinal de $\sigma_i - \sigma_j$ é oposto ao de $i - j$. No caso de uma transposição, esse número é, sempre, ímpar.

Exercício 17.21 *Prove isso.*

A ideia acima vai nos levar ao seguinte resultado: toda permutação σ de $\{1, \ldots, n\}$ tem um sinal, $sgn\sigma$, que é 1 ou -1, dado pela paridade do número de trocas entre dois elementos para, partindo da identidade, se chegar à permutação desejada (1, se o número for par; -1, se for ímpar).

Proposição 2: Existe uma função $sgn : S_n \to \{1, -1\}$ tal que

1. $sgn(\sigma\tau) = sgn\sigma sgn\tau$, quaisquer que sejam as permutações σ e τ;
2. $sgn\sigma = -1$ para toda transposição σ.

Demonstração: A ideia é, dada σ, ver quantas inversões tem σ. Seja

$$sgn : S_n \to \{1, -1\}$$

dado por

$$sgn(\sigma) = \frac{\prod_{1 \le i < j \le n}(\sigma_j - \sigma_i)}{\prod_{1 \le i < j \le n}(j - i)} = \prod_{1 \le i < j \le n} \frac{\sigma_j - \sigma_i}{j - i}.$$

Como σ é uma bijeção, os números que aparecem nos numeradores são os mesmos que temos nos denominadores, a menos do sinal, o que garante que $sgn\sigma \in \{-1, 1\}$. Além disso, se σ é a transposição que troca i com j, o número de fatores negativos no numerador é $2(j - i) - 1$, o que prova que $sgn\sigma = -1$ se σ é transposição. Resta mostrar que, para quaisquer σ e τ em S_n, $sgn(\sigma\tau) = sgn(\sigma)sgn(\tau)$. Ora,

$$\begin{aligned} sgn(\sigma\tau) &= \textstyle\prod_{1 \le i < j \le n} \frac{(\sigma\tau)_j - (\sigma\tau)_i}{j - i} = \\ &= \textstyle\prod_{1 \le i < j \le n} \frac{\sigma_{\tau_j} - \sigma_{\tau_i}}{\tau_j - \tau_i} \prod_{1 \le i < j \le n} \frac{\tau_j - \tau_i}{j - i} = sgn(\sigma)sgn(\tau), \end{aligned}$$

embora a igualdade

$$\prod_{1 \le i < j \le n} \frac{\sigma_{\tau_j} - \sigma_{\tau_i}}{\tau_j - \tau_i} = \prod_{1 \le i < j \le n} \frac{\sigma_j - \sigma_i}{j - i}$$

até mereça uma pensada... ■

Exercício 17.22 *Seja $X = \{x_1, \ldots, x_n\} \subset \mathbb{R}$, com $x_1 < x_2 < \ldots < x_n$. Mostre que também podemos, em $S(X)$ definir o* ***sinal da permutação*** *σ, com as mesmas propriedades, por*

$$sgn(\sigma) = \frac{\prod_{1\leq i<j\leq n}(\sigma(x_j) - \sigma(x_i))}{\prod_{1\leq i<j\leq n}(x_j - x_i)} = \prod_{1\leq i<j\leq n} \frac{\sigma(x_j) - \sigma(x_i)}{x_j - x_i}.$$

Provado esse resultado algébrico fundamental, podemos seguir em frente.

17.5 O determinante como forma de volume

Além das permutações de n elementos, trabalharemos também, fixado n, com as ***p*-listas**, que são elementos de $\{1, \ldots, n\}^p$. Uma p-lista será designada por J, sendo $J = (J_1, \ldots, J_p)$. Também trabalharemos com um conjunto especial de p-listas, designado pela letra $\mathcal{J}$, definido por:

$$\boldsymbol{\mathcal{J} = \left\{J \in \{1, \ldots, n\}^p \mid J_1 < J_2 < \ldots < J_p\right\}.}$$

Lema 1: Seja $\mathbf{V}$ um espaço vetorial de dimensão n, real ou complexo. Fixada uma base de $\mathbf{V}$, cujos elementos designaremos por $e_1, \ldots, e_n$, seja, para cada p-lista J, $e_J \in \mathbf{V}^p$ dado por $e_J = (e_{J_1}, \ldots, e_{J_p})$. Então duas formas p-lineares ω e η de $\mathbf{V}$ são iguais sempre que $\omega(e_J) = \eta(e_J)$ para todo J em $\{1, \ldots, n\}^p$.

Demonstração: Fazemos por indução sobre p. O caso $p = 1$ é a bijeção entre transformações lineares e matrizes, fixada uma base. Supondo o resultado válido para p, consideremos duas formas $(p+1)$-lineares ω e η tais que $\omega(e_J) = \eta(e_J)$ para todo J em $\{1, \ldots, n\}^{p+1}$ e $p+1$ vetores $v_0, \ldots, v_p$. Escrevendo $v_0 = a_1 e_1 + \ldots + a_n e_n$, temos:

$$\omega(v_0, \ldots, v_p) = \sum_{j=1}^{n} a_j \omega(e_j, v_1 \ldots, v_p) = \sum_{j=1}^{n} a_j \eta(e_j, v_1 \ldots, v_p) = \eta(v_0, \ldots, v_p)$$

(note que a igualdade central decorre da hipótese de indução). ∎

Lema 2: Sejam $\mathbf{V}$ um espaço vetorial, real ou complexo, e ω uma forma p-linear alternada em $\mathbf{V}$. Para quaisquer $(v_1, \ldots, v_p)$ em $\mathbf{V}^p$ e σ em S_p, vale

$$\omega\left(v_{\sigma_1}, \ldots, v_{\sigma_p}\right) = sgn\sigma\ \omega(v_1, \ldots, v_p).$$

Demonstração: Usando a Proposição 1 da página 137, podemos escrever σ como produto de transposições: $\sigma = \tau_1 \ldots \tau_k$. Da Proposição 2 da página 138, segue $sgn\sigma = (-1)^k$. Como ω é alternada, temos

$$\omega\left(v_{\sigma_1}, \ldots, v_{\sigma_p}\right) = (-1)^k \omega(v_1, \ldots, v_p) = sgn\sigma\ \omega(v_1, \ldots, v_p).$$

∎

Observação 1: Dos lemas 1 e 2 acima, segue que duas formas p lineares alternadas ω e η são iguais se

$$\omega(e_J) = \eta(e_J)\ \forall J \in \mathcal{J}.$$

Em particular, se V é de dimensão n e $e_1, \ldots, e_n$ formam base de V, então a forma n linear alternada ω é identicamente nula se, e somente se, $\omega(e_1, \ldots, e_n) = 0$.

Lema 3: Seja $\mathbf{V}$ um espaço vetorial, real ou complexo. Se ω é p-linear e ω_A é definida por

$$\omega_A(v_1, \ldots, v_p) = \sum_{\sigma \in S_p} sgn\sigma\ \omega(v_{\sigma_1}, \ldots, v_{\sigma_p}),$$

então ω_A é p-linear e alternada.

Demonstração: Basta notar que trocar a ordem dos vetores v_i e v_j equivale a aplicar a transposição τ que troca i com j (note que $sgn\tau = -1$). Ora, dados v_1,...,v_i,...,v_j, ...,v_p em $\mathbf{V}$, temos, para cada τ em S_p,

$$\omega(v_{\tau_1}, \ldots, v_{\tau_p}) = \sum_{\sigma \in S_p} sgn\sigma\ \omega(v_{(\tau\sigma)_1}, \ldots, v_{(\tau\sigma)_p}) =$$

$$= -\sum_{\sigma \in S_p} sgn(\tau\sigma)\ \omega(v_{(\tau\sigma)_1}, \ldots, v_{(\tau\sigma)_p}) = -\omega(v_{\sigma_1}, \ldots, v_{\sigma_p})$$

(note que, para τ fixo, quando σ percorre S_p, o mesmo ocorre com $\tau\sigma$). ∎

Observação 2: Note que, se ω é p-linear alternada, então, dos lemas 2 e 3, segue que $\omega_A = p!\,\omega$.

Passemos, finalmente, à definição do determinante (numa primeira versão, como uma forma n-linear alternada). Fixando uma base de $\mathbf{V}$, cujos elementos designaremos por $e_1, \ldots, e_n$, podemos identificar $\mathbf{V}$ com K^n ($K = \mathbb{R}$ ou $\mathbb{C}$, conforme o caso). Cada elemento v de $\mathbf{V}$ pode, pois, ser identificado com uma n-upla $(v_1, \ldots, v_n)$ de K^n. Assim, basta-nos definir o determinante em $\mathbb{R}^n$ ou $\mathbb{C}^n$. Comecemos definindo uma forma n-linear

$$\begin{array}{cccc} \delta_n : & (K^n)^n & \longrightarrow & K \\ & (v_1, \ldots, v_n) & \longmapsto & v_{11}v_{22}...v_{nn}, \end{array}$$

entendido que

$$v_j = \begin{pmatrix} v_{1j} \\ \vdots \\ v_{nj} \end{pmatrix}.$$

Exercício 17.23 *Note que, tomando a matriz $n \times n$ $\left[v_{ij}\right]$, a forma δ_n calcula o produto dos elementos da diagonal principal.*

Exercício 17.24 *Mostre que, se $J \in \{1, \ldots, n\}^n$, então $\delta_n(e_J) = 1$, se $J = (1, 2, \ldots, n)$ e $\delta_n(e_J) = 0$, se $J \neq (1, 2, \ldots, n)$.*

Definição:A forma n-linear alternada det_n, definida em $\mathbb{R}^n$ ou $\mathbb{C}^n$, seguindo o lema 3, por

$$det_n(v_1,\dots,v_n) = \sum_{\sigma\in S_n} sgn\sigma\delta_n(v_{\sigma_1},\dots,v_{\sigma_n}) = \sum_{\sigma\in S_n} sgn\sigma\; v_{1\sigma_1}\dots v_{n\sigma_n},$$

é chamada de **determinante**.[5]

Observação 3: Note que, na expressão $\sum_{\sigma\in S_n} sgn\sigma\; v_{1\sigma_1}\dots v_{n\sigma_n}$, podemos usar a comutatividade do produto para, a nosso bel prazer, trocar a ordem dos fatores em cada uma das parcelas. Se decidirmos, na parcela $v_{1\sigma_1}\dots v_{n\sigma_n}$, permutar os fatores usando σ^{-1}, teremos

$$v_{1\sigma_1}\dots v_{n\sigma_n} = v_{\sigma^{-1}1\sigma_{\sigma^{-1}1}}\dots v_{\sigma^{-1}n\sigma_{\sigma^{-1}n}} = v_{\sigma^{-1}{}_1 1}\dots v_{\sigma^{-1}{}_n n}.$$

Assim, podemos escrever:

$$det_n(v_1,\dots,v_n) = \sum_{\sigma\in S_n} sgn\sigma\; v_{1\sigma_1}\dots v_{n\sigma_n} =$$

$$= \sum_{\sigma\in S_n} sgn\sigma v_{\sigma^{-1}{}_1 1}\dots v_{\sigma^{-1}{}_n n} = \sum_{\sigma\in S_n} sgn(\sigma^{-1}) v_{\sigma^{-1}{}_1 1}\dots v_{\sigma^{-1}{}_n n},$$

já que $sgn\sigma = sgn(\sigma^{-1})$. Ora, quando σ percorre $\S_n$, é claro que σ^{-1} também percorre S_n. Concluímos, então, que

$$det_n(v_1,\dots,v_n) = \sum_{\sigma\in S_n} sgn\sigma\; v_{1\sigma_1}\dots v_{n\sigma_n} = \sum_{\sigma\in S_n} sgn\sigma v_{\sigma_1 1}\dots v_{\sigma_n n},$$

o que, no caso de matrizes, significará que o determinante de toda matriz quadrada é igual ao de sua transposta.

Exercício 17.25 *Mostre que $det_n(e_1,\dots,e_n) = 1$.*

Exercício 17.26 *Defina, em $\mathbb{R}^n$, a **norma** do vetor $v = (v_1,\dots,v_n)$, $|v|$, por*

$$|v| = \sqrt{v_1^2 + \cdots + v_n^2}.$$

Mostre que, para quaisquer vetores $u_1,\dots,u_n$ em $\mathbb{R}^n$, vale

$$|det_n(u_1,\dots,u_n)| \le |u_1|\cdots|u_n|.$$

[5] Estamos usando a seguinte notação: o vetor v_j tem por coordenadas $(v_{1j},\dots,v_{nj})$

17.6 O determinante de transformação linear

Examinemos agora o espaço das n-formas lineares alternadas num espaço $\mathbf{V}$ de dimensão n, que será notado $\mathcal{A}_n(\mathbf{V})$. Se $\omega \in \mathcal{A}_n(\mathbf{V})$, então ω é determinada por seu valor em $(v_1, \ldots, v_n)$, onde $\{v_1, \ldots, v_n\}$ é base de $\mathbf{V}$ (isto segue da Observação 1, página 140). Assim, $\mathcal{A}_n(\mathbf{V})$ tem dimensão 1, isto é, se $\omega, \eta \in \mathcal{A}_n(\mathbf{V})$, $\omega \neq 0$, então existe $\lambda \in \mathbb{R}$ tal que $\eta = \lambda\omega$ (ou seja, a menos de fixação da unidade de medida, só existe uma forma de medir coisas de dimensão n em $\mathbf{V}$).

Proposição 1: Se o espaço vetorial $\mathbf{V}$, real ou complexo, tem dimensão n, então o espaço vetorial $\mathcal{A}_n$ das formas n-lineares alternadas em V tem dimensão 1.

Demonstração: Já demonstramos a existência de um elemento não nulo em $\mathcal{A}_n$ (o determinante, que notamos det_n). Seja ω outro elemento de $\mathcal{A}_n$; faça $\alpha = \omega(e_1, \ldots, e_n)$. Dos lemas 1 e 2 da seção anterior, segue que, como ω e αdet_n coincidem em $(e_1, \ldots, e_n)$, então são iguais. ■

Seja agora $T : \mathbf{V} \to \mathbf{V}$ linear. Para cada $\omega \in \mathcal{A}_n(\mathbf{V})$, seja $\omega_T \in \mathcal{A}_n(\mathbf{V})$ dada por

$$\omega_T(v_1, \ldots, v_n) = \omega(Tv_1, \ldots, Tv_n).$$

A aplicação $\omega \to \omega_T$ é claramente uma transformação linear de $\mathcal{A}_n(\mathbf{V})$ em $\mathcal{A}_n(\mathbf{V})$. Sendo $\mathcal{A}_n(\mathbf{V})$ de dimensão 1, existe um único escalar $\mathbf{d}etT$ tal que

$$\omega(Tv_1, \ldots, Tv_n) = detT\ \omega(v_1, \ldots, v_n) \forall \omega \in \mathcal{A}_n(\mathbf{V}) \ .$$

Definição: O número $detT$, definido acima, é chamado **determinante de** T.

Proposição 2: Se $T_1, T_2 : \mathbf{V} \to \mathbf{V}$ são lineares, então,

$$det(T_1T_2) = detT_1.detT_2 \ .$$

Demonstração: Suponhamos $T_1, T_2 : \mathbf{V} \to \mathbf{V}$ lineares. Então, para qualquer ω em $\mathcal{A}_n(\mathbf{V})$, temos

$$\begin{aligned} det(T_1T_2)\omega(v_1, \ldots, v_n) &= \omega(T_1T_2v_1, \ldots, T_1T_2v_n) = \\ &= \omega_{T_1}(T_2v_1, \ldots, T_2v_n) = detT_1\omega(T_2v_1, \ldots, T_2v_n) = \\ &= detT_1.detT_2.\omega(v_1, \ldots, v_n) \quad \forall (v_1, \ldots, v_n) \in \mathbf{V}^n \ , \end{aligned}$$

■

17.7 Determinante de matriz

Dada a matriz $n \times n$ $[a_{ij}]$, temos, pelo menos, três maneiras naturais para definir $det\,[a_{ij}]$:

1. usando a definição de det_n como forma linear, fazemos

$$det\,[a_{ij}] = det_n(a_1, \ldots, a_n),$$

sendo a_j o j-ésimo vetor coluna de $[a_{ij}]$;

2. mesma coisa, usando os vetores linha;

3. consideramos a transformação linear A definida por $\left[a_{ij}\right]$, na base canônica, e fazemos $det\left[a_{ij}\right] = detA$.

Curiosamente, as três dão o mesmo número. Há, ainda, uma possibilidade (infinitas, na verdade, uma para cada possível base):

4. Fixamos uma base β e consideramos a transformação linear T definida por $\left[a_{ij}\right]$, na base β, e fazemos $det\left[a_{ij}\right] = detT$.

Lema: As quatro definições de determinante de matriz propostas acima são equivalentes.

Demonstração: Comecemos provando que 1 e 3 dão o mesmo número. Como det_n é uma forma n-linear alternada, temos $det_n(Ae_1, \ldots, Ae_n) = detAdet_n(e_1, \ldots, e_n) = detA$. A equivalência entre 3 e 4 segue do fato de que a transformação T definida em 4 é dada por $B^{-1}AB$, sendo B a transformação que leva, na ordem, os vetores de β nos da base canônica. Assim, $detT = det(B^{-1}AB) = det(B{-}1)detAdetB = detA$ (já que $B^{-1}B = I$, e $detI = 1$). Resta provar que 1 e 2 também fornecem o mesmo resultado. Usaremos a fórmula, correspondente a 1:

$$det\left[a_{ij}\right] = \sum_{\sigma \in S_n} sgn\sigma\ a_{1\sigma_1} a_{2\sigma_2} \ldots a_{n\sigma_n}.$$

Observemos que cada parcela consiste em multiplicar o sinal de σ pelo produto dos números $a_{i\sigma_i}$ com i em $\{1, \ldots, n\}$. Podemos, é claro, alterar a ordem dos fatores, reescrevendo cada parcela como

$$sgn\sigma\ a_{(\sigma^{-1})_1\sigma_{(\sigma^{-1})_1}} a_{(\sigma^{-1})_2\sigma_{(\sigma^{-1})_2}} \ldots a_{(\sigma_n^{-1})\sigma_{(\sigma^{-1})_n}}.$$

Como $\sigma_{(\sigma^{-1})_j} = j$ para todo j (é pura notação), segue

$$det\left[a_{ij}\right] = \sum_{\sigma \in S_n} sgn\sigma\ a_{(\sigma^{-1})_1 1} a_{(\sigma^{-1})_2 2} \ldots a_{(\sigma_n^{-1})n}.$$

Como $sgn\sigma = sgn(\sigma^{-1})$ para todo σ em S_n e, quando σ percorre S_n, σ^{-1} faz o mesmo, o termo à direita corresponde, exatamente, à expressão que obtemos de 2. ■

Escolhendo uma das definições equivalentes acima, podemos definir o determinante de uma matriz e dar por provadas umas tantas coisas.

Definição: O **determinante** da matriz $n \times n$ $\left[a_{ij}\right]$, real ou complexa, é o número $det\left[a_{ij}\right]$, também notado por $\left|a_{ij}\right|$, e definido por

$$det\left[a_{ij}\right] = \sum_{\sigma \in S_n} sgn\sigma\ a_{1\sigma_1} a_{2\sigma_2} \ldots a_{n\sigma_n}.$$

Teorema: O determinante da matriz $n \times n$ $\left[a_{ij}\right]$, real ou complexa, é uma forma n-linear alternada dos vetores coluna de $\left[a_{ij}\right]$ e satisfaz as seguintes propriedades:

1. $det\,[I] = 1$, sendo $[I]$ a matriz identidade;

2. $det\left([A]\,[B]\right) = det\,[A]\,det\,[B]$, para quaisquer matrizes $n \times n$ $[A]$ e $[B]$

3. $det\,[A]^T = det\,[A]$, para qualquer matriz $n \times n$ $[A]$.

Demonstração: ■

Os resultados a seguir merecem um certo destaque.

Proposição 1: A matriz $n \times n$ $\left[a_{ij}\right]$ tem inversa se, e somente se, $det\left[a_{ij}\right] \neq 0$.

Demonstração: Se $\left[a_{ij}\right]$ tem inversa, então

$$det\left[a_{ij}\right]\; det\left(\left[a_{ij}\right]^{-1}\right) = det\left(\left[a_{ij}\right]\left[a_{ij}\right]^{-1}\right) = det\,[I] = 1;$$

logo, $det\left[a_{ij}\right] \neq 0$. Reciprocamente, dizer que $det\left[a_{ij}\right] \neq 0$ é o mesmo que dizer que $det_n(a_1,\ldots,a_n) \neq 0$, sendo a_j o j-ésimo vetor coluna de $\left[a_{ij}\right]$. Mas isso só ocorre se os vetores coluna de $\left[a_{ij}\right]$ são linearmente independentes, o que significa que $\left[a_{ij}\right]$ tem inversa. ■

Proposição 2 (Regra de Cramer): Se o vetor coluna $[x_i]$ é solução do sistema $n \times n$ $\left[a_{ij}\right][x_i] = [b_i]$, ou seja,

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

então, para cada $i = 1,\ldots,n$, vale $x_j\, det\left[a_{ij}\right] = det\left[\tilde{a}_{ij}\right]$, sendo $\left[\tilde{a}_{ij}\right]$ a matriz $n \times n$ que se obtém substituindo a j-ésima coluna de $\left[a_{ij}\right]$ por $[b_i]$. Em particular se $det\left[a_{ij}\right] \neq 0$ então a solução $[x_i]$ é única e, se representamos por a_j os vetores coluna de $\left[a_{ij}\right]$ e por b o vetor correspondente a $[b_i]$, é dada por

$$x_j = \frac{det_n(a_1,\ldots,a_{j-1},b,a_{j+1},\ldots,a_n)}{det_n(a_1,\ldots,a_n)}.$$

Demonstração: Basta escrever $b = x_1 a_1 + \ldots + x_n a_n$ e substituir em

$$det_n(a_1,\ldots,a_{j-1},b,a_{j+1},\ldots,a_n).$$

■

Falta uma fórmula de cálculo "efetiva". A fórmula que define o determinante de matriz $n \times n$ não é muito alentadora: prevê a soma de $n!$ parcelas, cada uma com $(n-1)$ multiplicações (sem contar o cálculo do sinal da permutação). Mesmo para n bem pequeno, relativamente às matrizes que aparecem em problemas reais, esse número de operações exige, mesmo com os mais modernos processadores, tempo que nenhum ser humano tem para esperar.[6] Modestamente, não vamos aqui discutir algoritmos

[6] Podemos atuar nas três frentes: melhorar o algoritmo, fabricar processadores mais rápidos e aumentar o tempo de vida. Matemáticos podem se engajar em qualquer uma das três

mais eficientes; apenas apresentaremos um método tradicional, que reduz o cálculo de um determinante $(n+1) \times (n+1)$, essencialmente, ao de $n+1$ determinantes $n \times n$.

O ponto de partida é a observação, simples, de que, no cálculo do determinante de $[a_{ij}]$, a entrada a_{ij} só multiplica entradas da matriz que se obtém *riscando* a linha i e a coluna j. Logo vemos que, com um pouco de sorte, a_{ij} deve multiplicar, talvez a menos do sinal, o determinante dessa matriz um pouco menor. O resultado preciso é a chamada **fórmula dos cofatores**.

Proposição 3: Seja $[a_{ij}]$ matriz $n \times n$ $(n > 1)$. Fixada a coluna j, o determinante de $[a_{ij}]$ é dado pela **fórmula dos cofatores**:

$$det\,[a_{ij}] = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij},$$

sendo M_{ij} o determinante da matriz $(n-1) \times (n-1)$ $[\tilde{a}_{ij}]$ obtida de $[a_{ij}]$ *riscando* a linha i e a coluna j.[7] Equivalentemente, fixada a linha i, temos

$$det\,[a_{ij}] = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}.$$

Demonstração: Vamos provar apenas a primeira fórmula, já que para a segunda, basta tomar a transposta da matriz. Para simplificar, começamos com $j=1$. Chamando, como de hábito, os vetores coluna de $a_1, \ldots, a_n$ e recorrendo a nossa definição de det_n, temos

$$det\,[a_{ij}] = det_n(\sum_{i=1}^{n} a_{i1} e_i, a_2, \ldots, a_n) = \sum_{i=1}^{n} a_{i1} det_n(e_i, a_2, \ldots, a_n).$$

Vejamos, agora, o que nos dá, para i fixo, $det_n(e_i, a_2, \ldots, a_n)$. Usando a notação de barras para o determinante, trata-se de

$$\begin{vmatrix} 0 & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ 1 & a_{i2} & \ldots & a_{in} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} & \ldots & a_{nn} \end{vmatrix} = \begin{vmatrix} 0 & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} & \ldots & a_{nn} \end{vmatrix} = (-1)^{i-1} \begin{vmatrix} 1 & 0 & \ldots & 0 \\ 0 & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ \hat{0} & \hat{a}_{i2} & \ldots & \hat{a}_{in} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

(os chapéus sobre as entradas da i-ésima linha indicam que essa linha foi suprimida). O último determinante faz aparecer, se riscarmos a primeira linha e a primeira coluna, a matriz $[\tilde{a}_{ij}]$. Se definirmos a aplicação que, a cada matriz $(n-1) \times (n-1)$ $[b_{ij}]$, associa o número

$$\begin{vmatrix} 1 & 0 & \ldots & 0 \\ 0 & b_{11} & \ldots & b_{1(n-1)} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & b_{(n-1)2} & \ldots & b_{(n-1)(n-1)} \end{vmatrix},$$

[7] O número $(-1)^{i+j}Mij$ é chamado **cofator** de a_{ij}; M_{ij} é dito **determinante menor** associado a a_{ij}

teremos uma forma $(n-1)$-linear e alternada nas colunas de $[b_{ij}]$ que, além disso, assumirá o valor 1 quando $[b_{ij}]$ for a matriz identidade. Ora, só há uma função com tais características: o determinante. Logo, se tomarmos $[b_{ij}] = [\tilde{a}_{ij}]$, teremos M_{ij}. Isto nos dá

$$det_n(e_i, a_2, \ldots, a_n) = (-1)^{i-1} M_{ij} = (-1)^{i+1} M_{ij},$$

Nossa fórmula está provada, para $j = 1$. Para j qualquer, basta notar que, trazendo para a extrema esquerda a coluna j, estaremos efetuando $j-1$ transposições, de modo que

$$\begin{vmatrix} a_{11} & \ldots & \ldots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \ldots & \ldots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \ldots & \ldots & a_{nn} \end{vmatrix} = (-1)^{j-1} \begin{vmatrix} a_{1j} & a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{ij} & a_{i1} & \ldots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{nj} & a_{n1} & \ldots & a_{nn} \end{vmatrix}$$

Aplicando ao termo da direita a fórmula dos cofatores para a primeira coluna, já demonstrada, obtemos a fórmula geral. ■

17.8 Orientação

Definição: Diremos que duas bases ordenadas

$$\begin{aligned} \alpha &= (u_1, \ldots, u_n) \quad e \\ \beta &= (v_1, \ldots, v_n) \end{aligned}$$

de $\mathbb{R}^n$ têm a mesma **orientação** se existem funções $f_1, \ldots, f_n : [0,1] \longrightarrow \mathbf{V}$ tais que:

(i)f_i é contínua $\forall i = 1 \ldots n$;
(ii)$f_i(0) = u_i,\ f_i(1) = v_i\ \forall i = 1 \ldots n$;
(iii)$f_1(t), \ldots, f_n(t)$ são linearmente independentes $\forall t \in [0,1]$.

Observação: O leitor, provavelmente, há de se perguntar o que significa dizer que determinada função $g : I \to \mathbb{R}^n$ é contínua. Para feitos de consumo imediato, diremos que g é contínua se, escrevendo $g(t) = (g_1(t), \ldots, g_n(t))$, cada uma das $g_j : I \to \mathbb{R}$ é contínua.

Exercício 17.27 *Fixe em $\mathbb{R}^n$ o produto interno canônico,*

$$\langle u, v \rangle = \sum_{j=1}^{n} u_j v_j,$$

*e defina a **norma** de u por $|u| = \sqrt{\langle u, u \rangle}$. Mostre que $g : I \to \mathbb{R}^n$ é contínua em t_o se, e somente se,*

$$\lim_{t \to t_o} |g(t) - g(t_o)| = 0.$$

Exercício 17.28 *Mostre que "ter a mesma orientação"é uma relação de equivalência no conjunto das bases de $\mathbb{R}^n$.*

Exercício 17.29 *Seja $m_2 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ como definida há pouco. Seja $\omega_2 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ dada por $\omega_2(u,v) = 0$ se u e v são linearmente dependentes, $\omega_2(u,v) = m_2(u,v)$ se (u,v) tem a mesma orientação que (e_1,e_2) e $\omega_2(u,v) = -m_2(u,v)$ se (u,v) não tem a mesma orientação que (e_1,e_2). Mostre que*

$$\begin{array}{rl} (i) & \omega_2(u,v) = 0 \text{ se, e somente se, } u \text{ e } v \text{ são linearmente dependentes} \\ (ii) & \omega_2(u,v) = -\omega_2(v,u), \ \forall u,v \\ (iii) & \omega_2(\lambda u + w, v) = \lambda\omega_2(u,v) + \omega_2(w,v) \ \forall \lambda, u, w, v \end{array}$$

Exercício 17.30 *Seja ω uma forma $(n+1)$-linear alternada em $\mathbb{R}^{n+1}$. Suponha que ω não é identicamente nula. Mostre que duas bases ordenadas*

$$(v_1,\ldots,v_i,u,v_{i+1},\ldots,v_n) \text{ e } (v_1,\ldots,v_i,w,v_{i+1},\ldots,v_n)$$

têm a mesma **orientação** *se e somente se*

$$\omega(v_1,\ldots,v_i,u,v_{i+1},\ldots,v_n) \text{ e } \omega(v_1,\ldots,v_i,w,v_{i+1},\ldots,v_n)$$

têm o mesmo sinal.

Proposição: Seja ω uma forma $(n+1)$-linear alternada em $\mathbb{R}^{n+1}$, não identicamente nula. Então duas bases ordenadas $(u_1,\ldots,u_{n+1})$ e $(v_1,\ldots,v_{n+1})$ de $\mathbb{R}^{n+1}$ têm a mesma orientação se e somente se $\omega(u_1,\ldots,u_{n+1})$ e $\omega(v_1,\ldots,v_{n+1})$ têm o mesmo sinal.

Demonstração: Supondo que as duas bases tenham a mesma orientação, considere as funções contínuas $f_1,\ldots,f_{n+1} : [0,1] \to \mathbb{R}^{n+1}$ que transformam uma na outra e faça $\alpha : [0,1] \to \mathbb{R}$, $\alpha(t) = \omega(f_1(t),\ldots,f_{n+1}(t))$. A função α é contínua e não pode se anular (pois, nesse caso, ω seria identicamente nula - ver a Observação 1, página 140). O resultado segue do Teorema do Valor Intermediário.

Para a recíproca, comecemos observando que o processo de ortonormalização de Gram-Schmidt nos fornece uma deformação de $(u_1,\ldots,u_{n+1})$ em uma base ortonormal com a mesma orientação, mantendo o sinal de ω. O mesmo acontece com $(v_1,\ldots,v_{n+1})$. Assim, podemos supor que as duas bases são ortonormais e que $\omega(u_1,\ldots,u_{n+1})$ e $\omega(v_1,\ldots,v_{n+1})$ têm o mesmo sinal. Vamos agora, passo a passo, deformar cada u_i em cada v_i.

Se $u_1 = v_1$ ou $u_1 = -v_1$, nada fazemos; caso contrário, tomamos θ tal que $\cos\theta = < u_1, v_1 >$, fazemos $e_1 = u_1$, $\bar{v}_1 = v_1 - < v_1, u_1 > u_1$, $e_2 = (1/|\bar{v}_1|)\bar{v}_1$ e, para $t \in [0,1]$, consideramos a transformação T_t de $\mathbf{V}$ em $\mathbf{V}$ dada por $T_t e_1 = \cos(t\theta)e_1 + \sin(t\theta)e_2$, $T_t e_2 = -\sin(t\theta)e_1 + \cos(t\theta)e_2$, mantendo fixos os vetores ortogonais ao espaço gerado por e_1 e e_2. Assim, como T_t preserva a ortonormalidade, a antiga base $(u_1,\ldots,u_{n+1})$ se deforma em uma nova (que continuaremos a chamar de $(u_1,\ldots,u_{n+1})$), com o novo u_1 igual a v_1. Fazemos o mesmo com u_2 (note que, agora, como $u_1 = \pm v_1$, que é ortogonal a v_2 e ao novo u_2, u_1 não será alterado). Prosseguimos da mesma forma, até u_n, observando que os passos posteriores ao k-ésimo não alteram mais $u_1,\ldots,u_k$. Teremos, no fim, uma nova base ortonormal, que continuamos chamando $(u_1,\ldots,u_{n+1})$, em que $u_i = v_i$ ou $u_i = -v_i$, $i = 1,\ldots,n$. Daí decorre que também temos $u_{n+1} = v_{n+1}$ ou $u_{n+1} = -v_{n+1}$. Durante todo o processo, o sinal de $\omega(u_1,\ldots,u_{n+1})$ não se alterou, continuando igual ao de $\omega(v_1,\ldots,v_{n+1})$. Logo, o número de índices i para os quais $u_i = -v_i$ é par; $2k$, digamos. Mas, se $u_i = -v_i$ e $u_j = -v_j$, podemos fazer, no espaço W gerado por u_i e u_j, uma rotação de 180^0, deixando parados os vetores de $\mathbf{w}^\perp$. Com k rotações como essa, transformamos finalmente uma base na outra. ■

Corolário: Duas bases ordenadas $(u_1,\ldots,u_{n+1})$ e $(v_1,\ldots,v_{n+1})$ de $\mathbb{R}^{n+1}$ têm a mesma orientação se e somente se $det(u_1,\ldots,u_{n+1})$ e $det(v_1,\ldots,v_{n+1})$ têm o mesmo sinal.

17.9 A dimensão do espaço das formas p-lineares alternadas

Dedicaremos os próximos parágrafos a provar que, se **V** é espaço vetorial de dimensão n, real ou complexo, então

$$dim\ \mathcal{A}_p(\mathbf{V}) = \binom{n}{p}.$$

Admitiremos fixada uma base de **V**, cujos elementos designaremos por $\mathbf{e_1}, \ldots, \mathbf{e_n}$, e identificaremos V com $\mathbf{K^n}$ ($\mathbf{K} = \mathbb{R}$ ou $\mathbf{K} = \mathbb{C}$) por meio de $\mathbf{v} = v_1 e_1 + \ldots + v_n e_n$.

Recordamos que as ***p*-listas** são elementos de $\{1, \ldots, n\}^p$. Uma p-lista será designada por J, sendo $J = (J_1, \ldots, J_p)$. Também trabalharemos com as p-listas crescentes, que estão em $\mathcal{J}$ definido por

$$\mathcal{J} = \{J \in \{1, \ldots, n\}^p \mid J_1 < J_2 < \ldots < J_p\}.$$

Seja, para cada p-lista J, π_J a projeção de K^n (ou **V**) sobre K^p, dada por:

$$\pi_J(\mathbf{v}) = (v_{J_1}, \ldots, v_{J_p}).$$

Observemos agora que, se $\mathbf{v_1}, \ldots, \mathbf{v_p}$ são elementos de **V**, faz sentido calcular o determinante det_p nos p vetores de K^p obtidos aplicando π_J a $\mathbf{v_1}, \ldots, \mathbf{v_p}$. O lema 2 nos permite definir as formas elementares dx_J, para as p-listas J.

Definição: As **formas alternadas elementares dx_J** são definidas, para cada p-lista J, por

$$dx_J(\mathbf{v_1}, \ldots, \mathbf{v_p}) = det_p(\pi_J(\mathbf{v_1}), \ldots, \pi_J(\mathbf{v_p})).$$

Para cada p-lista J, definiremos $\mathbf{e_J} \in \mathbf{V}^p$ por $e_J = (\mathbf{e_{J_1}}, \ldots, \mathbf{e_{J_p}})$.

Exercício 17.31 *Note que dx_J é identicamente nula se J tem índices repetidos (já que, nesse caso, $\pi_J(\mathbf{v_1}), \ldots, \pi_J(\mathbf{v_p})$ são linearmente dependentes). Note, também, que, se os índices são todos diferentes, $dx_J(\mathbf{e_J}) = 1$. Em particular, se $J \in \mathcal{J}$, $dx_J(\mathbf{e_J}) = 1$.*

Exercício 17.32 *Suponha que I e J são tais que $\{I_1, \ldots, I_p\} \neq \{J_1, \ldots, J_p\}$. Mostre que $dx_I(\mathbf{e_J}) = 0$.*

Exercício 17.33 *Suponha que I e J são tais que $\{I_1, \ldots, I_p\} = \{J_1, \ldots, J_p\}$. Seja $\tau \in S_p$ tal que $(I_1, \ldots, I_p) = (J_{\tau_1}, \ldots, J_{\tau_p})$. Mostre que, para toda forma p-linear alternada ω e para quaisquer $\mathbf{v_1}, \ldots, \mathbf{v_p}$ em **V**, vale $\omega(\mathbf{v_{I_1}}, \ldots, \mathbf{v_{I_p}}) = sgn\tau\omega(\mathbf{v_{J_1}}, \ldots, \mathbf{v_{J_p}})$.*

Finalmente podemos demonstrar nosso resultado principal. Embora não seja necessário, vamos explicitar $\mathbf{V} = \mathbb{R}^n$ ou $\mathbf{V} = \mathbb{C}^n$.

Proposição: Seja $\mathbf{V} = \mathbb{R}^n$ ou $\mathbf{V} = \mathbb{C}^n$. Então $\{dx_J,\ J \in \mathcal{J}\}$ é uma base para o espaço das formas p-lineares alternadas em $\mathbf{V}$. Consequentemente, a dimensão de $\mathcal{A}_p(\mathbf{V})$ é dada por

$$dim\ \mathcal{A}_p(\mathbf{V}) = \binom{n}{k}.$$

Demonstração: Comecemos observando que $\{dx_J,\ J \in \mathcal{J}\}$ é um conjunto linearmente independente. De fato, se

$$\sum_{J \in \mathcal{J}} a_J dx_J = 0,$$

então, para todo $I \in \mathcal{J}$,

$$a_I = \sum_{J \in \mathcal{J}} a_J dx_J(\mathbf{e_I}) = 0.$$

Por outro lado, dada ω em $\mathcal{A}_p(\mathbf{V})$, seja ω_o dada por

$$\omega_o = \sum_{J \in \mathcal{J}} \omega(\mathbf{e_J}) dx_J.$$

É imediato que $\omega_o(\mathbf{e_J}) = \omega(\mathbf{e_J})$ para todo J em $\mathcal{J}$. A igualdade desejada ($\omega = \omega_o$) está, então, garantida pela Observação 1 da página 140. ∎

Índice Remissivo