Introdução a Álgebra Linear com álgebra e geometria

André Caldas

2 de julho de 2022

Sumário

Este Livro é Livre

Este livro pode ser copiado à vontade. Se você recebeu em formato digital, fique à vontade para copiá-lo e compartilhá-lo quantas vezes quiser. Você pode também imprimí-lo e fotocopiá-lo o tanto que quiser. Claro que é sempre importante pensar na natureza e no impacto ambiental. Procure não desperdiçar recursos. ;-)

Quer imprimir e vender este livro para os seus colegas? Fique à vontade, também! Você é dono de uma editora e quer imprimir sua própria versão, vender e ficar rico com esse livro sem precisar pagar nenhum tostão em direitos autorais? Pois na minha opinião, se você o fizer estará contribuindo para um mundo melhor. Uma das poucas restrições é que você não tire dos outros essa liberdade que lhe foi concedida. Se você passar esse livro pra frente, não poderá proibir aquele que o recebeu de usufruir dessas mesmas liberdades. Ao repassar esse livro, é preciso deixar claro que quem recebeu pode copiá-lo e compartilhá-lo à vontade.

Este livro está licenciado sob os termos da licença "Creative Commons Attribution Share Alike 3.0". Os termos estão disponíveis na internet através do endereço
http://creativecommons.org/licenses/by-sa/3.0/.

A versão mais atual deste livro pode ser encontrada na internet no endereço

http://andrec.mat.unb.br/ensino/.
Neste endereço também podem ser encontrados os arquivos XƎLaTeX, que você pode alterar e usar para criar sua própria versão deste livro.

Por Quê?

Eu (André Caldas) NÃO acredito que "coletar taxas" seja a melhor maneira de se sobreviver da produção cultural. Na minha opinião, as pessoas devem receber para produzir; e não produzir na esperança de "coletar taxas" para o resto da vida.

Não vejo sentido em um professor de uma universidade pública, que já recebe um salário do governo para que produza conhecimento, ter monopólios sobre o fruto do seu trabalho. Vejo menos sentido ainda quando esse professor vende tal monopólio para ser explorado por alguma empresa privada. Perceba que isso não é uma crítica ao lucro ou à exploração da produção científica. É uma crítica ao monopólio sobre os direitos de uso daquilo que foi produzido com dinheiro público. Não faz sentido que a sociedade faça esse tipo de investimento e depois não possa ter acesso ao que foi produzido. Livros é o menor dos problemas. Vemos que pesquisas para o desenvolvimento de medicamentos ou de tecnologia são feitas com dinheiro público em *"parceria"* com instituições privadas de modo que a sociedade acaba sendo privada de seus frutos. Não acredito que seja errado fazer parceria com instituições privadas. O que não se pode é privar a sociedade dos frutos do trabalho no qual investiu. Vemos instituições de pesquisa se valerem de recursos públicos durante as pesquisas e depois correrem atrás de registros de patentes e coisas do tipo que servem apenas ao propósito de privar a sociedade dos frutos dessa mesma pesquisa. O **monopólio** não é necessário. Todos os que prestam algum serviço na cadeia, desde que haja demanda, serão beneficiados.

Se o governo, por exemplo, a uns 10 anos atrás tivesse começado a financiar a produção de livros e exigisse que o fruto desse trabalho fosse verdadeiramente livre (como é este livro), hoje não faríamos licitações para

a “compra” de livros; faríamos licitações para a impressão e a distribuição desses livros. Falta visão de longo prazo. Deveríamos investir na produção de livros livres. O autor deve sim receber por seu trabalho. Só que deve receber enquanto faz o trabalho, e não depois através do recolhimento de taxas e de mecanismos de opressão. O leitor não é um ladrão ou criminoso. Gostaria de sugerir a leitura de *“o direito de ler”*:
https://www.gnu.org/philosophy/right-to-read.html

Tenho vários amigos que fazem cópias de livros. Sempre lembro a eles que quando virarem autores, não devem “virar a casaca” e começar a perseguir os que fazem cópias. Gostaria de fazer o mesmo pedido ao leitor! :-)

Brasília, 29 de fevereiro de 2016,
André Caldas

Parte I

Pré-Requisitos

Aula 1

Conjuntos

Um *conjunto* A é um objeto matemático a respeito do qual estamos interessados em estabelecer se determinado *elemento* x *pertence* ou *não pertence* a A. Escrevemos

$$x \in A$$

para dizer que x *pertence* a A. E para dizer que *não pertence*, escrevemos

$$x \notin A.$$

Uma forma expressar um conjunto A é listando todos os seus elementos. Por exemplo, escrevendo

$$A = \{5, 9, 13, 52\}$$

para dizer que $x \in A$ se, e somente se, $x = 5$, ou $x = 9$, ou $x = 13$, ou $x = 52$. Às vezes, assumimos que determinado *"padrão"* é suficientemente evidente para, com alguns exemplos, deixarmos claro quais são os elementos de A. Por exemplo,

$$\begin{aligned} A &= \{1, 2, 3, 4, \ldots\} \\ A &= \{2, 4, 6, 8, \ldots\} \\ A &= \{1, 2, \ldots, 10\} \\ A &= \{10, 20, 30, 40, \ldots\}. \end{aligned}$$

Dizemos que dois conjuntos A e B são iguais e escrevemos $A = B$, quando
$$x \in A \Leftrightarrow x \in B.$$
Ou seja, A e B são formados pelos mesmos elementos.

É importante notar que não estamos interessados na *"quantidade de cópias"* de x que estão em A. Queremos saber apenas se $x \in A$ ou $x \notin A$. Dessa forma,
$$\{1, 2, 3\} = \{1, 1, 1, 2, 3\}.$$
Também não estamos interessados na ordem. De modo que
$$\{5, 7, 15\} = \{5, 15, 7\}.$$

Quando o conjunto tem finitos elementos, dizemos que é *finito*. Nem todos os conjuntos são finitos. E nem todos podem ser escritos com uma regra simples. Um exemplo é o conjunto $\mathbb{R}$, dos números reais. Ah... e já que estamos falando dos reais... tem também os naturais: $\mathbb{N}$. Nesse livro, $0 \in \mathbb{N}$:
$$\mathbb{N} = \{0, 1, 2, \ldots\}.$$
Existe também o conjunto $\emptyset$, que é tal que, qualquer que seja x,
$$x \notin \emptyset.$$
Chamamos $\emptyset$ de *conjunto vazio*.

Usamos a expressão *cardinalidade* de A, e escrevemos $\#A$, para nos referirmos à quantidade de elementos em A. Se o conjunto A é infinto, escrevemos $\#A = \infty$. E se é finito, $\#A < \infty$. Por exemplo,
$$A = \{1, 5, 12\} \Rightarrow \#A = 3.$$

1.1 Subconjuntos

Se A e B são conjuntos, dizemos que A é é um subconjunto de B, e escrevemos $A \subset B$, quando
$$x \in A \Rightarrow x \in B.$$

Ou seja, todo elemento de A é também um elemento de B. Note que, qualquer que seja o conjunto A,

$$A \subset A.$$

E de fato, dois conjuntos A e B são iguais se, e somente se,

$$A \subset B \quad \text{e} \quad B \subset A.$$

Pode até parecer estranho, mas para mostrar que $A = B$, é muito comum fazer uma demonstração em duas etapas. Primeiro, mostra-se que $A \subset B$. Depois, mostra-se que $B \subset A$. Ou seja, primeiro mostra-se que

$$x \in A \Rightarrow x \in B.$$

Depois mostra-se o contrário... que $x \in B \Rightarrow x \in A$.

Exemplo 1.1 (Múltiplos de 6)**.** Seja A o conjunto dos números naturais que são múltiplos de 6. E seja B o conjunto dos números naturais que são múltiplos de 2 e de 3. Então, $A = B$.

De fato, todos os múltiplos de 6 são múltiplos de 2 e de 3. Portanto,

$$A \subset B.$$

Agora é fácil! É só mostrar que $B \subset A$... ooops! Na verdade, essa segunda parte não é tão fácil! :-P

Em termos de *subconjuntos*... um exemplo simples, mas também muito estranho, é o conjunto vazio.

Exemplo 1.2. Se A é um conjunto, então

$$\emptyset \subset A.$$

Dado um conjunto A, uma maneira muito usual de se construir um subconjunto B, é especificando uma propriedade que os elementos de B devem satisfazer. Os elementos de B são aqueles de A que satisfazem essa tal propriedade. Os que não satisfazem, não pertencem a B. Por exemplo,

$$B = \left\{ x \in \mathbb{N} \,\middle|\, \exists n \in \mathbb{N}\, x = 2n \right\}$$

é o conjunto dos números naturais pares. São os elementos de $\mathbb{N}$ que podem ser escritos como o dobro de um outro elemento de $\mathbb{N}$.

1.2 União e Interseção

Se A e B são dois conjuntos, escrevemos

$$A \cup B$$

para representar o conjunto tal que

$$x \in A \cup B \Leftrightarrow x \in A \text{ ou } x \in B.$$

Chamamos $A \cup B$ de *união* dos conjuntos A e B.

Exemplo 1.3. Se

$$A = \{1, 3, 5\} \quad \text{e} \quad B = \{2, 3, 15\},$$

então

$$A \cup B = \{1, 2, 3, 5, 15\}.$$

Note que $\#A = 3$ e $\#B = 3$, mas que no entanto, a cardinalidade de $A \cup B$ não é 6, mas sim, 5.

A *interseção*, denotada por $A \cap B$, é a parte comum a A e B. É o conjunto tal que

$$x \in A \cap B \Leftrightarrow x \in A \text{ e } x \in B.$$

A *interseção* pode ser expressa por

$$A \cap B = \{x \in A \mid x \in B\}.$$

Ou seja, são os elementos de A que também estão em B. Ou então,

$$A \cap B = \{x \in A \cup B \mid x \in A \text{ e } x \in B\}.$$

Onde $x \in A$, $x \in B$ siginfica $x \in A$ e $x \in B$.

Podemos falar da *união* ou *interseção* de 3 ou mais conjuntos da mesma maneira. Se $A_1, \dots, A_n$ são conjuntos, escrevemos

$$A_1 \cup \cdots \cup A_n$$

para a *união*, e

$$A_1 \cap \cdots \cap A_n$$

para a *interseção*. Também escrevemos

$$\bigcup_{j=1}^{n} A_j \quad \text{e} \quad \bigcap_{j=1}^{n} A_j.$$

1.3 Diferença de Conjuntos

Se A e B são dois conjuntos, a *diferença* $A \setminus B$ é o conjunto dos elementos de A que não estão em B. Ou seja,

$$A \setminus B = \{a \in A \mid a \notin B\}.$$

Note que

$$\#(A \setminus B) = (\#A) - (\#(A \cap B)).$$

Exemplo 1.4. Sejam

$$A = \{x \mid x \text{ é professor de matemática}\}$$
$$B = \{x \mid x \text{ é uma pessoa que não é careca}\}.$$

Então,
$$\text{“André Caldas”} \in A \setminus B.$$

Exemplo 1.5. Se $A = \{1, 2, 3, 4\}$ e $B = \{2, 4, 6, 8\}$, então
$$A \setminus B = \{1, 3\}.$$

1.4 Produto Cartesiano

Se A e B são conjuntos, o *produto cartesiano* de A e B é o conjunto formado pelos pares (a, b), onde $a \in A$ e $b \in B$.

Da mesma forma, podemos definir $A_1 \times \cdots \times A_n$ como sendo o conjunto das *n-uplas ordenadas* $(a_1, \ldots, a_n)$, onde $a_1 \in A_1, \ldots, a_n \in A_n$. Quando todos os A_j são iguais a um mesmo A, então, ao invés de escrevermos $A \times \cdots \times A$, escrevemos A^n.

Exemplo 1.6. O conjunto $\mathbb{R}^3$ é formado pelas *ternas ordenadas* de números reais. Por exemplo, alguns elementos de $\mathbb{R}^3$ são: $(5, 4, 2)$, $(\pi, \mathrm{e}, 7.5)$ e $(5000, \pi^{\mathrm{e}}, \sqrt{2})$.

Notação 1.7. *Neste livro, quando $a \in A^n$, vamos usar o símbolo a_j como referência ao j-ésimo termo de a. Ou seja, vamos escrever apenas $a \in A^n$, e estará implícito que $a = (a_1, \ldots, a_n)$. No caso de $\mathbb{R}^n$, será comum usarmos uma* setinha *em cima, pra lembrar que estamos falando de um elemento de $\mathbb{R}^n$. Nesse caso, também estará implícito que $\vec{a} = (a_1, \ldots, a_n)$.*

Exemplo 1.8 (Muito importante!!! Aprenda bem este exemplo!!!)**.**

Sabemos que

$$(x+y)^3 = xxx + xxy + xyx + yxx + xyy + yxy + yyx + yyy.$$

Na escolha acima, aparecem todas as combinações possíveis de uma sequência de três elementos de $\{x, y\}$. Ou seja, aparecem todos os elementos de $\{x, y\}^3$. Assim, podemos escrever

$$(x+y)^3 = \sum_{a \in \{x,y\}^3} a_1 a_2 a_3.$$

Onde o símbolo $\sum$ está indicando que devemos somar uma vez para cada elemento $a \in \{x, y\}^3$.

Mas atenção! Aqui, x não é uma variável. É apenas uma letra… um símbolo. A letra x é diferente da letra y. Estaria errado escrever algo como

$$(1+1)^3 \overset{!}{=} \sum_{a \in \{1,1\}^3} a_1 a_2 a_3 = 1.$$

Pois $\{1, 1\} = \{1\}$. E portanto, $\{1, 1\}^3 = \{(1, 1, 1)\}$ só tem um elemento!

1.5 Alguns Conjuntos Famosos

Alguns conjuntos conhecidos, utilizados neste curso, estão listados a seguir. Temos os números naturais, os números naturais não nulos e os inteiros.

$$\begin{aligned} \mathbb{N} &= \{0, 1, 2, \ldots\} \\ \mathbb{N}^* &= \{1, 2, \ldots\} \\ \mathbb{Z} &= \{\ldots, -2, -1, 0, 1, 2, \ldots\}. \end{aligned}$$

Algumas pessoas usam o símbolo $\mathbb{N}$ para se referir ao que estamos chamando de $\mathbb{N}^*$.

Neste curso, vamos usar a notação

$$[n] = \{1, \dots, n\}$$

para representar os primeiros n naturais não nulos.

Assumimos que o leitor já entende mais ou menos o que é o conjunto $\mathbb{R}$, dos números reais. Algumas variações do conjunto dos números reais são

$$\begin{aligned}\mathbb{R}^+ &= \left\{x \in \mathbb{R} \,\middle|\, x \geq 0\right\} \\ \mathbb{R}^- &= \left\{x \in \mathbb{R} \,\middle|\, x \leq 0\right\},\end{aligned}$$

que são os reais positivos e negativos. Algumas pessoas não gostam de chamar o número 0 de positivo ou negativo. Por isso, costumam preferir dizer reais não negativos para se referir a $\mathbb{R}^+$ e reais não positivos para se referir a $\mathbb{R}^-$. Outros vão usar a notação $\mathbb{R}^+$ e $\mathbb{R}^-$ para denotar os reais estritamente positivos e os estritamente negativos. Ou seja, excluem o 0 da definição.

Por fim, tem o conujunto dos números complexos

$$\mathbb{C} = \left\{a + bi \,\middle|\, a, b \in \mathbb{R}\right\}.$$

Não vamos precisar muito dos números complexos. Apenas em um ou outro exemplo.

1.6 Exercícios

Exercício 1.1. Escreva todos os elementos de A^n.

1. $A = \{1, 2, 3\}$, $n = 2$.

   Solução.

   $$A^n = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)\}.$$

2. $A = \{1, 2\}$, $n = 3$.
3. $A = \{$🐮, 🐱$\}$, $n = 3$.
4. $A = \{1, 2, 3, 4, 5\}$, $n = 1$.
5. $A = \{1,$ 🐮$, 1, 7, 7,$ 🐱$\}$, $n = 2$.

Exercício 1.2. Escreva o somatório dos elementos do conjunto finito $A \subset \mathbb{R}$:

$$\sum_{\boxed{}} \boxed{}.$$

Exercício 1.3. Escreva o produtório dos elementos do conjunto finito $A \subset \mathbb{R}$:

$$\prod_{\boxed{}} \boxed{}.$$

Exercício 1.4. Escreva usando somatório.

1. $$(a_1 + a_2)^3 = \sum_{\boxed{}} \boxed{}.$$

2. $$(x + y + z)^5 = \sum_{\boxed{}} \boxed{}.$$

3. $$(x + y + z + w)^{25} = \sum_{\boxed{}} \boxed{}.$$

Exercício 1.5. Escreva usando somatório e produtório.

1.
$$(x+y+z)^{23} = \sum_{\boxed{}} \left(\prod_{\boxed{}} \boxed{} \right).$$

2.
$$\left(\sum_{a \in A} a\right)^{23} = \sum_{\boxed{}} \left(\prod_{\boxed{}} \boxed{} \right).$$

3.
$$\left(\sum_{a \in A} a\right)^{n} = \sum_{\boxed{}} \left(\prod_{\boxed{}} \boxed{} \right).$$

4.
$$\left(\sum_{j=1}^{25} a_j\right)^{37} = \sum_{\boxed{}} \left(\prod_{\boxed{}} \boxed{} \right).$$

Exercício 1.6. Suponha que $\mathscr{F}$ é um conjunto finito de símbolos que podem ser multiplicados uns com os outros, somados e multiplicados por números reais (assim como matrizes). Também como no caso de matrizes, **não** assuma que o produto é comutativo. Ou seja, dados $X, Y \in \mathscr{F}$, pode ser que $XY \neq YX$. Além disso, se $\alpha, \beta \in \mathbb{R}$, $\alpha X \beta Y = \alpha\beta XY$.

Escreva com somatório:

1.
$$(\alpha_A A + \alpha_B B + \alpha_C C)^2 = \sum_{\boxed{}} \alpha_{\boxed{}} \alpha_{\boxed{}} \boxed{}.$$

2.

$$\left(\sum_{A\in\mathscr{F}}\alpha_A A\right)^n = \sum_{\boxed{}}\left(\left(\prod_{\boxed{}}\alpha_{\boxed{}}\right)\left(\prod_{\boxed{}}\boxed{}\right)\right).$$

Aula 2

Funções

Dados dois conjuntos A e B, uma função $f : A \to B$ é um objeto matemático tal que para cada $x \in A$ podemos perguntar:

> Qual é o elemento de B que corresponde a x?

Em outras palavras, quem é $f(x)$? A resposta deve depender apenas do elemento escolhido $x \in A$. Não pode depender de quem perguntou, nem de quantas vezes perguntou, nem do momento em que perguntou, ou de quais perguntas foram feitas antes, e nem do humor de quem está respondendo. Sempre que perguntada, para um mesmo $x \in A$, a resposta deve ser sempre a mesma: $f(x)$.

Exemplo 2.1. Quando A é um conjunto finito. Uma função $f : A \to B$ pode ser representada por uma tabela. Por exemplo, se

$$A = \{🐮, 🐵, 🐱\} \quad \text{e} \quad B = \mathbb{N},$$

Para descrevermos uma função f tal que $f(🐮) = 2$ e $f(🐵) = f(🐱) = 0$, podemos, por exemplo, usar uma tabela.

🐮	🐵	🐱
2	0	0

E se for uma função $g : A \to \mathbb{R}^3$, nossa tabela pode ter três colunas:

🐮	🐵	🐱
2.5	1	3.1
−5	7.1	3.14
3	0.1	3.141

Neste caso, por exemplo, $g($🐮$) = (2.5, -5, 3)$.

Uma maneira comum de se definir uma função é utilizando uma fórmula.

Exemplo 2.2. Podemos definir uma função $f : \mathbb{N} \to \mathbb{N}$ que associa a cada $n \in \mathbb{N}$ o triplo do seu valor. Escrevemos

$$\begin{array}{rccc} f: & \mathbb{N} & \to & \mathbb{N} \\ & x & \mapsto & 3x \end{array}.$$

A função f responde, para cada $x \in \mathbb{N}$, a pergunta:

> Qual é o triplo de x?

Da mesma maneira, poderíamos definir $f : \mathbb{N} \to \mathbb{R}$, ou $f : \mathbb{R} \to \mathbb{R}$. Mas seria incorreto definir

$$\begin{array}{rccc} f: & \mathbb{R} & \to & \mathbb{N} \\ & x & \mapsto & 3x \end{array},$$

pois, em geral, o triplo de um número real é um número natural!

Exemplo 2.3 (Identidade)**.** Para cada conjunto A, podemos definir a função *identidade*, que é a função que "não faz nada". Que associa um elemento a ele mesmo.

$$\begin{array}{rccc} \mathrm{id}_A : & A & \to & A \\ & x & \mapsto & x \end{array}.$$

Quando o contexto deixar claro a qual conjunto nos referimos, costumamos omitir o índice e escrever apenas id no lugar de id_A.

2.1 Domínio, contradomínio e imagem

Seja $f : A \to B$ uma função. O conjunto A é chamado de *domínio* de f, e B é o *contradomínio*. Se $D \subset A$, então escrevemos $f(D)$ para representar o conjunto

$$f(D) = \left\{ f(a) \,\middle|\, a \in D \right\}.$$

Nem sempre temos que $f(A) = B$. O conjunto $f(A)$ é chamdo de *imagem* de f, e é denotado por

$$\mathrm{Im}(f) = f(A).$$

E quando $f(A) = B$, dizemos que f é *sobrejetiva*.

Exemplo 2.4. Considere a função

$$\begin{array}{rccc} f : & \mathbb{N} & \to & \mathbb{N} \\ & x & \mapsto & 2x \end{array}.$$

O *domínio* de f é $\mathbb{N}$. O *contradomínio* também é $\mathbb{N}$. Mas a imagem, $f(\mathbb{N})$, é o conjunto dos naturais pares. A função f é injetiva, mas não é sobrejetiva.

2.2 Composição de funções

É importante que o leitor tenha em mente, que uma função não é uma fórmula, mas uma correspondência. Por vezes, podemos expressar essa correspondência usando uma fórmula.

A rigor, não deveríamos dizer que $f(x)$ é uma função. A função é f. Enquanto que $f(x)$ é o valor que f toma em x. Podemos imaginar uma função $f : A \to B$ como uma máquina, um dispositivo... uma *"caixinha"*!

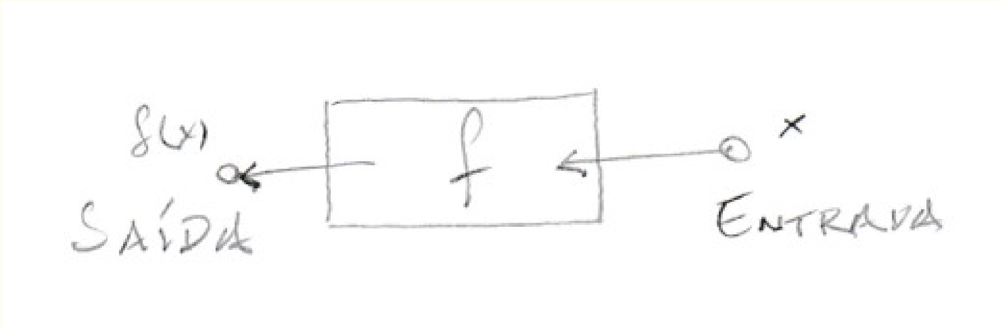


Figura 2.1: A função f pode ser vista como um dispositivo que tem uma entrada (input) e uma saída correspondente (output).

Um dipositivo que sempre que recebe uma entrada (input) x, emite uma resposta (output) $f(x)$. Se é, por exemplo, uma máquina onde você coloca uma fruta e sai suco dessa fruta, não podemos colocar uma bola de tênis! Os únicos *inputs* permitidos são as frutas. Ou seja, o *domínio* de f é o conjunto das frutas. É bom também deixarmos claro quem é o contradomínio. Poderia ser o conjunto de todos os tipos de bebida, ou o conjunto de todos os tipos de líquido, ou simplesmente, o conjunto de todos os tipos de suco de fruta.

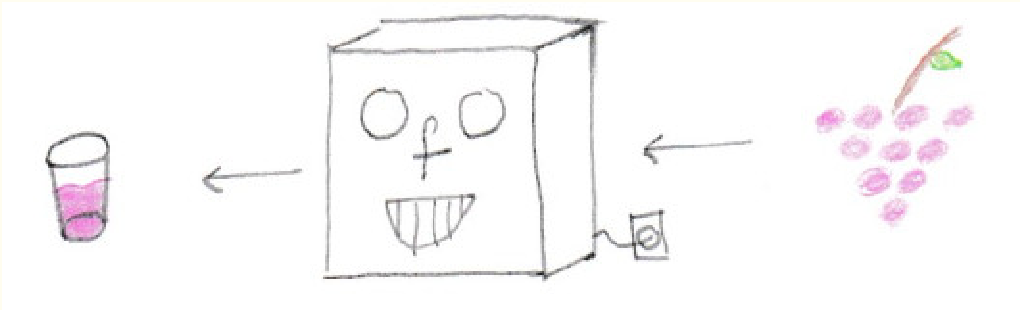


Figura 2.2: Poderíamos descrever a função *suco de fruta* por uma fórmula:

$$\begin{array}{rcl} f: \text{FRUTAS} & \to & \text{SUCOS} \\ x & \mapsto & \text{suco de ``}x\text{''} \end{array}.$$

Imagine uma função

$$g : \text{SUCOS} \to \text{PICOLÉS}.$$

Se ligássemos a saída de f com a entrada de g, teríamos uma máquina

$$g \circ f : \text{FRUTAS} \to \text{PICOLÉS}.$$

Veja a figura 2.3.

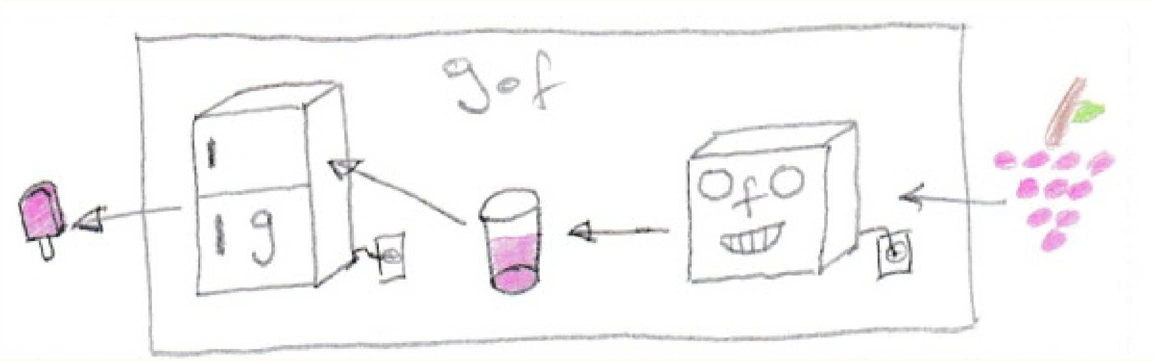


Figura 2.3: A *"máquina"* $g \circ f$ é a composição de g com f, e leva a fruta "x" em "picolé de x".

De um modo geral, dadas duas funções $f : A \to B$ e $g : B \to C$, definimos a composição $g \circ f$, como sendo a função de A em C que toma $x \in A$, e leva em $g(f(x))$. De modo mais rigoroso,

$$\begin{array}{rccc} g \circ f : & A & \to & C \\ & x & \mapsto & g(f(x)) \end{array}.$$

O nome da função é $g \circ f$. Usamos a notação $(g \circ f)(x)$ para aplicarmos $x \in A$ na função $g \circ f$. No final das contas,

$$(g \circ f)(x) = g(f(x)).$$

Podemos complicar ainda mais, e pensar na composição de mais funções… por exemplo, dadas as funções

$$\begin{aligned} h &: A \to B \\ g &: B \to C \\ f &: C \to D, \end{aligned}$$

qual é o significado de $f \circ g \circ h$? Veja a figura 2.4. É como se ligássemos a saída de h na entrada de g, e a saída de g na entrada de f, para formarmos uma "caixa"

$$\begin{array}{rccl} f \circ g \circ h: & A & \to & D \\ & x & \mapsto & f\left(g\left(h\left(x\right)\right)\right) \end{array}.$$

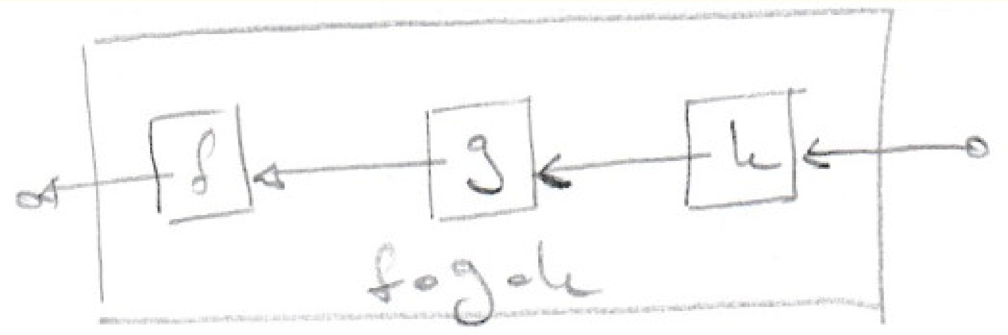


Figura 2.4: Da mesma forma que definimos a composição de duas funções, podemos definir a composição de três ou mais.

Por outro lado, $g \circ h$ é o nome da "caixa" que leva x em $g(h(x))$. Vamos chamar essa "caixa" de $G = g \circ h$. A notação $f \circ (g \circ h)$ representa — como ilustrado na figura 2.5 — a composição de f com a "caixa" $g \circ h$. Temos

$$\begin{array}{rccl} f \circ (g \circ h): & A & \to & D \\ & x & \mapsto & f\left((g \circ h)(x)\right) \end{array}.$$

Da mesma forma, se compusermos a caixa $f \circ g$ com h, temos

$$\begin{array}{rccl} (f \circ g) \circ h: & A & \to & D \\ & x & \mapsto & \left(f \circ g\right)\left(h(x)\right) \end{array}.$$

Mas no final das contas, é tudo a mesma coisa!

$$\begin{aligned} (f \circ g \circ h)(x) &= f\left(g\left(h(x)\right)\right) \\ (f \circ (g \circ h))(x) = f\left((g \circ h)(x)\right) &= f\left(g\left(h(x)\right)\right) \\ ((f \circ g) \circ h)(x) = \left(f \circ g\right)\left(h(x)\right) &= f\left(g\left(h(x)\right)\right). \end{aligned}$$

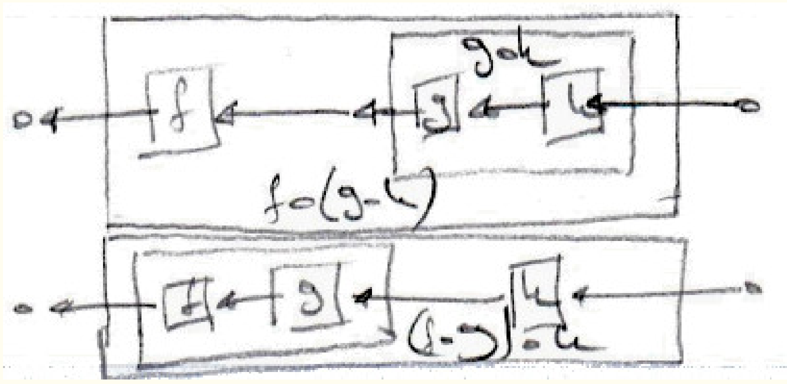


Figura 2.5: A composição de três funções pode ser feita em etapas. Primeiro fazemos a composição de g e h para formarmos a "caixinha" $g \circ h$. Depois fazemos a composição de f com $g \circ h$ para formarmos $f \circ (g \circ h)$. Da mesma forma, podemos fazer $f \circ g$ e depois compor com h. O resultado final é o mesmo: $f \circ g \circ h$.

Com ou sem as "caixas intermediárias" $f \circ g$ e $g \circ h$, o resultado corresponde a tomar x, passar por h, depois por g e depois por f. Ou seja, a operação de *composição* é *associativa*:

$$f \circ (g \circ h) = f \circ g \circ h = (f \circ g) \circ h. \tag{2.2.1}$$

Observação 2.5. A composição de funções é *associativa*:

$$(f \circ g) \circ h = f \circ (g \circ h).$$

Mas a composição não é *comutativa*! Ou seja, pode acontecer de $f \circ g$ ser diferente de $g \circ f$.

Exemplo 2.6. Sejam

$$\begin{array}{rccl} f: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & |x| \\ g: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & -x \end{array}.$$

Então,

$$\begin{array}{rccl} f \circ g: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & |x| \\ g \circ f: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & -|x| \end{array}.$$

O leitor familiarizado com produto de matrizes deve lembrar que o produto de matrizes é associativo, mas não é comutativo. E pode ser que, dadas duas matrizes A e B, o produto AB esteja definido, mas o produto BA não esteja.

Observação 2.7. Sejam $f : \mathbb{R}^3 \to \mathbb{R}^2$ e $g : \mathbb{R}^4 \to \mathbb{R}^3$. Então, $f \circ g$ está definido, mas $g \circ f$ não está. O primeiro caso está definido porque a saída de g (dimensão 3) é igual à entrada de f. O segundo caso não está definido porque a saída de f (dimensão 2) NÃO é igual à entrada de g (dimensão 4).

Da mesma forma, se F é uma matriz 2×3 e G é uma matriz 3×4, então o produto FG está definido porque o número de linhas de G (dimensão 3) é igual ao número de colunas de F. Mas GF não está definido porque o número de linhas de F (dimensão 2) NÃO é igual ao número de colunas de G (dimensão 4).

Ainda vamos falar sobre produto de matrizes no curso. Lembre-se desta semelhança entre o produto de matrizes e a composição de funções.

2.3 Injetividade, sobrejetividade e etc.

Já sabemos que quando a imagem $\mathrm{Im}(f)$ de uma função $f : A \to B$ é igual ao contradomínio B, dizemos que a função é *sobrejetiva*. Uma forma mais complicada de dizer isso é assim:

Para todo $b \in B$, existe $a \in A$ tal que $f(a) = b$.

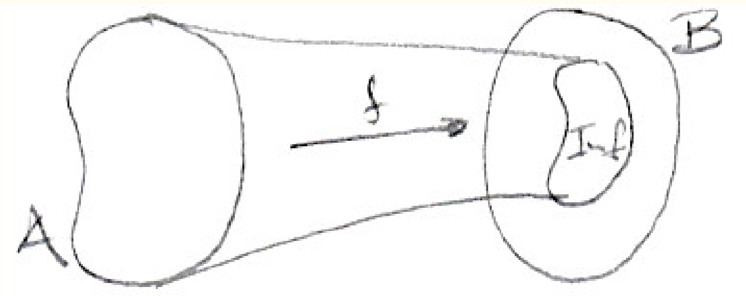


Figura 2.6: A função f, representada esquematicamente na figura, **não** é sobrejetiva.

Por outro lado, se $b = f(a)$, será que esse $a \in A$ é o único tal que $b = f(a)$? Ou será que existem outros elementos de A que tem a mesma imagem b? Para cada $x \in A$, quando não existe mais nenhum outro $y \in A$ tal que $f(x) = f(y)$, dizemos que f é injetiva. Matematicamente, dizemos que f é injetiva quando para todo $x, y \in A$,

$$f(x) = f(y) \Rightarrow x = y.$$

Ou então,

$$x \neq y \Rightarrow f(x) \neq f(y).$$

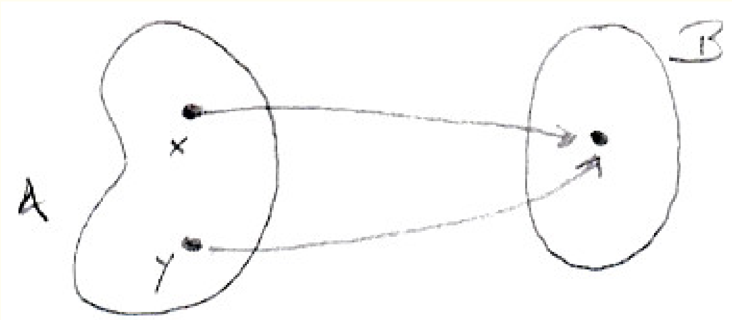


Figura 2.7: A função f, representada esquematicamente na figura, **não** é injetiva.

Quando uma função $f : A \to B$ é injetiva e sobrejetiva, dizemos que é *bijetiva*. A *sobrejetividade* de f significa que o sistema

$$f(x) = b \tag{2.3.1}$$

sempre pode ser resolvido, qualquer que seja $b \in B$. *Injetividade* significa que a equação (2.3.1) é tal que textbfquando a solução existir, será única. Juntando as duas coisas, *bijetividade* significa que o sistema **sempre pode ser resolvido** de maneira **única**.

Exemplo 2.8. A função

$$\begin{array}{rccc} f: & \mathbb{N} & \to & \mathbb{N} \\ & n & \mapsto & n+1 \end{array}$$

é *injetiva* mas não é *sobrejetiva*.

Se estamos trabalhando com conjuntos infinitos, podemos ter funções que são *sobrejetivas* mas não são *injetivas* e vice-versa. Mas quando os conjuntos são finitos e tem o mesmo número de elementos, isso não acontece.

Proposição 2.9. *Sejam A e B conjuntos finitos com o mesmo número de elementos. Então, dada uma função $f : A \to B$, são equivalentes:*

1. f *é* injetiva.
2. f *é* sobrejetiva.
3. f *é* bijetiva.

2.4 Funções bem definidas

O título desta subseção deveria ser *"funções **mal** definidas"*!!! Suponha que queiramos definir uma função f, especificando seu domínio, seu contradomínio, e uma *fórmula*. Existem vários erros que podemos cometer, que fazem com que nossa definição não faça sentido. Vamos analisar esses casos.

Exemplo 2.10 (Sem correspondente no contra-domínio)**.** Já vimos que

$$\begin{array}{rccc} f: & \mathbb{R} & \to & \mathbb{N} \\ & x & \mapsto & 3x \end{array}$$

não está bem definida. Isso porque, $3x$ pode não ser um número natural. No entanto,

$$\begin{array}{rccc} f: & \mathbb{N} & \to & \mathbb{N} \\ & x & \mapsto & 3x \end{array}$$

está bem definida.

Exemplo 2.11 (Definição ambígua)**.** Sabemos que todos os números racionais podem ser escritos da forma $\frac{a}{b}$, onde a e b são números inteiros.

Sabendo disso, poderíamos tentar definir uma função do tipo

$$\begin{array}{rccc} f: & \mathbb{Q} & \to & \mathbb{Z} \\ & \frac{a}{b} & \mapsto & a+b \end{array}.$$

Esta f está mal definida! Isso porque, por exemplo, $\frac{1}{1} = \frac{2}{2}$. Ou seja, existem duas maneiras diferentes de escrever um mesmo número racional como uma fração. Assim,

$$2 = 1 + 1 = f\left(\frac{1}{1}\right) \overset{!}{=} f(1) \overset{!}{=} f\left(\frac{2}{2}\right) = 2 + 2 = 4.$$

O problema aqui, é que existem duas maneiras de escrever o racional 1 como uma fração de números inteiros, e cada uma resulta em um valor diferente para $f(1)$.

Exemplo 2.12 (Não está definida para todo mundo)**.** Numa tentativa de definir a função *raiz quadrada*, poderíamos fazer assim:

$$\begin{array}{rccc} f: & \mathbb{R} & \to & \mathbb{R}^+ \\ & x^2 & \mapsto & x \end{array}.$$

Ou seja, pra calcular a raiz de 4, por exemplo, escrevemos $4 = 2^2$ e utilizamos a *fórmula*

$$f(2^2) = 2.$$

Mas essa função não está bem definida!

Por exemplo, quanto vale $f(-1)$?

O problema com essa definição é que nem todo elemento do domínio pode ser escrito como como um número real ao quadrado! Não estamos definindo o valor de f para todo o domínio!

Ao definirmos uma função, precisamos ficar atentos para

1. Definir f para TODOS os elementos do domínio.
2. Que a fórmula leve os elementos do domínio em algum elemento do contra domínio.
3. Garantir que se houver várias maneiras de expressar determinado elemento do domínio, que elas levem sempre ao mesmo resultado. Uma maneira fácil de evitar esse problema é, por exemplo, garantindo que cada elemento do domínio possa ser escrito de uma única maneira.

2.5 Função inversa

Dada uma função $f : A \to B$, quando é que a função

$$\begin{array}{rrcl} f^{-1}: & B & \to & A \\ & f(x) & \mapsto & x \end{array}$$

está bem definida?

A função *"suco de fruta"* é inversível? E a função *"picolé"*, que pega um suco e leva num picolé? Bem... a função *"picolé"* parece ser inversível, já que basta derreter o picolé para ter o suco de volta.

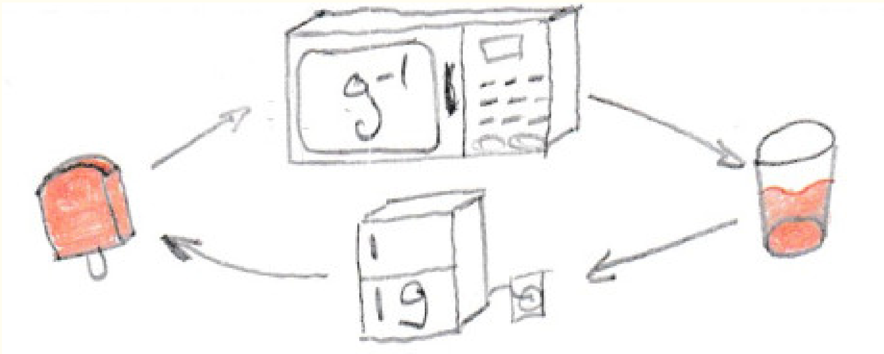


Figura 2.8: A inversa de g é

$$g^{-1}: \begin{array}{rcl} \text{PICOLÉS} & \to & \text{SUCOS} \\ \text{picolé de “}x\text{”} & \mapsto & \text{suco de “}x\text{”} \end{array}$$

Fisicamente, parece ser impossível inverter a função *“suco de fruta”*. Mas os matemáticos não estão muito interessados em saber se é *fisicamente* possível ou não. No nosso caso, inverter significa fazer a correspondência inversa. No caso do suco, significa dizer de qual fruta veio o tal suco. De qual fruta veio o *suco de limão*? Ora... veio do *limão*!!!

Exemplo 2.13. A função

$$f: \begin{array}{rcl} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & |x| \end{array}$$

não é inversível. Mas a função

$$g: \begin{array}{rcl} \{x \in \mathbb{R} \mid x \leq 0\} & \to & \{x \in \mathbb{R} \mid x \geq 0\} \\ x & \mapsto & |x| \end{array}$$

é inversível.

2.6 Exercícios

Exercício 2.1. Classifique como *injetiva*, *sobrejetiva*, *bijetiva* e *nem injetiva e nem sobrejetiva.*

1. $f: \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^2 \end{array}.$

2. $f: \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^3 \end{array}.$

3. $f: \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}^+ \\ x & \mapsto & x^2 \end{array}.$

4. $f: \begin{array}{ccc} \mathbb{R}^+ & \to & \mathbb{R}^- \\ x & \mapsto & -x^2 \end{array}.$

5. $f: \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^3 + x^2 \end{array}.$

6. $f: \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & 10 + x^3 \end{array}.$

7. $f: \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}^2 \\ x & \mapsto & (x, 2x) \end{array}.$

8. $f: \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x + y \end{array}.$

9. $f: \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (x,y) & \mapsto & (xy, x + y) \end{array}.$

Exercício 2.2. As funções a seguir **não** estão bem definidas. Justifique.

1. $f: \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}^+ \\ x & \mapsto & x^3 \end{array}.$

2. $$\begin{array}{rccl} f: & \mathbb{R}^+ & \to & \mathbb{R} \\ & x^2 & \mapsto & x \end{array}.$$

3. $$\begin{array}{rccl} f: & \mathbb{Q} & \to & \mathbb{R}^2 \\ & \frac{a}{b} & \mapsto & (a,b) \end{array}, \text{ onde } a \in \mathbb{Z} \text{ e } b \in \mathbb{N}^*.$$

Exercício 2.3. Considere as funções

$$\begin{array}{rcclrccl} f: & \mathbb{R} & \to & \mathbb{R} & g: & \mathbb{R} & \to & \mathbb{R}^2 \\ & x & \mapsto & x^3 + 10x & & x & \mapsto & (x^2, x^3) \\ h: & \mathbb{R}^2 & \to & \mathbb{R} & k: & \mathbb{R} & \to & \mathbb{R} \\ & (x,y) & \mapsto & x + y & & x & \mapsto & |x| \end{array}.$$

As seguintes composições estão bem-definidas? Para as que estiverem bem-definidas, escreva: domínio, contra-domínio e uma fórmula.

$$(f \circ g) \qquad (h \circ g) \qquad (h \circ h) \qquad (f \circ f) \qquad (h \circ g) \circ f \qquad (k \circ h) \circ (g \circ f) \circ f$$

Exercício 2.4. Escreva com somatórios e produtórios usando funções como variáveis do somatório e do produtório.

1. $(a_1 + a_2)^3$.

   Solução.
   $$(a_1 + a_2)^3 = \sum_{f:[3]\to[2]} a_{f(1)} a_{f(2)} a_{f(3)}.$$

2. $(a_1 + a_2)(b_1 + b_2)(c_1 + c_2)$.

3. $(a_1 + \cdots + a_m)^n$.

4. $(a_1 + \cdots + a_m)(b_1 + \cdots + b_m)$.

5. $(a_{11} + \cdots + a_{1n})(a_{21} + \cdots + a_{2n})$.

6. Troque a ordem do somatório e do produtório:

$$\prod_{j=1}^{n}\left(\sum_{k=1}^{m} a_{jk}\right) = \sum_{\sigma:\boxed{}\to\boxed{}}\left(\prod_{j=1}^{n}\boxed{}\right).$$

Exercício 2.5. Faça como no item 2 do exercício 1.6, mas utilizando funções $f : [n] \to \mathscr{F}$ ao invés de usar o conjunto $\mathscr{F}^n$:

$$\left(\sum_{A\in\mathscr{F}} \alpha_A A\right)^n = \sum_{\boxed{}}\left(\left(\prod_{\boxed{}}\alpha_{\boxed{}}\right)\left(\boxed{}\right)\right).$$

Exercício 2.6. No exercício 2.5, se além de tudo, soubermos que qualquer produto que tenha ao menos um símbolo repetido é sempre igual a zero. Por exemplo, $ABCDEFAGH = 0$, pois o símbolo A aparece duas vezes. Então, se $n = \#\mathscr{F}$, podemos usar funções bijetivas $\sigma : [n] \to \mathscr{F}$, e também, colocar o produto dos números reais em evidência:

$$\left(\sum_{A\in\mathscr{F}} \alpha_A A\right)^n = \left(\prod_{\boxed{}}\alpha_{\boxed{}}\right)\sum_{\sigma:[n]\to\mathscr{F}\text{bijetiva}}\left(\boxed{}\right).$$

Exercício 2.7. Assim como no exercício 2.6, se soubermos que qualquer produto que tenha ao menos um símbolo repetido é sempre igual a zero. Por exemplo, $ABCDEFAGH = 0$, pois o símbolo A aparece duas vezes. Mas dessa vez, suponha que $n > \#\mathscr{F}$. Justifique porque nesse caso,

$$\left(\sum_{A\in\mathscr{F}} \alpha_A A\right)^n = 0.$$

Exercício 2.8. Seja

$$P = \left\{a_1X^1 + \cdots + a_{10}X^{10} \;\middle|\; a_1, \cdots, a_{10} \in \mathbb{R}\right\}$$

o conjunto dos polinômios de grau até 10. Sabendo que determinada função $f : P \times P \to \mathbb{R}$ satisfaz a lei de distributividade (*"chuveirinho"*)

$$\begin{aligned} f(p(x)+r(x),q(x)) &= f(p(x),q(x)) + f(r(x),q(x)) \\ f(p(x),q(x)+s(x)) &= f(p(x),q(x)) + f(p(x),s(x)). \end{aligned}$$

Escreva $f(a_1X^1 + \cdots + a_{10}X^{10}, b_1X^1 + \cdots + b_{10}X^{10})$ na forma

$$\sum_{\sigma:\boxed{}\to\boxed{}} \boxed{}.$$

Exercício 2.9. No exercício 2.8, se ainda por cima, f satisfaz

$$\begin{aligned} f(\alpha p(x),q(x)) &= \alpha f(p(x),q(x)) \\ f(p(x),\beta q(x)) &= \beta f(p(x),q(x)), \end{aligned}$$

então, podemos colocar os a_j e b_j pra "fora":

$$\sum_{\sigma:\boxed{}\to\boxed{}} \boxed{}.$$

Aula 3

Arranjos e Permutações

O conteúdo desta seção será necessário apenas quando estudarmos *determinantes*, na seção 13.3. Em uma primeira leitura, o conteúdo é um pouco difícil e o leitor não deve esperar compreendê-lo por completo. É necessário um pouco de tempo, paciência e maturidade para se acostumar com a notação. Por isso é importante que o leitor tente fazer os exercícios e entender os exemplos.

3.1 Escolhendo em A^n

No exemplo 1.7, vimos como o conjunto A^n representa todas as escolhas possíveis de uma sequência de n elementos de A. Vamos exercitar um outra maneira de expressar o conjunto A^n.

Quando o conjunto tem k elementos, podemos enumerá-los:

$$A = \{x_1, x_2, \dots, x_k\}.$$

Enumerar os elementos de A equivale a escolher uma *bijeção*

$$p : \{1, \dots, k\} \to A.$$

Exemplo 3.1. Se

$$A = \{🐮, 🐵, 🐱\},$$

uma enumeração de A é dada pela seguinte tabela.

1	2	3
🐮	🐵	🐱

Esta mesma enumeração é dada pela função

$$\begin{aligned} p:\ \{1,2,3\} &\to A \\ n &\mapsto \begin{cases} 🐮, & n=1 \\ 🐵, & n=2 \\ 🐱, & n=3. \end{cases} \end{aligned}$$

Assim, fixada uma enumeração, cada $a \in A^n$ corresponde a uma função

$$f : \{1, \ldots, n\} \to \{1, \ldots, k\}.$$

Para cada f, temos o elemento de A^n

$$a^f = (x_{f(1)}, \ldots, x_{f(n)}).$$

Exemplo 3.2. Vamos continuar o exemplo 3.1, onde o conjunto

$$A = \{🐮, 🐵, 🐱\}$$

está enumerado pela função p, na ordem em que aparecem listados.

Com respeito a $a = (🐵, 🐱) \in A^2$, podemos dizer que

1. a entrada $a_1 =$ 🐵 corresponde ao **segundo** elemento de A; e
2. a entrada $a_2 =$ 🐱 corresponde ao **terceiro** elemento de A.

Dada a enumeração p, podemos pensar em $a \in A^2$ como equivalente à função

$$\begin{array}{rccl} f_a: & \{1,2\} & \to & \{1,2,3\} \\ & n & \mapsto & \begin{cases} 2, & n = 1 \\ 3, & n = 2. \end{cases} \end{array}$$

Já o elemento $b = ($🐱$,$🐱$) \in A^2$, corresponde à função constante

$$\begin{array}{rccl} f_b: & \{1,2\} & \to & \{1,2,3\} \\ & n & \mapsto & 3 \end{array}.$$

Para facilitar, vamos denotar por

$$[n] = \{1, \ldots, n\}$$

o conjunto dos números naturais de 1 até n.

Exemplo 3.3. Continuando o exemplo 1.7, ao invés de escrevermos

$$(x+y)^n = \sum_{a \in \{x,y\}^n} a_1 \cdots a_n,$$

podemos escrever

$$(x_1 + x_2)^n = \sum_{f:[n] \to [2]} x_{f(1)} \cdots x_{f(n)},$$

onde a soma é feita sobre todas as funções $f : [n] \to [2]$.

O objetivo desta seção é que o leitor se acostume com esse tipo de notação do exemplo 3.3.

3.2 Permutando Conjuntos

Permutar um conjunto A consiste em *"embaralhar"* seus elementos. Imaginamos que os elementos de A estão enfileirados — ou posicionados em locais específicos, em algum lugar — e então, *permutamos* os elementos de A, trocando uns pelos outros.

Falar sobre permutação se dá, então, em duas etapas:

1. Posicionar/organizar/ordenar os elementos de A.
2. Trocar/substituir os elementos uns pelos outros.

Matematicamente, é bastante simples falar sobre *"substituir os elementos de A uns pelos outros"*. É o mesmo que falar de uma *bijeção* $f : A \to A$. É isso que vamos chamar de *permutação* de um conjunto A.

Vamos denotar por

$$\mathcal{P}_A = \left\{ f : A \to A \,\middle|\, f \text{ é bijetiva} \right\}$$

o conjunto de todas as *permutações* de elementos de A.

Exemplo 3.4. Suponha que

$$A = \{🐮, 🐱, 🐵\}.$$

Estamos ignorando essa questão de *"onde estão posicionados inicialmente os elementos de A"*. Por isso, uma mesma permutação pode ser representada, em uma tabela, de várias formas diferentes.

🐮	🐵	🐱
🐱	🐮	🐵

é o mesmo que

🐵	🐮	🐱
🐮	🐱	🐵

Ambas as tabelas representam a mesma *permutação*, que associa 🐮 $\mapsto$ 🐱, 🐱 $\mapsto$ 🐵 e 🐵 $\mapsto$ 🐮.

Uma *permutação* de A também pode ser representada graficamente, como mostra o exemplo a seguir.

Exemplo 3.5. No exemplo 3.4, listávamos os elementos de A em colunas de uma tabela e, na linha seguinte, listávamos os elementos *permutados*. Agora, ao invés de permutarmos os elementos de lugar, vamos usar uma representação com linhas ligando os elementos que são permutados.

Figura 3.1: Escolhendo (🐮, 🐵, 🐱) ou (🐵, 🐮, 🐱) como "referência", podemos expressar a mesma permutação de modos diferentes.

3.3 Permutando Índices

Quando estudarmos *determinante* (seção 13.3), vamos precisar falar sobre *"permutação das colunas de uma matriz"*. Normalmente, as colunas da matriz estarão numeradas: primeira coluna, segunda coluna, etc.

Quando

$$A = \{x_1, x_2, \dots, x_n\}$$

tem n elementos que estão ordenados (numerados), temos uma identificação de A com $[n]$. Então, ao invés de permutarmos o conjunto A, podemos permutar o conjunto de índices $[n]$.

Trabalhar com números naturais tem uma vantagem: os números naturais são naturalmente ordenados. Ao invés de escrever

$$\sigma(1) = 2, \quad \sigma(2) = 3 \quad \text{e} \quad \sigma(3) = 1,$$

ou falar da tabela

$$\begin{array}{ccc} \hline 1 & 2 & 3 \\ \hline 2 & 3 & 1 \\ \hline \end{array},$$

vamos simplesmente escrever $\sigma = (2, 3, 1)$. Dessa forma, estamos identificando os elementos de $\mathcal{P}_3$ com os elementos de $\{1, 2, 3\}^3$ que não possuem entradas repetidas.

Com a representação gráfica, fica fácil determinar a composição de duas permutações.

Exemplo 3.6. Vamos compor a função $\sigma = (2, 3, 1)$ com a função $\gamma = (1, 3, 2)$. No caso de permutações, vamos usar a notação $\gamma\sigma$ no lugar de $\gamma \circ \sigma$.

Pela figura 3.2, é fácil ver que

$$\gamma\sigma = (3, 2, 1).$$

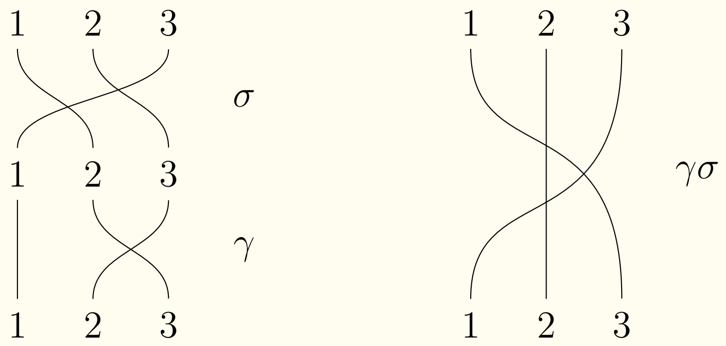


Figura 3.2: Determinar a composição $\gamma\sigma$ é fácil. Basta "escorregar" pelas linhas.

3.4 Sinal da Permutação

Neste curso, será importante estudar permutações para que possamos entender com facilidade o conceito de *determinante* que será apresentado na seção 13.3. Uma propriedade muito importante das permutações é que a cada permutação podemos associar um *sinal*.

Depois de compor várias permutações, a representação gráfica pode ficar bem "trançada". A figura 3.3, mostra como podemos simplificar o gráfico permutando os pontos intermediários. Os cruzamentos em excesso podem ser eliminados de dois em dois, como mostrado na figura 3.3.

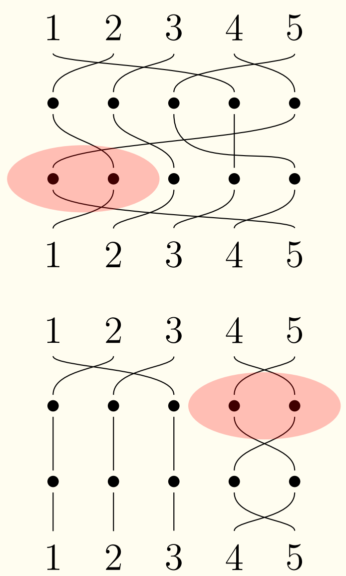


Figura 3.3: Depois de fazer a composição de várias permutações, a representação gráfica resultante pode ser cheia de "cruzamentos". Podemos simplificar o gráfico manipulando permutando pontos vizinhos nos passos intermediários. O número de cruzamentos é reduzido de dois em dois.

Assim, fazendo a composição de várias permutações, o número de cruzamentos em uma representação gráfica simplificada para a composição pode não ser igual à soma do número de cruzamentos em cada permutação. Mas terá sempre a mesma *paridade*: se uma quantidade é par, a outrea também é par; e se uma for ímpar, a outra será impar, também.

No caso específico da figura 3.3, a representação original tinha 13 cruzamentos. No segundo passo, se tornaram 11. Removemos mais dois e ficaram 9. A seguir, fizemos dois passos de uma vez e ficamos com 5 cruzamentos. Por fim, fizemos a última simplificação e terminamos com 3 cruzamentos.

Para que esse procedimento de contagem funcione, precisamos sempre tomar o cuidado de não permitir que mais de duas linhas se cruzem em um mesmo ponto. Tomando esse cuidado, ao eliminar ou produzir novos cruzamentos, os cruzamentos sempre aparecem ou desaparecem em **pares**.

Sendo assim, podemos definir o sinal de uma permutação de acordo com sua paridade. Dizemos que a permutação é *par* quando o número de cruzamentos for *par*, e que a permutação é *ímpar* quando o número e cruzamentos for *ímpar*.

$$\operatorname{sgn}(\sigma) = \begin{cases} 1, & \sigma \text{ tem um número par de cruzamentos} \\ -1, & \sigma \text{ tem número ímpar de cruzamentos} \end{cases}$$

Uma outra maneira de escrever isso é com a fórmula

$$\operatorname{sgn}(\sigma) = (-1)^{\text{Nº de cruzamentos em qualquer representação gráfica}}.$$

Pelas considerações feitas até aqui,

$$\begin{aligned} \operatorname{sgn}(\sigma\gamma) &= (-1)^{\text{paridade de } \sigma + \text{paridade de } \gamma} \\ &= (-1)^{\text{paridade de } \sigma}(-1)^{\text{paridade de } \gamma} \\ &= \operatorname{sgn}(\sigma)\operatorname{sgn}(\gamma). \end{aligned}$$

Ou seja, se soubermos o sinal das permutações γ e σ, então também sabemos o sinal de $\sigma\gamma$. Quando somamos dois números pares, o resultado é par. Então, a composição de duas permutações pares é uma permutação par. Uma par e uma ímpar dá ímpar. Já a composição de duas permutações ímpares dá uma permutação par.

3.5 Transposições

Uma permutação σ que troca dois números distintos i e j, e que deixa os demais fixos, é denotada por $\sigma = (i\,j)$ e é chamada de *permutação simples* ou *transposição*. As permutações simples tem duas propriedades especialmente interessantes.

Lema 3.7. *Se σ é simples, então*

$$\operatorname{sgn}(\sigma) = -1.$$

Demonstração. Assumindo que $i < j$, basta notar — observando a figura 3.4 — que as linhas saindo de i e j se cruzam uma vez (ímpar) e depois cada uma das duas cruza as retas $i+1$ até $j-1$ (par). Assim, o número de cruzamentos é ímpar. Ou seja, $\operatorname{sgn}(\sigma) = -1$.

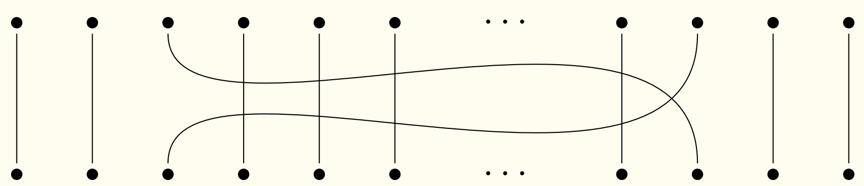


Figura 3.4: A representação gráfica de qualquer transposição $\sigma = (i\,j)$ mostra que $\operatorname{sgn}(\sigma) = -1$, pois os cruzamentos aparecemaos pares, com excessão de um.

Lema 3.8. *Toda permutação σ, de um conjunto finito A, pode ser escrita como composição de transposições.*

Observação 3.9. Normalmente, com respeito à permutação identidade — que não permuta ninguém, mantendo todos os elementos no mesmo lugar — entendemos que ela é um produto de *zero* transposições.

Demonstração. Vamos mostrar por indução no número de elementos que não são mantidos fixos. Se σ não permuta nenhum elemento, então o enunciado do lema é verdadeiro.

Assumindo que o lema vale quando o número de elementos que não são mantidos fixos é menor que n, vamos mostrar que vale quando esse número é igual a n. Seja $j \in A$ um elemento que não é mantido fixo por σ. Faça $k = \sigma(j)$. Pela escolha de j, k é diferente de j. Além disso, como σ é injetiva, k também não é mantido fixo.

Seja $\tau = (j\,k)$ a transposição que permuta j e k. Observe que a permutação $\gamma = \sigma\tau$ mantém k fixo. De fato,

$$\gamma(k) = \sigma(\tau(k)) = \sigma(j) = k.$$

Além disso, γ mantém fixos todos os elementos que já eram mantidos fixos por σ. De fato, se p é um tal elemento,

$$\sigma(\tau(p)) = \sigma(p) = p.$$

Aqui, a segunda igualdade se deve ao fato de p ser mantido fixo por σ. E a primeira igualdade se deve ao fato de p ser distinto de j e k, já que p é mantido fixo por σ, mas j e k não são.

Pela hipótese de indução, γ é um produto de transposições. Então,

$$\sigma = \sigma\tau\tau = \gamma\tau$$

também é, pois τ é uma transposição.

Agora, com os lemas 3.7 e 3.8, podemos interpretar o *sinal* de uma permutação como sendo a paridade do número de transposições necessárias para escrevê-la. A permutação σ pode ser escrita como o produto de um número **ímpar** de transposições exatamente quando $\operatorname{sgn}(\sigma) = -1$. E pode ser escrita com um número **par** de transposições exatamente quando $\operatorname{sgn}(\sigma) = +1$.

Corolário 3.10. *Escolhido $n \in \mathbb{N}^*$, existe uma única função*

$$\operatorname{sgn} : \mathcal{P}_n \to \{-1, 1\}$$

tal que

$$\begin{aligned} \operatorname{sgn}(\sigma) = 1 &\Leftrightarrow \sigma \textit{ é o produto de um núemro } \textbf{\textit{par}} \textit{ de transposições} \\ \operatorname{sgn}(\sigma) = -1 &\Leftrightarrow \sigma \textit{ é o produto de um núemro } \textbf{\textit{ímpar}} \textit{ de transposições.} \end{aligned}$$

Demonstração. O lema 3.8 mostra que

$$\sigma = \tau_1 \cdots \tau_n,$$

onde todas as τ_j são transposições. Pelo lema 3.7,

$$\begin{aligned} \operatorname{sgn}(\sigma) &= \operatorname{sgn}(\tau_1 \cdots \tau_n) \\ &= \operatorname{sgn}(\tau_1) \cdots \operatorname{sgn}(\tau_n) \\ &= (-1)^n. \end{aligned}$$

Essa igualdade conclui a demonstração.

A existência da função sgn tem implicações intrigantes!

Exemplo 3.11 (Identidade é par)**.** A permutação *identidade*

$$\begin{array}{rccc} \operatorname{id} : & [n] & \to & [n] \\ & x & \mapsto & x \end{array}$$

é o elemento de $\mathcal{P}_n$ que "não faz nada". Se

$$\operatorname{id} = \tau_1 \cdots \tau_k,$$

onde $\tau_1, \ldots, \tau_k$ são transposições, então k é necessariamente par.

3.6 Jogo do 15

Existe um quebra-cabeça chamado de *jogo do* 15 ou *racha-cuca*, que é formado por 15 quadrados deslizantes dispostos dentro de uma caixa de tamanho 4×4, como na figura 3.5. Veja o artigo

https://pt.wikipedia.org/wiki/O_jogo_do_15

da Wikipedia.

Figura 3.5: O *Racha-Cuca* é um quebra-cabeça famoso tal que o objetivo é colocar os quadrados numerados em uma determinada sequência, normalmente a sequência inicial.

O desafio é a partir de um quebra-cabeça embaralhado, ordená-lo como mostrado no canto direito inferior da figura 3.5. Será que o leitor consegue fazer isso com a configuração mostrada na figura 3.6?

Figura 3.6: Qual é a sequência de movimentos válidos que transforma o estado inicial no estado mostrado na figura?

Associando o quadrado vazio ao número 16, cada movimento do jogo pode ser representado como uma transposição do tipo $(j\,16)$. Cada transposição modifica a linha ou a coluna do quadrado 16, somando ou subtraindo 1. Dessa forma, uma sequência de transposições que mantenha o quadrado 16 imóvel deve ser necessariamente par, pois o número de vezes que o quadrado subir, deve também descer; e o número de vezes que for para a direita, deve também ir para a esquerda. Assim, todos os movimentos possíveis que comecem e terminem com o quadrado 16 na mesma posição são necessariamente permutações pares.

Por outro lado, a permutação σ que deva do estado inicial ao estado da figura 3.6 é uma permutação ímpar. De fato, $\sigma = (14\,15)$ é uma transposição.

Ou seja, é impossível levar o estado inicial do quebra-cabeça ao estado da figura 3.6 ou vice-versa!!! :-)

3.7 Exercícios

Exercício 3.1. Calcule o sinal das seguintes permutações de [5].

1. $(1, 3, 4, 5, 2)$
2. $(3, 2, 4, 5, 1)$
3. $(3, 2, 4, 1, 5)$
4. $(5, 4, 3, 2, 1)$

Exercício 3.2. Escreva as seguintes permutações de [6] como produto de transposições.

1. (2,3,6,1,4,5)
2. (1,3,2,5,4,6)
3. (6,5,4,3,2,1)
4. (3,2,1,6,5,4)

Exercício 3.3. O **??** explica como podemos conluir que a configuração onde aparecem torcados apenas os números 14 e 15 é impossível de se conseguir. Mas naquela explicação... eu não entendi nada!... :-(

Explique melhor o argumento, com suas palavras e de modo bem mais didático e convincente que o autor.

Exercício 3.4. Suponha que

$$\mathscr{F} = \{A_1, \dots, A_n\}$$

é um conjunto de n símbolos que podem ser multiplicados uns com os outros, somados e multiplicados por números reais (assim como matrizes). Também como no caso de matrizes, **não** assuma que o produto é comutativo. E também como com as matrizes, assuma que para $\alpha, \beta \in \mathbb{R}$, $\alpha X \beta Y = \alpha\beta XY$. Usando permutações $\sigma \in \mathcal{P}_n$, escreva

$$\begin{aligned}&(\alpha_{1,1}A_1 + \cdots + \alpha_{1,n}A_n)\,(\alpha_{2,1}A_1 + \cdots)\cdots(\alpha_{n,1}A_1 + \cdots + \alpha_{n,n}A_n)\\ &= \sum_{\boxed{}} \left(\boxed{}\right)\left(\boxed{}\right).\end{aligned}$$

Exercício 3.5. Por fim, como última variação do exercício 3.4, assuma que sempre que trocamos a ordem de dois símbolos no produto, o resultado se multiplica por -1. Por exemplo,

$$ABCDE = -BACDE \quad \text{e} \quad ABCDE = -ABEDC.$$

Nesse caso, podemos colocar o produto $A_1 A_2 \cdots A_n$ em evidência:

$$\begin{aligned}&(\alpha_{1,1}A_1 + \cdots + \alpha_{1,n}A_n)\,(\alpha_{1,1}A_1 + \cdots)\cdots(\alpha_{1,1}A_1 + \cdots + \alpha_{1,n}A_n)\\ &= \left(\sum_{\boxed{}} \left(\boxed{}\right)\right) A_1 A_2 \cdots A_n.\end{aligned}$$

Parte II

Introdução à Álgebra Linear

Aula 4

Espaço Vetorial $\mathbb{R}^n$

Em problemas físicos, falamos, por exemplo, em posição e velocidade. Quanto à posição, representamos por um *ponto*. A velocidade é um *vetor*.

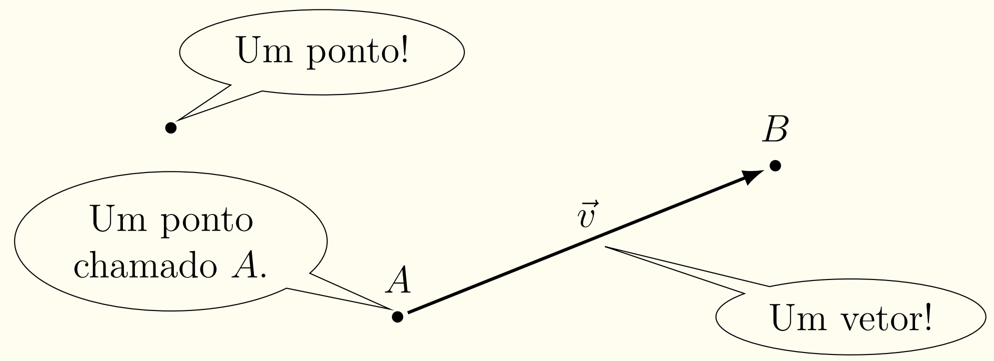


Figura 4.1: *Pontos* e um *vetor*.

Fixado um sistema de coordenadas, tanto os pontos do plano quanto os vetores de $\mathbb{R}^2$ podem ser representados por um par ordenado. Ou seja, por um elemento $(a, b) \in \mathbb{R}^2$. No caso tridimensional, por um elemento $(a, b, c) \in \mathbb{R}^3$. Neste capítulo, antes de interpretarmos fisicamente ou geometricamente o significado dos elementos de $\mathbb{R}^2$ ou $\mathbb{R}^3$, vamos simplesmente tratar de manipulações algébricas (cálculos) que podem ser feitas com duplas ou ternas ordenadas. Ou de modo mais geral, manipulações algébricas que podem ser feitas com elementos de $\mathbb{R}^n$. Esperamos que o leitor adquira habilidade suficiente para utilizar a linguagem vetorial para representar uma série de conceitos em diferentes contextos, que podem ir muito além da física e da geometria.

Exemplo 4.1 (batata e cenoura)**.** Certa vez, fui visitar a cidade *fimdomundo* (veja [Ban13]). Lá, descobri que os cidadãos de *fimdomundo* adoram fazer sopa de batata misturada com cenoura. Na loja de sopas e caldos, o cliente pede sua sopa assim:

— Gostaria de uma sopa $(20, 30)$.

No *fimdomundo*, todos sabem que isso significa 20 gramas de batata e 30 gramas de cenoura.

E se você misturar uma sopa $(20, 30)$ com uma sopa $(15, 28)$? Vai dar uma sopa $(35, 58)$! Se você tomar 2.5 tigelas da sopa $(20, 30)$ e 1.5 tigela da sopa $(15, 28)$, quanto de batata e quanto de cenoura vai ter ingerido? Fazendo as contas...

$$\begin{aligned} 2.5 \times (20, 30) + 1.5 \times (15, 28) &= (50, 75) + (22.5, 52) \\ &= (77.5, 127). \end{aligned}$$

Ou seja, 77.5 gramas de batata e 127 gramas de cenoura.

A multiplicação é feita *coordenada a coordenada*. A soma também é feita *coordenada a coordenada*.

Poderíamos fazer o mesmo, com mais ingredientes. Por exemplo: batata, cenoura, tomate e couve. Se você já ouviu alguém dizer que a quarta dimensão é o tempo... nesse exemplo, a quarta dimensão é a quantidade de couve!!! :-)

Exemplo 4.2 (cidades no mapa)**.** Escolhido um sistema de coordenadas num mapa, podemos associar coordenadas a cada ponto do mapa.

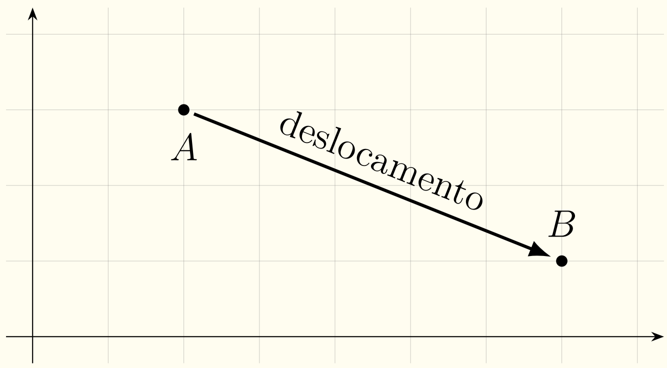


Figura 4.2: O ponto A tem coordenadas $(2, 3)$, e B tem coordenadas $(7, 1)$. Para saírmos de A e chegarmos em B, precisamos fazer um deslocamento correspondente à diferença

$$B - A = (5, -2).$$

Pra nós, vetores são simplesmente um elemento de $\mathbb{R}^n$: uma n-upla ordenada. Podem ser usados para representar posições e velocidades... mas também podem ser usados pra falar de quantidades de ingredientes, quantidades de peças, quantidades de nutrientes, etc. Qualquer coisa que possamos tratar como uma n-upla de números, que podemos *somar* e *multiplicar* por um outro número.

4.1 Operações Básicas em $\mathbb{R}^n$

Assim como com as batatas e cenouras no exemplo 4.1, ou a soma dos deslocamentos no exemplo 4.2, dados $\vec{a}, \vec{b} \in \mathbb{R}^n$, a *soma* de dois vetores $\vec{a}$

e $\vec{b}$ é o vetor $\vec{c}$ dado por

$$\vec{c} = (a_1 + b_1, \ldots, a_n + b_n).$$

Ou seja, para calcular a soma de dois vetores, fazemos o cálculo coordenada a coordenada. Escrevemos $\vec{c} = \vec{a} + \vec{b}$.

A soma de vetores é evidentemente comutativa e associativa:

$$\vec{a} + \vec{b} = \vec{b} + \vec{a}$$
$$(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c}).$$

Chamamos os elementos de $\mathbb{R}^n$ de *vetores* e os elementos de $\mathbb{R}$ de *escalares*. Dado um número $\alpha \in \mathbb{R}$ e $\vec{a} \in \mathbb{R}^n$, o *produto* de $\vec{a}$ pelo *escalar* α é o vetor

$$\alpha\vec{a} = (\alpha a_1, \ldots, \alpha a_n).$$

Tanto a *soma* quanto o *produto por escalar* são feitos *termo a termo*. Além das propriedades já mencionadas, a *soma* e o *produto por escalar* também satisfazem:

1. $0\vec{a} = \vec{0}$, onde $\vec{0} = (0, \ldots, 0)$.
2. $1\vec{a} = \vec{a}$.
3. $(\alpha\beta)\vec{a} = \alpha(\beta\vec{a})$.
4. $(\alpha + \beta)\vec{a} = \alpha\vec{a} + \beta\vec{a}$.
5. $\alpha(\vec{a} + \vec{b}) = \alpha\vec{a} + \alpha\vec{b}$.

O leitor certamente não terá dificuldades em interpretar essas propriedades em termos de batatas e cenouras!

A diferença entre dois vetores

$$\vec{c} = \vec{a} - \vec{b}$$

é o vetor $\vec{c}$ tal que

$$\vec{b} = \vec{a} + \vec{c}.$$

Geometricamente, fica claro que $\vec{c}$ é o vetor que sai da pontinha de $\vec{b}$ e vai até a pontinha de $\vec{a}$. Como a operação de soma é feita coordenada a coordenada, a diferença também é.

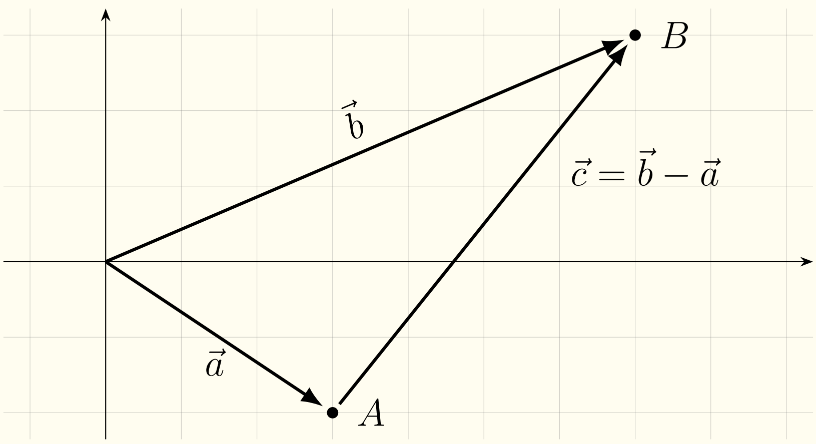


Figura 4.3: O vetor $\vec{c}$ é tal que $\vec{b} = \vec{a} + \vec{c}$. Para desenhá-lo, basta ligar a pontinha do vetor $\vec{a}$ à pontinha do vetor $\vec{b}$. No caso, $\vec{c} = (7, 3) - (3, -2) = (4, 5)$.

O conjunto $\mathbb{R}^n$, juntamente com as operações de soma e produto por escalar formam o que chamamos de *espaço vetorial*.

Definição 4.3. *Neste curso, um* espaço vetorial *é um subconjunto não vazio $V \subset \mathbb{R}^n$ (para algum $n \in \mathbb{N}$), onde é possível restringir as operações de* soma *e* produto por escalar*. Ou seja,*

$$\vec{a}, \vec{b} \in V \Rightarrow \vec{a} + \vec{b} \in V$$
$$\alpha \in \mathbb{R} \text{ e } \vec{a} \in V \Rightarrow \alpha\vec{a} \in V.$$

Quando $V \subset W$ são espaços vetoriais, dizemos que V é subespaço *de* W*. Note que as operações de* soma *e* produto por escalar *de* V *são a restrição a* V *das operações de* W*.*

O fato de um *espaço vetorial* V não ser vazio significa que $\vec{0} \in V$, pois

$$\vec{v} \in V \Rightarrow \vec{0} = 0\vec{v} \in V.$$

Exemplo 4.4. O *plano xy*

$$P = \left\{ (x, y, 0) \,\middle|\, x, y \in \mathbb{R} \right\}$$

é um *espaço vetorial.* Um *subespaço* de $\mathbb{R}^3$.

Observação 4.5. A definição de espaço vetorial é mais geral do que a que apresentamos na definição 4.3. O que é importante é ter um conjunto não vazio V onde estão definidas operações de *soma* e *produto por escalar* que tenham propriedades semelhantes às de $\mathbb{R}^n$.

4.2 Sistemas de Coordenadas

Vetores de $\mathbb{R}^2$ são simplesmente pares de números reais que podem ser somados (e subtraídos), e multiplicados por um número real (escalar). Quando falamos de *pontos e vetores no plano*, a princípio, não estamos falando de $\mathbb{R}^2$. O ponto é esse *"pontinho"* que a gente desenha.

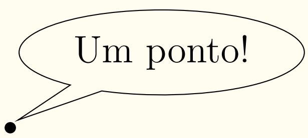


Figura 4.4: Um ponto! O que é um ponto? Deixa pra lá... :-)

O vetor é essa *"setinha"* que liga dois pontos.

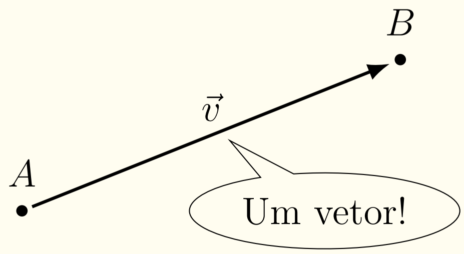


Figura 4.5: O vetor $\vec{v}$ representa o "deslocamento necessário" pra sair do ponto A e ir ao ponto B. Ou seja, $B = A + \vec{v}$.

Se A e B são pontos e $\vec{v}$ é um vetor, faz sentido falar do ponto $A + \vec{v}$, e faz sentido falar do vetor $B - A$. Mas não faz muito sentido falar de $A + B$. A princípio, pontos e vetores são coisas diferentes, de natureza diferente, e muitas vezes é melhor deixá-los assim. Mas se quisermos, podemos escolher um *ponto base* O, e então identificar os pontos do plano com os vetores do plano. Se falarmos do ponto (correspondente ao vetor) $\vec{v}$, na verdade, estamos falando do ponto $O + \vec{v}$. E se falarmos do vetor (correspondente ao ponto) A, na verdade, estamos nos referindo ao vetor $A - O$.

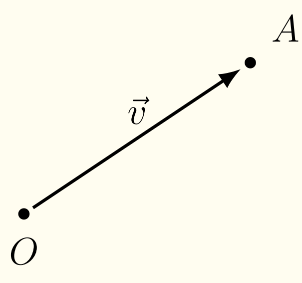


Figura 4.6: Escolhida a origem O, podemos tratar o ponto A e o vetor $\vec{v}$ como se fossem a mesma coisa.

Com essa interpretação geométrica feita até agora, não faz muito sentido somar um vetor $B-A$, que sai de A e vai pra B, com um vetor $D-C$, que sai de um outro ponto $C \neq B$, e vai até D!

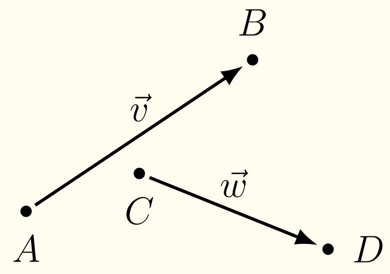


Figura 4.7: Faz sentido somarmos os vetores $\vec{v}$ e $\vec{w}$?

Para evitar esse problema, quando duas *setas* forem paralelas, tiverem o mesmo tamanho e apontarem no mesmo sentido, vamos dizer que são o mesmo *vetor*.

Tudo o que fizemos para pontos no plano, poderia ter sido feito para pontos no espaço. Estamos então, com duas noções diferentes de *vetores*. Se pensarmos em $\mathbb{R}^2$ — ou de modo mais geral, $\mathbb{R}^n$ — temos pares de números, com operação de soma e produto por escalar feitos *coordenada a coordenada*. Se pensarmos no plano — ou no espaço — temos *"pontinhos"* e *"setinhas"*. O coneceito que vai fazer a ligação entre o plano e $\mathbb{R}^2$, é o *sistema de coordenadas*. O *eixo coordenado* serve para especificarmos um vetor $\vec{v}$ em termos de suas *coordenadas*: $\vec{v} = (v_1, v_2)$.

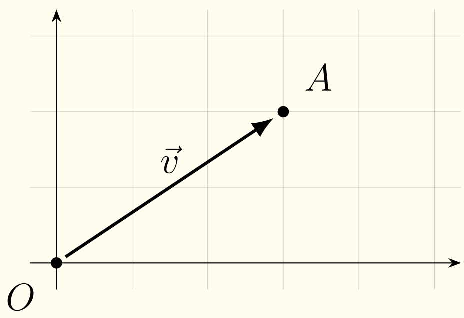


Figura 4.8: Escolhida a origem O e um sistema de coordenadas, podemos especificar tanto o ponto A como o vetor $\vec{v}$, utilizando coordenadas: $(3, 2)$.

Com a interpretação geométrica usando *"pontos"* e *"setas"*, mesmo sem eixos coordenados, podíamos falar de adição e produto por escalar.

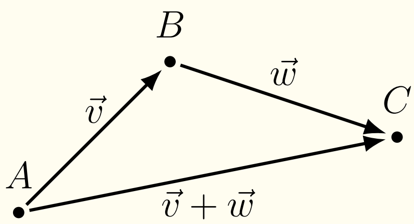


Figura 4.9: Suponha que v representa o "deslocamento necessário" para ir de A a B, e que $\vec{w}$ representa o "deslocamento" de B até C. Então, denotamos por $\vec{v} + \vec{w}$ o "deslocamento necessáiro" para ir de A até C.

O vetor $\vec{v}$, saindo do ponto A com coordenadas (a_1, a_2) e indo para o ponto B de coordenadas (b_1, b_2) será identificado com o vetor $(b_1 - a_1, b_2 - a_2) \in \mathbb{R}^2$. Usando essa identificação com $\mathbb{R}^2$, fica mais fácil falar de soma de vetores do plano sem nos preocuparmos com as tais *setinhas*.

Temos então, a princípio, duas maneiras de somar vetores. Temos a maneira geométrica, utilizando *setas*. E temos a maneira algébrica, fazendo a soma coordenada a coordenada. A figura 4.10 deve ser suficiente para convencer o leitor de que ambas as maneiras são equivalentes.

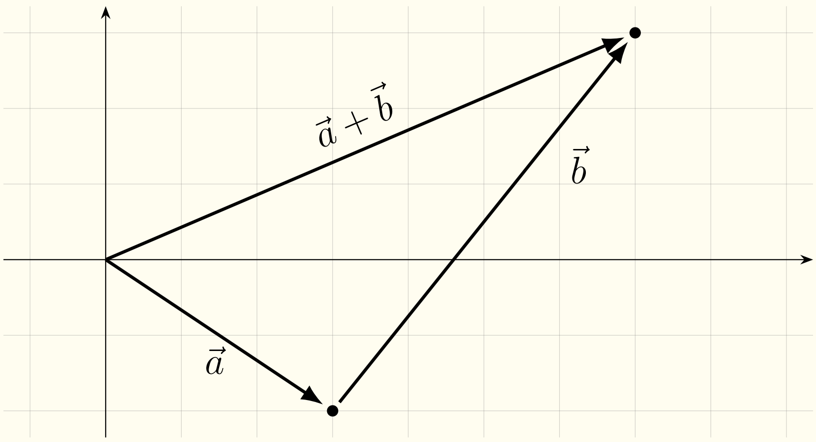


Figura 4.10: É fácil perceber que se somarmos $\vec{a}$ e $\vec{b}$ pelo método geométrico, o resultado será igual à soma coordenada a coordenada. E esse último é o método algébrico. O deslocamento de $\vec{a}$ na direção x é igual a 3. O deslocamento de $\vec{b}$ na direção x é igual a 4. Então, o deslocamento de $\vec{a}+\vec{b}$ na direção x é igual à soma: $3+4=7$. O mesmo ocorre na direção y.

Assim, para somar dois vetores, não precisamos que a ponta de um esteja no bumbum do outro. O mais comum é desenhar todos os vetores partindo de um mesmo ponto. Geometricamente, a soma é feita como mostrado na figura 4.11.

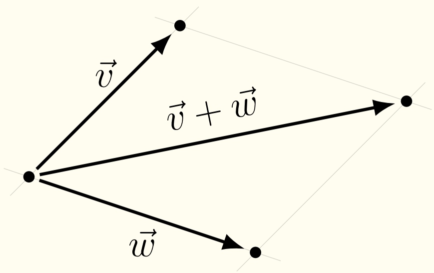


Figura 4.11: Depois de identificar os vetores através de translações, escrevemos $\vec{v}$ e $\vec{w}$ em um mesmo *ponto base*, e determinamos $\vec{v}+\vec{w}$ através do paralelogramo mostrado na figura.

Sabendo que os pontos de vista *geométrico* e *algébrico* são equivalentes, podemos escolher o que for mais conveniente. A seguir, vejamos um exemplo de um problema geométrico que é mais fácil resolvido com álgebra. E depois, um problema algébrico que é mais facilmente resolvido com o ponto de vista geométrico.

Exemplo 4.6 (adição é comutativa)**.** Do ponto de vista geométrico, como podemos mostrar que a adição de vetores é *comutativa*? Será que a situação da figura 4.12 nunca ocorre?

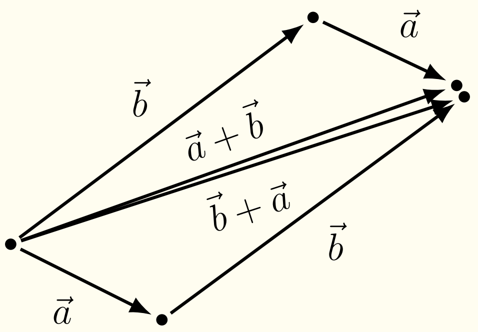


Figura 4.12: Se desenharmos $\vec{a}$ e juntarmos com $\vec{b}$, encontramos $\vec{a}+\vec{b}$. Mas será que se desenharmos primeiro $\vec{b}$ e juntarmos com $\vec{a}$ vamos chegar no mesmo lugar? Será que $\vec{a}+\vec{b}=\vec{b}+\vec{a}$?

Por outro lado, do ponto de vista algébrico é muito fácil ver que adição de vetores de $\mathbb{R}^n$ é comutativa:

$$\begin{aligned}\vec{v}+\vec{w} &= (v_1,\dots,v_n)+(w_1,\dots,w_n)\\ &= (v_1+w_1,\dots,v_n+w_n)\\ &= (w_1+v_1,\dots,w_n+v_n)\\ &= (w_1,\dots,w_n)+(v_1,\dots,v_n)=\vec{w}+\vec{v}.\end{aligned}$$

Ou seja, sabendo da equivalência entre os pontos de vista *geométrico* e *algébrico*, concluímos que, de fato, a situação da figura 4.12 nunca ocorre!

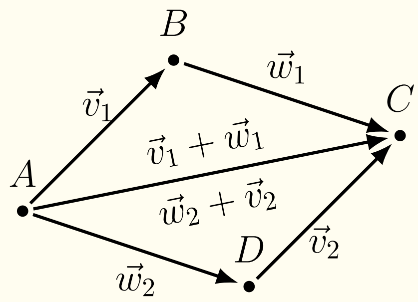


Figura 4.13: Com o uso de coordenadas, é fácil ver que o vetor $\vec{v}_2$ é uma *translação* de $\vec{v}_1$, e que o mesmo ocorre com $\vec{w}_1$ e $\vec{w}_2$. Por isso, tratamos $\vec{v}_1$ e $\vec{v}_2$, e também, $\vec{w}_2$ e $\vec{w}_2$ como sendo o mesmo vetor.

O exemplo a seguir mostra um problema cuja formulação é algébrica, e cuja solução é facilmente compreendida utilizando um ponto de vista geométrico.

Exemplo 4.7. Por algum motivo, estamos interessados em investigar combinações de dois vetores $\vec{v}$ e $\vec{w}$ que sejam do tipo

$$\alpha\vec{v} + \beta\vec{w},$$

onde $\alpha \geq 0$ e $\beta \geq 0$. Geometricamente é bastante fácil ver quais vetores podem ou não podem serem escritos dessa maneira. Veja a figura 4.14.

Figura 4.14: Geométricamente é fácil ver quais vetores do plano podem ser escritos na forma $\alpha\vec{v} + \beta\vec{w}$, com $\alpha \geq 0$ e $\beta \geq 0$. Também é fácil ver quais podem ser escritos com a restrição extra $0 \leq \alpha + \beta \leq 1$.

Agora temos intuição *geométrica* (desenho) e intuição *algébrica* (cálculos) sobre um mesmo conceito. A intuição *geométrica* não é melhor que a *algébrica*, ou vice-versa. É importante conhecermos as duas e lançarmos mão da que for mais adequada ao problema sendo tratado.

O que tratamos aqui, vale também para $\mathbb{R}^3$. Com um pouco de esforço, podemos extrapolar essas ideias para $\mathbb{R}^4$, ou mesmo para $\mathbb{R}^n$. Cada pessoa gosta de enxergar vetores em $\mathbb{R}^n$ de uma maneira diferente. Mas quando usamos apenas *álgebra*, não precisamos nos preocupar com isso. Afinal de contas, estamos falando simplesmente de n-uplas de números.

Muita gente gosta de pensar na quarta dimensão como se fosse o *"tempo"*. O leitor pode pensar no tempo, se preferir. Mas lembre-se que $\mathbb{R}^n$ é simplesmente uma n-upla ordenada com operações de *soma* e *produto por escalar* definidos. E, como no exemplo 4.1, é bem possível que a quarta dimensão seja a *"couve"*!!! Quando o número de dimensões aumenta muito, a geometria continua nos fornecendo ideias e intuição. Mas às vezes, o mais fácil é pensar nos elementos de $\mathbb{R}^4$ ou $\mathbb{R}^{23}$ como sendo sequências de 4 ou 23 números, respectivamente.

4.3 Geometria sem Coordenadas

Provavelmente, no ensino médio o leitor aprendeu a *equação da reta* (ou as equações), a *equação do círculo*, a *equação do plano*, etc. Também aprendeu fórmula para calcular a distância entre duas retas no plano e no espaço ($\mathbb{R}^2$ e $\mathbb{R}^3$), entre ponto e reta, reta e plano, etc. Provavelmente aprendeu (ou não!) um monte de fórmulas com x, y, z, a, b, c, t, etc.

Um sistema de coordenadas pode ser muito útil para expressarmos conceitos geométricos de uma forma concreta. No entanto, utilizar a notação vetorial sem o uso de coordenadas pode ser uma maneira mais limpa e mais enxuta de fazê-lo. Vejamos alguns exemplos.

Ponto Médio

Sejam $A = (a_1, a_2)$ e $B = (b_1, b_2)$ pontos no plano. O ponto M que fica, no segmento de reta que liga A e B, equidistante a ambos, é dado por

$$M = \left(\frac{a_1 + b_1}{2}, \frac{a_2 + b_2}{2} \right). \tag{4.3.1}$$

Como sabemos disso? Como chegamos nessa fórmula?

Em notação vetorial, é fácil perceber que

$$M = A + \frac{1}{2}\vec{u},$$

onde $\vec{u} = (B - A)$. Portanto,

$$M = A + \frac{1}{2}(B - A) = A - \frac{1}{2}A + \frac{1}{2}B = \frac{1}{2}(A + B).$$

Isso mostra que a fórmula (4.3.1) vale até mesmo quando o sistema de coordenadas não é formado por eixos ortogonais!

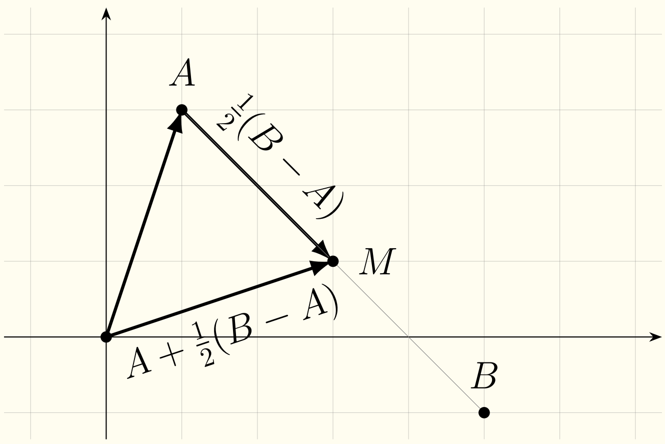


Figura 4.15: Dados os pontos A e B, o ponto médio M é dado por

$$M = A + \frac{1}{2}(B - A) = \frac{1}{2}(A + B).$$

Não é difícil se convencer deste fato.

Repare na figura 4.16, que utiliza o paralelogramo. Alguém poderia dizer que não é evidente que o ponto médio é exatamente o ponto de interseção das diagonais do paralelogramo. Mas quando levamos tudo para notação vetorial, tanto o ponto médio quanto o ponto de interseção tem a mesma expressão

$$A + \frac{1}{2}(B - A) = \frac{1}{2}(A + B).$$

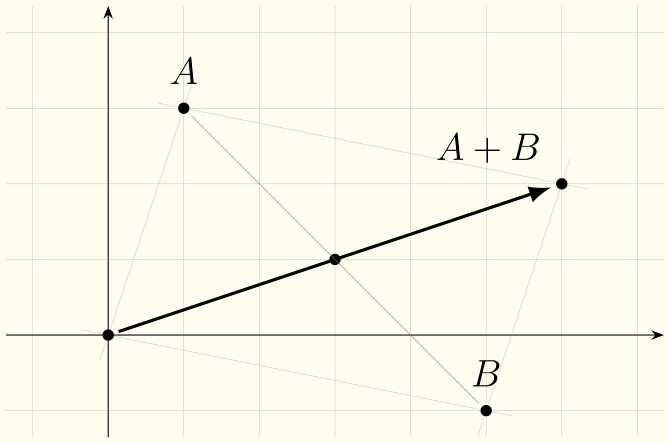


Figura 4.16: Dados os pontos A e B, o ponto médio é dado por $\frac{1}{2}(A+B)$. Por um argumento geométrico, com semelhança de triângulos, por exemplo, podemos nos convencer de que o ponto de interseção das duas diagonais é de fato o ponto médio, e que a distância da origem até o ponto médio é metade da distância da origem até $A + B$. Mas o arugmento vetorial na figura 4.15 parece bem mais simples.

Nada do que fizemos depende do fato de estarmos em $\mathbb{R}^2$. Os mesmos argumentos valem em $\mathbb{R}^3$. Depois que já estamos acostumados com a notação vetorial, podemos até mesmo falar de ponto médio em $\mathbb{R}^4$, ou $\mathbb{R}^n$. Você vai até o ponto A, e de lá, avança mais $\frac{1}{2}(B - A)$. Você não precisa *"enxergar"* um espaço de quatro dimensões! Basta enxergar a *soma* de dois vetores — que se dá em um plano passando pela origem que contém os pontos A e B — e o *produto por escalar*. Depois você vai lá e ***define*** o ponto médio (em qualquer espaço vetorial!) como sendo o ponto

$$M = \frac{1}{2}(B - A).$$

Círculo e Norma

Usando coordenadas cartesianas, o círculo de raio r e centro $A = (a, b)$ é dado por

$$C = \left\{ (x, y) \in \mathbb{R}^2 \,\middle|\, (x - a)^2 + (y - b)^2 = r^2 \right\}.$$

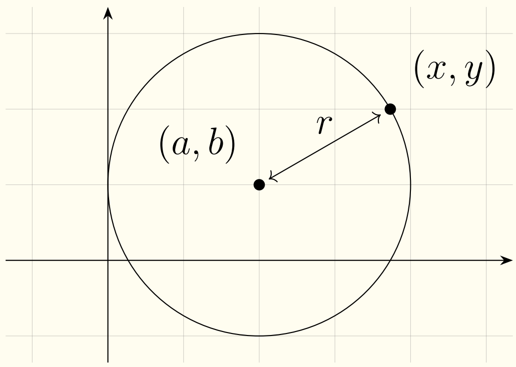


Figura 4.17: Os pontos do círculo centrado em $A = (a, b)$, de raio r, são os pontos (x, y) que satisfazem a equação

$$(x - a)^2 + (y - b)^2 = r^2.$$

Se dispusermos de uma operação de *norma*, que nos diz qual é o tamanho (*módulo* ou *norma*) $\|\vec{v}\|$ de um vetor $\vec{v}$, podemos dizer que o círuclo C é formado pelos pontos que consguimos alcançar, partindo de A e fazendo um deslocamento $\vec{v}$ de tamanho r. Ou seja,

$$C = \left\{ A + \vec{v} \,\middle|\, \vec{v} \in \mathbb{R}^2,\ \|\vec{v}\| = r \right\}.$$

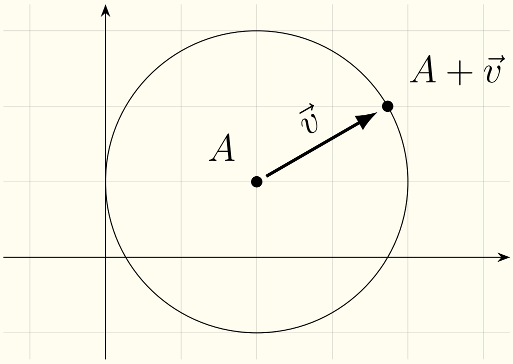


Figura 4.18: Mesmo sem saber que a fórmula para $\|\cdot\|$ é

$$\|\vec{v}\| = v_1^2 + v_2^2,$$

sabemos que o círculo centrado em A, de raio r é formado pelos pontos da forma $A + \vec{v}$, onde $\|\vec{v}\| = r$.

Ou então, ainda utilizando a norma, podemos escrever

$$C = \left\{ P \in \mathbb{R}^2 \,\middle|\, \|P - A\| = r \right\}.$$

Basicamente, isso nada mais é do que a fórmula $(x-a)^2 + (y-b)^2 = r^2$ da primeira equação, só que mais enxuta.

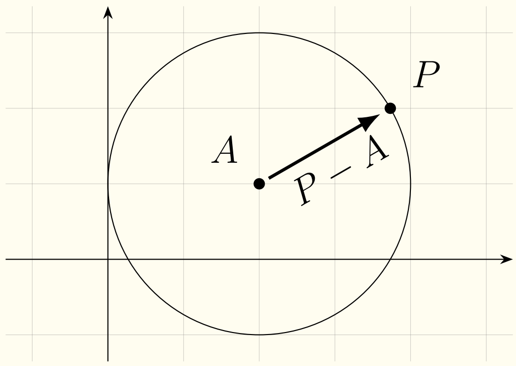


Figura 4.19: Um pouco mais parecido com a fórmula

$$(x-a)^2 + (y-b)^2 = r^2,$$

é quando escrevemos

$$\|P - A\| = r.$$

Ao invés do círuclo em $\mathbb{R}^2$, podemos falar da esfera em $\mathbb{R}^3$. É exatamente a mesma coisa!!!

$$S = \left\{ A + \vec{v} \;\middle|\; \vec{v} \in \mathbb{R}^3,\ \|\vec{v}\| = r \right\}.$$

E, agora que você já se acostumou com essa ideia, podemos *definir* em $\mathbb{R}^n$ a esfera n-dimensional centrada em A, de raio r, como sendo o conjunto dos pontos que estão a uma distância r de A:

$$S = \left\{ A + \vec{v} \;\middle|\; \vec{v} \in \mathbb{R}^n,\ \|\vec{v}\| = r \right\}.$$

Para isso, tudo o que precisamos é da norma $\|\cdot\|$.

A *norma* nos permite dividir o problema em dois. Um deles é o cálculo da norma de um vetor. Como a gente calcula a norma de um vetor? O outro é o uso das propriedades da norma. Separar essas duas coisas é muito útil. Afinal de contas, ninguém precisa calcular raiz quadrada pra usar um compasso! Nesse sentido, usar a norma sem de fato fazer a conta, é como usar um compasso. :-)

A norma possui propriedades que nos serão úteis.

1. Para $\alpha \in \mathbb{R}$, $\|\alpha\vec{v}\| = |\alpha|\|\vec{v}\|$.
2. $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$.
3. $\|\vec{v}\| = 0 \Rightarrow \vec{v} = \vec{0}$.

Reta

Provavelmente o leitor já conhece a *equação da reta* em alguma forma. Vamos primeiramente discutir várias maneiras **horríveis** de falar sobre a *reta*.

Uma reta pode ser especificada como um conjunto de pontos

$$r = \left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; y = ax + b \right\},$$

onde $a, b \in \mathbb{R}$ são constantes. A fórmula $y = ax + b$ nos permite algumas interpretações geométricas: a é o coeficiente angular da reta, e b é o ponto que a reta corta o eixo das ordenadas.

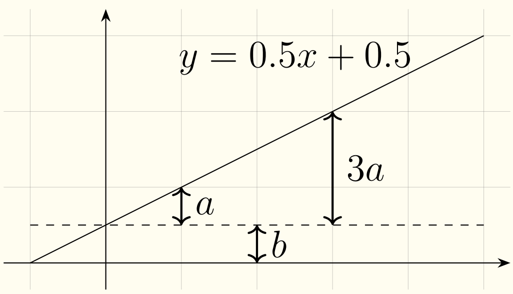


Figura 4.20: O gráfico da função $f(x) = ax + b$ é uma reta. O gráfico é formado pelos pontos (x, y) da forma $y = ax + b$.

Mas espere! Existem retas que não cortam o eixo das ordenadas. Existe também o próprio eixo, que é uma reta. As retas $x = a$ não estão representadas, pois não são o gráfico de uma função da forma $f(x) = ax + b$.

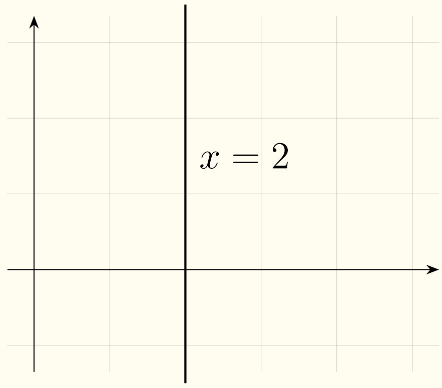


Figura 4.21: A reta $x = 2$ não é gráfico de uma função $f(x)$, e não pode ser representada na forma $y = ax + b$.

Uma maneira melhor, seria

$$r = \left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; ax + by = c \right\}, \tag{4.3.2}$$

onde $a, b, c \in \mathbb{R}$ são constantes. Vamos retornar a essa equação, na seção 9.3, depois que falarmos de produto interno.

Outra maneira (*horrível*) de especificar uma reta, é utilizando um parâmetro t e várias funções de $\mathbb{R}$ em $\mathbb{R}$:

$$\begin{aligned} x(t) &= at + p \\ y(t) &= bt + q, \end{aligned}$$

onde $a, b, p, q \in \mathbb{R}$ são constantes. Parece bastante claro que o uso excessivo de coordenadas torna a geometria difícil de ser compreendida. Repare que o sistema de equações que determina os pontos da reta a partir do parâmetro t pode ser escrita na forma vetorial.

$$r(t) = P + t\vec{v},$$

onde $P = (p, q)$ e $\vec{v} = (a, b) \neq (0, 0)$. É como se no *"instante"* $t = 0$, estivéssemos no ponto $P = (p, q)$, nos movendo com velocidade (constante)

$\vec{v} = (a, b)$. Ou seja, a cada unidade de tempo, nos deslocamos na direção e no sentido do vetor $\vec{v}$, com velocidade igual a $\|\vec{v}\|$.

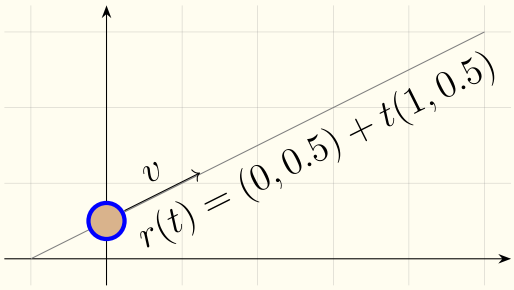


Figura 4.22: Um objeto se deslocando com velocidade constante $\vec{v} = (1, 0.5)$.

Repare que ao utilizarmos a linguagem vetorial, não precisamos mais do sistema de coordenadas! Em coordenadas, a equação ficaria

$$r(t) = (p, q) + t(a, b).$$

O que equivale a

$$\begin{aligned} x(t) &= p + at \\ y(t) &= q + bt. \end{aligned}$$

A equação vetorial continuará valendo, mesmo quando não estivermos restritos ao plano. Em um espaço vetorial V qualquer, uma reta que passa pelo ponto $P \in V$, com direção igual à do vetor não nulo $\vec{v} \in V$, é a o conjunto

$$r = \left\{ P + t\vec{v} \,\middle|\, t \in \mathbb{R} \right\}.$$

Se $A, B \in \mathbb{R}^2$ são pontos distintos do plano, então, a reta que passa em A e B pode ser parametrizada usando-se $P = A$ e $\vec{v} = (B - A)$. Mas existem outras parametrizações que também funcionam. Por exemplo,

$$\begin{aligned}
r(t) &= B + t^3(A - B),\\
r(t) &= A + t(B - A),\\
r(t) &= 2B - A + 5t(B - A),\\
r(t) &= \ldots
\end{aligned}$$

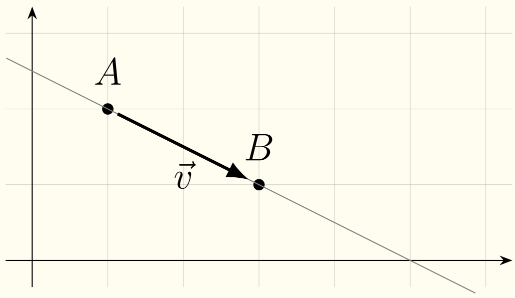


Figura 4.23: Dados dois pontos distintos A e B, a reta que passa em A e B pode ser dada pela equação $r(t) = A + t\vec{v}$, onde $\vec{v} = B - A$.

Pra ser rigoroso, $r(t)$ é uma função de $\mathbb{R}$ em $\mathbb{R}^2$ que leva cada $t \in \mathbb{R}$ em um ponto da reta.

$$\begin{array}{rccc} r: & \mathbb{R} & \to & \mathbb{R}^2 \\ & t & \mapsto & P + t\vec{v} \end{array}.$$

Segmentos de reta e semi retas podem ser especificados restringindo o domínio de definição. Por exemplo, o segmento de reta que liga os pontos distintos $A, B \in \mathbb{R}^2$ pode ser parametrizado por

$$\begin{array}{rccc} r: & [0,1] & \to & \mathbb{R}^2 \\ & t & \mapsto & A + t(B - A) \end{array}.$$

Que é quase a mesma coisa que já haviamos escrito para a reta. Mas agora, a função r está definida apenas para $t \in [0,1]$.

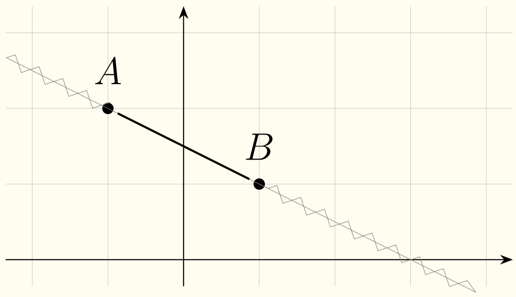


Figura 4.24: Se restringirmos o domínio de definição de $r(t) = A + t(B - A)$ ao intervalo $[0, 1]$, temos uma parametrização para o segmento de reta que liga os pontos A e B.

Triângulo e Outros Convexos

O leitor pode ficar à vontade para pular esta subseção.

Sejam $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathbb{R}^2$ distintos. Vamos procurar uma maneira de expresar, em notação vetorial, o *triângulo sólido* com vértices $\vec{v}_1$, $\vec{v}_2$ e $\vec{v}_3$. O conjunto formado pelos pontos do triângulo sólido é *convexo*. Um conjunto *convexo* é um conjunto C tal que se $\vec{a}, \vec{b} \in C$, então todo o segmento de reta que liga os pontos $\vec{a}$ e $\vec{b}$ está contido em C. Assim como no caso do *ponto médio*, o leitor pode verificar facilmente que os pontos do segmento de reta que liga $\vec{a}$ e $\vec{b}$ são aqueles que podem ser expressos na forma

$$\vec{p} = \alpha\vec{a} + \beta\vec{b}, \tag{4.3.3}$$

com $\alpha, \beta \in [0, 1]$ e $\alpha + \beta = 1$. No caso do ponto médio, $\alpha = \beta = \frac{1}{2}$.

Figura 4.25: Exemplos de conjuntos que não são convexos.

O círculo, o triângulo, o quadrado, etc. são exemplos de conjuntos convexos.

Figura 4.26: Exemplos de conjuntos que são convexos.

Os pontos $\vec{p}$ da forma da equação 4.3.3 são chamados de combinação convexa dos vetores $\vec{a}$ e $\vec{b}$. Um conjunto C é convexo, quando é fechado por combinação convexa. Resumidamente, C é convexo quando para todo $\alpha, \beta \in [0,1]$, com $\alpha + \beta = 1$,

$$\vec{v}, \vec{w} \in C \Rightarrow \alpha\vec{v} + \beta\vec{w} \in C.$$

O triângulo que estamos tratando é o menor convexo que contém os vértices $V = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. As arestas do triângulo são formadas pelos pontos que são combinação convexa de dois vértices. Se A é o conjunto das arestas, então

$$A = \left\{ \alpha\vec{v} + \beta\vec{w} \;\middle|\; \alpha, \beta \in [0,1],\, \alpha + \beta = 1,\, \vec{v}, \vec{w} \in V \right\}.$$

Suponha que $\vec{p} = \alpha\vec{v}_1 + \beta\vec{v}_2 \in A$ seja um ponto da aresta que liga $\vec{v}_1$ e $\vec{v}_2$. Então, qualquer combinação convexa de $\vec{p}$ e $\vec{v}_3$ também deve estar em T. Qual é a forma de uma combinação convexa $\vec{q}$, de $\vec{p}$ e $\vec{v}_3$? Por exemplo,

$$\begin{aligned} \vec{q} &= x\vec{p} + y\vec{v}_3 \\ &= x(\alpha\vec{v}_1 + \beta\vec{v}_2) + y\vec{v}_3 \\ &= x\alpha\vec{v}_1 + x\beta\vec{v}_2 + y\vec{v}_3, \end{aligned}$$

onde $x, y \in [0,1]$ e $x + y = 1$. Note que $x\alpha, x\beta, y \in [0,1]$ e $x\alpha + x\beta + y = x + y = 1$. Dessa forma,

$$T = \left\{ \alpha\vec{v}_1 + \beta\vec{v}_2 + \gamma\vec{v}_3 \;\middle|\; \alpha, \beta, \gamma \in [0,1],\, \alpha + \beta + \gamma = 1 \right\}.$$

Repare que, ao invés de dizermos $\alpha, \beta, \gamma \in [0,1]$, bastava dizer que $\alpha, \beta, \gamma \geq 0$, já que juntamente com a condição $\alpha + \beta + \gamma = 1$, não podemos ter α, β ou γ maiores que 1.

De um modo geral, combinações convexas de combinações convexas de combinações... de combinações convexas de vetores de um conjunto V assumem a seguinte forma

$$\alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n,$$

onde $\vec{v}_1, \ldots, \vec{v}_n \in V$, $\alpha_1, \ldots, \alpha_n \geq 0$ e $\sum_{j=1}^{n} \alpha_j = 1$. Esses pontos formam o *envoltório convexo* de V. Note que conjunto V não precisa nem mesmo ter finitos elementos!

Figura 4.27: O *fecho convexo* dos conjuntos da figura 4.25.

4.4 Bases

A princípio, não precisamos que os eixos sejam ortogonais pra especificar um vetor em *"coordenadas"*. Quando os eixos não são ortogonais, temos ao menos duas maneiras de relacionar cada ponto do plano a um par de coordenadas. Veja a figura 4.28. Poderíamos fazer *projeções ortogonais* em cada eixo, como na figura 4.29.

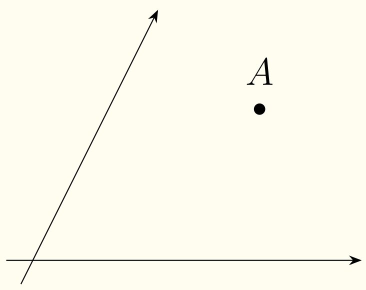


Figura 4.28: Podemos usar eixos não ortogonais. Mas qual seria a coordenada do ponto A?

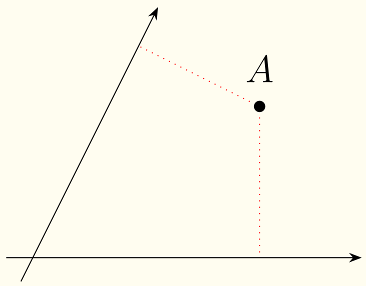


Figura 4.29: Podemos definir as coordenadas do ponto A como sendo as projeções ortogonais nos eixos coordenados.

Ou então, poderíamos fazer desse jeito "complicado" da figura 4.30.

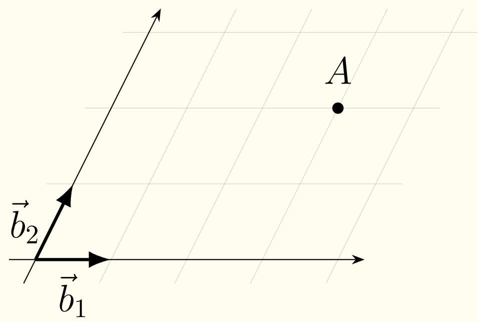


Figura 4.30: Podemos definir as coordenadas do ponto A como sendo os coeficientes necessários para escrever A como *combinação linear* dos vetores $\vec{b}_1$ e $\vec{b}_2$. No caso específico do ponto A,

$$A = 3\vec{b}_1 + 2\vec{b}_2.$$

Cada um desses métodos tem suas vantagens e desvantagens. Neste curso, vamos nos limitar a trabalhar com o esquema da figura 4.30. Primeiramente, repare nos vetores $\vec{b}_1$ e $\vec{b}_2$ da figura. Vamos chamar $\mathscr{B} = (\vec{b}_1, \vec{b}_2)$, de *base ordenada* $\mathscr{B}$. Qualquer vetor $\vec{v} \in \mathbb{R}^2$ pode ser escrito da forma

$$\vec{v} = a\vec{b}_1 + b\vec{b}_2.$$

Neste caso, escrevemos

$$\vec{v} = (a, b)_{\mathscr{B}},$$

para indicar que a e b são as coordenadas de $\vec{v}$ na *base ordenada* $\mathscr{B}$.

Ao invés de trabalharmos com os eixos, trabalhamos com os vetores $\vec{b}_1$ e $\vec{b}_2$. Escolher um *eixo coordenado* para o plano consiste em:

1. Escolher um ponto do plano para coresponder à *origem* dos eixos.

2. Escolher uma *base ordenada*.

Em álgebra linear, costumamos tratar de *bases ordenadas*, e não de *eixos coordenados*.

Quando temos vetores $\vec{v}_1, \ldots, \vec{v}_n$, uma soma do tipo

$$\vec{v} = \alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n,$$

com $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ é chamada de *combinação linear* dos vetores $\vec{v}_1, \ldots, \vec{v}_n$. O que queremos saber, é se dados os vetores $\mathscr{B} = (\vec{b}_1, \ldots b_n)$, qualquer vetor do espaço em questão pode ou não ser escrito de maneira única como uma combinação linear dos vetores em $\mathscr{B}$. Este assunto será tratado na seção 8.3.

Exemplo 4.8 (um é pouco!)**.** Em $\mathbb{R}^2$, apenas um vetor (eixo) não é suficiente para escrever todos os outros vetores.

Em $\mathbb{R}^3$, dois vetores nunca são suficientes.

Exemplo 4.9 (três é demais!)**.** Por outro lado, se tivermos 3 vetores (3 eixos) em $\mathbb{R}^2$, nenhuma combinação linear será *única*.

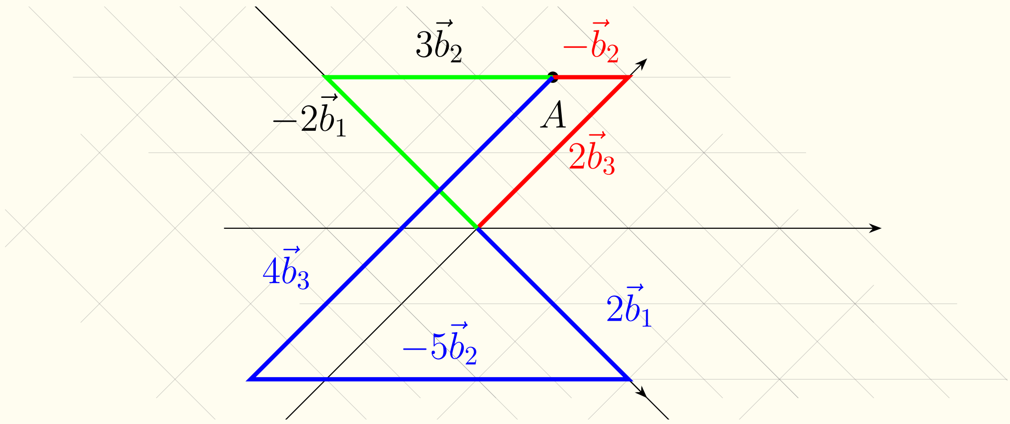


Figura 4.31: Três eixos são um pouco demais para representar pontos em $\mathbb{R}^2$. Em termos de combinação linear, existem várias maneiras diferentes de escrever, por exemplo, o vetor $\vec{a} = (1, 2)$ como combinação linear de $\vec{b}_1 = (1, -1)$, $\vec{b}_2 = (1, 0)$ e $\vec{b}_3 = (1, 1)$:

$$\vec{a} = -1\vec{b}_2 + 2\vec{b}_3 = 2\vec{b}_1 - 5\vec{b}_2 + 4\vec{b}_3 = -2\vec{b}_1 + 3\vec{b}_2.$$

Os espaços $\mathbb{R}^n$ possuem uma *base ordenada canônica* a saber:

$$\begin{aligned} \vec{e}_1 &= (1, 0, 0, \ldots, 0) \\ \vec{e}_2 &= (0, 1, 0, \ldots, 0) \\ &\vdots \\ \vec{e}_n &= (0, 0, 0, \ldots, 1). \end{aligned}$$

Uma questão interessante, é saber como um vetor

$$\vec{v} = \beta_1 \vec{b}_1 + \cdots + \beta_n \vec{b}_n,$$

expresso em uma *base ordenada* $\mathscr{B}$, é escrito em outra *base ordenada* $\mathscr{C}$:

$$\vec{v} = \gamma_1 \vec{c}_1 + \cdots + \gamma_m \vec{c}_m.$$

Como determinar $\gamma_1, \ldots, \gamma_m$, conhecendo $\beta_1, \ldots, \beta_n$? Este assunto será tratado no capítulo 17.

4.5 Distribuição em n Estados

Dados n grupos de alguma coisa, podemos usar um vetor de $\mathbb{R}^n$ pra indicar a quantidade de itens ou indivíduos em cada grupo. Será que o leitor consegue encontrar utilidade para as operações de soma e produto por escalar neste contexto?

Exemplo 4.10 (o *impeachment* e o **golpe**)**.** Em 2016, a população brasileira se divide em 3 grupos:

1. As pessoas que acham que o *impeachment* é golpe.
2. As pessoas que acham que o golpe é *impeachment*.
3. E a grande maioria, formada por pessoas que não acham nada e também não se importam.

Podemos usar um vetor de $\mathbb{R}^3$ para representar a maneira como a população brasileira se divide nesses três grupos.

A cada vez que Sérgio Moro, DD e o resto de sua trupe produzem uma *delação forjada* e repassam para *jornalistas amigos*, por exemplo, essa distribuição varia. Seja $\vec{v}(t) \in \mathbb{R}^3$ a distribuição no instante t. A diferença entre a distribuição de dois momentos distintos $t = a$ e $t = b$ é a *variação* da distribuição: $\vec{v}(b) - \vec{v}(a)$. A taxa de variação é o *vetor*

$$\frac{1}{b-a}(\vec{v}(b) - \vec{v}(a)).$$

4.6 Sistemas Lineares

De modo impreciso, um *sistema de equações lineares* é uma lista de equações do tipo

$$\begin{cases} 2x + 3y + 4z = 5 \\ 3x + 2y - 9z = 7 \\ 7x - 4y - 5z = -1 \end{cases} . \tag{4.6.1}$$

Estamos interessados em descobrir os vetores $(x, y, z) \in \mathbb{R}^3$ que satisfazem o sistema de equações. Se é que existe algum.

Escrevendo com *matrizes coluna*, o sistema fica assim:

$$\begin{bmatrix} 5 \\ 7 \\ -1 \end{bmatrix} = x \begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} + y \begin{bmatrix} 3 \\ 2 \\ -4 \end{bmatrix} + z \begin{bmatrix} 4 \\ -9 \\ -5 \end{bmatrix}.$$

Em notação vetorial, o sistema da equação (4.6.1) corresponde ao sistema

$$\vec{d} = x\vec{a} + y\vec{b} + z\vec{c},$$

onde

$$\begin{aligned} \vec{a} = (2, 3, 7) \quad &\text{e} \quad \vec{b} = (3, 2, -4) \\ \vec{c} = (4, -9, -5) \quad &\text{e} \quad \vec{d} = (5, 7, -1). \end{aligned}$$

Solucionar o sistema da equação (4.6.1) corresponde a descobrir como o vetor $\vec{d}$ pode ser escrito como combinação linear dos vetores $\vec{a}$, $\vec{b}$ e $\vec{c}$.

Se o sistema tem n linhas, então os vetores estarão em $\mathbb{R}^n$. Se o sistema tem m incógnitas, então haverá m vetores a serem usados pra fazer combinações lineares. Enxergando sistemas lineares dessa forma nos permite interpretar o problema de modo mais geométrico, como na figura 4.32.

Figura 4.32: Os *sistemas de equações lineares* são problemas que podem ser interpretados como: descobrir maneiras possíveis de se escrever um vetor $\vec{w}$ como combinação linear dos vetores $\vec{v}_1, \ldots, \vec{v}_n$.

4.7 Exercícios

Exercício 4.1. Sejam $\vec{a} = (1, 2, 3)$, $\vec{b} = (4, 5, 6)$ e $\vec{c} = (3, 2, 2)$. Calcule:

1. $\vec{a} + \vec{b} + \vec{c}$
2. $2\vec{a} + 3\vec{b} + 4\vec{c}$
3. $\vec{a} - 2\vec{b} + (-3)\vec{c}$
4. $(-1)\vec{a} + (-1)\vec{b} + (-1)\vec{c}$

Exercício 4.2. Sejam $\vec{a} = (1, 2, 3)$, $\vec{b} = (4, 5, 6)$ e $\vec{c} = (3, 2, 1)$. Escreva os vetores a seguir como combinação linear de $\vec{a}$, $\vec{b}$ e $\vec{c}$.

1. $(1, 1, 1)$

   Solução.
$$(1, 1, 1) = -\frac{1}{3}\vec{a} - -\frac{1}{3}\vec{b}$$

2. $(1, 1, 2)$
3. $(2, 2, 4)$
4. $(3, 3, 6)$
5. $(2, 2, 3)$
6. $(2, 5, 3)$
7. $(1, 0, 0)$
8. $(0, 1, 0)$
9. $(0, 0, 1)$

Exercício 4.3. Usando os itens 7 a 9 do exercício 4.2, e observando, por exemplo, que $(2, 2, 3) = 2\vec{e}_1 + 2\vec{e}_2 + 3\vec{e}_3$, recalcule de uma forma mais "inteligente" as combinações $(3, 3, 6)$, $(2, 2, 3)$ e $(2, 5, 3)$ do exercício 4.2.

Exercício 4.4. Será que existem outras formas de escrever os vetores do exercício 4.2 como combinação linear dos vetores dados? Discuta.

Exercício 4.5. Mostre que o vetor $(3, 2, 1)$ não pode ser escrito como combinação linear de $(1, 2, 3)$ e $(4, 5, 6)$.

Exercício 4.6. Escreva os seguintes sistemas lineares usando a notação de combinação linear.

1. $\begin{cases} 5x + 3y = 2 \\ 7x - 4y = 1 \end{cases}$
2. $\begin{cases} 4x + 3y + 2z = 1 \\ -x - y + 4z = 2 \end{cases}$
3. $\begin{cases} 3x + 8y = 7 \\ 6x - 5y = 8 \\ -2x - y = 9 \end{cases}$

Exercício 4.7. Os pontos do triângulo (incluindo o interior, e não apenas as arestas) de vértices $(1, 3)$, $(3, 5)$ e $(-1, 2)$ são dados pelo conjunto

$$T = \left\{ \boxed{} \,\middle|\, \boxed{} \right\}.$$

Exercício 4.8. Os pontos do triângulo (incluindo o interior, e não apenas as arestas) de vértices $(1, 2, 3)$, $(-1, 3, 5)$ e $(-1, 2, 4)$ são dados pelo conjunto

$$T = \left\{ \boxed{} \,\middle|\, \boxed{} \right\}.$$

Exercício 4.9. Usando a norma $\|\cdot\|$ em $\mathbb{R}^3$, o cone sólido oblíquo cuja base é o círculo no plano x-y de raio 3 centrado no ponto $(1, 2, 0)$, e cujo vértice é dado pelo ponto $(3, 4, 5)$ é dado pelo conjunto

$$C = \left\{ \boxed{} \,\middle|\, \boxed{} \right\}.$$

Exercício 4.10. Em $\mathbb{R}^3$, o fecho convexo do conjunto formado pela esfera de raio 2 centrada no ponto $(1, 1, 1)$ com mais o ponto $(0, 0, -3)$ é um sólido parecido com um "sorevete". Os pontos desse sólido formam o conjunto

$$S = \left\{ \boxed{} \,\middle|\, \boxed{} \right\}.$$

Aula 5

Matrizes

É muito provável que o leitor já tenha se deparado com *matrizes* anteriormente. Um monte de regras "sem sentido" pra fazer uma porção de contas que parecem não ter nenhum propósito. O principal objetivo deste livro é mostrar ao leitor que *matrizes* não apenas fazem sentido, mas também são muito úteis.

No entanto, não vá se animando, não!... :-p
Para o que vamos aprender, é muito útil se o leitor já possuir habilidade com as operações envolvendo matrizes. Por isso, por mais doloroso que seja, vamos primeiramente treinar nossas habilidades com matrizes. Durante esse processo, esperamos também levantar certas questões a serem respondidas no desenrolar do curso. Por exemplo:

> Por que o *produto de matrizes* é definido dessa maneira complicada e não de um modo simples, como no caso da *soma de matrizes*?

5.1 Notação e Definição

Uma *matriz* $q \times p$ é uma tabela com q linhas e p colunas. Escrevemos, por exemplo,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{q1} & a_{q2} & \cdots & a_{qp} \end{bmatrix}.$$

Para sermos mais concisos, escrevemos, por exemplo,

$$A = \big[a_{ij}\big].$$

Exemplo 5.1. Os cidadãos do *fimdomundo* acreditam que existem nutrientes nos alimentos. Os mais importantes seriam os nutrientes X, Y e Z. Portanto, eles costumam associar a cada alimento um vetor $(x, y, z) \in \mathbb{R}^3$ que representa o número de miligramas de nutrientes X, Y e Z em cada grama deste alimento. Suponha que $\vec{b} = (3, 10, 8)$ é a quantidade de nutrientes na batata, e $\vec{c} = (15, 7, 11)$ é a quantidade de nutrientes na cenoura. Essa informação pode ser resumida em uma tabela.

	batata	**cenoura**
nutriente X	3	15
nutriente Y	10	7
nutriente Z	5	12

Mas se concordarmos que ao falar de batata e cenoura utilizando um elemento de $\mathbb{R}^2$, o primeiro número é a quantidade de batata e o segundo, de cenoura. E se concordarmos que para nutrientes o primeiro número é a quantidade de nutrientes X o segundo é a quantidade de nutrientes Y, e o terceiro, a de nutrientes Z, então podemos eliminar a legenda da tabela e escrever simplesmente a matriz

$$N = \begin{bmatrix} 3 & 15 \\ 10 & 7 \\ 5 & 12 \end{bmatrix}.$$

Cada coluna da *matriz* pode ser vista como um vetor escrito "em pé". Vamos utilizar muito essa ideia. Por isso, dado um vetor $\vec{v} \in \mathbb{R}^p$, vamos utilizar a notação

$$\begin{bmatrix} | \\ \vec{v} \\ | \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_p \end{bmatrix}$$

para representar o vetor $\vec{v}$ como uma *matriz coluna*. As barras verticais servem pra lembrar que não estamos falando de uma matriz 1×1. Quando já estivermos mais acostumados, ou para não usar muito espaço, vamos escrever $\begin{bmatrix} \vec{v} \end{bmatrix}$, sem as barras verticais.

Por fim, quando as p colunas de uma matriz A são dadas pelos vetores $\vec{a}_1, \dots, \vec{a}_p$, escrevemos

$$A = \begin{bmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_p \\ | & & | \end{bmatrix}.$$

Ou então, simplesmente $A = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_p \end{bmatrix}$.

Exemplo 5.2. A matriz do exemplo 5.1 poderia ser escrita

$$\begin{bmatrix} 3 & 15 \\ 10 & 7 \\ 5 & 12 \end{bmatrix} = \begin{bmatrix} \vec{b} & \vec{c} \end{bmatrix},$$

onde aparecem os já conhecidos vetores $\vec{b} = (3, 10, 5)$ e $\vec{c} = (15, 7, 12)$, que indicam as quantidades de nutrientes X, Y e Z em cada grama de batata e cada grama de cenoura.

E algumas vezes, vamos querer falar das linhas da matriz. Nesse caso, se as linhas da matriz A forem os vetores $\vec{a}^1, \dots, \vec{a}^q$, escrevemos

$$A = \begin{bmatrix} — & \vec{a}^1 & — \\ & \vdots & \\ — & \vec{a}^q & — \end{bmatrix}.$$

5.2 Operações com Matrizes

Soma e Produto por Escalar

As operações mais simples com matrizes são a *soma de matrizes* e o *produto por escalar*. Uma matriz $q \times p$ pode ser multiplicada por um escalar $\alpha \in \mathbb{R}$, e duas matrizes $q \times p$ podem ser somadas da mesma forma que vetores são multiplicados por escalar ou somados. De fato, uma matriz $q \times p$ é como se fosse um vetor de $\mathbb{R}^{qp}$, só que ao invés de escvevermos em uma linha, escrevemos em um quadro, que é mais organizado! Ou então, pode-se pensar na matriz $q \times p$ como se fossem p vetores de $\mathbb{R}^q$. Neste caso, o *produto por escalar* e a *soma de matrizes* correspondem ao *produto por escalar* e à *soma* de cada um dos vetores.

$$\alpha \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_p \end{bmatrix} = \begin{bmatrix} \alpha\vec{v}_1 & \dots & \alpha\vec{v}_p \end{bmatrix}$$
$$\begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_p \end{bmatrix} + \begin{bmatrix} \vec{w}_1 & \dots & \vec{w}_p \end{bmatrix} = \begin{bmatrix} \vec{v}_1 + \vec{w}_1 & \dots & \vec{v}_p + \vec{w}_p \end{bmatrix}$$

Podemos também, entender uma matriz $q \times p$ como sendo um vetor. Como se fosse um elemento de $\mathbb{R}^{qp}$. Mas ao invés de escrevermos um linhão comprido, escrevemos os números organizados em uma tabela. Mas assim como os elementos de $\mathbb{R}^{qp}$, a soma e o produto por escalar são feitos entrada por entrada (coordenada a coordenada).

Exemplo 5.3. Sejam

$$A = \begin{bmatrix} 23 & 38 \\ 15 & 7 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 32 & 55 \\ 52 & 11 \end{bmatrix}.$$

Então,

$$A + B = \begin{bmatrix} 55 & 93 \\ 67 & 18 \end{bmatrix} \quad \text{e} \quad 3A = \begin{bmatrix} 69 & 114 \\ 45 & 21 \end{bmatrix}.$$

Produto de Matriz por Matriz Coluna

O leitor provavelmente já sabe, ou pelo menos já viu em algum lugar, que o produto de uma matriz $A = \begin{bmatrix} a_{jk} \end{bmatrix}$ de tamanho $q \times p$ por uma matriz $B = \begin{bmatrix} b_{jk} \end{bmatrix}$ de tamanho $p \times r$ é a matriz $AB = \begin{bmatrix} c_{jk} \end{bmatrix}$ de tamanho $q \times r$ onde

$$c_{jk} = \sum_{s=1}^{p} a_{js} b_{sk}. \tag{5.2.1}$$

Quando B é uma matriz coluna, então AB também será uma matriz coluna. Escrevendo

$$A = \begin{bmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_q \\ | & & | \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} b_1 \\ \vdots \\ b_p \end{bmatrix},$$

a fórmula do produto de matrizes fica assim:

$$AB = b_1 \begin{bmatrix} | \\ \vec{a}_1 \\ | \end{bmatrix} + \cdots + b_p \begin{bmatrix} | \\ \vec{a}_p \\ | \end{bmatrix}.$$

Exemplo 5.4. Vamos calcular o produto de

$$M = \begin{bmatrix} 1 & 5 \\ -2 & 7 \end{bmatrix} \quad \text{e} \quad N = \begin{bmatrix} 3 \\ 8 \end{bmatrix}.$$

Acompanhe pela figura 5.1.

$$\begin{bmatrix} 3 \\ 8 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 5 \\ -2 & 7 \end{bmatrix} = 3\begin{bmatrix} 1 \\ -2 \end{bmatrix} + 8\begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

Figura 5.1: Multiplicando $M = \begin{bmatrix} 1 & 5 \\ -2 & 7 \end{bmatrix}$ pela matriz coluna $N = \begin{bmatrix} 3 \\ 8 \end{bmatrix}$.

Basta então fazermos como mostrado na figura 5.1.

$$\begin{aligned} MN &= 3\begin{bmatrix} 1 \\ -2 \end{bmatrix} + 8\begin{bmatrix} 5 \\ 7 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ -6 \end{bmatrix} + \begin{bmatrix} 40 \\ 56 \end{bmatrix} \\ &= \begin{bmatrix} 43 \\ 50 \end{bmatrix}. \end{aligned}$$

Produto de Matrizes

Já sabemos que existe uma fórmula para calcular

$$C = AB,$$

desde que o número de colunas da matriz A seja igual ao número de linhas de B. É a fórmula da equação (5.2.1). Vamos escrevê-la de um modo mais limpo, partindo do princípio que sabemos calcular o produto de uma matriz qualquer por uma matriz coluna.

Vamos denotar por B_j a j-ésima coluna da matriz B. E por C_j, a j-ésima coluna da matriz C. A fórmula do produto de matrizes fica simples assim:

$$C_j = AB_j.$$

A j-ésima coluna de C é o resultado de se aplicar a j-ésima coluna de B na matriz A.

O que o leitor deve fazer agora, é multiplicar umas matrizes pra treinar/relembrar. Mais pra frente, vamos entender melhor o significado do produto de matrizes.

Pergunta Intrigante 5.5. Se A e B são matrizes $n \times n$. Será que $AB = BA$?

Pergunta Intrigante 5.6. Por que não estamos interessados em definir o produto AB de duas matrizes quando A é $p \times q$, B é $r \times s$, e $r \neq q$?

Pergunta Intrigante 5.7. Se A, B e C são matrizes $p \times q$, $q \times r$ e $r \times s$ respectivamente, podemos calcular $M = BC$, e depois fazer o produto AM. Ou então, podemos fazer $N = AB$, e depois fazer o produto NC. Dizem que dá na mesma! Que $AM = NC$. Em outras palavras,

$$A(BC) = (AB)C.$$

Será que é fácil mostrar que sempre dá igual?

Matriz Diagonal

A matriz *diagonal* é a matriz dos sonhos de qualquer estudante de *Introdução à Álgebra Linear*!!!

- Ah... se todxs fossem iguais a você!!!

Uma matriz quadrada, $n \times n$ é diagonal quando todas as entradas que não estão na *diagonal principal* são iguais a zero. Neste caso, costumamos omitir as entradas que não são zero:

$$\begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{bmatrix}.$$

Vejamos o que aconteceria se, como nos sonhos dos estudantes, todas as matrizes fossem diagonais... 😍

Exemplo 5.8. Produto de duas matrizes diagonais...

$$\begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{bmatrix} \begin{bmatrix} \beta_1 & & & \\ & \beta_2 & & \\ & & \ddots & \\ & & & \beta_n \end{bmatrix} = \begin{bmatrix} \alpha_1\beta_1 & & & \\ & \alpha_2\beta_2 & & \\ & & \ddots & \\ & & & \alpha_n\beta_n \end{bmatrix}.$$ 😱

Faça as contas!

Mas nem todas precisam ser diagonais! As matrizes diagonais se relacionam bem com todas as outras. 😉

Exemplo 5.9. O efeito prático de se multiplicar uma matriz A à direita por uma matriz diagonal com entradas $\alpha_1, \ldots, \alpha_p$, é que as colunas de A são multiplicadas por esses números:

$$\begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_p \\ | & | & & | \end{bmatrix} \begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_p \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \alpha_1\vec{a}_1 & \alpha_2\vec{a}_2 & \cdots & \alpha_p\vec{a}_p \\ | & | & & | \end{bmatrix}.$$

A matriz a esquerda não precisa ser quadrada! O número de linhas é determinado pela dimensão dos vetores. Serão q linhas quando $\vec{a}_j \in \mathbb{R}^q$.

Faça as contas!

Exemplo 5.10. O efeito prático de se multiplicar uma matriz A à direita por uma matriz diagonal com entradas $\alpha_1, \ldots, \alpha_q$, é que as linhas de A são multiplicadas por esses números:

$$\begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_q \end{bmatrix} \begin{bmatrix} — & \vec{a}_1 & — \\ — & \vec{a}_2 & — \\ & \vdots & \\ — & \vec{a}_q & — \end{bmatrix} = \begin{bmatrix} — & \alpha_1\vec{a}_1 & — \\ — & \alpha_2\vec{a}_2 & — \\ & \vdots & \\ — & \alpha_q\vec{a}_q & — \end{bmatrix}.$$

A matriz a esquerda não precisa ser quadrada! O número de colunas é determinado pela dimensão dos vetores. Serão p colunas quando $\vec{a}_j \in \mathbb{R}^p$.

Faça as contas!

Matriz Identidade

Para cada $n = 1, 2, \ldots$, existe uma matriz *diagonal* $n \times n$ que é ainda mais especial. É a chamada *matriz identidade*:

$$I_n = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix},$$

que tem todas as entradas da *diagonal principal* iguais a 1, e as demais iguais a 0. Por comodidade, é comum omitirmos o índice n e escrevermos apenas I para denotar a matriz identidade de tamanho adequado ao contexto.

A matriz identidade pode ser caracterizada como a única matriz tal que

$$[\vec{v}] = I_n[\vec{v}]$$

para todo $\vec{v} \in \mathbb{R}^n$. Temos que

$$A = I_n A \quad \text{e} \quad B = BI_n$$

para toda matriz A de tamanho $n \times p$, e toda matriz B de tamanho $q \times n$, para p e q quaisquer.

Você pode fazer as contas para verificar que, de fato, multipliacar à direita ou à esquerda por I_n é o mesmo que não fazer nada. Para quem se lembra dos exemplos 5.9 e 5.10, também é fácil perceber isso, porque multiplicar por uma matriz diagonal à direita é o mesmo que multiplicar as colunas pelos valores que estão na matriz diagonal, e multiplicar à esquerda é o mesmo que multiplicar as linhas.

Exemplo 5.11. Sejam

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 1 & 2 \\ 5 & 6 \\ 8 & 9 \end{bmatrix}.$$

Então,

$$I_2 A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}.$$

Da mesma forma,

$$BI_2 = \begin{bmatrix} 1 & 2 \\ 5 & 6 \\ 8 & 9 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 5 & 6 \\ 8 & 9 \end{bmatrix}.$$

Faça as contas... :-)

Matriz Inversa

Quando A é uma matriz $n \times n$, por vezes existe uma matrix X, $n \times n$, tal que

$$AX = XA = I_n.$$

Quando isso acontece, dizemos que X é a *inversa* de A e escrevemos $X = A^{-1}$. Da mesma forma, A é a *inversa* de X:

$$A = \left(A^{-1}\right)^{-1}.$$

Muitos problemas podem ser expressos na linguagem matricial. E muitos podem ser resolvidos se conseguirmos inverter uma determinada matriz. Veja a seção 5.4.

Exemplo 5.12. A matriz

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

é a inversa dela mesma:

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

Exemplo 5.13. Faça as contas para perceber que

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

E, portanto,

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$$

são inversas uma da outra.

Existem matrizes que não tem inversa.

Exemplo 5.14. As seguintes matrizes não tem inversa.

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 1 \\ 5 & 5 \end{bmatrix} \qquad \begin{bmatrix} 3 & -2 \\ 6 & -4 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & -1 \\ 4 & 8 & 5 \end{bmatrix} \qquad \begin{bmatrix} 3 & 1 & 2 \\ 1 & 5 & -4 \\ 2 & 1 & 1 \end{bmatrix}.$$

Pergunta Intrigante 5.15. Quando uma matriz A tem, e quando não tem inversa? Tem uma maneira simples de saber?

Pergunta Intrigante 5.16. Por que não falamos da inversa de uma matriz que não é *quadrada*?

Pergunta Intrigante 5.17. Que técnicas podemos usar para encontrar a inversa de uma determinada matriz?

Pergunta Intrigante 5.18. Dada uma matriz quadrada A, $n \times n$. Se eu encontrar M tal que

$$MA = I,$$

já posso concluir que A é inversível e $M = A^{-1}$? Ou eu preciso ainda verificar que $AM = I$?

Pergunta Intrigante 5.19. Dada uma matriz quadrada A, $n \times n$. E se eu encontrar duas matrizes **diferentes** M e N, tais que

$$MA = I \quad \text{e} \quad AN = I?$$

Exemplo 5.20 (inversa só por um lado)**.** A matriz

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

tem *inversa pela esquerda*:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Mas essa **não é a única**!!!

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Compare com o exemplo 5.13.

Da mesma forma,

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

tem ao menos uma inversa à direita: A.

Pergunta Intrigante 5.21. No exemplo 5.20, a matriz A tinha uma inversa à esquerda, e B tinha uma inversa à direita. Será que elas possuem inversa pelo outro lado, também?

Matriz Transposta

Dada uma matriz $q \times p$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & & a_{2p} \\ \vdots & & \ddots & \vdots \\ a_{q1} & a_{q2} & \dots & a_{qp} \end{bmatrix},$$

a *transposta* de A — denotada por A^t — é a matriz $p \times q$ dada por

$$A^t = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{q1} \\ a_{12} & a_{22} & & a_{q2} \\ \vdots & & \ddots & \vdots \\ a_{1p} & a_{2p} & \dots & a_{qp} \end{bmatrix}.$$

Em termos de *linhas* e *colunas*, na matriz transposta, as *linhas* viram *colunas* e as *colunas* viram *linhas*:

$$\begin{bmatrix} | & & | \\ \vec{a}_1 & \dots & \vec{a}_p \\ | & & | \end{bmatrix}^t = \begin{bmatrix} — & \vec{a}_1 & — \\ & \vdots & \\ — & \vec{a}_p & — \end{bmatrix}.$$

Exemplo 5.22. Um exemplo é melhor do que mil palavras! :-)

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^t = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

Transposta é isso!!!

Pergunta Intrigante 5.23. A *matriz transposta* é muito fácil de definir... também é muito fácil de *calcular*... mas que utilidade tem?

Uma propriedade da transposta é a seguinte:

$$(PQ)^t = Q^t P^t.$$

5.3 Matrizes como Funções

Uma matriz $q \times p$

$$A = \begin{bmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_p \\ | & & | \end{bmatrix}$$

pode ser vista como uma função

$$\begin{array}{rccl} T: & \mathbb{R}^p & \to & \mathbb{R}^q \\ & (x_1, \dots, x_p) & \mapsto & x_1\vec{a}_1 + \cdots + x_p\vec{a}_p \end{array}.$$

Funções de $\mathbb{R}^p$ em $\mathbb{R}^q$ que são desse tipo, são chamadas de *transformações lineares* e serão estudadas com mais detalhes no capítulo 6. As *transformações lineares* são *máquinas* de fazer combinações lineares. Você escolhe os coeficientes $x_1, \dots, x_p$, e T faz a combinação linear dos vetores $\vec{a}_1, \dots, \vec{a}_p$ usando esses coeficientes:

$$T(x_1, \dots, x_p) = x_1\vec{a}_1 + \cdots + x_p\vec{a}_p.$$

Assim, T pode ser vista como um *dispositivo*... uma *máquina*... uma *caixinha* com *input* e *output*. Se pensarmos nos vetores $\vec{x} \in \mathbb{R}^p$ e $T\vec{x} \in \mathbb{R}^q$ como matrizes coluna

$$\begin{bmatrix} | \\ \vec{x} \\ | \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} | \\ T\vec{x} \\ | \end{bmatrix},$$

a transformação T fica ainda mais parecida com as *"caixinhas"*:

$$\begin{bmatrix} | \\ T\vec{x} \\ | \end{bmatrix} = \begin{bmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_q \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}. \tag{5.3.1}$$

Do lado direito entra o vetor $(x_1, \ldots, x_p)$, e do lado esquerdo sai o resultado: $T(x_1, \ldots, x_p)$.

De fato, quando definimos o produto de uma matriz por uma matriz coluna na seção 5.2, o fizemos de modo que a equação (5.3.1) seja verdadeira.

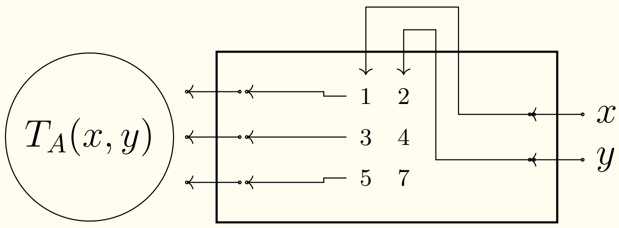


Figura 5.2: A matriz $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 7 \end{bmatrix}$ pode ser vista como uma função $T : \mathbb{R}^2 \to \mathbb{R}^3$, onde

$$\begin{bmatrix} | \\ T(x, y) \\ | \end{bmatrix} = x \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + y \begin{bmatrix} 2 \\ 4 \\ 7 \end{bmatrix}.$$

Exemplo 5.24. A matriz da tabela nutricional do exemplo 5.1 pode ser usada para construir um dispositivo... um *APP*!!! Veja o projeto do *app* na figura 5.3.

Você coloca a quantidade de batatas e cenouras e o *App da Nutrição* mostra para você a quantidade de nutrientes X, Y e Z.

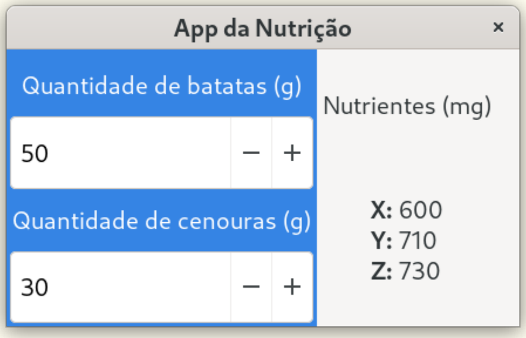


Figura 5.3: *Nutrition app*: a revolução da nutrição!!!

Em python, por exemplo, o nosso *app* poderia ser assim... (se você não tem interesse em programação, apenas ignore!)

Os cálculos (produto de matrizes) são feitos nas linhas 41–43. Lá está o produto de matrizes

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 & 15 \\ 10 & 7 \\ 8 & 11 \end{bmatrix} \begin{bmatrix} p \\ c \end{bmatrix} = p \begin{bmatrix} 3 \\ 10 \\ 8 \end{bmatrix} + c \begin{bmatrix} 15 \\ 7 \\ 11 \end{bmatrix}.$$

```
1 import gi
2 gi.require_version('Gtk', '3.0')
3 from gi.repository import Gtk
4
5 potato_nutrients = {'x': 3, 'y': 10, 'z': 8}
6 carrot_nutrients = {'x': 15, 'y': 7, 'z': 11}
7
8 class NutritionApp(Gtk.Window):
9   def __init__ (self):
10    super().__init__(title='App da Nutrição')
11
12    fbox = Gtk.FlowBox()
```

```
    input_box = Gtk.VBox()
    output_box = Gtk.VBox()
    fbox.add(input_box)
    fbox.add(output_box)

    self.potatoes = Gtk.SpinButton(value=0)
    self.carrots = Gtk.SpinButton(value=0)

    for x in self.potatoes, self.carrots:
        x.set_range(0, 1000)
        x.set_increments(10,100)
        x.connect('value-changed', self.update_total_nutrients)

    input_box.add(Gtk.Label(label='Quantidade de batatas (g)'))
    input_box.add(self.potatoes)
    input_box.add(Gtk.Label(label='Quantidade de cenouras (g)'))
    input_box.add(self.carrots)

    self.result = Gtk.Label()
    output_box.add(Gtk.Label(label='Nutrientes (mg)'))
    output_box.add(self.result)

    self.update_total_nutrients()
    self.add(fbox)

  def update_total_nutrients (self, arg=None):
      p = self.potatoes.get_value()
      c = self.carrots.get_value()
      x = p * potato_nutrients['x'] + c * carrot_nutrients['x']
      y = p * potato_nutrients['y'] + c * carrot_nutrients['y']
      z = p * potato_nutrients['z'] + c * carrot_nutrients['z']

      self.result.set_markup(f'<b>X:</b> {x:0.0f}\n<b>Y:</b> {y:0.0f}\n<b>Z:</b> {z:0.0f}')

app = NutritionApp()
app.connect("destroy", Gtk.main_quit)
app.show_all()
Gtk.main()
```

Pergunta Intrigante 5.25. Dada uma função $T : \mathbb{R}^p \to \mathbb{R}^q$, existe um critério simples para saber se esta função pode ser representada por uma matriz $q \times p$?

Suponha que por algum motivo saibamos que a função $T : \mathbb{R}^p \to \mathbb{R}^q$ pode ser representada por uma matriz. Nesse caso, chamamos T de *transformação linear*. Como podemos determinar a matriz correspondente à transformação T?

Repare que dada a matriz $A = \begin{bmatrix} \vec{a}_1 \cdots \vec{a}_p \end{bmatrix}$, a primeira coluna de A é

$$\begin{bmatrix} | \\ \vec{a}_1 \\ | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_p \\ | & | & & | \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

o produto de A pelo primeiro vetor da base canônica: $\begin{bmatrix} \vec{e}_1 \end{bmatrix}$. De modo geral, obtemos a j-ésima coluna de A quando aplicamos em A o vetor $\vec{e}_j$:

$$\begin{bmatrix} | \\ \vec{a}_j \\ | \end{bmatrix} = A \begin{bmatrix} | \\ \vec{e}_j \\ | \end{bmatrix}.$$

Assim, a matriz correspondente à transformação T é dada por

$$\begin{bmatrix} T \end{bmatrix} = \begin{bmatrix} | & & | \\ T\vec{e}_1 & \dots & T\vec{e}_p \\ | & & | \end{bmatrix}.$$

Em outras palavras, é a matriz cuja j-ésima coluna é dada pelo vetor $T\vec{e}_j$. Dessa forma, sabendo que T é uma *transformação linear*, **é *"fácil"* determinar a matriz correspondente** $\begin{bmatrix} T \end{bmatrix}$: é a matriz cuja primeira coluna é dada por $T\vec{e}_1$, a segunda é dada por $T\vec{e}_2$, etc.

Exemplo 5.26. Suponha agora que lá na cidade de *fimdomundo* tem uma loja que vende dois tipos de sopa. A sopa de verão, com mais cenoura (para ajudar no bronzeado), e a sopa de inverno, com mais batata (para reserva de lipídios, para proteção contra o frio). Do exemplo 4.1, sabemos a quantidade de nutrientes em cada grama de batata e cada grama de cenoura. A quantidade de batata e cenoura por tijela de sopa também pode ser representada em uma tabela.

	verão	**inverno**
batata	10	20
cenoura	30	15

Ou então, uma matriz:

$$S = \begin{bmatrix} 10 & 20 \\ 30 & 15 \end{bmatrix}.$$

Podemos encarar S como uma função de $\mathbb{R}^2$ em $\mathbb{R}^2$, que nos diz a quantidade de batatas e cenouras correspondente a determinadas quantidade de sopas de verão e inverno. Agora, se você quisesse saber a quantidade total de cada nutriente nas sopas que você comprou, você pode abrir o *soup app*, ver a quantidade de batatas e cenouras, copiar–e–colar o resultado no *nutri app* para saber a quantidade de nutrientes.

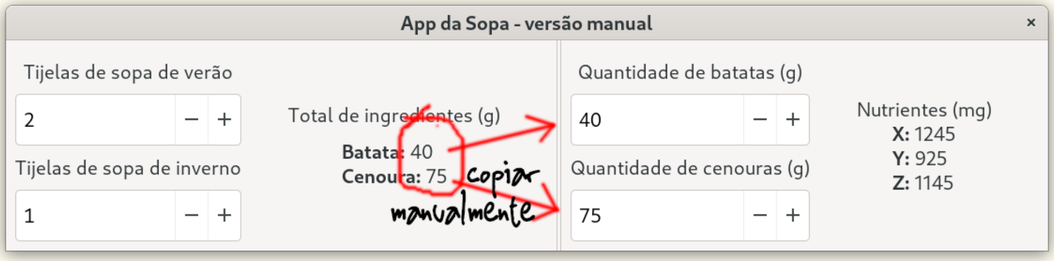


Figura 5.4: *Soup app*: revolução em matéria de sopa!!!
Você coloca o número de tijelas de sopa, o aplicativo lhe diz a quantidade total de batatas e cenouras. Aí você usa esses números, pra descobrir as quantidades de nutrientes, X, Y e Z.

Assumimos que o cliente não está interessado nas quantidades de batata e cenoura. O cliente tem sim, interesse em saber as quantidades dos nutrientes X, Y e Z. O aplicativo, convenhamos, não é lá muito inteligente.

Da mesma maneira, talvez o proprietário do restaurante não devesse colocar duas tabelas, uma com as quantidades de ingredientes por sopa e outra com as quantidades de nutrientes por ingrediente. Faria muito mais sentido, uma tabela com as quantidades de nutrientes por sopa! Precisamos saber como podemos mesclar as tabelas

	verão	**inverno**
batata	10	20
cenoura	30	15

	batata	**cenoura**
nutriente X	3	15
nutriente Y	10	7
nutriente Z	5	12

para formar a seguinte tabela:

	verão	**inverno**
nutriente X	v_x	i_x
nutriente Y	v_y	i_y
nutriente Z	v_z	i_z

A primeira coluna da tabela são as quantidades de nutrientes na *sopa de verão*. Basta calcularmos quantos nutrientes tem em 10 gramas de batata e 30 gramas de cenoura. Ou seja, basta aplicar, em N (matriz do exemplo 4.1), a primeira coluna da matriz S:

$$\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} 3 & 15 \\ 10 & 7 \\ 5 & 12 \end{bmatrix} \begin{bmatrix} 10 \\ 30 \end{bmatrix} = 10 \begin{bmatrix} 3 \\ 10 \\ 5 \end{bmatrix} + 30 \begin{bmatrix} 15 \\ 7 \\ 12 \end{bmatrix}.$$

O mesmo vale para a segunda coluna. O resultado:

	verão	**inverno**
nutriente X	480	285
nutriente Y	310	305
nutriente Z	410	280

No exemplo 5.26, vimos que o produto NS nada mais é do que a tabela que representa o *dispositivo* resultante de se concatenar os *dispositivos* S e N. O produto é a *composição* das duas tabelas. Uma pergunta que precisa ser respondida é:

Pergunta Intrigante 5.27. Considere as transformações T_N e T_S correspondentes às matrizes N e S. Será que $T_N \circ T_S$ é sempre uma função que pode ser representada por uma matriz? Em outras palavras, quando T_1 e T_2 são transformações lineares que podem ser compostas, será que $T_1 \circ T_2$ também é uma transformação linear?

Exemplo 5.28. A função

$$\begin{array}{rccc} f: & \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 \\ & (x,y) & \mapsto & (x+2, x+y) \end{array}$$

não é uma *transformação linear*. Ou seja, não pode ser representada por uma matriz 2×2, da maneira que estamos fazendo. Para qualquer matriz M de tamanho 2×2, é verdade que

$$M \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Ou seja,

$$T_M(0,0) = (0,0) \neq (2,0) = f(0,0).$$

De modo geral, toda *transformação linear* f tem que satisfazer $f(\vec{0}) = \vec{0}$.

Exemplo 5.29. A função

$$\begin{array}{rccc} f: & \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 \\ & (x,y) & \mapsto & (xy, x+y) \end{array}$$

não pode ser representada por uma matriz.

Note que toda matriz M de tamanho 2×2 satisfaz o seguinte:

$$M \begin{bmatrix} 2 \\ 2 \end{bmatrix} = M \left(2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = 2M \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

E portanto,

$$T_M(2,2) = 2T_M(1,1).$$

No entanto,

$$f(2,2) = (4,4) \neq 2(1,2) = 2f(1,1).$$

Produto de matrizes como composição

Nem toda função $f : \mathbb{R}^p \to \mathbb{R}^q$ é *linear*. Nem toda f pode ser representada por uma matriz da maneira que estamos fazendo até agora. No exemplo 5.26, vimos um exemplo em que faziamos a composição de duas matrizes. Tinhamos uma tabela que dizia as quantidades de ingredientes de cada tipo de sopa. E uma tabela com as quantidades de nutrientes em cada ingrediente. Com isso, montamos uma tabela que nos dizia as quantidades de nutrientes em cada tipo de sopa.

Mais pra frente, na proposição 7.8, vamos mostrar que isso sempre ocorre. A composição de duas *transformações lineares* também é uma transformação linear. Assim, se as matrizes P e Q corresponde a transformações lineares $T_P : \mathbb{R}^p \to \mathbb{R}^q$ e $T_Q : \mathbb{R}^q \to \mathbb{R}^k$, então a composição $T_Q \circ T_P$ também pode ser representada por uma matriz. o produto QP é exatamente a matriz que representa essa composição:

$$QP = \left[T_Q \circ T_P\right].$$

De fato, para determinarmos a j-ésima coluna de $\left[T_Q \circ T_P\right]$, tudo o que temos que fazer, é encontrar $T_Q(T_P\vec{e}_j)$. Mas $T_P\vec{e}_j$ é justamente a j-ésima coluna da matriz P.

Então, tudo o que temos que fazer é aplicar a j-ésima coluna de P, diretamente na matriz Q. Mas isso é justamente o produto de matrizes, como descrito nas seção 5.2.

Com essa interpretação de produto de matrizes como composição das transformações lineares correspondentes, podemos responder, por exemplo, às perguntas intrigantes 5.6 e 5.7. Só multiplicamos QP quando o número de colunas de Q é igual ao número de linhas de P, justamente porque só fazemos a composição $T_Q \circ T_P$ quando o domínio de Q é igual ao contradomínio de P. E o produto de matrizes é associativo, pois a composição de funções é associativa. Veja a seção 2.2.

Exemplo: matriz de permutação

Uma função que *permuta* as entradas de um vetor de $\mathbb{R}^p$ é uma *transformação linear*. Melhor do que ficar de blá, blá, blá... é mostrar um exemplo!

Exemplo 5.30. Tomar um vetor (x, y, z) e transformá-lo em (z, x, y), é o que faz a transformação

$$\begin{array}{rccc} T: & \mathbb{R}^3 & \to & \mathbb{R}^3 \\ & (x, y, z) & \mapsto & x\vec{e}_2 + y\vec{e}_3 + z\vec{e}_1 \end{array}.$$

Ela leva $\vec{e}_1 \mapsto \vec{e}_2$, $\vec{e}_2 \mapsto \vec{e}_3$ e $\vec{e}_3 \mapsto \vec{e}_1$. Assim, temos a representação matricial

$$[T] = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Em cada linha tem apenas um número 1. Os demais são zero. O mesmo ocorre em cada coluna: apenas uma entrada tem o número 1, e as demais são zero.

E, de fato,

$$\begin{bmatrix} z \\ x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

E se multiplicarmos uma matriz de *permutação* por uma outra matriz?

Exemplo 5.31. Continuando o exemplo 5.30, vamos calcular

$$\begin{bmatrix} | & | & | \\ \vec{b}_1 & \vec{b}_2 & \vec{b}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Para encontrar $\vec{b}_1$, tudo o que precisamos fazer é usar $\vec{e}_1$ como *input* na matriz de permutação; e depois tomar o resultado e aplicá-lo na matriz seguinte.

O vetor $\vec{e}_1$ aplicado na matriz de permutação retorna $\vec{e}_2$. Este, quando aplicado na matriz seguinte, retorna $\vec{a}_2$. Assim, $\vec{b}_1 = \vec{a}_2$.

Da mesma forma, $\vec{e}_2$ é levado em $\vec{e}_3$, que por sua vez é levado em $\vec{a}_3$. E, ao final, $\vec{e}_3$ é levado em $\vec{a}_1$.

De um modo geral, se

$$\sigma : [n] \to [n]$$

é uma permutação, podemos imaginar uma transformação T_σ, que leva a j-ésima coordenada do vetor para a posição $\sigma(j)$. Sendo assim,

$$\begin{array}{rrcl} T_\sigma : & \mathbb{R}^n & \to & \mathbb{R}^n \\ & (x_1, \dots, x_n) & \mapsto & (x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)}) \end{array}$$

é uma transformação linear. Repare que na fórmula de T_σ utilizamos a *permutação inversa* σ^{-1}! Para entender o porquê, basta notar que σ nos diz em que posição a j-ésima coordenada vai parar. Mas para escrever o resultado em ordem, precisamos saber

- Quem é que vai parar na primeira posição?
- Quem é que vai parar na segunda posição?
Etc...

Quem vai parar na j-ésima posição é a coordenada $\sigma^{-1}(j)$, pois $\sigma\left(\sigma^{-1}(j)\right) = j$.

No entanto, se pensarmos nos vetores da base canônica, $\vec{e}_1, \dots, \vec{e}_n$, é fácil ver que

$$T_\sigma \vec{e}_j = \vec{e}_{\sigma(j)}.$$

E portanto, T_σ é representada pela matriz

$$[T_\sigma] = \begin{bmatrix} | & & | \\ \vec{e}_{\sigma(1)} & \dots & \vec{e}_{\sigma(n)} \\ | & & | \end{bmatrix}.$$

Quanto à multiplicação de uma matriz A, à direita, por uma permutação, temos que

$$\begin{bmatrix} | & & | \\ \vec{a}_{\sigma(1)} & \dots & \vec{a}_{\sigma(n)} \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \vec{a}_1 & \dots & \vec{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} | & & | \\ \vec{e}_{\sigma(1)} & \dots & \vec{e}_{\sigma(n)} \\ | & & | \end{bmatrix}.$$

Isso é claro, pois com a composição das duas matrizes, $\vec{e}_j$ é levado em $\vec{e}_{\sigma(j)}$, que por sua vez é levado em $\vec{a}_{\sigma(j)}$.

Estamos olhando para as *colunas* da matriz A como sendo o *input*, e as *linhas* como sendo o *output*. Quando permutamos à direita, ou seja, antes de aplicar a matriz A. Estamos permutando o *input*. E por isso, a operação afeta as colunas de A. Se multiplicarmos à esquerda, estamos permutando o *output*.

Exemplo 5.32. O leitor está convidado a calcular o resultado de se multiplicar uma matriz A à esquerda por uma matriz de permutação T_σ. O resultado prático é o de se permutar as linhas da matriz, com a permutação inversa.

Por exemplo,

$$\begin{bmatrix} — & \vec{b}_1 & — \\ — & \vec{b}_2 & — \\ — & \vec{b}_3 & — \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} — & \vec{a}_1 & — \\ — & \vec{a}_2 & — \\ — & \vec{a}_3 & — \end{bmatrix},$$

é tal que

$$\vec{b}_1 = \vec{a}_3, \quad \vec{b}_2 = \vec{a}_1 \quad \text{e} \quad \vec{b}_3 = \vec{a}_2.$$

De um modo geral, se σ é uma permutação de $[n]$,

$$\begin{bmatrix} — & \vec{b}_1 & — \\ & \vdots & \\ — & \vec{b}_n & — \end{bmatrix} = [T_\sigma] \begin{bmatrix} — & \vec{a}_1 & — \\ & \vdots & \\ — & \vec{a}_n & — \end{bmatrix}$$

é tal que a coordenada j é levada na coordenada $\sigma(j)$. Assim,

$$\vec{b}_{\sigma(j)} = \vec{a}_j.$$

Ou, utilizando a permutação inversa,

$$\vec{b}_j = \vec{a}_{\sigma^{-1}(j)}.$$

Esse argumento todo com as linhas pode ser feito utilizando a matriz transposta. Basta observar que

$$\left([T_\sigma]A\right)^t = A^t [T_\sigma]^t.$$

Ou seja,

$$[T_\sigma]A = \left(A^t [T_\sigma]^t\right)^t.$$

A operação de transposição transforma linhas em colunas e vice-versa. É preciso, também, notar que $[T_\sigma]^t = T_{\sigma^{-1}}$.

Exemplo: multiplicando muitas matrizes

Por vezes, é mais fácil trabalhar com produto de matrizes observando o que é que elas fazem como funções: *transformações lineares.*

Só por diversão, vamos determinar

$$M = M_1 M_2 M_3 M_4,$$

onde

$$M_1 = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \quad M_2 = \begin{bmatrix} 5 & 7 & 1 \\ 4 & 1 & 0 \\ 1 & 3 & 0 \end{bmatrix} \quad M_3 = \begin{bmatrix} 2 & 0 & 0 \\ 5 & 0 & 0 \\ 4 & 1 & 3 \end{bmatrix} \quad M_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Para determinar M, vamos passar os vetores $\vec{e}_1$, $\vec{e}_2$ e $\vec{e}_3$, um por um pelas matrizes M_4, M_3, M_2 e M_1, e registrar os resultados, que correspondem às colunas de M. O resultado está descrito na figura 5.5

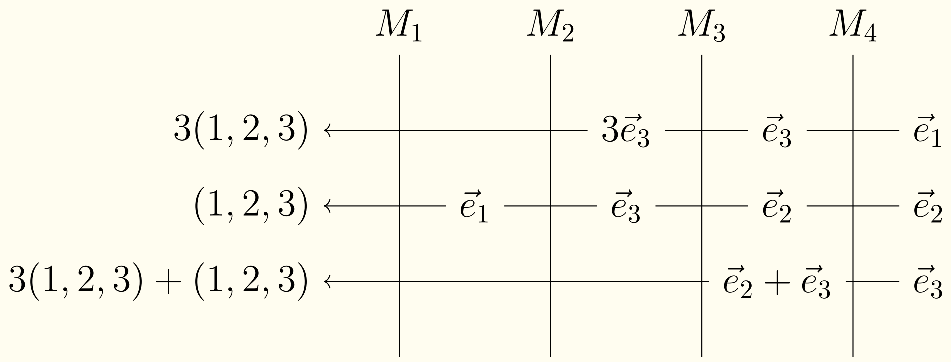


Figura 5.5: Aplicando os vetores $\vec{e}_1$, $\vec{e}_2$ e $\vec{e}_3$ na sequência de matrizes M_4, M_3, M_2 e M_1, descobrimos a j-ésima coluna de $M = M_1M_2M_3M_4$. Em alguns momentos, já sabemos alguns resultados, que não precisamos calcular de novo.

A matriz inversa como função

Uma função pode ter inversa à esquerda, quando é *injetiva*; e pode ter inversa à direta, quando é *sobrejetiva*. Quando a função é *bijetiva*, ela tem apenas uma inversa à esquerda, que é também a única inversa à direita.

Mais adiante, veremos que uma transformação linear $T : \mathbb{R}^p \to \mathbb{R}^q$ só pode ser *bijetiva* quando $p = q$. E é por isso que só falamos em inversa de matrizes *quadradas*.

Vamos refazer o exemplo 5.12.

Exemplo 5.33 (permutação de dois elementos)**.** A matriz

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

é a inversa dela mesma. Afinal de contas, vista como função, ela permuta $\vec{e}_1$ e $\vec{e}_2$:

$$T_M\vec{e}_1 = \vec{e}_2 \quad \text{e} \quad T_M\vec{e}_2 = \vec{e}_1.$$

Portanto,

$$MM = \big[T_M \circ T_M\big] = \big[\mathrm{id}\big] = I_2.$$

Um outro caso que já vimos, é o exemplo 5.13.

Exemplo 5.34. A matriz

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

leva $\vec{e}_2$ em $\vec{e}_1$. Portanto, **se a inversa existir**, deve levar $\vec{e}_1$ em $\vec{e}_2$. Assim, a primeira coluna da matriz inversa é formada pelo vetor $\vec{e}_2$:

$$\begin{bmatrix} 0 & * \\ 1 & * \end{bmatrix}.$$

Por outro lado, $(1, -1)$ é levado em $\vec{e}_2$:

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

A inversa, se existir, deve levar $\vec{e}_2$ em $(1, -1)$. Assim, a segunda coluna da inversa é formada pelo vetor $(1, -1)$. Portanto, se tem uma matriz que é a inversa de $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, tem que ser a matriz $\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$.

No exemplo 5.34, enfatizamos bastante a expressão **se a inversa existir**. Isso, porque o que de fato mostramos é que tal matriz **desfaz** a operação que a matriz original fez. Portanto, o que determinamos foi a *inversa à esquerda*. Esta questão está posta na pergunta intrigante 5.18.

Já sabemos que existem matrizes que não tem inversa...

Exemplo 5.35. Já vimos no exemplo 5.14 que as matrizes a seguir não possuem inversa. Você consegue demonstrar isso de uma maneira simples?

$$\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix} \qquad \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}$$
$$\begin{bmatrix}1 & 1\\ 5 & 5\end{bmatrix} \qquad \begin{bmatrix}3 & -2\\ 6 & -4\end{bmatrix}$$
$$\begin{bmatrix}1 & 2 & 3\\ 2 & 4 & -1\\ 4 & 8 & 5\end{bmatrix} \qquad \begin{bmatrix}3 & 1 & 2\\ 1 & 5 & -4\\ 2 & 1 & 1\end{bmatrix}.$$

Por exemplo, as duas primeiras matrizes

$$\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix} \quad \text{e} \quad \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}$$

levam qualquer múltiplo do vetor $\vec{e}_2$ em $(0, 0)$. E portanto, não são *injetivas*. Já a matriz

$$\begin{bmatrix}1 & 1\\ 5 & 5\end{bmatrix}$$

leva qualquer vetor em um múltiplo do vetor $(1, 5)$. Assim, vetores que não são múltiplos de $(1, 5)$ não estão na imagem da transformação correspondente, que portanto, não é *sobrejetiva*.

E as outras? Você consegue dizer porque não são inversíveis?

5.4 Exemplos

Resolvendo sistemas lineares

Um sistema linear pode ser escrito na forma matricial

$$A[\vec{v}] = [\vec{a}],$$

onde a matrix $m \times n$ A e $\vec{a} \in \mathbb{R}^m$ são dados, e $\vec{v} \in \mathbb{R}^n$ é o que queremos descobrir.

Se A é uma matriz inversível, e $\vec{v}$ uma solução do sistema, então

$$\vec{v} = A^{-1}A\big[\vec{v}\big] = A^{-1}\big[\vec{a}\big].$$

Usamos o fato de que $A^{-1}A = I$ para concluir que **se $\vec{v}$ é solução do sistema**, então podemos determinar

$$\vec{v} = A^{-1}\big[\vec{a}\big].$$

Por outro lado, o fato de que $AA^{-1} = I$ implica que $\vec{v} = A^{-1}\big[\vec{a}\big]$ é solução do sistema, pois

$$A\vec{v} = AA^{-1}\big[\vec{a}\big] = \big[\vec{a}\big].$$

Repare como foi importante o fato de A^{-1} ser a inversa de A tanto quando multiplicado à esquerda, quanto quando multiplicado à direita.

Vamos fazer um caso mais concreto, com números. Considere o sistema linear

$$\begin{cases} x + 2y + 3z = 7 \\ \quad + \;\; y + 4z = 8 \\ 5x + 6y + \quad\; = 9 \end{cases}. \tag{5.4.1}$$

Em termos de matrizes, o sistema assume a forma

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 5 & 6 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}. \tag{5.4.2}$$

Queremos determinar $x, y, z \in \mathbb{R}$.

Seja

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

A inversa de A é

$$A^{-1} = \begin{bmatrix} -24 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{bmatrix}.$$

Portanto, a **única solução** do sistema da equação (5.4.2) é dada por

$$\begin{aligned} [\vec{v}] &= \begin{bmatrix} -24 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \\ &= 7 \begin{bmatrix} -24 \\ 20 \\ -5 \end{bmatrix} + 8 \begin{bmatrix} 18 \\ -15 \\ 4 \end{bmatrix} + 9 \begin{bmatrix} 5 \\ -4 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -168 \\ 140 \\ -35 \end{bmatrix} + \begin{bmatrix} 144 \\ -120 \\ 32 \end{bmatrix} + \begin{bmatrix} 45 \\ -36 \\ 9 \end{bmatrix} \\ &= \begin{bmatrix} 21 \\ -16 \\ 6 \end{bmatrix}. \end{aligned}$$

Assim, descobrimos que $\vec{v} = (21, -16, 6)$. De fato,

$$\begin{cases} \phantom{5 \times {}} 21 - 2 \times 16 + 3 \times 6 = 7 \\ - \phantom{2 \times {}} 16 + 4 \times 6 = 8 \\ 5 \times 21 - 6 \times 16 \phantom{{} + 4 \times 6} = 9 \end{cases}.$$

Por vezes, nem mesmo da matriz nós precisamos. O sistema da equação (5.4.1) pode ser pensado em termos de uma *transformação linear*. Se $T : \mathbb{R}^3 \to \mathbb{R}^3$ é a *transformação linear* representada pela matriz

$$[T] = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 5 & 6 & 0 \end{bmatrix},$$

o sistema pode ser escrito na forma

$$T(x, y, z) = (7, 8, 9).$$

Um exemplo será dado a seguir, quando falarmos sobre *sistemas lineares homogêneos*.

Sistemas homogêneos

Um *sistema linear homogêneo* é um sistema da forma

$$T(\vec{v}) = \vec{0},$$

onde $T : \mathbb{R}^p \to \mathbb{R}^q$ é uma *transformação linear* e $\vec{v}$ são as *incógnitas*.

O primeiro fato interessante sobre os *sistemas homogêneos* é que ele sempre tem ao menos uma solução $\vec{v} = \vec{0}$. Pois para uma transformação linear é sempre verdade que

$$T(\vec{0}) = \vec{0}.$$

Esta é a *solução trivial*. Será que existem outras? A resposta depende, claro, do sistema específico.

Considere um sistema não necessariamente homogêneo

$$T(\vec{v}) = \vec{a}. \tag{5.4.3}$$

Suponha que esse sistema tenha uma solução que não é única. Ou seja, existem $\vec{s}, \vec{t} \in \mathbb{R}^p$ **distintos** tais que $T\vec{s} = T\vec{t}$ $(= \vec{a})$. Neste caso, o sistema homogêneo associado tem solução não trivial, pois $\vec{s} - \vec{t} \neq \vec{0}$. E pela *linearidade* de T,

$$T(\vec{s} - \vec{t}) = T\vec{s} - T\vec{t} = \vec{0}.$$

Por outro lado, quando $\vec{z}$ é uma solução qualquer do *sistema homogêneo* e $\vec{s}$ é uma solução de (5.4.3), então $\vec{s} + \vec{z}$ também é solução.

Assim, denotando por

$$Z = T^{-1}(\vec{0})$$

o conjunto de todas as soluções do sistema homogêneo, as soluções de (5.4.3) são dadas pelo conjunto

$$\vec{s} + Z = \left\{ \vec{s} + \vec{z} \;\middle|\; \vec{z} \in Z \right\}.$$

Em particular, para um sistema linear, ou todas as soluções são únicas ou nenhuma solução é única. Quando Z é um conjunto unitário, todas as soluções são únicas; e quando não é unitário, nenhuma solução é única. Mas é importante notar que **mesmo que as soluções não sejam únicas**, podem haver escolhas para $\vec{a}$ tais que o sistema não tenha solução.

Exemplo 5.36. Seja $T : \mathbb{R}^3 \to \mathbb{R}^2$ a *transformação linear* cuja representação matricial seja dada por

$$[T] = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}.$$

O *sistema homogêneo* correpondente

$$T\vec{v} = \vec{0}$$

tem mais de uma solução. Por exemplo,

$$\begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} - \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

E, portanto,

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - 2 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Ou seja, $T(1, -2, 1) = (0, 0, 0)$.

O leitor está convidado a mostrar que

$$T\vec{v} = (0, 1, 0)$$

não tem solução. Mas, no entanto, para $\vec{a} = (5,7,9)$, além da solução

$$T(1,1,0) = (5,7,9),$$

temos também

$$\begin{aligned} T(1+1, 1-2, 0+1) &= T(2,-1,1) \\ &= (2*1-4+7, 2*2-5+8, 2*3-6+9) \\ &= (5,7,9). \end{aligned}$$

Por fim, podemos também concluir que um sistema homogêneo pode

1. não ter solução;
2. ter uma única solução; ou
3. ter infinitas soluções.

Isso, porque o conjunto $Z = T^{-1}(\vec{0})$ ou é unitário com

$$Z = \left\{\vec{0}\right\};$$

ou é inifinito, pois, para todo $\alpha \in \mathbb{R}$,

$$T(\alpha\vec{z}) = \alpha T\vec{z} = \alpha\vec{0} = \vec{0}.$$

Portanto, se $\vec{z} \in Z$ não é o vetor nulo, então $\alpha\vec{z} \in Z$ $(\alpha \in \mathbb{R})$ são infinitas soluções distintas para o *sistema homogêneo*.

Inversa à esquerda e à direita

xxx

Multiplicidade de caminhos

xxx

Cadeias de Markov

xxx

5.5 Exercícios

Exercício 5.1. Calcule:

1. $\begin{bmatrix}1 & 1\\ 5 & 5\end{bmatrix}\begin{bmatrix}0.7\\ 0.3\end{bmatrix}$
2. $\begin{bmatrix}1 & 1\\ 5 & 5\end{bmatrix}\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$
3. $\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 5 & 5\end{bmatrix}$
4. $\begin{bmatrix}1 & 1\\ 5 & 5\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix}$
5. $\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}^{-1}$
6. $\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\begin{bmatrix}-2\\ 1\end{bmatrix}$
7. $\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\begin{bmatrix}\frac{-2}{-2}\\ \frac{1}{-2}\end{bmatrix}$
8. $\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\begin{bmatrix}-4\\ 3\end{bmatrix}$
9. $\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\begin{bmatrix}\frac{-4}{2}\\ \frac{3}{2}\end{bmatrix}$
10. $\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\begin{bmatrix}\frac{-4}{-2} & 1\\ \frac{3}{2} & \frac{-1}{2}\end{bmatrix}$

Exercício 5.2. Insumos e custos.

Exercício 5.3. Insumos e custos.

Exercício 5.4. Escreva os sistemas lineares do exercício 4.6 usando matrizes coluna.

Exercício 5.5. Escreva os sistemas lineares do exercício 4.6 usando produto de matrizes.

Exercício 5.6. Encontre a Matriz M tal que

$$M\begin{bmatrix}1\\ 2\end{bmatrix} = \begin{bmatrix}1\\ 0\end{bmatrix} \quad \text{e} \quad M\begin{bmatrix}2\\ 1\end{bmatrix} = \begin{bmatrix}0\\ 1\end{bmatrix}.$$

Exercício 5.7. Encontre a inversa de

1. $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$
2. $\begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}$
3. $\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$

Exercício 5.8. Resolva os sistemas lineares.

1. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -5 & 4 & 0 \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

Aula 6

Transformações Lineares com Matrizes

As transformações lineares são o principal objeto de estudo da disciplina de *álgebra linear*. Nos capítulos anteriores aprendemos sobre *vetores* e *transformações lineares* de um ponto de vista mais voltado para os cálculos. Agora, vamos entender o que é uma transformação linear de um ponto de vista mais *geométrico*.

6.1 Motivação Geométrica

A figura 6.1 ilustra uma maneira de transformarmos um desenho do plano. A ideia é quadricular o desenho original e fazer uma correspondência entre esse quadriculado e o reticulado de destino, formado de pequenos paralelogramos.

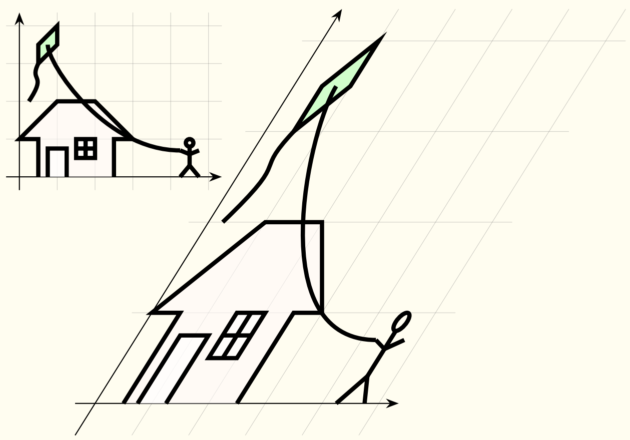


Figura 6.1: Qual é a maneira *"correta"* de se transportar um desenho no quadriculado em cima, para um desenho no reticulado em baixo?

Cada ponto do desenho original é levado em um ponto no novo reticulado. Qual é a regra que determina onde determinado ponto $P = (\alpha, \beta)$ é levado? As coordenadas de P indicam que, partindo da origem, para chegarmos ao ponto P, devemos andar α unidades na direção e no sentido do vetor $\vec{e}_1 = (1, 0)$, e β unidades na direção e no sentido do vetor $\vec{e}_2 = (0, 1)$. Estamos apenas dizendo que $P = \alpha\vec{e}_1 + \beta\vec{e}_2$. Pela figura, é evidente que queremos que os pontos $\vec{e}_1 = (1, 0)$ e $\vec{e}_2 = (0, 1)$ sejam transportados para os pontos correspondentes aos vetores $\vec{v}_1$ e $\vec{v}_2$, representados na figura 6.2. Já o ponto P, ao ser transportado, deverá corresponder ao vetor $\alpha\vec{v}_1 + \beta\vec{v}_2$. O que queremos é que nesse segundo sistema de coordendadas $(\vec{v}_1, \vec{v}_2)$, o ponto P seja transportado para o ponto $(\alpha, \beta)_{\mathscr{R}}$, onde $\mathscr{R} = (\vec{v}_1, \vec{v}_2)$ é a base ordenada formada pelos vetores $\vec{v}_1$ e $\vec{v}_2$.

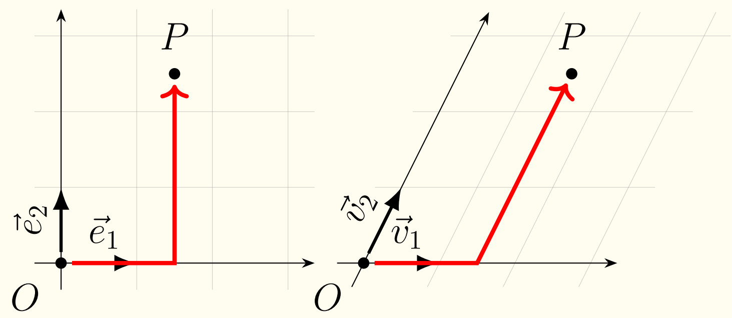


Figura 6.2:

Note que pra fazermos os cálculos, não estamos de fato trabalhando com pontos. Não cosumamos *somar* pontos, ou multiplicá-los por um escalar. Estamos trabalhando com vetores. Se escolhermos um ponto base O, podemos então falar do ponto $O + \vec{v}$, ao invés de falarmos do vetor $\vec{v}$. Com essa identificação, podemos usar todo o ferramental disponível ao se trabalhar com vetores: soma e produto por escalar, combinações lineares, conjuntos geradores, base, transformação linear, etc. Daqui por diante, não vamos mais fazer teoria com pontos, apenas vetores. Vamos deixar os pontos para os exemplos ou aplicações.

Estamos então, falando de uma transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$. O vetor $\vec{p}$ (vamos usar letra minúscula, já que não é mais um *ponto*) é levado no vetor $T(\vec{p})$. É costume omitir os parênteses e escrever $T\vec{p}$. Acompanhando pela figura 6.3, fica claro que a imagem de $\vec{e}_1$ é o vetor $\vec{v}_1$, e que a imagem de $\vec{e}_2$ é o vetor $\vec{v}_2$. Como já explicado, a imagem do vetor $\vec{p} = p_1\vec{e}_1 + p_2\vec{e}_2$ é alcançada se andarmos p_1 unidades do vetor $\vec{v}_1$ e em seguida, p_2 unidades do vetor $\vec{v}_2$. Ou seja,

$$T\vec{p} = p_1\vec{v}_1 + p_2\vec{v}_2.$$

Ou, numa notação mais precisa,

$$\begin{array}{rrcl} T: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (p_1, p_2) & \mapsto & p_1\vec{v}_1 + p_2\vec{v}_2 \end{array}.$$

Como já sabemos, as transformações desse tipo são as já conhecidas *transformações lineares*. Não esqueça que

$$\vec{v}_1 = T\vec{e}_1 \quad \text{e} \quad \vec{v}_2 = T\vec{e}_2.$$

Assim, nossa transformação original pode ser expressa pela matriz

$$[T] = [\vec{v}_1 \ \ \vec{v}_2].$$

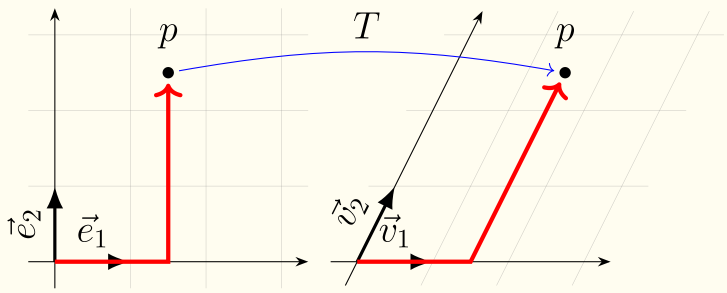


Figura 6.3:

6.2 Casa Inclinada

Vamos desenhar a *"casa inclinada"* da figura 6.1 em um papel quadriculado. Suponha que nesse papel quadriculado, os vetores $\vec{v}_1$ e $\vec{v}_2$ tem coordenadas $\vec{v}_1 = (1.5, 0)$ e $\vec{v}_2 = (0.5, 1)$. Qual é, por exemplo, a coordenada da quina superior direita da janela? Se no desenho original a coordenada da quina é $\vec{p} = (2, 1)$, então, no desenho inclinado, a quina corresponde ao ponto

$$\begin{aligned} T\vec{p} &= 2\vec{v}_1 + \vec{v}_2 \\ &= (3, 0) + (0.5, 1) = (3.5, 1). \end{aligned}$$

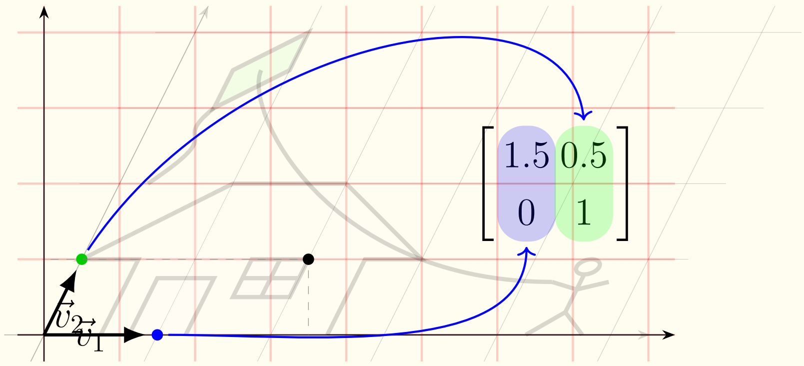


Figura 6.4:

Se fossemos desenhar a *"casa inclinada"* da figura 6.1 na tela de um *tablet*, provavelmente precisamos passar ao tablet a informação de quais são as coordenadas *horizontal* e *vertical* do ponto da tela que queremos manipular. O leitor que souber alguma linguagem de programação não terá dificuldade em fazer um programa que faça esse tipo de transformação com imagens! O programa *ImageMagick* faz isso e um pouco mais! Veja:

https://www.imagemagick.org/Usage/distorts/affine/

6.3 Rotação de $90°$

Vamos aprender um *"macete"* muito útil...

Quando giramos o plano um quarto de volta no sentido anti-horário, $\vec{e}_1$ é levado em $\vec{e}_2$, e $\vec{e}_2$ é levado em $-\vec{e}_1$. Isso significa que para rotacionarmos um vetor (a, b) por um ângulo $\frac{\pi}{2}$, basta trocar a ordem de a e b, e colocar um sinal de menos na primeira entrada: $(-b, a)$. Ou seja, essa rotação é a transformação linear $R_{\frac{\pi}{2}}$, representada pela matriz

$$[R_{\frac{\pi}{2}}] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

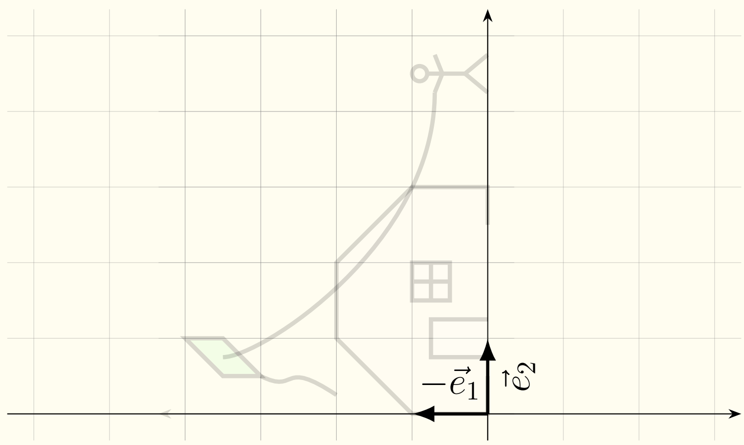


Figura 6.5: Rotacionar uma imagem 90° no sentido anti-horário é o mesmo que aplicar $T(x,y) = (-y,x)$. Dito de outro modo,

$$\begin{array}{rccc} T: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & x\vec{e}_1 + y\vec{e}_2 & \mapsto & x\vec{e}_2 - y\vec{e}_1 \end{array},$$

pois a rotação é *linear*, $T\vec{e}_1 = \vec{e}_2$ e $T\vec{e}_2 = -\vec{e}_1$.

Da mesma forma, se girarmos no sentido horário, teremos $(b,-a)$. Em outras palavras,

$$[R_{-\frac{\pi}{2}}] = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Sem fazer muita conta, sabemos que $R_{\frac{\pi}{2}}^{-1} = R_{-\frac{\pi}{2}}$. Sendo assim, se tivéssemos paciência de fazer as contas... poderíamos veririficar que:

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

6.4 Rotação

A transformação $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$, que leva um vetor $\vec{v}$ em sua rotação por um ângulo α, é uma transformação linear. Você consegue se convencer, ou convencer alguém disso?

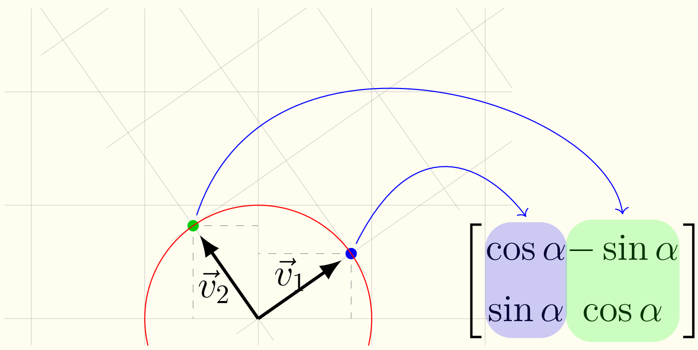


Figura 6.6: A rotação R_α é uma transformação linear que leva $\vec{e}_1$ em $(\cos(\alpha), \sin(\alpha))$, e $\vec{e}_2$ em $(-\sin(\alpha), \cos(\alpha))$.

De fato, a rotação é uma transformação muito mais simples que a da *"casa inclinada"* (da seção 6.2), já que os eixos são apenas rotacionados. Os eixos não são esticados e o ângulo entre eles continua sendo reto. Na seção 6.2, os *quadradinhos* do quadriculado foram transformados em paralelogramos. Agora, os *quadradinhos* continuam *quadradinhos*, só que rotacionados. Vamos chamar essa transformação de R_α. A representação matricial de R_α é

$$[R_\alpha] = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}.$$

De fato, como $\vec{v}_1$ faz um ângulo α com o eixo das abcissas,

$$\vec{v}_1 = (\cos(\alpha), \sin(\alpha)).$$

Agora, é só observar que $\vec{v}_2$ é o $\vec{v}_1$ rotacionado 90°, como na seção 6.3, para concluir que

$$\vec{v}_2 = (-\sin(\alpha), \cos(\alpha)).$$

Para o leitor que ainda não está convencido da utilidade de tudo isso que tem sido discutido, o próximo exemplo é, na opinião do autor, bastante impressionante.

Exemplo 6.1 (Soma de ângulos)**.** Dados dois ângulos, α e β, existe uma fórmula para o cálculo de

$$\sin(\alpha+\beta) \quad \text{e} \quad \cos(\alpha+\beta).$$

Vamos deduzir essas fórumlas. :-)

Por um lado (lado esquerdo da igualdade) sabemos que ao rotacionar por um ângulo α o vetor $\vec{v} = (\cos(\beta), \sin(\beta))$, obtemos o vetor $(\cos(\alpha+\beta), \sin(\alpha+\beta))$. Por outro lado (lado direito), usando a *linearidade* da aplicação R_α,

$$\begin{aligned}
\begin{bmatrix} \cos(\alpha+\beta) \\ \sin(\alpha+\beta) \end{bmatrix} &= [R_\alpha \vec{v}] \\
&= \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} \cos(\beta) \\ \sin(\beta) \end{bmatrix} \\
&= \cos(\beta) \begin{bmatrix} \cos(\alpha) \\ \sin(\alpha) \end{bmatrix} + \sin(\beta) \begin{bmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{bmatrix} \\
&= \begin{bmatrix} \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \\ \cos(\beta)\sin(\alpha) + \sin(\beta)\cos(\alpha) \end{bmatrix}.
\end{aligned}$$

Ou seja,

$$\begin{aligned}
\cos(\alpha+\beta) &= \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \\
\sin(\alpha+\beta) &= \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta).
\end{aligned}$$

Agora, o leitor é convidado a procurar a demonstração dessas fórmulas em qualquer livro do ensino médio!

Composição de rotação... é rotação!

Exemplo 6.2. Uma outra maneira de expressar a soma de ângulos em termos de matrizes, é calculando $R_{\alpha+\beta} = R_\alpha \circ R_\beta$.

$$\begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{bmatrix}.$$

Os cálculos são essencialmente os mesmos do exemplo 6.1.

Exemplo 6.3 (Inversa da rotação)**.** Às vezes, usamos a álgebra de matrizes para calcular coisas geométricas, como a soma de ângulos. E às vezes usamos geometria para não precisarmos calcular as matrizes…

Mesmo sem fazer cálculos com matrizes, sabemos que

$$\begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Isso, porque $R_\alpha^{-1} = R_{-\alpha}$. E também, porque $\cos(-\alpha) = \cos(\alpha)$ e $\sin(-\alpha) = -\sin(\alpha)$.

6.5 Reflexão

Por fim, uma transformação linear simples que não é uma rotação, é o *espelhamento* pelo eixo das ordenadas. Igualmente simples é o espelhamento pelo eixo das abscissas.

Exemplo 6.4 (Espelho)**.** Se o eixo das ordenadas fosse um *"espelho"*, transformaria um desenho o desenho da casa em uma versão espelhada…

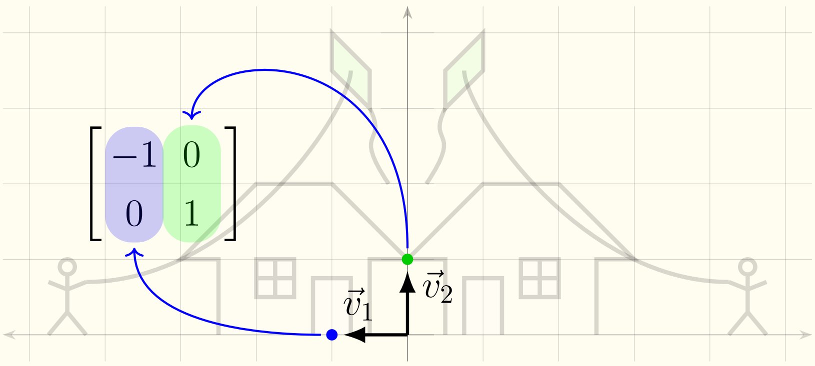


Figura 6.7:

A matriz que representa a transformação linear correspondente é

$$[T] = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

E se fosse pra fazer um espelho com o eixo das abscissas? Então, seria

$$[T] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

E se quiséssemos um espelho ortogonal ao vetor $\vec{v}$?

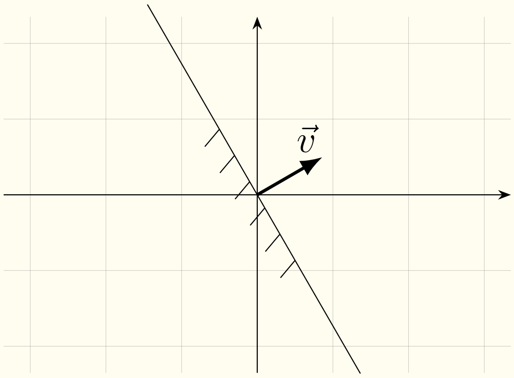


Figura 6.8:

6.6 Composição de Transformações Lineares

Sejam $T : \mathbb{R}^2 \to \mathbb{R}^2$ e $S : \mathbb{R}^2 \to \mathbb{R}^2$ transformações lineares. Já sabemos da proposição 7.8 que a composição

$$\begin{array}{rccc} S \circ T : & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & \vec{x} & \mapsto & S(T\vec{x}) \end{array}$$

também é uma transformação linear.

Exemplo 6.5 (Duas rotações)**.** Seja $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ a rotação por um ângulo α, e $R_\beta : \mathbb{R}^2 \to \mathbb{R}^2$ a rotação por um ângulo β. Então, $R_\alpha \circ R_\beta$ é a rotação pelo ângulo $\alpha + \beta$.

Assim, obtemos, como já foi visto no exemplo 6.1,

$$\begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{bmatrix}.$$

Exemplo 6.6 (Espelho)**.** Considere a transformação dada no final do exemplo 6.4.

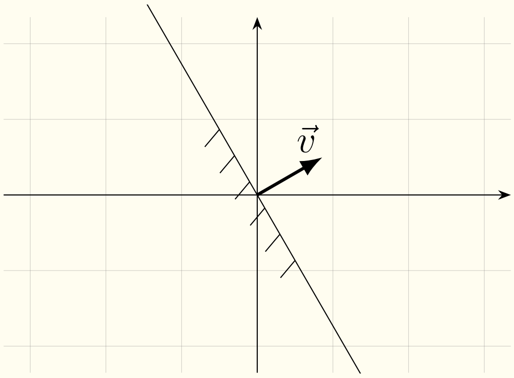


Figura 6.9: Se $\vec{v} = (\cos(\alpha), \sin(\alpha))$, então, ao rotacionar toda a figura por um ângulo $-\alpha$, transformamos o problema no prolema da figura 6.7. Ou seja, basta fazermos uma rotação R_α^{-1}, seguida de uma reflexão pelo eixo das ordenadas. Aí então, precisamos desfazer a rotação aplicando R_α.

Ou seja, considere um vetor $\vec{v} = (\cos(\alpha), \sin(\alpha))$. A transformação $E_{\vec{v}}$, que age como um espelho normal ao vetor $\vec{v}$, pode ser representada como

$$E_{\vec{v}} = R_{\vec{v}} \circ E_{\vec{e}_1} \circ R_{\vec{v}}^{-1},$$

onde $R_{\vec{v}}$ é a rotação que leva $\vec{e}_1$ em $\vec{v}$, $R_{\vec{v}}^{-1}$ é a *inversa* de $R_{\vec{v}}$ (a rotação que leva $\vec{v}$ em $\vec{e}_1$) e $E_{\vec{e}_1}$ é o espelhamento pelo eixo das ordenadas. Ou seja,

$$[E_{\vec{e}_1}] = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Vamos determinar a representação matricial de $E_{\vec{v}}$. Usando a representação matricial, e escrevendo a e b no lugar de $\cos(\alpha)$ e $\sin(\alpha)$,

$$[E_{\vec{v}}] = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

Usando a proposição 7.9, e fazendo primeiro o produto das duas matrizes à esquerda,

$$\begin{aligned} [E_{\vec{v}}] &= \begin{bmatrix} -a & -b \\ -b & a \end{bmatrix} \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \\ &= \begin{bmatrix} -a^2 + b^2 & -2ab \\ -2ab & -b^2 + a^2 \end{bmatrix}. \end{aligned}$$

Agora, o leitor é convidado a quadricular um desenho qualquer, fazer as contas e transportá-lo por um espelhamento como o desse exemplo. Que tal tentar com $\vec{v} = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$?

6.7 Transformações em $\mathbb{R}^3$

Para representarmos uma *transformação linear* em $\mathbb{R}^2$, desenhavamos um quadriculado e um reticulado no plano, e *"transportávamos"* o desenho do quadriculado para o reticulado. Em $\mathbb{R}^3$ é muito parecido, mas agora não é mais um quadriculado. Quem sabe um *"cubiculado"*!!! :-)

Exemplo 6.7. Vamos desenhar uma casinha em 3 dimensões, aplicando várias transformações.

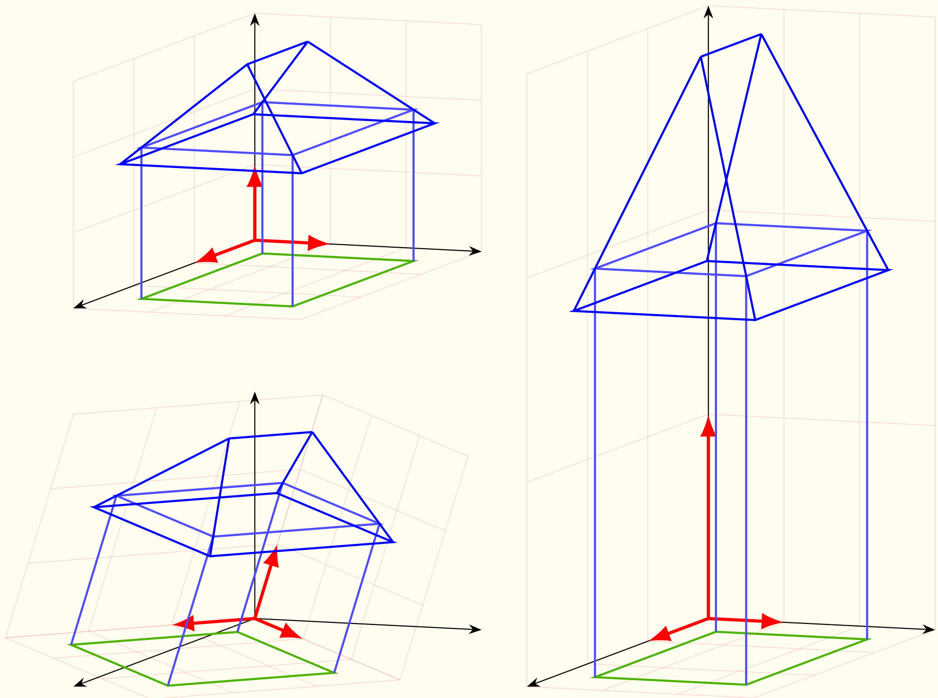


Figura 6.10:

Para especificarmos as *transformações lineares* da figura 6.10, precisamos apenas dizer qual é a imagem dos vetores $\vec{e}_1$, $\vec{e}_2$ e $\vec{e}_3$. A imagem desses vetores estão indicadas em vermelho. Para *"esticar"* a casinha como na parte direita da figura, foi aplicada a transformação

$$\begin{array}{rrcl} T: & \mathbb{R}^3 & \to & \mathbb{R}^3 \\ & (a,b,c) & \mapsto & (a, 2.7b, c) \end{array}.$$

O outro caso é mais complicado. É dado pela transformação

$$\begin{array}{rrcl} T: & \mathbb{R}^3 & \to & \mathbb{R}^3 \\ & (a,b,c) & \mapsto & (1.2a + 0.3b - 0.3c, b + 0.2c, 0.7a + c) \end{array}.$$

Repare que os vetores $T\vec{e}_1 = (1.2, 0, 0.7)$, $T\vec{e}_2 = (0.3, 1, 0)$ e $T\vec{e}_3 = (-0.3, 0.2, 1)$ é que determinam os números *"mágicos"* que aparecem na expressão $(1.2a + 0.3b - 0.3c, b + 0.2c, 0.7a + c)$.

6.8 Projetando $\mathbb{R}^3$ em $\mathbb{R}^2$

Na seção 6.7, tem uma porção de desenhos de uma casinha em 3D que foi transformada de várias maneiras. Mas a verdade, é que os desenhos foram feitos no papel, que é 2D. :-)

Projetar um desenho 3D num sistema de coordenadas 2D é muito simples. No desenho original, cada ponto tem três coordenadas:

$$x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3.$$

Tudo o que precisamos fazer, é especificar, no plano, um ponto para ser a origem, e três vetores no plano para corresponderem a $\vec{e}_1$, $\vec{e}_2$ e $\vec{e}_3$. Estamos basicamente desenhando, no plano, os eixos x, y e z.

Exemplo 6.8. Se escolhermos

$$\vec{e}_1 = (1, 0), \quad \vec{e}_2 = (0, 1) \quad \text{e} \quad \vec{e}_2 = (1, 0.5),$$

o desenho da casinha fica como mostrado na figura 6.11.

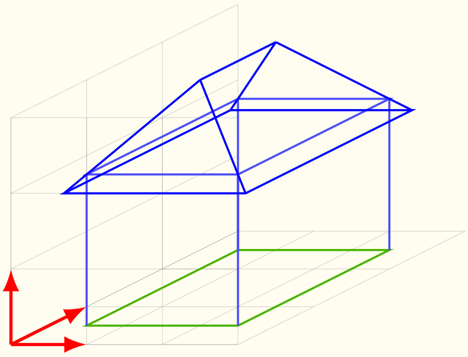


Figura 6.11: Escolhendo, por exemplo, $P\vec{e}_1 = (1,0)$, $P\vec{e}_2 = (0,1)$ e $P\vec{e}_3 = (1, 0.5)$, projetamos $\mathbb{R}^3$ em $\mathbb{R}^2$.

6.9 Números Complexos

Esta seção é um exemplo de como os conceitos aprendidos neste capítulo podem nos ajudar a entender melhor os números complexos.

Não existe um número real x tal que $x^2 = -1$. Mas podemos acrescentar a $\mathbb{R}$ um novo símbolo i, tal que $i^2 = -1$. Podemos fazer isso *preservando as propriedades* das operações de soma e produto. Fazemos soma e produto exatamente como fazemos com polinômios, tratando i como se fosse uma variável. Mas ao final, substituímos i^2 por -1.

$$(a+bi)+(x+yi) = (a+x)+(b+y)i \tag{6.9.1}$$

$$\begin{aligned}(a+bi)(x+yi) &= ax + (ay+xb)i + byi^2 \\ &= (ax-by) + (ay+xb)i.\end{aligned} \tag{6.9.2}$$

E como tudo na vida, na matemática a gente também precisa de sorte! :-)

O objetivo deste capítulo é mostrar ao leitor como essa construção algébrica resulta em um objeto com muitas propriedades geométricas. **Sorte?**

Os números complexos são o conjunto

$$\mathbb{C} = \left\{ a + bi \;\middle|\; (a, b) \in \mathbb{R}^2 \right\},$$

junto com as operações de soma e produto. A notação $a+bi$ é uma questão de gosto. Utilizando a identificação $(a, b) \sim a+bi$, também é muito comum falarmos do número complexo (a, b).

No entanto, há que se tomar cuidado! Existe uma operação de soma em $\mathbb{R}^2$, e uma operação de soma em $\mathbb{C}$. Mas, por sorte, são a mesma coisa: soma coordenada a coordenada. Ou seja, a equação (6.9.1) mostra que tratando (a, b) e (x, y) como números complexos

$$(a + bi) + (x + yi) \sim (a, b) + (x, y).$$

Em $\mathbb{R}^2$, temos o produto por escalar

$$(z, w) = \alpha(x, y),$$

com $\alpha \in \mathbb{R}$. Em $\mathbb{C}$, também podemos multiplicar um número real por um número complexo:

$$\alpha(a + bi) = \alpha a + \alpha bi.$$

Com a identificação entre $\mathbb{C}$ e $\mathbb{R}^2$,

$$\alpha(a + bi) \sim \alpha(a, b).$$

Portanto, ao identificarmos $\mathbb{C}$ e $\mathbb{R}^2$, estamos identificando mais do que os conjuntos $\mathbb{C}$ e $\mathbb{R}^2$. Estamos também identificando-os como *espaços vetoriais*: soma e produto por escalar (produto por número real) também são identificados.

Mas e o produto de dois números complexos? Em $\mathbb{C}$, podemos multiplicar dois números, como mostra a equação (6.9.2). Mas em $\mathbb{R}^2$, a princípio não temos um produto

$$(z, w) = (a, b)(x, y).$$

Seja $k = a + bi$ um número complexo. Vamos denotar por

$$\begin{array}{rccc} M_k : & \mathbb{C} & \to & \mathbb{C} \\ & z & \mapsto & kz \end{array}$$

a transformação que toma um $z \in \mathbb{C}$ qualquer e multiplica por k. As propriedades de *associatividade*, *comutatividade* e *distributividade* em $\mathbb{C}$ mostram que M_k é uma transformação linear. De fato, para $\alpha \in \mathbb{R}$

$$M_k(\alpha z) = k(\alpha z) = \alpha(kz) = \alpha M_k(z)$$
$$M_k(z + w) = k(z + w) = kz + kw = M_k(z) + M_k(w).$$

Assim, identificando $\mathbb{C}$ com $\mathbb{R}^2$, M_k pode ser representado como uma matriz. Podemos descobrir as colunas da matriz, calculando M_k em $(1, 0)$ e $(0, 1)$. Lembre-se que $(1, 0) \sim 1$ e $(0, 1) \sim i$. Portanto, as colunas de $\big[M_k\big]$ são dadas por

$$M_k(1) = k1 = k = a + bi \tag{6.9.3}$$
$$M_k(i) = ki = -b + ai. \tag{6.9.4}$$

Ou seja,

$$\big[M_k\big] = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Vamos analisar melhor essa matriz, que é parecida com a *matriz de rotação* da seção 6.4.

Produto por i

Quando $k = i$, a matriz de M_k assume a forma

$$\big[M_k\big] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

E essa é exatamente a *rotação de* 90°, da seção 6.3.

Multiplicar um número complexo por i é o mesmo que rotacioná-lo 90° no sentido anti-horário.

Produto de complexos é rotação e dilatação

Assim como fizemos com vários exemplos de transformações lineares no capítulo 6, vejamos de que forma a transformação linear M_k transforma $\mathbb{R}^2$. Vamos denotar os números complexos tanto na forma $x + yi$, quanto como um par (x, y) sem fazer maiores ressalvas. Lembre-se que misturando essas duas notações, $1 = (1, 0)$ e $i = (0, 1)$. Primeiro, vamos fazer um caso específico, com um exemplo.

Exemplo 6.9. No caso $k = 4 + 3i$,

$$M_k(1,0) = (4,3) \quad \text{e} \quad M_k(0,1) = (-3,4).$$

Portanto, $\mathbb{R}^2$ é transformado como mostrado na figura 6.12.

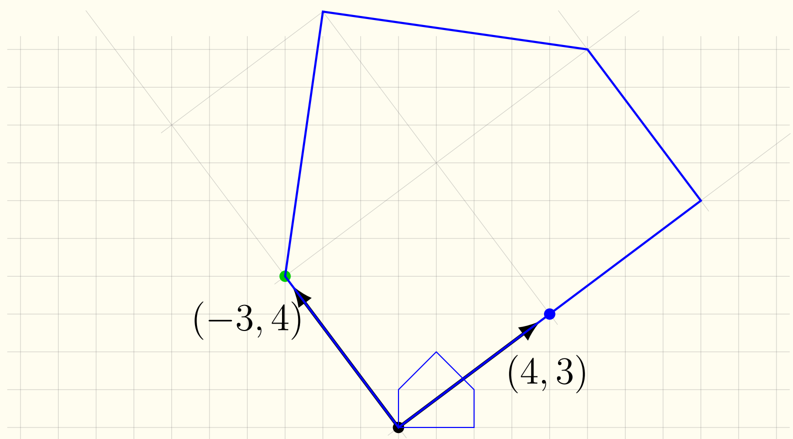


Figura 6.12: Os novos eixos são dados pelos vetores $(4, 3)$ e $(-3, 4)$, que são *ortogonais* e ambos tem tamanho igual a 5. Veja como a figura, além de rotacionada, também tem seu tamanho multiplicado por 5.

Como no exemplo 6.9, o vetor $(-b, a)$ é o vetor (a, b) rotacionado 90° no sentido anti-horário. E o tamanho de ambos é $\sqrt{a^2 + b^2}$. Quando multiplicamos um número complexo por $k = a + bi$, estamos fazendo duas coisas:

1. multiplicando pelo número real $\sqrt{a^2 + b^2}$; e
2. rotacionando pelo ângulo θ que k faz com o eixo x.

De fato, em termos de matrizes,

$$[M_k] = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \sqrt{a^2 + b^2} \begin{bmatrix} \frac{a}{\sqrt{a^2+b^2}} & \frac{-b}{\sqrt{a^2+b^2}} \\ \frac{b}{\sqrt{a^2+b^2}} & \frac{a}{\sqrt{a^2+b^2}} \end{bmatrix}.$$

E a matriz

$$[R_\theta] = \begin{bmatrix} \frac{a}{\sqrt{a^2+b^2}} & \frac{-b}{\sqrt{a^2+b^2}} \\ \frac{b}{\sqrt{a^2+b^2}} & \frac{a}{\sqrt{a^2+b^2}} \end{bmatrix}$$

é a matriz de rotação da seção 6.4.

6.10 Exercícios

Exercício 6.1. resolva os exercícios anteriores que pedem para "resolver", "calcular" e "encontrar", **sem fazer muita conta**, com argumentos **mais interessantes** do que simplesmente aplicar os métodos tradicionais para solução de sistemas lineares.

Exercício 6.2. Sem ficar fazendo muita conta, calcule:

1. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 8 & 3 \\ 6 & 9 & 3 \\ 7 & 10 & 3 \end{bmatrix}$
2. $\begin{bmatrix} 5 & 8 & 3 \\ 6 & 9 & 3 \\ 7 & 10 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
3. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}^2$
4. $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$
5. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
6. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x & z \\ y & w \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

7. $\left(\left(\begin{bmatrix}1&4&7\\2&5&8\\3&6&9\end{bmatrix}\begin{bmatrix}0&0&0\\0&1&0\\1&0&0\end{bmatrix}\right)\begin{bmatrix}1&0&0\\0&0&0\\0&1&3\end{bmatrix}\right)\begin{bmatrix}0&0&0\\0&1&1\\1&0&1\end{bmatrix}$

8. $\left(\left(\begin{bmatrix}0&x&y\\0&1&3\\0&2&4\end{bmatrix}\begin{bmatrix}5&6&1\\0&0&2\\0&0&3\end{bmatrix}\right)\begin{bmatrix}4&w&1\\5&z&2\\0&0&3\end{bmatrix}\right)\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$

Exercício 6.3. Sabemos que do ponto de vista nutricional, a única coisa importante no universo, são os nutrientes X e Y. Seja $\vec{b} = (b_x, b_y) \in \mathbb{R}^2$ o vetor que representa a quantidade b_x de nutriente X, e a quantidade b_y de nutriente Y que contém uma batata. E seja $\vec{c} = (c_x, c_y) \in \mathbb{R}^2$ o vetor que representa a quantidade c_x de nutriente X, e a quantidade c_y de nutriente Y que contém uma cenoura. Sabemos que a função

$$\begin{array}{rrcl} T: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (b, c) & \mapsto & b\vec{b} + c\vec{c} \end{array}$$

que indica a quantidade de nutrientes X e Y contidos em b batatas e c cenouras é uma transformação linear. Determine a representação matricial de T.

Exercício 6.4. Continuando o exercício 6.3, suponha que existam dois tipos de sopa, de verão e de inverno, feitas apenas de batata e cenoura, em proporções diferentes. Se cada cumbuca da sopa de inverno tem $\vec{i} = (i_b, i_c)$ de batata e cenoura, e se cada cumbuca de sopa de verão tem $\vec{v} = (v_b, v_c)$ de batata e cenoura, então

$$\begin{array}{rrcl} S: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (v, i) & \mapsto & v\vec{v} + i\vec{i} \end{array}$$

indica a quantidade de batata e cenoura contidos em v cumbucas de sopa de verão e i cumbucas de sopa de inverno.

Nesse contexto, qual é o significado de $T \circ S$? E $S \circ T$?

Utilizando o produto de matrizes, determine o sistema de equações lineares que relaciona a quantidade de nutrientes X e Y contidos em v cumbucas de sopa de verão e i cumbucas de sopa de inverno.

Exercício 6.5. Determine se $T : \mathbb{R}^2 \to \mathbb{R}^2$ é ou não bijetiva quando $\begin{bmatrix} T \end{bmatrix}$ é igual a

1. $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$
2. $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$
3. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
4. $\begin{bmatrix} 3 & 9 \\ 4 & 12 \end{bmatrix}$
5. $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

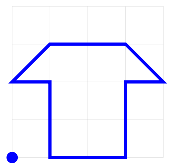
Exercício 6.6. Determine se as aplicações $T : \mathbb{R}^2 \to \mathbb{R}^2$ dadas a seguir são *lineares*.

1. $T(x, y) = (2x + y, 3x - y)$.
2. $T(x, y) = (2|x + y|, y)$.
3. $T(x, y) = (x + y, 0)$.
4. $T(x, y) = (y^2, x^2)$.
5. $T(x, y) = (x + y, x^{-1})$.
6. $T(x, y) = (-x, -x)$.

Exercício 6.7. Dadas as aplicações $T : \mathbb{R}^2 \to \mathbb{R}^2$, determine T^{-1}, quando existir.

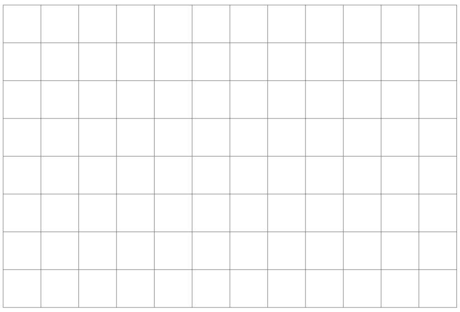
1. $T(x, y) = (4x + y, 2x)$.
2. $T(x, y) = (x - 4y, -5x + 20y)$.
3. $T(x, y) = (2x + 2y, 2x - y)$.
4. $T(x, y) = (y, 2x - y)$.

Exercício 6.8. Aplique ao desenho

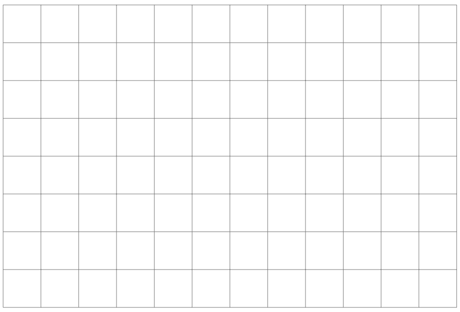


as transformações indicadas. Primeiro, escolha o ponto que representa o vetor $\vec{0}$.

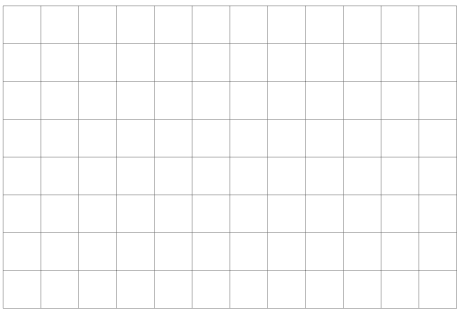
$$\begin{bmatrix} 1 & 1 \\ -0.5 & 2 \end{bmatrix}$$



$$\begin{bmatrix} 2 & 1 \\ -0.5 & 2 \end{bmatrix}$$



$$\begin{bmatrix} -1 & 1 \\ -0.5 & 2 \end{bmatrix}$$



Aula 7

Transformações Lineares SEM Matrizes

As transformações lineares são o principal objeto de estudo da disciplina de *álgebra linear*. Nos capítulos anteriores aprendemos sobre *vetores* e *transformações lineares* de um ponto de vista mais voltado para os cálculos. Agora, vamos entender o que é uma transformação linear de um ponto de vista mais *abstrato*.

7.1 Funções Representadas por Matrizes

> Por que acreditamos que basta uma tabela para representar as quantidades de nutrientes em uma determinada quantidade de batatas e cenouras?

> Por que acreditamos que basta uma tabela para sabermos os custos de produção de uma certa quantidade de bicicletas e motos?

Exemplo 7.1. O preço de certos produtos pode ser representado por uma tabela? Por exemplo...

	batata	**cenoura**	**inhame**	**couve**
Preço / Kg (R$)	4	5	6	10

Chegando lá na quitanda, tem uma plaquinha de desconto no preço da batata:

> Levando 1Kg (ou mais) de batata, você tem 50% de desconto.

Se levar 500g de batata, 500g de cenoura, 500g de inhame e 250g de couve, sai tudo por **R$ 10**. E se levar 1Kg de batata, 1Kg de cenoura, 1Kg de inhame e 500g de couve, sai tudo por **R$ 18**.

Neste caso, a função preço não pode ser representada por uma matriz no espírito do que estamos fazendo até agora.

O que torna o exemplo 7.1 diferente daquilo que temos feito até agora, é a falta de consistência entre os preços dos diferentes *"pacotes"* de legumes e verduras. Por consistência, queremos dizer que o preço de uma combinação qualquer de legumes e verduras deve ser igual à combinação dos preços. A combinação de x quilos de batata, y quilos de cenoura, z quilos de inhame e w quilos de couve pode ser expressa pelo vetor $(x, y, z, w) \in \mathbb{R}^4$. Mas pode ser expressa também na base canônica

$$(x, y, z, w) = x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3 + w\vec{e}_4.$$

Se $T : \mathbb{R}^4 \to \mathbb{R}^1$ é uma função que tem a *"consistência desejada"*, então,

$$T(x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3 + w\vec{e}_4) = xT\vec{e}_1 + yT\vec{e}_2 + zT\vec{e}_3 + wT\vec{e}_4.$$

O preço da combinação é a combinação dos preços! Nesse caso, dizemos que T é uma *transformação linear*.

Essa propriedade não vale apenas para os vetores $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$, etc. Podemos falar da combinação de quaisquer vetores.

Exemplo 7.2. Se o preço na quitanda não tem complicações como no exemplo 7.1. Ou seja, se o preço é *linear*. Então, se Vinícius encheu uma sacola com quantidade de itens representada pelo vetor $\vec{v}$, e se Walter encheu uma sacola $\vec{w}$, Vinícius pagará $T\vec{v}$ e Walter pagará $T\vec{w}$.

Se Gabriel quer comprar o equivalente a 5 sacolas de Vinícius e 3.5 sacolas de Walter, ou seja, $\vec{g} = 5\vec{v} + 3.5\vec{w}$, então Gabriel pagará

$$T(\vec{g}) = T(5\vec{v} + 3.5\vec{w}) = 5T\vec{v} + 3.5T\vec{w}.$$

Mas se Cinthia, a mãe de Gabriel, disser que tem muita coisa no carrinho de compras, e disser pra ele remover o equivalente a uma sacola de Walter, então Gabriel pagará

$$T(\vec{g} - \vec{w}) = T\vec{g} - T\vec{w}.$$

Ou seja, $T\vec{w}$ a menos do que pagaria se sua mãe não tivesse pedido pra remover uma sacola do tipo $\vec{w}$.

Definição 7.3 (linearidade em espaços $\mathbb{R}^n$)**.** *Uma função*

$$T : \mathbb{R}^p \to \mathbb{R}^q$$

é uma transformação linear *quando*

1. $\vec{v}, \vec{w} \in \mathbb{R}^p \Rightarrow T(\vec{v} + \vec{w}) = T\vec{v} + T\vec{w}$.
2. $\alpha \in \mathbb{R}$, $\vec{v} \in \mathbb{R}^p \Rightarrow T(\alpha\vec{v}) = \alpha T\vec{v}$.

Existem várias maneiras de entender o que é uma *transformação linear*. Vamos resumir isso em uma *proposição*. Mas antes, precisamos de algumas definições.

Definição 7.4 (combinação linear)**.** *Seja $B \subset \mathbb{R}^n$ um conjunto de vetores. O vetor $\vec{v}$ é uma* combinação linear *de elementos de B quando existirem $\vec{b}_1, \ldots, \vec{b}_k \in B$ e $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tais que*

$$\vec{v} = \alpha_1 \vec{b}_1 + \cdots + \alpha_k \vec{b}_k.$$

Definição 7.5 (conjunto gerador)**.** *Um subconjunto $B \subset \mathbb{R}^n$ é chamado de* gerador *quando* todo *$\vec{v} \in \mathbb{R}^n$ pode ser escrito como combinação linear de elementos de B.*

Proposição 7.6 (caracterizações de linearidade)**.** *Considere $T : \mathbb{R}^p \to \mathbb{R}^q$. As condições a seguir são equivalentes.*

1. *T é linear (definição 7.3).*
2. *$a \in \mathbb{R}$, $\vec{v}, \vec{w} \in \mathbb{R}^p \Rightarrow T(a\vec{v} + \vec{w}) = aT\vec{v} + T\vec{w}$.*
3. *Para toda combinação linear $\vec{v} = \alpha_1 \vec{v}_1 + \cdots + \alpha_k \vec{v}_k$ de elementos de $\mathbb{R}^p$, vale que*

   $$T(\vec{v}) = \alpha_1 T\vec{v}_1 + \cdots + \alpha_k T\vec{v}_k.$$

4. *Para algum gerador $B \subset \mathbb{R}^p$, vale que*

   $$T(\alpha_1 \vec{b}_1 + \cdots + \alpha_k \vec{b}_k.) = \alpha_1 T\vec{b}_1 + \cdots + \alpha_k T\vec{b}_k$$

   para toda combinação linear de elementos de B.
5. *Para qualquer $(v_1, \ldots, v_p) \in \mathbb{R}^p$,*

   $$T(v_1, \ldots, v_p) = v_1 T\vec{e}_1 + \cdots + v_p T\vec{e}_p.$$

6. Existem $a_1, \dots, a_p \in \mathbb{R}^q$ *tais que*

$$T(v_1, \dots, v_p) = v_1\vec{a}_1 + \cdots + v_p\vec{a}_p.$$

TODO: comentários.

Demonstração. TODO: demonstrar!

Exemplo 7.7. Se $A = \begin{bmatrix}\vec{a}_1 \cdots \vec{a}_q\end{bmatrix}$ é uma matriz $p \times q$,

$$\begin{array}{rrcl} T_A : & \mathbb{R}^q & \to & \mathbb{R}^p \\ & (v_1, \dots, v_q) & \mapsto & v_1\vec{a}_1 + \cdots + v_q\vec{a}_q \end{array}$$

é a *transformação linear* associada a A. Assim, o item 6 da proposição 7.6 afirma essencialmente que dizer que uma função $f : \mathbb{R}^q \to \mathbb{R}^p$ é linear é a mesma coisa que dizer que pode ser representada por uma matriz $p \times q$.

Dessa forma, a proposição 7.6 responde à pergunta intrigante 5.25.

E respondendo a uma pergunta intrigante...

Proposição 7.8. *Se* $f : \mathbb{R}^p \to \mathbb{R}^q$ *e* $g : \mathbb{R}^r \to \mathbb{R}^p$ *são lineares, então* $f \circ g$ *também é linear.*

Demonstração. Vamos verificar o item 2 da proposição 7.6.

Sejam $a \in \mathbb{R}$ e $\vec{v}, \vec{w} \in \mathbb{R}^r$. Então,

$$\begin{aligned} (f \circ g)(a\vec{v} + \vec{w}) &= f(g(a\vec{v} + \vec{w})) \\ &= f(ag(\vec{v}) + g(\vec{w})) \\ &= af(g(\vec{v})) + f(g(\vec{w})) \\ &= a(f \circ g)(\vec{v}) + (f \circ g)(\vec{w}). \end{aligned}$$

E portanto, $f \circ g$ é linear.

A proposição 7.8 nos permite responder à pergunta intrigante 5.27.

Considere duas matrizes P e

$$Q = \begin{bmatrix} | & & | \\ \vec{q}_1 & \dots & \vec{q}_n \\ | & & | \end{bmatrix},$$

onde o número de colunas de P é igual ao número de linhas de Q. Vamos denotar por T_P e T_Q, as transformações lineares correspondentes a essas duas matrizes. Veja a seção 5.3.

Pelo que foi discutido acima, a composição $T = T_P \circ T_Q$ também é linear, e pode ser representada por uma matriz. Vamos denotar tal matriz por

$$M = \begin{bmatrix} | & & | \\ \vec{m}_1 & \dots & \vec{m}_n \\ | & & | \end{bmatrix},$$

Para determinar a matriz M, tudo o que precisamos fazer é calcular todos os $T(\vec{e}_j)$. Esse vetor corresponde à j-ésima coluna de M, $\vec{m}_j$.

Assim, a j-ésima coluna de M corresponde ao vetor

$$T(\vec{e}_j) = T_P(T_Q(\vec{e}_j)) = T_P(\vec{q}_j)$$

Ou seja,

$$\vec{m}_j = P\vec{q}_j.$$

E portanto, $M = PQ$.

Concluímos que o produto de duas matrizes P e Q é justamente a matriz correspondente à composição das transformações lineares:

$$T_{PQ} = T_P \circ T_Q.$$

Proposição 7.9 (o produto de matrizes é associativo)**.** *Se as matrizes P, Q e R podem ser multiplicadas, então*

$$(PQ)R = P(QR).$$

Demonstração. Lembre-se que o produto de matriz representa a composição das transformações lineares. Na seção 2.2, a equação (2.2.1) mostra que composição de funções é *associativa*. Assim,

$$\begin{aligned} T_{(PQ)R} &= T_{PQ} \circ T_R \\ &= (T_P \circ T_Q) \circ T_R \\ &= T_P \circ (T_Q \circ T_R) \\ &= T_P \circ T_{QR} = T_{P(QR)}. \end{aligned}$$

O produto de matrizes não é *comutativo*.

Exemplo 7.10 (Produto não é comutativo)**.** Assim como a composição de funções não é *comutativa*, o produto de matrizes também não é. Por exemplo,

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \neq \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

O leitor é convidado a fazer as contas! Que interpretação você pode dar a esse exemplo específico?

7.2 Núcleo da Transformação

O *núcleo* de uma transformação linear é o conjunto de todos os vetores que são levados por T no vetor nulo.

Definição 7.11 (Núcleo)**.** *O* núcleo *ou* kernel *da transformação linear* $T : V \to W$ *é o conjunto*

$$\ker(T) = \left\{\vec{v} \in V \,\middle|\, T\vec{v} = \vec{0}\right\}.$$

É fácil ver que o $\ker(T)$ é um subsepaço de V. De fato, $\ker(T) \neq \emptyset$, pois $T\vec{0} = \vec{0}$. E se $\vec{a}, \vec{b} \in \ker(T)$, então $\alpha\vec{a} + \beta\vec{b} \in \ker(T)$, pois

$$\begin{aligned} T(\alpha\vec{a} + \beta\vec{b}) &= \alpha T\vec{a} + \beta T\vec{b} \\ &= \alpha\vec{0} + \beta\vec{0} \\ &= \vec{0}. \end{aligned}$$

Notação 7.12 (Imagem inversa)**.** *É comum usarmos a notação*

$$T^{-1}(\vec{a}) = \left\{\vec{v} \in V \,\middle|\, T\vec{v} = \vec{a}\right\}.$$

Nesse caso, $\ker(T) = T^{-1}(\vec{0})$.

Já vimos que resolver um sistema de equações lineares com m linhas e n incógnitas, corresponde a descobrir todos os vetores $\vec{v} \in \mathbb{R}^n$ tais que $T\vec{v} = \vec{w}$, onde T e $\vec{w} \in \mathbb{R}^m$ são dados. Com a notação 2, resolver um tal sistema linear corresponde a determinar o conjunto $T^{-1}(\vec{w})$. E o núcleo $\ker(T)$ nada mais é do que a solução do sistema homogêneo.

Injetividade

Se a transformação linear $T : V \to W$ é injetiva, então é evidente que $\ker(T) = \left\{\vec{0}\right\}$, pois só pode haver um vetor de V levado em $\vec{0} \in W$.

Por outro lado, suponha que T não seja injetiva. Ou seja, suponha que existam vetores distintos $\vec{a}, \vec{b} \in V$ tais que $T\vec{a} = T\vec{b}$. Então,

$$\begin{aligned} T(\vec{a} - \vec{b}) &= T\vec{a} - T\vec{b} \\ &= \vec{0}. \end{aligned}$$

Mas como $\vec{a} \neq \vec{b}$, temos que

$$\vec{0} \neq \vec{a} - \vec{b} \in \ker(T).$$

Ou seja, T é injetiva exatamente quando $\ker(T) = \left\{\vec{0}\right\}$. Resumindo...

Proposição 7.13. *Uma transformação linear $T : V \to W$ é injetiva se, e somente se,* $\ker(T) = \left\{\vec{0}\right\}$.

7.3 Geradores e as Transformações Lineares

Considere n vetores $\vec{v}_1, \dots, \vec{v}_n \in V$ do espaço vetorial V. Podemos construir a transformação linear

$$\begin{array}{rrcl} T: & \mathbb{R}^n & \to & V \\ & (\alpha_1, \dots, \alpha_n) & \mapsto & \alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n \end{array}.$$

Essa é a transformação que faz com que cada n-upla $(\alpha_1, \dots, \alpha_n)$ corresponda a uma combinação linear de $\vec{v}_1, \dots, \vec{v}_n$.

Qual é a imagem de T?

A imagem de T são todas as possíveis combinações lineares dos vetores $\vec{v}_1, \dots, \vec{v}_n$. Ou seja, a imagem de T é o subespaço de V gerado pelos vetores $\vec{v}_1, \dots, \vec{v}_n$.

Quando é que esses vetores são geradores de V?

Por definição, são geradores quando o subespaço gerado por eles é todo o espaço V. Ou seja, quando $T(\mathbb{R}^n) = V$. Ou seja, quando T é sobrejetiva!!!

Se $V = \mathbb{R}^m$, então a matriz de T é

$$\left[T\right] = \left[\vec{v}_1 \quad \cdots \quad \vec{v}_n\right],$$

a matriz com m linhas e n colunas onde a j-ésima coluna corresponde ao vetor $\vec{v}_j$ *"escrito em pé"*.

Sendo assim, em termos de matrizes, os vetores $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^m$ são geradores de $\mathbb{R}^m$ quando a matriz $\left[\vec{v}_1 \cdots \vec{v}_n\right]$ representa uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ sobrejetiva.

Repare que a transformação T leva $\vec{e}_1, \dots, \vec{e}_n$, que são geradores de $\mathbb{R}^n$ nos vetores $\vec{v}_1, \dots, \vec{v}_n$, que geram a imagem de T.

Proposição 7.14. *Seja*

$$T : E \to F$$

uma transformação linear. E seja $S \subset E$ um conjunto gerador. Então, $T(S)$ é um gerador da imagem de T. Em particular, T é sobrejetiva exatamente quando $T(S)$ é um conjunto gerador de F.

Demonstração. Basta verificar que

$$\begin{aligned} L_{T(S)} &= \left\{\alpha_1\vec{w}_1 + \cdots + \alpha_k\vec{w}_k \;\middle|\; \alpha_1, \dots, \alpha_k \in \mathbb{R},\, \vec{w}_1, \dots, \vec{w}_k \in T(S)\right\} \\ &= \left\{\alpha_1 T\vec{v}_1 + \cdots + \alpha_k T\vec{v}_k \;\middle|\; \alpha_1, \dots, \alpha_k \in \mathbb{R},\, \vec{v}_1, \dots, \vec{v}_k \in S\right\} \\ &= \left\{T(\alpha_1\vec{v}_1 + \cdots + \alpha_k\vec{v}_k) \;\middle|\; \alpha_1, \dots, \alpha_k \in \mathbb{R},\, \vec{v}_1, \dots, \vec{v}_k \in S\right\} \\ &= T\left(L_S\right). \end{aligned}$$

Observação 7.15. Ao demonstrarmos a proposição 7.13, acabamos mostrando um pouco mais. Mostramos que a operação *"gerador"* comuta com a *transformação linear*: $T\,(\operatorname{gen}\{S\}) = \operatorname{gen}\{T(S)\}$.

7.4 Transformações Lineares e Independência Linear

Novamente, considere n vetores $\vec{v}_1, \ldots, \vec{v}_n \in V$ do espaço vetorial V, e a transformação linear

$$\begin{array}{rrcl} T: & \mathbb{R}^n & \to & V \\ & (\alpha_1, \ldots, \alpha_n) & \mapsto & \alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n \end{array}.$$

Essa é a transformação que faz com que cada n-upla $(\alpha_1, \ldots, \alpha_n)$ corresponda a uma combinação linear de $\vec{v}_1, \ldots, \vec{v}_n$.

Quando é que T é injetiva?

A transformação T é injetiva quando

$$T\vec{a} = T\vec{b} \Rightarrow \vec{a} = \vec{b}$$

Mas isso é o mesmo que dizer que

$$a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = b_1\vec{v}_1 + \cdots + b_n\vec{v}_n \Rightarrow a_1 = b_1, \ldots, a_n = b_n.$$

Ou seja, T é injetiva exatamente quando cada combinação linear dos vetores $\vec{v}_1, \ldots, \vec{v}_n$ for única: exatamente como no item 4 da proposição 8.15.

Resumindo, a transformação T é injetiva exatamente quando os vetores $\vec{v}_1, \ldots, \vec{v}_n$ são linearmente independentes. Repare que uma transformação injetiva $T : \mathbb{R}^n \to V \subset \mathbb{R}^m$ pode ser vista como uma transformação de $\mathbb{R}^n$ em $\mathbb{R}^m$, sem comprometer a injetividade.

Em termos de matrizes, os vetores $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^m$ são linearmente independente quando a matriz $\begin{bmatrix}\vec{v}_1 \cdots \vec{v}_n\end{bmatrix}$ representa uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ injetiva.

7.5 Transformações Lineares e Bases

Já sabemos que quando $T : V \to W$ é uma transformação linear, então T fica completamente determinada se conhecermos o valor de T em um conjunto gerador. Agora, se ao invés de já sabermos que $T : V \to W$ é uma transofrmação linear… se estivermos tentando **construir** uma transformação linear a partir de um conjunto gerador $\vec{v}_1, \cdots, \vec{v}_n \in V$. Será que basta escolher vetores $\vec{w}_1, \cdots, \vec{w}_n \in W$ para corresponderem a $T\vec{v}_1, \cdots, T\vec{v}_n V$? Ou seja, será que a transformação

$$\begin{array}{rrcl} T: & V & \to & W \\ & \alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n & \mapsto & \alpha_1\vec{w}_1 + \cdots + \alpha_n\vec{w}_n \end{array}$$

está bem definida?

Exemplo 7.16. Considere os vetores $\vec{v}_1 = (1,0)$, $\vec{v}_2 = (0,1)$ e $\vec{v}_3 = (1,1)$ de $\mathbb{R}^2$. Esses três vetores geram o espaço $\mathbb{R}^2$.

A transformação

$$\begin{array}{rrcl} T: & \mathbb{R}^2 & \to & \mathbb{R}^3 \\ & a\vec{v}_1 + b\vec{v}_2 + c\vec{v}_3 & \mapsto & (a,b,c) \end{array}$$

não está bem definida! De fato, o vetor $(1,1)$ pode ser escrito de duas maneiras:

$$\begin{aligned} (1,1) &= 1\vec{v}_1 + 1\vec{v}_2 + 0\vec{v}_3 \\ (1,1) &= 0\vec{v}_1 + 0\vec{v}_2 + 1\vec{v}_3. \end{aligned}$$

Dessa forma, assim como no exemplo 2.11, se calcularmos T usando a primeira combinação linear, temos

$$T(1,1) = T\left(1\vec{v}_1 + 1\vec{v}_2 + 0\vec{v}_3\right) = (1,1,0).$$

Mas se usarmos a segunda,

$$T(1,1) = T\left(0\vec{v}_1 + 0\vec{v}_2 + 1\vec{v}_3\right) = (0,0,1).$$

O problema que ocorre no exemplo 7.15, é que a maneira de representar os vetores de $\mathbb{R}^2$ como combinação linear de $\vec{v}_1$, $\vec{v}_2$ e $\vec{v}_3$ não é *única*. Se fosse única, não haveria ambiguidade.

Proposição 7.17. *Se $\vec{v}_1, \dots, \vec{v}_n$ formam uma base de V, e se $\vec{w}_1, \dots, \vec{w}_n$ são vetores quaisquer de W, então*

$$\begin{array}{rrcl} T: & V & \to & W \\ & \alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n & \mapsto & \alpha_1\vec{w}_1 + \cdots + \alpha_n\vec{w}_n \end{array}$$

é uma transformação linear bem definida.

Demonstração. O fato de $\vec{v}_1, \dots, \vec{v}_n$ serem linearmente independentes garante que não há dois valores distintos associados ao mesmo elemento de V. E o fato de serem geradores de V garante que T está definido para todo elemento de V.

xxxxxxxxxxxxxxx A FAZER: LEVA BASE EM BASE. xxxxxxxxxxxxxxxxx

7.6 Exercícios

Exercício 7.1. Determine o núcleo de T.

1. $$T: \begin{array}{rcl} \mathbb{R}^3 & \to & \mathbb{R} \\ (x,y,z) & \mapsto & x+y \end{array}$$

2. $$T: \begin{array}{rcl} \mathbb{R}^3 & \to & \mathbb{R} \\ (x,y,z) & \mapsto & 2x+3y+4z \end{array}$$

3. $$T: \begin{array}{rcl} \mathbb{R}^3 & \to & \mathbb{R}^2 \\ (x,y,z) & \mapsto & (2x+3y+4z, x+y) \end{array}$$

4. $$T: \begin{array}{rcl} \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x,y,z) & \mapsto & (2x+3y+4z, x+y, 2x+2y) \end{array}$$

Exercício 7.2. Encontre todos os vetores $\vec{v} \in \mathbb{R}^3$ tais que $T\vec{v} = (0,2,4)$, onde

$$T: \begin{array}{rcl} \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x,y,z) & \mapsto & (2x+3y+4z, x+y, 2x+2y) \end{array}.$$

Exercício 7.3. Encontre uma bijeção linear entre $\ker(T)$ e $\ker(S)$, onde

$$T: \begin{array}{rcl} \mathbb{R}^3 & \to & \mathbb{R} \\ (x,y,z) & \mapsto & x+y \end{array} \quad \text{e} \quad T: \begin{array}{rcl} \mathbb{R}^3 & \to & \mathbb{R} \\ (x,y,z) & \mapsto & 2x+3y+4z \end{array}.$$

Exercício 7.4. Classifique $T : \mathbb{R}^n \to \mathbb{R}^m$ como *bijetiva*, *apenas injetiva*, *apenas sobrejetiva* ou *nem injetiva nem sobrejetiva*.

1. $\left[T\right] = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$

2. $\left[T\right] = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

3. $\left[T\right] = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 0 \end{bmatrix}$

4. $\left[T\right] = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 7 & 7 \end{bmatrix}$

5. $\left[T\right] = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -1 & 2 & -3 \end{bmatrix}$

Exercício 7.5. Quantas transformações lineares $T : \mathbb{R}^3 \to \mathbb{R}^5$ existem satisfazendo as condições dos itens a seguir? Escreva a representação matricial de alguma T desse tipo, se existir.

1. $\begin{cases} T(1,1,1) = (1,2,3,4,5) \\ T(1,1,0) = (5,4,3,2,1) \end{cases}$

2. $\begin{cases} T(1,1,1) = (1,2,3,4,5) \\ T(1,1,0) = (5,4,3,2,1) \\ T(0,0,1) = (1,1,1,1,1) \end{cases}$

3. $\begin{cases} T(1,1,1) = (1,2,3,4,5) \\ T(1,1,0) = (5,4,3,2,1) \\ T(0,0,1) = (-4,-2,0,2,4) \end{cases}$

Exercício 7.6. Para cada item a seguir, escreva a representação matricial da transformação $T : \mathbb{R}^3 \to \mathbb{R}^3$ que satisfaz as condições especificadas.

1. $\begin{cases} T(1,1,1) = (1,0,0) \\ T(1,1,0) = (0,1,0) \\ T(1,0,0) = (0,0,1) \end{cases}$

2. $\begin{cases} T(1,2,3) = (1,0,0) \\ T(4,5,6) = (0,1,0) \\ T(1,1,2) = (0,0,1) \end{cases}$

Exercício 7.7. Para cada item a seguir, encontre a inversa das matrizes apresentadas.

1. $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 4 & 1 \\ 2 & 5 & 1 \\ 3 & 6 & 2 \end{bmatrix}$

Exercício 7.8. Seja $\vec{a}$, $\vec{b}$ e $\vec{c}$ uma base de $\mathbb{R}^3$. Então, a transformação

$$\begin{array}{rrcl} T: & \mathbb{R}^3 & \to & \mathbb{R}^3 \\ & \alpha\vec{a} + \beta\vec{b} + \gamma\vec{c} & \mapsto & (a,b,c) \end{array}$$

é a transformação linear que dado um vetor $\vec{v} \in \mathbb{R}^3$, nos diz como se escreve $\vec{v}$ como combinação linear de $\vec{a}$, $\vec{b}$ e $\vec{c}$. Por exemplo, se $T\vec{v} = (-1,5,7)$, então

$$\vec{v} = -\vec{a} + 5\vec{b} + 7\vec{c}.$$

Encontre a representação matricial de T para os casos a seguir, e depois escreva os vetores $(3,7,8)$, $(1,1,1)$, $(-2,-3,8)$ e $(9,8,7)$ como combinação linear de $\vec{a}$, $\vec{b}$ e $\vec{c}$.

1. $\begin{cases} \vec{a} = (1,1,1) \\ \vec{b} = (1,1,0) \\ \vec{c} = (1,0,0) \end{cases}$

2. $\begin{cases} \vec{a} = (1,2,3) \\ \vec{b} = (4,5,6) \\ \vec{c} = (1,1,2) \end{cases}$

Aula 8

Subespaços Vetoriais

8.1 Operações com Vetores em $\mathbb{R}^n$

Assim como em $\mathbb{R}^2$, a soma e o produto por escalar em $\mathbb{R}^n$ são feitos coordenada a coordenada:

$$\begin{aligned}(a_1, \dots, a_n) + (b_1, \dots, b_n) &= (a_1 + b_1, \dots, a_n + b_n)\\ \alpha(a_1, \dots, a_n) &= (\alpha a_1, \dots, \alpha a_n).\end{aligned}$$

Vejamos que, ao menos para o caso $\mathbb{R}^3$, as operações coordenada a tem um sentido geométrico.

Exemplo 8.1 (Soma em $\mathbb{R}^3$)**.** Para ilustrarmos a soma de vetores em $\mathbb{R}^3$ feita coordenada à coordenada, considere os vetores $\vec{v}$ e $\vec{w}$ da figura 8.1. Assim como no caso de $\mathbb{R}^2$, se observarmos apenas a *"altura"*, ou seja, a segunda coordenada, vemos que a *"altura"* de $\vec{v}+\vec{w}$ é igual à altura de $\vec{v}$ somada à altura de $\vec{w}$. Os vetores são *decompostos* em *"altura"* e *"sombra"*. Ou seja, na componente $\vec{e}_2$, e em uma componente no plano xz.

$$\begin{aligned}\vec{v} &= v_2\vec{e}_2 + \vec{v}_s\\ \vec{w} &= w_2\vec{e}_2 + \vec{w}_s.\end{aligned}$$

Ao calcular a soma $\vec{v}+\vec{w}$, somamos a *"altura"* $(v_2+w_2)\vec{e}_2$, e a *"sombra"*

$\vec{v}_s + \vec{w}_s$. Assim,

$$\vec{v} + \vec{w} = (v_2 + w_2)\vec{e}_2 + \vec{v}_s + \vec{w}_s.$$

No entanto, os vetores no plano xz também podem ser decompostos em suas coordenadas $\vec{e}_1$ e $\vec{e}_3$. E assim como em $\mathbb{R}^2$, a soma é feita coordenada a coordenada. (veja o exemplo 8.2). Portanto,

$$\begin{aligned}\vec{v} + \vec{w} &= (v_2 + w_2)\vec{e}_2 + \vec{v}_s + \vec{w}_s \\ &= (v_2 + w_2)\vec{e}_2 + (v_1 + w_1)\vec{e}_1 + (v_3 + w_3)\vec{e}_3.\end{aligned}$$

Ou seja, em $\mathbb{R}^3$, a soma é feita coordenada a coordenada.

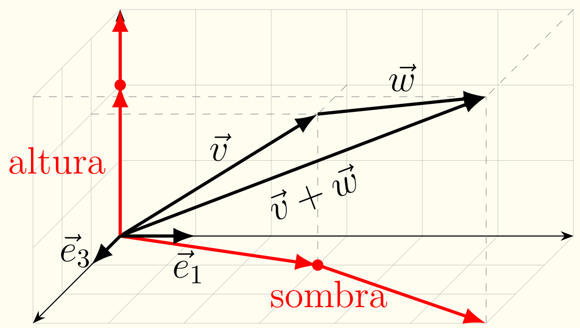


Figura 8.1: A *"altura"* (coordenada y) de $\vec{v} = (3, 2, 1)$ somada à de $\vec{w} = (3, 1, 2)$ corresponde à *"altura"* do vetor $\vec{v} + \vec{w}$. E a *"sombra"* de $\vec{v}$ projetada no plano xz, somada à *"sombra"* de $\vec{w}$ corresponde à *"sombra"* de $\vec{v} + \vec{w}$. Já vimos que, como a *"sobra"* está no plano xz, a soma pode ser feita coordenada a coordenada.

O mesmo raciocínio pode ser usado para o *produto por escalar*: $\alpha\vec{v}$. Mas é tão mais fácil ver que multiplicação coordenada a coordenada significa *"esticar"* / *"encolher"* / *"inverter o sentido"*, que nem precisa desenhar! ;-)

8.2 Subespaços

No exemplo 8.1, tratamos o plano xz em $\mathbb{R}^3$, como se fosse $\mathbb{R}^2$. Subconjuntos de $\mathbb{R}^n$ que com as mesmas operações de *soma* e *produto por escalar* se comportam como se fossem um $\mathbb{R}^m$ dentro de $\mathbb{R}^n$, são os *subespaços vetoriais* de $\mathbb{R}^n$. Dizemos que esses subconjuntos são *fechados com relação à soma* e *fechados com relação ao produto por escalar*. Antes de uma definição formal, vejamos alguns exemplos.

Exemplo 8.2 (Plano xz)**.** Seja $P \subset \mathbb{R}^3$, o *plano xz*. Ou seja,

$$P = \left\{ \alpha\vec{e}_1 + \beta\vec{e}_3 \,\middle|\, \alpha, \beta \in \mathbb{R} \right\}.$$

Em coordenadas,

$$P = \left\{ (\alpha, 0, \beta) \in \mathbb{R}^3 \,\middle|\, \alpha, \beta \in \mathbb{R} \right\}.$$

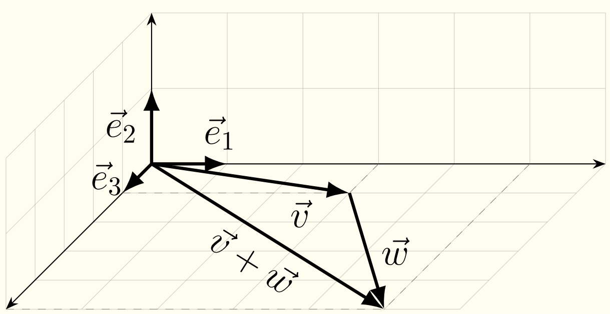


Figura 8.2: O *plano xz* é como se fosse um $\mathbb{R}^2$ dentro de $\mathbb{R}^3$. A soma de vetores do plano xz e o produto por escalar continuam confinados a este mesmo plano.

Restritas a P, as operações de soma e produto por escalar continuam bem definidas. Ou seja,

$$\begin{aligned} \vec{v}, \vec{w} \in P &\Rightarrow \vec{v} + \vec{w} \in P \\ \alpha \in \mathbb{R},\, \vec{v} \in P &\Rightarrow \alpha\vec{v} \in P. \end{aligned}$$

E temos uma bijeção natural entre $\mathbb{R}^2$ e P, que preserva a soma e o produto por escalar. Ou seja, preserva a estrutura de *espaço vetorial.*

$$\begin{array}{rrcl} T: & \mathbb{R}^2 & \to & P \\ & (\alpha, \beta) & \mapsto & \alpha\vec{e}_1 + \beta\vec{e}_3 \end{array}.$$

Todo subconjunto $P \subset \mathbb{R}^n$ não vazio que é *fechado por produto por escalar* deve necessariamente conter $\vec{0}$. De fato, se $\vec{v} \in P$, então

$$\vec{0} = 0\vec{v} \in P.$$

Para que um plano $P \subset \mathbb{R}^3$ seja um subespaço, é necessário e suficente que $\vec{0} \in P$. Ou seja, que P passe pela origem.

Exemplo 8.3 (Plano qualquer contendo $\vec{0}$)**.** Suponha que $P \subset \mathbb{R}^3$ seja um plano contendo $\vec{0}$. Sejam $\vec{v}, \vec{w} \in P$ tais que $\vec{0}$, $\vec{a}$ e $\vec{b}$ não são colineares. Então,

$$P = \left\{ \alpha\vec{a} + \beta\vec{b} \;\middle|\; \alpha, \beta \in \mathbb{R} \right\}$$

é o plano de $\mathbb{R}^3$ que contém os pontos $\vec{0}$, $\vec{a}$ e $\vec{b}$. O leitor é convidado a demonstrar que o conjunto P é *fechado com relação à soma* e *fechado com relação ao produto por escalar*. Na verdade, neste exemplo, em nenhum momento utilizamos o fato de o espaço em questão ser $\mathbb{R}^3$. Tudo funciona exatamente da mesma maneira se considerarmos $\mathbb{R}^n$.

Os conjuntos P dessa forma são os *subespaços de dimensão* 2. Para que estejamos de fato tratando de um plano, e não de uma reta, tivemos que assumir que os pontos $\vec{0}$, $\vec{a}$ e $\vec{b}$ não são colineares. Isso significa que os três pontos são distintos, e que não existe $\alpha \in \mathbb{R}$ tal que $\vec{a} = \alpha\vec{b}$.

Novamente, temos uma bijeção natural entre $\mathbb{R}^2$ e P, que preserva a soma e o produto por escalar. Ou seja, preserva a estrutura de *espaço vetorial*.

$$\begin{array}{rccl} T: & \mathbb{R}^2 & \rightarrow & P \\ & (\alpha, \beta) & \mapsto & \alpha\vec{a} + \beta\vec{b} \end{array}.$$

Um exemplo ainda mais simples é o subespaço unidimensional.

Exemplo 8.4 (Subespaço unidimensional)**.** Seja $\vec{v} \in \mathbb{R}^n$ um vetor não nulo. Então, o conjunto

$$R = \left\{ \alpha v \;\middle|\; \alpha \in \mathbb{R} \right\}$$

é um reta em $\mathbb{R}^n$ que passa em $\vec{0}$. O conjunto R é *fechado com relação à soma* e *fechado com relação ao produto por escalar*. Esses são os subespaços *unidimensionais*.

Assim como no caso do plano, temos uma bijeção natural entre $\mathbb{R}$ e R, que preserva a soma e o produto por escalar. Ou seja, preserva a estrutura de *espaço vetorial*.

$$\begin{array}{rccl} T: & \mathbb{R} & \rightarrow & R \\ & \alpha & \mapsto & \alpha\vec{v} \end{array}.$$

Definição 8.5 (Subespaço vetorial)**.** *Um subconjunto* ***não vazio*** $W \subset \mathbb{R}^n$ *é chamado de* subespaço vetorial *quando é fechado com relação à soma e fechado com relação ao produto por escalar. Ou seja,*

1. $\alpha \in \mathbb{R}$, $\vec{v} \in W \Rightarrow \alpha\vec{v} \in W$.

2. $\vec{v}, \vec{w} \in W \Rightarrow \vec{v} + \vec{w} \in W$*.*

Com uma definição precisa, agora podemos discutir um exemplo que ao mesmo tempo que é evidente, é também intrigante, já que sua utilidade parece, à primeira vista, muito duvidosa.

Exemplo 8.6 (Subespaços triviais)**.** O subconjunto de $\mathbb{R}^n$ dado por

$$X = \left\{\vec{0}\right\}$$

é um subespaço de $\mathbb{R}^n$ de dimensão 0. E o conjunto $Y = \mathbb{R}^n$ também é um subespaço de $\mathbb{R}^n$. Alguns autores dizem que X é o *subespaço trivial* de $\mathbb{R}^n$. Outros dizem que X e Y são os *subespaços triviais* de $\mathbb{R}^n$.

Observação 8.7. Daqui por diante, falaremos de várias propriedades que se aplicam tanto a $\mathbb{R}^n$ quanto a qualquer subespaço de $\mathbb{R}^n$. Apenas no **??** vamos definir *espaços vetorias* de um modo mais geral. Veja a definição 4.3.

Por enquanto, quando falarmos em *espaços vetoriais*, estaremos nos referindo apenas a $\mathbb{R}^n$ e a seus subespaços. Mas sempre que nos referirmos a *espaços vetoriais*, os resultados serão verdadeiros também para *espaços vetoriais* gerais, conforme a definição 4.3.

8.3 Bases e Independência Linear

Os exemplos da seção 8.2, em especial o exemplo 8.3, mostram de que forma podemos gerar um subespaço a partir de dois vetores, $\vec{v}$ e $\vec{w}$. O exemplo também mostra que condições precisamos impor a esses vetores para que o espaço gerado seja de fato um plano. Se não impusermos nenhuma coindição, o espaço gerado pode ser apenas uma reta, ou então, apenas o espaço trivial $\left\{\vec{0}\right\}$.

Subespaço Gerado

Dois vetores $\vec{v}, \vec{w} \in \mathbb{R}^3$ podem gerar um *plano*, quando $\vec{0}$, $\vec{v}$ e $\vec{w}$ não são colineares. Ou seja, quando $\vec{v}$ e $\vec{w}$ não são paralelos. Podem gerar o *espaço trivial* $\left\{\vec{0}\right\}$, quando $\vec{v} = \vec{w} = \vec{0}$. E podem gerar uma *reta*, quando não são paralelos e não são todos nulos. De qualquer forma, independentemente de que vetores $\vec{v}$ e $\vec{w}$ escolhemos, esses dois vetores *geram* um *subespaço* de $\mathbb{R}^n$. Pode não ser um plano, mas será um *subespaço*.

De modo um pouco mais geral, se

$$S = \{\vec{v}_1, \dots, \vec{v}_m\}$$

é um subconjunto finito de $\mathbb{R}^n$, o subconjunto

$$\operatorname{gen}\{S\} = \left\{\alpha_1\vec{v}_1 + \cdots + \alpha_m\vec{v}_m \;\middle|\; \alpha_1, \dots, \alpha_m \in \mathbb{R}\right\}$$

é o subespaço (de $\mathbb{R}^n$) gerado por S.

Podemos definir o *subespaço gerado* por um subconjunto $S \subset \mathbb{R}^n$ qualquer (mesmo que infinito). O leitor é convidado a fazer, por conta própria, definições alternativas. É sempre bom saber falar da mesma coisa de vários pontos de vista diferentes.

Definição 8.8 (Subespaço gerado)**.** *Seja $S \subset \mathbb{R}^n$ um subconjunto qualquer de $\mathbb{R}^n$. O* subespaço gerado *por S, denotado por* $\operatorname{gen}\{S\}$*, é o menor subespaço E de $\mathbb{R}^n$, tal que*

$$S \subset E.$$

Quando $S = \{\vec{v}_1, \dots, \vec{v}_n\}$, escrevemos $\operatorname{gen}\{\vec{v}_1, \dots, \vec{v}_n\}$.

Para que a definição 8.8 faça sentido é necessário dar um significado preciso para a expressão

> ... o menor subespaço $E \subset \mathbb{R}^n$ tal que $S \subset E$.

Queremos dizer três coisas:

1. Que E é um subespaço de $\mathbb{R}^n$.
2. Que $S \subset E$.
3. Que se F é um subespaço de $\mathbb{R}^n$ tal que $S \subset F$, então
$$E \subset F.$$

Mas será que tal coisa existe? Por exemplo,

> Qual é o menor número real maior do que α?

Queremos encontrar a satisfazendo três coisas:

1. $a \in \mathbb{R}$.
2. $\alpha < a$.
3. Se $b \in \mathbb{R}$ é maior que α, então
$$a \leq b.$$

Tal número **não** existe! Qual é o menor número real maior que 0? 0.1? 0.01? 0.001? 0.00000001? No entanto, para o subespaço da definição 8.8, a resposta é: **sim, existe**. Antes de mostrarmos que existe, vejamos um exemplo simples.

Exemplo 8.9. Seja $\vec{v} \in \mathbb{R}^3$ um vetor não nulo. Então,

$$\operatorname{gen}\{\vec{v}\} = \left\{\alpha\vec{v} \;\middle|\; \alpha \in \mathbb{R}\right\}.$$

De fato, se chamarmos o subespaço da direita de W, temos que $\vec{v} \in W$, e todo subespaço de $\mathbb{R}^3$ que contém $\vec{v}$ tem que conter todos os múltiplos de $\vec{v}$, e portanto, tem que conter W. Ou seja, W é o *menor* subespaço de $\mathbb{R}^3$ que contém $\vec{v}$. Dessa forma, $\operatorname{gen}\{\vec{v}\} = W$.

Uma outra forma de caracterizar o *subespaço gerado* é através de *combinações lineares*. Vamos recordar a definição 7.4.

Seja $S \subset \mathbb{R}^n$ um subconjunto qualquer de $\mathbb{R}^n$. Uma combinação linear de elementos de S é um vetor $\vec{v} \in \mathbb{R}^n$ da forma

$$\vec{v} = \alpha_1\vec{v}_1 + \cdots + \alpha_k\vec{v}_k,$$

onde $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ e $\vec{v}_1, \ldots, \vec{v}_k \in S$.

Vamos, por agora, denotar por

$$L_S = \left\{\alpha_1\vec{v}_1 + \cdots + \alpha_k\vec{v}_k \;\middle|\; \alpha_1, \ldots, \alpha_k \in \mathbb{R},\, \vec{v}_1, \ldots, \vec{v}_k \in S\right\}$$

o conjunto de todas as combinações lineares de elementos de S. Perceba que dizer que $S \subset V$ é um *subespaço* de V é o mesmo que dizer que $L_S = S$. A demonstração da proposição a seguir é muito semelhante ao exemplo 8.9.

Proposição 8.10. *Seja V um espaço vetorial. Se $S \subset V$ é não vazio, então* gen $\{S\}$ *é o conjunto formado por* ***todas*** *as* combinações lineares *de elementos de S.*

Demonstração. Se $W \subset V$ é um subespaço qualquer com $S \subset W$, então

$$S \subset L_S \subset L_W = W.$$

Ou seja, L_S é o **menor** subespaço de V que contém S. Assim,

$$\operatorname{gen}\{S\} = L_S.$$

Exemplo 8.11. O conjunto

$$S = \left\{ \left(\frac{1}{n+1}, 1, 0 \right) \in \mathbb{R}^3 \;\middle|\; n \in \mathbb{N} \right\}$$

é um conjunto infinito que gera o plano xy.

Independência Linear

No exemplo 8.3, precisamos impor condições para que gen $\{\vec{v}, \vec{w}\}$, o espaço gerado pelos vetores $\vec{v}$ e $\vec{w}$ fosse de fato um plano, e não uma reta, por exemplo. Mas pensando bem… o que é *um plano que passa na origem*? Não seria o mesmo que um subespaço bidimensional? Mas… *bidimensional* significa o quê? Que tem *dimensão* 2? E o que significa isso tudo? Uma reta passando pela origem, é um subespaço de dimensão 1. O subespaço $\left\{\vec{0}\right\}$ tem dimensão 0.

Podemos dar um significado preciso aos conceitos de plano, reta, *dimensão*? E se estivermos falando de $\mathbb{R}^n$? Podemos generalizar esses conceitos e falar dos subespaços de *dimensão* k?

Definição 8.12 (Independência Linear)**.** *Dizemos que um vetor* $\vec{v} \in \mathbb{R}^n$ *é* (linearmente) independente *de* $A \subset \mathbb{R}^n$*, quando* $\vec{v} \notin \operatorname{gen}\{A\}$*.*

Dizemos que A *é* linearmente independente *(ou simplesmente* LI*) quando todo* $\vec{v} \in A$ *é idependente de* $A \setminus \{\vec{v}\}$*. Quando* A *não é* linearmente independente*, dizemos que é* linearmente dependente*, ou simplesmente,* LD*.*

Diremos também que uma coleção indexada *de vetores* $\vec{v}_1, \ldots, \vec{v}_k$ *é* linearmente independente *quando forem todos distintos e o conjunto* $\{\vec{v}_1, \ldots, \vec{v}_k\}$ *for* LI*.*

Dizer que vetores são linearmente independentes, é dizer que um não pode ser escrito como combinação dos outros.

Exemplo 8.13 (dois vetores que não geram um plano)**.** Quando dois vetores distintos não nulos, $\vec{v}$ e $\vec{w}$ não geram um plano, é porque um é *linearmente dependente do outro*. Em um certo sentido, o conjunto $S = \{\vec{v}, \vec{w}\}$ tem elementos *"desnecessários"*. O espaço gerado por S é o espaço gerado pelo vetor $\vec{v}$.

Quando falamos que $\vec{0}$ e mais dois pontos $\vec{v}$ e $\vec{w}$, quando não são três pontos colineares, determinam um plano, estamos justamente falando da independência linear de $\vec{v}$ e $\vec{w}$. Quando tem *"desperdício"* no conjunto gerador, o espaço gerado acaba menor do que o *"esperado"*. É por isso que o estudante prudente não diz que dois vetores geram um plano que passa pela origem! Ele diz:

> Dois vetores **linearmente independentes** geram um plano que passa pela origem.

Observação 8.14. Se $\vec{v}_1 = \vec{v}_2 = \vec{e}_1$, então diremos que $\vec{v}_1$ e $\vec{v}_2$ são *LD*. No entanto, o conjunto $\{\vec{v}_1, \vec{v}_2\}$ **NÃO** é *LD*. Por quê?

Existem outras maneiras de se pensar em *independência linear*.

Proposição 8.15 (Caracterizações de Independência Linear)**.** *Seja $A \subset \mathbb{R}^n$. Então, são equivalentes:*

1. *A é* linearmente independente.
2. *O conjunto A é um gerador minimal para* $\operatorname{gen}\{A\}$*. Ou seja, se $B \subsetneq A$, então*
$$\operatorname{gen}\{B\} \subsetneq \operatorname{gen}\{A\}.$$

3. *Sempre que* $\vec{v}_1, \dots, \vec{v}_k \in A$ *são distintos e*

$$\alpha_1\vec{v}_1 + \cdots + \alpha_k\vec{v}_k = \vec{0},$$

então $\alpha_1 = \cdots = \alpha_k = 0$.

4. *Toda combinação linear de elementos de A é* única, *no sentido de que se* $\vec{v}_1, \dots, \vec{v}_k \in A$ *são distintos, então*

$$\alpha_1\vec{v}_1 + \cdots + \alpha_k\vec{v}_k = \beta_1\vec{v}_1 + \cdots + \beta_k\vec{v}_k$$

implica que $\alpha_1 = \beta_1, \dots, \alpha_k = \beta_k$.

5. *Toda combinação linear de elementos de A é* única, *no sentido de que se*

$$\vec{v} = \alpha_1\vec{v}_1 + \cdots + \alpha_k\vec{v}_k,$$
$$\vec{v} = \beta_1\vec{w}_1 + \cdots + \beta_m\vec{w}_m,$$

com todos os $\vec{v}_1, \dots, \vec{v}_k \in A$ *distintos entre si, todos os* $\vec{w}_1, \dots, \vec{w}_m \in A$ *distintos entre si, e com todos os* $\alpha_1, \dots, \alpha_k$ *e* $\beta_1, \dots, \beta_k$ *não nulos, então* $k = m$, *e os índices de* β_j *e* $\vec{w}_j$ *podem ser rearranjados de modo que*

$$\begin{aligned} \vec{v}_1 = \vec{w}_1 \quad &\text{e} \quad \alpha_1 = \beta_1 \\ &\vdots \\ \vec{v}_k = \vec{w}_k \quad &\text{e} \quad \alpha_k = \beta_k. \end{aligned}$$

Demonstração. TODO.

O item 2 da proposição 8.15 mostra que $A \subset \mathbb{R}^n$ é linearmente independente quando não possui nenhum elemento *"sobrando"*, ao contrário do exemplo a seguir.

Exemplo 8.16. Seja

$$A = \{(1, 1, 0), (1, 0, 1), (0, 1, -1)\}.$$

Então A é *linearmente dependente*. De fato,

$$(1, 1, 0) = (1, 0, 1) + (0, 1, -1).$$

O vetor $(4, 1, 3)$, por exemplo, pode ser escrito como infinitas combinações lineares diferentes. Algumas delas, são

$$\begin{aligned}
(4, 1, 3) &= (1, 1, 0) + 3(1, 0, 1) \\
(4, 1, 3) &= 4(1, 1, 0) - 3(0, 1, -1) \\
(4, 1, 3) &= 2(1, 1, 0) + 2(1, 0, 1) - (0, 1, -1).
\end{aligned}$$

Exemplo 8.17. Continuando o exemplo 8.11, o conjunto

$$A = \left\{ \left(\frac{1}{n+1}, 1, 0 \right) \in \mathbb{R}^3 \,\middle|\, n \in \mathbb{N} \right\}$$

possui vários subconjuntos de dois elementos que são geradores de do plano xy. De fato, qualquer subconjunto de A com dois elementos é um gerador para o mesmo subespaço. Por exemplo,

$$B = \left\{ (1, 1, 0), \left(\frac{1}{2}, 1, 0 \right) \right\}$$

gera o plano xy. O leitor é convidado a mostrar que $\vec{e}_1, \vec{e}_2 \in \operatorname{gen}\{B\}$.

O exemplo a seguir, é mais uma proposição do que um exemplo.

Exemplo 8.18. Os subconjuntos de $\mathbb{R}^3$ que são linearmente independentes tem cardinalidade 0, 1, 2 ou 3.

De fato, se $A \subset \mathbb{R}^3$ tiver 4 elementos distintos $\vec{a} = (a_1, a_2, a_3)$, $\vec{b} = (b_1, b_2, b_3)$, $\vec{c} = (c_1, c_2, c_3)$ e $\vec{d} = (d_1, d_2, d_3)$, então o sistema de equações

$$\begin{cases} a_1x + b_1y + c_1z + d_1w &= 0 \\ a_2x + b_2y + c_2z + d_2w &= 0 \\ a_3x + b_3y + c_3z + d_3w &= 0 \end{cases}$$

tem 3 equações e 4 incógnitas, e portanto, tem infinitas soluções. Em particular, existem $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$, nem todos nulos, tais que

$$\alpha_1\vec{a} + \alpha_2\vec{b} + \alpha_3\vec{c} + \alpha_4\vec{d} = \vec{0}.$$

Agora, como é que você sabe que um sistema homogêneo com 3 equações e 4 incógnitas tem infinitas soluções... isso é com você!!! Como é que você sabe disso? :-)

Já vimos no **??** que uma transformação linear sobrejetiva leva conjunto gerador em conjunto gerador.

Lema 8.19. *Seja $T : E \to F$ uma transformação linear injetiva. Se $S \subset E$ é linearmente independente, então $T(S) \subset F$ também é.*

Demonstração. Suponha que

$$\alpha_1 T(\vec{v}_1) + \cdots + \alpha_k T(\vec{v}_k) = \vec{0},$$

com $\vec{v}_1, \ldots, \vec{v}_k \in S$. Para concluir que $T(S)$ é linearmente independente, basta mostrar que $\alpha_1 = \cdots = \alpha_k = 0$.

Note que,

$$\begin{aligned} T\left(\alpha_1\vec{v}_1 + \cdots + \alpha_k\vec{v}_k\right) &= \alpha_1 T(\vec{v}_1) + \cdots + \alpha_k T(\vec{v}_k) \\ &= \vec{0}. \end{aligned}$$

Mas pela injetividade de T, isso implica que

$$\alpha_1 T(\vec{v}_1) + \cdots + \alpha_k T(\vec{v}_k) = \vec{0}.$$

Pela independência linear de S, isso implica que $\alpha_1 = \cdots = \alpha_k = 0$.

Bases e Dimensão

O falarmos de vetores em $\mathbb{R}^2$, podíamos descrevê-los como combinação linear dos vetores $\vec{e}_1$ e $\vec{e}_2$. Todo vetor $\vec{v} \in \mathbb{R}^2$ é da forma

$$\vec{v} = \alpha\vec{e}_1 + \beta\vec{e}_2.$$

Além disso, essa representação é única. Dizer que todo vetor é combinação linear de $\vec{e}_1$ e $\vec{e}_2$, é o mesmo que dizer que esses dois vetores formam um conjunto gerador do espaço $\mathbb{R}^2$. E dizer que a representação é única, é o mesmo que dizer que são linearmente independentes. O mesmo vale para

$$\vec{e}_1, \ldots, \vec{e}_n \in \mathbb{R}^n.$$

São geradores e linearmente independentes. Em outras palavras, formam uma *base* para $\mathbb{R}^n$.

Definição 8.20. *Uma* base *de um espaço vetorial é um conjunto de vetores linearmente independentes que gera todo o espaço.*

Exemplo 8.21. Em $\mathbb{R}^2$, para especificarmos uma *transformação linear*

$T : \mathbb{R}^2 \to \mathbb{R}^2$, bastava especificar $\vec{v}_1 = T\vec{e}_1$ e $\vec{v}_2 = T\vec{e}_2$. Nesse caso,

$$T : \begin{array}{rcl} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ \alpha\vec{e}_1 + \beta\vec{e}_2 & \mapsto & \alpha\vec{v}_1 + \beta\vec{v}_2 \end{array}.$$

Na verdade, poderíamos substituir $\vec{e}_1$ e $\vec{e}_2$ por uma base qualquer. De fato, se soubermos que

$$\vec{w}_1 = T\vec{b}_1 \quad \text{e} \quad \vec{w}_2 = T\vec{b}_2,$$

onde $\vec{b}_1$ e $\vec{b}_2$ formam uma base para $\mathbb{R}^2$, então sabemos calcular $T\vec{u}$ para qualquer $\vec{u} \in \mathbb{R}^2$, pois $\vec{u}$ pode ser escrito na forma $\alpha\vec{b}_1 + \beta\vec{b}_2$. Nesse caso,

$$T : \begin{array}{rcl} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ \alpha\vec{b}_1 + \beta\vec{b}_2 & \mapsto & \alpha\vec{w}_1 + \beta\vec{w}_2 \end{array}.$$

Exemplo 8.22. Considere os vetores $\vec{e}_1$, $\vec{e}_2$ e $\vec{b} = (1, 1)$ em $\mathbb{R}^2$. É possível definir uma *transformação linear* $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que

$$T\vec{e}_1 = \vec{0}, \quad T\vec{e}_2 = \vec{e}_1 \quad \text{e} \quad T\vec{b} = \vec{e}_2?$$

A resposta é: **Não!** Mas por que motivo isso não é possível?

E se fossem só os vetores $\vec{e}_1$ e $\vec{e}_2$, sem a condição no vetor $\vec{b}$? Ou seja, será que existe uma *transformação linear* T tal que

$$T\vec{e}_1 = \vec{0} \quad \text{e} \quad T\vec{e}_2 = \vec{e}_1?$$

E, pra complicar um pouco, e se fossem só os vetores $\vec{e}_1$ e $\vec{b}$, sem impor uma condição no vetor $\vec{e}_2$? Seria possível? Será que conseguimos definir, sem nenhuma inconsistência, uma *transformação linear* T tal que

$$T\vec{e}_1 = \vec{0} \quad \text{e} \quad T\vec{b} = \vec{e}_2?$$

É fácil explicar os casos em que sabemos que existe um problema na definição de T. Basta apontar a contradição e dizer: **impossível!** Por exemplo, não é possível fazer

$$T\vec{e}_1 = \vec{0}, \quad T\vec{e}_2 = \vec{e}_1 \quad \text{e} \quad T\vec{b} = \vec{e}_2.$$

Pois ao mesmo tempo que sabemos que

$$T\vec{b} = \vec{e}_2,$$

também é verdade que

$$T\vec{b} = T(\vec{e}_1 + \vec{e}_2) = T\vec{e}_1 + T\vec{e}_2 = \vec{e}_1.$$

Se encontrarmos uma contradição como essa, sabemos que T **não está bem definida**. Mas como podemos **garantir que não há uma contradição**? Como podemos gararntir que a transformação T está bem definida? Sabendo o valor de T em $\vec{e}_1$, $\vec{e}_2$ e $\vec{b}$, concluímos que T deve satisfazer:

$$\begin{array}{rccc} T: & \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 \\ & \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \alpha_3\vec{b} & \mapsto & \alpha_1 T\vec{e}_1 + \alpha_2 T\vec{e}_2 + \alpha_3 T\vec{b} \end{array}.$$

Para que dessa forma T esteja bem definida, é necessário e suficiente que:

1. Todo vetor $\vec{v} \in \mathbb{R}^2$ possa ser expresso da forma $\alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \alpha_3\vec{b}$, para que T esteja definido em todo domínio.

2. Não haja ambiguidade na definição do valor de $T\vec{v}$ para nenhum $\vec{v} \in \mathbb{R}^2$. Ou seja, mesmo que $\vec{v}$ possa ser expresso de duas ma-

neiras diferentes,

$$\vec{v} = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \alpha_3\vec{b}$$
$$\vec{v} = \beta_1\vec{e}_1 + \beta_2\vec{e}_2 + \beta_3\vec{b},$$

é importante que o valor de $T\vec{v}$ seja o mesmo se calculado utilizando qualquer uma das duas maneiras distintas.

A condição do item 1 equivale a dizer que os vetores $\vec{e}_1$, $\vec{e}_2$ e $\vec{b}$ *geram* o domínio de T.

Agora, segundo a proposição 8.15, uma maneira simples de garantir que o item 2 seja satisfeito, é fazendo com que $\vec{e}_1$, $\vec{e}_2$ e $\vec{b}$ sejam *linearmente independentes*.

Assim, concluímos que para construir uma *transformação linear* podemos simplesmente escolher o valor de T para cada elemento de uma base preestabelecida. Como a base é *geradora*, estaremos definindo a transformação para todos os elementos do domínio. E como é *linearmente independente*, não existe ambiguidade na definição.

8.4 Exercícios

Exercício 8.1. Demonstre que S é ou não subespaço de $\mathbb{R}^3$.

1. $S = \left\{(x, y, z) \in \mathbb{R}^3 \;\middle|\; 2x + 3y + 4z = 0\right\}$
2. $S = \left\{(x, y, z) \in \mathbb{R}^3 \;\middle|\; 2x + 3y + 4z = 4\right\}$
3. $S = \left\{(2x + y, 3x + y, 4x + y) \;\middle|\; (x, y) \in \mathbb{R}^2\right\}$
4. $S = \left\{(2x + 1, 3x + 1, 4x + 1) \;\middle|\; x \in \mathbb{R}\right\}$

Exercício 8.2. Mostre que $\vec{v}_1, \ldots, \vec{v}_m \in \mathbb{R}^m$ são geradores de $\mathbb{R}^n$.

1. $\vec{v}_1 = (1, 1)$ e $\vec{v}_2 = (-1, 1)$
2. $\vec{v}_1 = (1, 2)$, $\vec{v}_2 = (2, 4)$ e $\vec{v}_3 = (3, 4)$
3. $\vec{v}_1 = (1, 0, 0)$, $\vec{v}_2 = (1, 1, 0)$ e $\vec{v}_3 = (1, 1, 1)$
4.
$$\begin{aligned} \vec{v}_1 &= \vec{e}_1 \\ \vec{v}_2 &= \vec{e}_1 + \vec{e}_2 \\ &\vdots \\ \vec{v}_n &= \vec{e}_1 + \cdots + \vec{e}_n \end{aligned}$$

Exercício 8.3. Determine se os vetores $\vec{v}_1, \ldots, \vec{v}_m \in \mathbb{R}^m$ são linearmente independentes ou não.

1. $\vec{v}_1 = (1, 0, 0)$, $\vec{v}_2 = (1, 1, 0)$ e $\vec{v}_3 = (1, 1, 1)$
2. $\vec{v}_1 = (1, 2, 3)$, $\vec{v}_2 = (4, 5, 6)$ e $\vec{v}_3 = (7, 8, 9)$
3. $\vec{v}_1 = (1, 1)$ e $\vec{v}_2 = (-1, 1)$
4. $\vec{v}_1 = (1, 1)$, $\vec{v}_2 = (-1, 1)$ e $\vec{v}_3 = (1, 2)$
5. $\vec{v}_1 = (0, 0, 0)$ e $\vec{v}_3 = (1, 0, 0)$
6. $\vec{v}_1 = (0, 0, 0)$

Exercício 8.4. Dentre os vetores a seguir, encontre ao menos duas bases para o espaço que eles geram.

1. $\vec{v}_1 = (1, 2)$, $\vec{v}_2 = (2, 4)$ e $\vec{v}_3 = (5, 6)$
2. $\vec{v}_1 = (1, 2)$, $\vec{v}_2 = (2, 4)$ e $\vec{v}_3 = (3, 6)$
3. $\vec{v}_j = (j, j + 1, j + 2)$, $(j = 1, 2, 3, 4, 5, 6)$
4. $\vec{v}_j = (j, j + 1, j + 2)$, $(j = 7, 9, 11, 23, 15, 61)$

Exercício 8.5. Dado o subespaço $S \subset \mathbb{R}^n$, encontre uma base $\vec{b}_1, \ldots, \vec{b}_n$ para $\mathbb{R}^n$ e $m \in \mathbb{N}$, tais que $\vec{b}_1, \ldots, \vec{b}_m$ seja uma base para S.

1. $S = \left\{(x, y, z) \in \mathbb{R}^3 \,\middle|\, 3x - 7y + 5z = 0\right\}$
2. $S = \left\{(x, y, z, w) \in \mathbb{R}^4 \,\middle|\, 3x - 4y + z - w = 0\right\}$
3. $S = \left\{(t, t, t) \in \mathbb{R}^3 \,\middle|\, t \in \mathbb{R}\right\}$
4. $S = \left\{(t, 2t, s, 2s) \in \mathbb{R}^4 \,\middle|\, (s, t) \in \mathbb{R}^2\right\}$

Parte III

Conceitos Motivados Geometricamente

Aula 9

Produto Interno

9.1 Definição

Comecemos com uma definição chata!

Definição 9.1 (Produto interno (canônico))**.** *Dados dois vetores* $\vec{a}, \vec{b} \in \mathbb{R}^n$*, o* produto interno (canônico) *de* $\vec{a}$ *por* $\vec{b}$ *é o número*

$$\vec{a} \cdot \vec{b} = a_1 b_1 + \cdots + a_n b_n.$$

Observação 9.2. Dizemos *"canônico"* porque existem outros tipos de *produto interno.* Alguns autores chamam o *produto interno canônico* de *produto euclideano.* Outros chamam de *produto escalar.* Neste livro trataremos apenas do *produto interno canônico*, por isso, quando dissermos *produto interno* estamos nos referindo ao produto definido na **??**.

Notação 9.3. *Também é comum que o produto interno seja denotado usando a notação alternativa* $\left\langle \vec{a}, \vec{b} \right\rangle$*. Principalmente quando queremos* *enxergar o produto interno como uma aplicação*

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

que toma os vetores $\vec{a}, \vec{b} \in \mathbb{R}^n$ *e leva em* $\vec{a} \cdot \vec{b}$*.*

Bom, já definimos o *produto interno*. Mas... **e daí**??? De onde veio essa definição? Pra que serve? O que significa?

9.2 Projeção Ortogonal

Em $\mathbb{R}^2$, vamos fixar um vetor $\vec{w}$. Podemos pensar em projetar um outro vetor qualquer $\vec{v}$, na direção e no sentido de $\vec{w}$. Veja a figura 9.1. Estamos falando de uma *"projeção com sinal"*

$$P_{\vec{w}} : \mathbb{R}^2 \to \mathbb{R},$$

que toma um vetor $\vec{v}$, e nos diz qual é o *"tamanho com sinal"* da *projeção ortogonal* de $\vec{v}$ em $\vec{w}$.

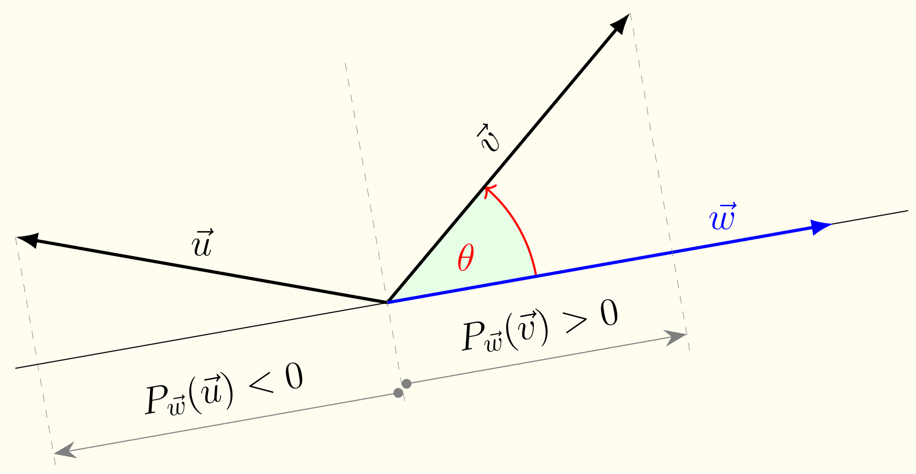


Figura 9.1: $P_{\vec{w}}$ é uma "projeção com sinal". A *direção* do vetor $\vec{w}$, nos dá o *"tamanho da projeção"*. E o *sentido* nos dá o sinal.

O vetor $\vec{w}$ **não pode ser nulo**. No mais, o tamanho de $\vec{w}$ não tem nenhuma influência em $P_{\vec{w}}$. Só o que interessa em $\vec{w}$ é sua *direção* e o seu *sentido*. Por isso, é suficiente tratar o caso em que $\vec{w}$ é um vetor unitário. Se conhecermos o caso unitário, quando formos tratar do caso geral, basta observar que

$$P_{\vec{w}} = P_{\frac{\vec{w}}{\|\vec{w}\|}}.$$

Usando trigonometria e a figura 9.1, é fácil ver que

$$P_{\vec{w}}(\vec{v}) = \|\vec{v}\| \cos(\theta).$$

Mas nós não queremos usar trigonometria. Nós vamos descobrir como usar as coordenadas dos vetores $\vec{v}$ e $\vec{w}$ pra encontrar $P_{\vec{w}}(\vec{v})$.

A projeção é *linear*

A transformação $P_{\vec{w}} : \mathbb{R}^2 \to \mathbb{R}$ tem algumas propriedades interessantes. A primeira delas diz respeito ao produto por escalar.

$$P_{\vec{w}}(\alpha\vec{v}) = \alpha P_{\vec{w}}(\vec{v}). \tag{9.2.1}$$

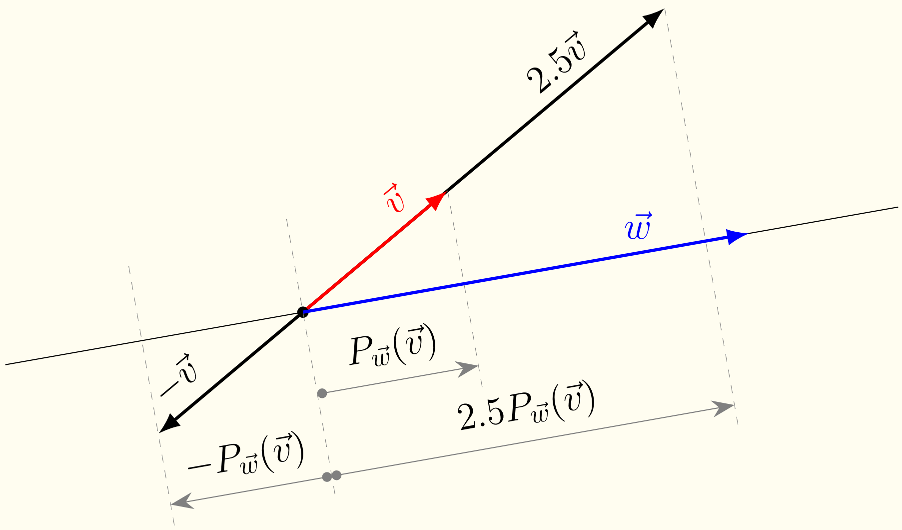


Figura 9.2: Quando multiplicamos o vetor $\vec{v}$ por um número real α, então $P_{\vec{w}}(\alpha\vec{v}) = \alpha P_{\vec{w}}(\vec{v})$. Quando definimos uma *projeção* ***com sinal***, foi, também, para que essa propriedade valha quando $\alpha < 0$.

Foi importante termos definido uma *projeção* ***com sinal***. Não fosse pelo sinal, teríamos que nos preocupar com o valor absoluto de α. Além do produto por escalar, tem também a projeção da soma de dois vetores: $P_{\vec{w}}(\vec{u} + \vec{v})$.

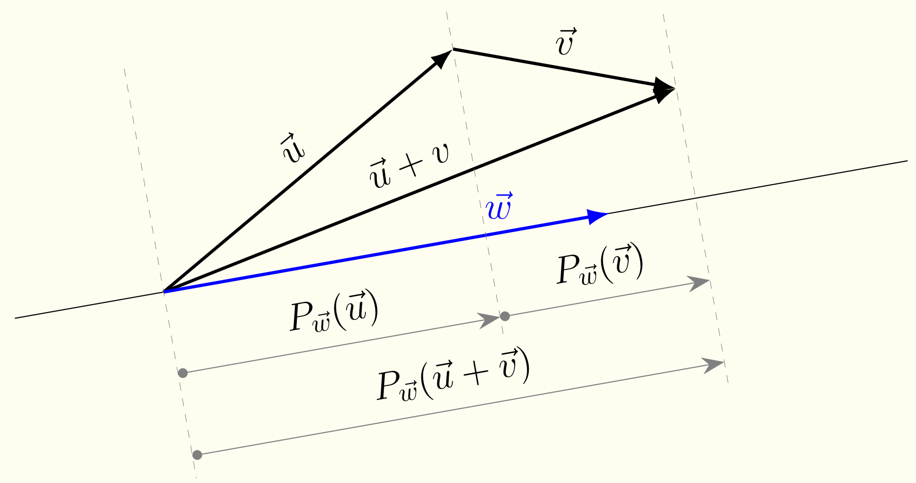


Figura 9.3: É fácil ver que $P_{\vec{w}}(\vec{u}+\vec{v}) = P_{\vec{w}}(\vec{u}) + P_{\vec{w}}(\vec{v})$. Ao menos quando $\vec{u}$ e $\vec{v}$ se projetam em $\vec{w}$ no mesmo sentido...

E graças ao fato de termos definido uma *projeção* ***com sinal***, a fórmula da soma vale mesmo que os vetores $\vec{u}$ e $\vec{v}$ tenham sentidos opostos quando projetados em $\vec{w}$.

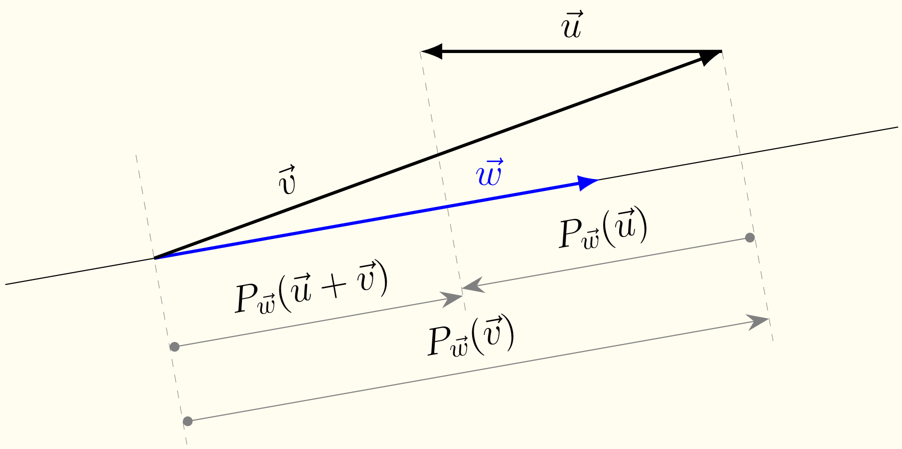


Figura 9.4: Mesmo quando $\vec{u}$ e $\vec{v}$ se projetam em $\vec{w}$ em sentidos opostos, graças ao fato de $P_{\vec{w}}$ ser uma *projeção* ***com sinal***, a fórmula $P_{\vec{w}}(\vec{u}+\vec{v}) = P_{\vec{w}}(\vec{u}) + P_{\vec{w}}(\vec{v})$ continua valendo.

Assim, quaisquer que sejam $\vec{u}, \vec{v} \in \mathbb{R}^2$,

$$P_{\vec{w}}(\vec{u}+\vec{v}) = P_{\vec{w}}(\vec{u}) + P_{\vec{w}}(\vec{v}). \tag{9.2.2}$$

As equações (9.2.1) e (9.2.2) mostram que $P_{\vec{w}}$ é uma *transformação linear*.

O produto interno em $\mathbb{R}^2$

Já sabemos que $P_{\vec{w}}$ é *linear*. Vamos utilizar esta linearidade para fazer uma parte das contas.

$$\begin{aligned} P_{\vec{w}}(\vec{v}) &= P_{\vec{w}}(v_1\vec{e}_1 + v_2\vec{e}_2) \\ &= v_1 P_{\vec{w}}(\vec{e}_1) + v_2 P_{\vec{w}}(\vec{e}_2). \end{aligned} \tag{9.2.3}$$

Resta agora, determinar $P_{\vec{w}}(\vec{e}_j)$.

Perceba que $P_{\vec{e}_j}(\vec{w}) = w_j$. Existem várias formas de verificar que

$$P_{\vec{w}}(\vec{e}_j) = \frac{w_j}{\|\vec{w}\|}.$$

Uma maneira, é observar que quando $\|\vec{v}\| = \|\vec{w}\| \neq 0$, então

$$P_{\vec{w}}(\vec{v}) = P_{\vec{v}}(\vec{w}).$$

Veja a figura 9.5.

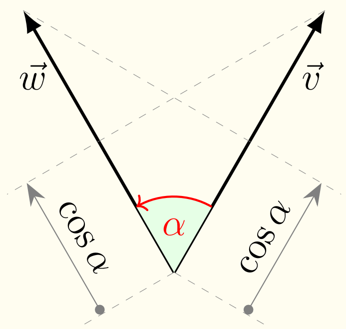


Figura 9.5: Quando $\|\vec{v}\| = \|\vec{w}\| = 1$, a projeção de $\vec{v}$ na direção (e sentido) de $\vec{w}$ é igual à projeção de $\vec{w}$ na direção (e sentido) de $\vec{v}$. Mais do que isso, $P_{\vec{w}}(\vec{v}) = \cos\alpha = P_{\vec{w}}(\vec{v})$.

Assim,

$$\begin{aligned} P_{\vec{w}}(\vec{e}_j) &= P_{\frac{\vec{w}}{\|\vec{w}\|}}(\vec{e}_j) \\ &= P_{\vec{e}_j}\left(\frac{\vec{w}}{\|\vec{w}\|}\right) \\ &= \frac{1}{\|\vec{w}\|} P_{\vec{e}_j}(\vec{w}) \\ &= \frac{w_j}{\|\vec{w}\|}. \end{aligned}$$

E portanto, continuando a equação (9.2.3),

$$P_{\vec{w}}(\vec{v}) = \frac{v_1 w_1 + v_2 w_2}{\|\vec{w}\|}. \tag{9.2.4}$$

Em especial, quando $\|\vec{w}\| = 1$,

$$P_{\vec{w}}(\vec{v}) = v_1 w_1 + v_2 w_2.$$

A quantidade que aparece no numerador da equação (9.2.4) é justamente o *produto interno* da definição 9.1. Assim,

$$P_{\vec{w}}(\vec{v}) = \frac{\vec{v} \cdot \vec{w}}{\|\vec{w}\|}.$$

E em particular, quando $\|\vec{w}\| = 1$, $P_{\vec{w}}(\vec{v}) = \vec{v} \cdot \vec{w}$.

Produto interno em $\mathbb{R}^n$

xxx

$$\tag{9.2.5}$$

Ortogonalidade

Geometricamente, dois vetores não nulos $\vec{a}, \vec{b} \in \mathbb{R}^n$ são ortogonais quando

$$P_{\vec{a}}(\vec{b}) = 0.$$

Pela equação (9.2.5), isso equivale a

$$\vec{a} \cdot \vec{b} = 0. \tag{9.2.6}$$

Essa última equação tem a vantagem de poder ser usada mesmo quando $\vec{a}$ e $\vec{b}$ são nulos.

Por isso, vamos dizer que dois vetores $\vec{a}$ e $\vec{b}$ são ortogonais, e vamos denotar tal fato por $\vec{a} \perp \vec{b}$, quando satisfazem a equação (9.2.6): $\vec{a} \cdot \vec{b} = 0$. Com esta definição, o vetor nulo é ortogonal a todos os vetores, inclusive a si mesmo.

A base canônica de $\mathbb{R}^n$

Quanto à base canônica $\vec{e}_1$, $\vec{e}_2$ e $\vec{e}_3$, note que

$$P_{\vec{e}_j}(\vec{e}_k) = \delta_{jk},$$

onde

$$\delta_{jk} = \begin{cases} 1, & j = k \\ 0, & j \neq k. \end{cases}$$

Dizemos que a base canônica é *ortonormal*, pois cada elemento tem norma 1, e elementos distintos são ortogonais.

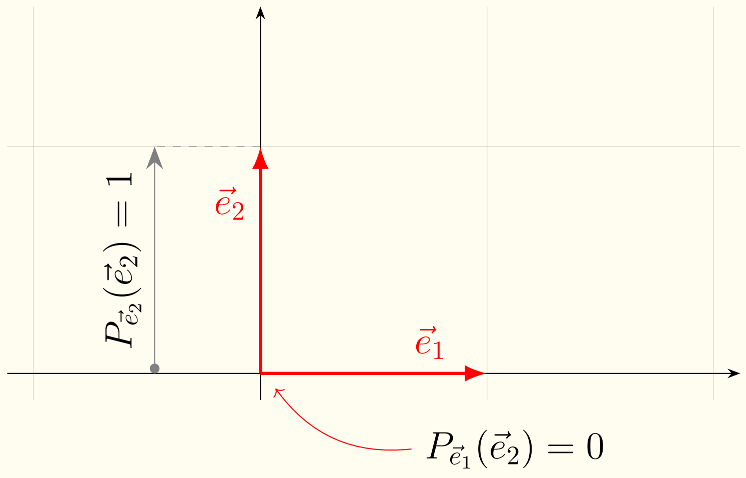


Figura 9.6: A base canônica é ortonormal: os vetores são unitários e ortogonais entre si.

Propriedades do Produto Interno

O produto interno é mais versátil do que a projeção $P_{\vec{w}}$. Por exemplo, $P_{\vec{w}}$ não é linear em $\vec{w}$.

$$P_{\alpha\vec{w}}(\vec{v}) \neq \alpha P_{\vec{w}}(\vec{v})$$
$$P_{\vec{w}+\vec{z}}(\vec{v}) \neq P_{\vec{w}}(\vec{v}) + P_{\vec{z}}(\vec{v}).$$

Vejamos algumas propriedades úteis do produto interno. A primeira delas é muito fácil perceber a partir da definição. O produto interno é simétrico:

$$\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$$

Também é claro que

$$\vec{a} \cdot \vec{0} = 0 = \vec{0} \cdot \vec{b}.$$

Além disso,

$$(\alpha\vec{a}) \cdot \vec{b} = \alpha \left(\vec{a} \cdot \vec{b}\right) = \vec{a} \cdot (\alpha\vec{b})$$

E também:

$$\begin{aligned}(\vec{a} + \vec{b}) \cdot \vec{v} &= \vec{a} \cdot \vec{v} + \vec{b} \cdot \vec{v}\\ \vec{a} \cdot (\vec{v} + \vec{w}) &= \vec{a} \cdot \vec{v} + \vec{a} \cdot \vec{w}\end{aligned}$$

Juntando essas últimas, podemos fazer o "chuveirinho":

$$(\vec{a} + \vec{b}) \cdot (\vec{v} + \vec{w}) = \vec{a} \cdot \vec{v} + \vec{a} \cdot \vec{w} + \vec{b} \cdot \vec{v} + \vec{b} \cdot \vec{w}.$$

9.3 Aplicações Geométricas

Munidos do significado geométrico do produto interno e de suas propriedades básicas, vejamos alguns exemplos simples do que conseguimos fazer.

Norma de um Vetor

A norma de um vetor é a grosso modo, seu "tamanho". A projeção ortogonal de um vetor não nulo em si mesmo, é a sua norma. Sendo assim,

$$\begin{aligned}\|\vec{a}\| &= P_{\vec{a}}(\vec{a})\\ &= \frac{\vec{a} \cdot \vec{a}}{\|\vec{a}\|}.\end{aligned}$$

Portanto,

$$\|\vec{a}\|^2 = \vec{a} \cdot \vec{a}.$$

E é evidente que essa igualdade também vale quando $\vec{a} = \vec{0}$.

Esse resultado, nada mais é do que o famoso *Teorema de Pitágoras* para o caso n-dimensional:

$$\|\vec{a}\|^2 = a_1^2 + \cdots + a_n^2.$$

Em particular, em $\mathbb{R}^2$, escrevendo $\vec{a} = (x, y)$, temos

$$\|\vec{a}\|^2 = x^2 + y^2.$$

Veja a figura 9.7.

Figura 9.7: Desenhando o vetor $\vec{a} = (x, y)$ e suas coordenadas, vemos que $\vec{a}$ é a hipotenusa, enquanto que $x\vec{e}_1$ e $y\vec{e}_2$ são os catetos de um triângulo retângulo. Os catetos tem tamanho $|x|$ e $|y|$. O *Teorema de Pitágoras* consiste na igualdade

$$\|\vec{a}\|^2 = x^2 + y^2.$$

Ângulo

Se $\vec{a}, \vec{b} \in \mathbb{R}^n$ são vetores não nulos, então

$$P_{\vec{a}}(\vec{b}) = \|b\| \cos\theta,$$

onde θ é o ângulo entre $\vec{a}$ e $\vec{b}$.

Figura 9.8: Se soubermos o tamanho do vetor $\vec{b}$ e a *projeção ortogonal* de $\vec{b}$ em $\vec{a}$, podemos calcular o cosseno do ângulo entre $\vec{a}$ e $\vec{b}$, pois

$$P_{\vec{a}}(\vec{b}) = \left\|\vec{b}\right\| \cos\theta.$$

Usando o produto interno, podemos calcular

$$\cos\theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \left\|\vec{b}\right\|}.$$

O Círculo e a Esfera

Como já vimos na figura 4.19, em $\mathbb{R}^2$, o círculo centrado em $\vec{a}$ de raio r é o conjunto

$$C = \left\{ \vec{v} \in \mathbb{R}^2 \,\middle|\, \|\vec{v} - \vec{a}\|^2 = r^2 \right\}.$$

Se quisermos fazer os cálculos com coordenadas, denotando o vetor $\vec{a}$ por (a, b), ficamos com

$$C = \left\{ (x, y) \in \mathbb{R}^2 \,\middle|\, (x - a)^2 + (y - b)^2 = r^2 \right\}.$$

Podemos fazer o mesmo pra esfera. E por analogia, podemos fazer o mesmo em $\mathbb{R}^n$!!! Assim, a *esfera n-dimensional* de raio r e centrada em $\vec{a} \in \mathbb{R}^n$ é o conjunto

$$S = \left\{ \vec{v} \in \mathbb{R}^n \,\middle|\, \|\vec{v} - \vec{a}\|^2 = r^2 \right\}.$$

Em termos de *coordenadas*,

$$S = \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n \,\middle|\, \sum_{j=1}^{n} (x_j - a_j)^2 = r^2 \right\}.$$

É melhor evitar coordenadas sempre que possível. Mas...

Plano e Hiperplano em $\mathbb{R}^n$

Em $\mathbb{R}^3$, um plano pode ser definido de diversas maneiras:

1. O plano que passa por três pontos distintos dados $\vec{a}$, $\vec{b}$ e $\vec{c}$.
2. O plano que passa no ponto $\vec{p}$ e que é tangente às direções $\vec{v}$ e $\vec{w}$.
3. O plano que passa no ponto $\vec{p}$ e que é *normal* à direção dada pelo vetor $\vec{n}$.

Os itens 1 e 2 são facilmente intercambiáveis. De fato, fazendo

$$\vec{p} = \vec{a}, \quad \vec{b} = \vec{a} + \vec{v} \quad \text{e} \quad \vec{c} = \vec{a} + \vec{w},$$

passamos do item 2 ao item 1. Podemos voltar fazendo

$$\vec{a} = \vec{p}, \quad \vec{v} = \vec{b} - \vec{a} \quad \text{e} \quad \vec{w} = \vec{c} - \vec{a}.$$

Com o ponto $\vec{a}$ e as direções $\vec{v}$ e $\vec{w}$, o plano é dado pelo conjunto

$$P = \left\{ \vec{a} + s\vec{v} + t\vec{w} \,\middle|\, s, t \in \mathbb{R} \right\}. \tag{9.3.1}$$

Essa construção pode ser feita também em outras dimensões. Não precisamos nos restringir a $\mathbb{R}^3$. Em qualquer $\mathbb{R}^n$, se $\vec{v}, \vec{w} \in \mathbb{R}^n$ são linearmente independentes e $\vec{a} \in \mathbb{R}^n$ é um vetor qualquer, então a equação (9.3.1) define um plano em $\mathbb{R}^n$.

Exemplo 9.4. O conjunto

$$P = \left\{ (s, 0, t, 0) \,\middle|\, s, t \in \mathbb{R} \right\}$$

é um plano em $\mathbb{R}^4$. É o que provavelmente chamaríamos de *plano x-z* em $\mathbb{R}^4$.

Por outro lado, o item 3 pode ser usado em $\mathbb{R}^n$ para falar de um *hiperplano*. Dado um ponto qualquer em $\vec{p} \in \mathbb{R}^n$ e uma direção não nula $\vec{n} \in \mathbb{R}^n$, o *hiperplano* que passa em $\vec{p}$ e é normal ao vetor $\vec{n}$ é o conjunto

$$\begin{aligned} H &= \left\{ \vec{p} + \vec{v} \,\middle|\, \vec{v} \perp \vec{n} \right\}. \\ &= \left\{ \vec{p} + \vec{v} \,\middle|\, \vec{v} \in \mathbb{R}^n,\ \vec{v} \cdot \vec{n} \right\}. \end{aligned}$$

Ou então,

$$\begin{aligned} H &= \left\{ \vec{q} \in \mathbb{R}^n \,\middle|\, (\vec{q} - \vec{p}) \cdot \vec{n} = 0 \right\} \\ &= \left\{ \vec{q} \in \mathbb{R}^n \,\middle|\, \vec{q} \cdot \vec{n} = \vec{p} \cdot \vec{n} \right\}. \end{aligned}$$

Em $\mathbb{R}^3$, *plano* e *hiperplano* é a mesma coisa.

Exemplo 9.5 (Plano em $\mathbb{R}^3$)**.** O plano P que passa em $(1,2,3)$ e é normal ao vetor $(4,5,6)$ é formado pelos pontos (x,y,z) que satisfazem a equação

$$(x,y,z)\cdot(4,5,6) = (1,2,3)\cdot(4,5,6).$$

Ou seja,

$$(x,y,z) \in P \Leftrightarrow 4x + 5y + 6z = 32.$$

Em $\mathbb{R}^2$, um hiperplano é o mesmo que uma reta.

Exemplo 9.6 (A reta em $\mathbb{R}^2$)**.** Em $\mathbb{R}^2$, a reta

$$r = \left\{ (1,2) + t(3,4) \;\middle|\; t \in \mathbb{R} \right\}$$

pode ser vista como a reta que passa no ponto $(1,2)$ e é normal ao vetor $(-4,3)$.

De fato, se s é a reta que passa em $(1,2)$ e é normal a $(-4,3)$, então

$$\begin{aligned}
s &= \left\{(x,y)\in\mathbb{R}^2 \;\middle|\; (x,y)\cdot(-4,3) = (1,2)\cdot(-4,3)\right\}\\
&= \left\{(x,y)\in\mathbb{R}^2 \;\middle|\; -4x+3y = 2\right\}\\
&= \left\{(x,y)\in\mathbb{R}^2 \;\middle|\; -4x+3y = 2\right\}\\
&= \left\{(x,y)\in\mathbb{R}^2 \;\middle|\; y = \frac{2}{3}+\frac{4}{3}x\right\}\\
&= \left\{\left(0,\frac{2}{3}\right) + t\left(1,\frac{4}{3}\right) \;\middle|\; t\in\mathbb{R}\right\}\\
&= \left\{\left(0,\frac{2}{3}\right) + \left(1,\frac{4}{3}\right) + (t-1)\left(1,\frac{4}{3}x\right) \;\middle|\; t\in\mathbb{R}\right\}\\
&= \left\{(1,2) + \frac{t-1}{3}(3,4) \;\middle|\; t\in\mathbb{R}\right\}\\
&= \left\{(1,2) + s(3,4) \;\middle|\; s\in\mathbb{R}\right\}\\
&= r.
\end{aligned}$$

Distâncias

Dados um ponto $\vec{p}$ e uma reta r em $\mathbb{R}^n$. A distância de $\vec{p}$ até o ponto $\vec{q}\in r$ mais próximo de $\vec{p}$. O processo para descobrir essa distância depende de como a reta r é dada. Se r é a reta que passa pelo ponto $\vec{a}\in\mathbb{R}^n$ e tem direção $\vec{v}\in\mathbb{R}^n$, então, a figura 9.9 nos mostra que

$$\begin{aligned}
\|\vec{p}-\vec{a}\|^2 &= \left(P_{\vec{v}}(\vec{p}-\vec{a})\right)^2 + d^2\\
&= \left(\frac{\vec{p}\cdot\vec{v}-\vec{a}\cdot\vec{v}}{\|\vec{v}\|}\right)^2 + d^2\\
&= \frac{(\vec{p}\cdot\vec{v}-\vec{a}\cdot\vec{v})^2}{\|\vec{v}\|^2} + d^2,
\end{aligned}$$

onde d é a distância de $\vec{p}$ a r.

Figura 9.9: Se d é a distância de r até $\vec{p}$, então

$$\|\vec{p} - \vec{a}\|^2 = P_{\vec{v}}(\vec{p} - \vec{a})^2 + d^2.$$

Pra quem gosta de fórmulas cheias de letras e coordenadas,

$$d = \sqrt{\sum_{j=1}^{n} (p_j - a_j)^2 - \frac{\left(\sum_{j=1}^{n} (p_j - a_j) v_j\right)^2}{\sum_{j=1}^{n} v_j^2}}.$$

Exemplo 9.7. A distância do ponto $(1, 2, 3)$ até a reta $r = \left\{(0, 0, 1) + t(1, 1, 1) \;\middle|\; t \in \mathbb{R}\right\}$ é

$$\begin{aligned} d &= \sqrt{1^2 + 2^2 + (3-1)^2 - \frac{(1+2+2)^2}{1+1+1}} \\ &= \sqrt{9 - \frac{25}{3}} \\ &= \sqrt{\frac{2}{3}}. \end{aligned}$$

Mas por outro lado, também é comum ao falar de um plano em $\mathbb{R}^3$ — ou de modo mais geral, de um híperplano em $\mathbb{R}^n$ —, termos um ponto do plano $\vec{a} \in \mathbb{R}^3$ e uma normal $\vec{n} \in \mathbb{R}^3$. Pela figura 9.10, é fácil ver que a distância de $\vec{p}$ ao plano é

$$d = |P_{\vec{n}}(\vec{p} - \vec{a})|,$$

onde $|\cdot|$ é usado porque $P_{\vec{n}}$ é uma projeção *"com sinal"*.

Figura 9.10: A distância de um ponto $\vec{p}$ a um híperplano com normal $\vec{n}$ passando pelo ponto $\vec{a}$ é simplesmente a projeção de $\vec{p} - \vec{a}$ na direção de $\vec{n}$.

XXXXXX Exercício: distância de plano a plano XXXXXXX

Triângulo Retângulo no Círculo

xxx

Aula 10

Componentes Tangente e Ortogonal

10.1 Decompondo um Vetor

Dado um vetor $\vec{v} \in \mathbb{R}^n$, os vetores paralelos a $\vec{v}$ são os vetores da forma $t\vec{v}$ para $t \in \mathbb{R}$. É muito comum, queremos decompor $\vec{a} \in \mathbb{R}^n$ em uma componente paralela a $\vec{v}$ e uma componente normal. Ou seja, queremos escrever

$$\vec{a} = t\vec{v} + \vec{n}, \tag{10.1.1}$$

onde $\vec{n} \perp \vec{v}$.

Figura 10.1: Dados $\vec{a}, \vec{v} \in \mathbb{R}^n$, como podemos encontrar as componentes de $\vec{a}$ que são paralela e ortogonal a $\vec{v}$? Se $\|\vec{v}\| = 1$, então $t = P_{\vec{v}}(\vec{a})$.

Argumento Geométrico

Se você é do tipo de pessoa que não gosta de argumentos geométicos, retas, planos, ângulos, projeções ortogonais, então você pode deixar essa subseção pra lá, e seguir direto para o *argumento algébrico*.

Pela figura 10.1, é fácil ver que quando $\|\vec{v}\| = 1$, então

$$t = P_{\vec{v}}(\vec{a}).$$

Mas quando a norma de $\vec{v}$ não é 1, basta *normalizá-lo*. Assim, de um modo geral, para qualquer $\vec{v}$ não nulo,

$$\vec{a} = P_{\vec{v}}(\vec{a})\frac{\vec{v}}{\|\vec{v}\|} + \vec{n},$$

com $\vec{n} \perp \vec{v}$. Ou seja, o valor de t que satisfaz a equação (10.1.1) é

$$\begin{aligned} t &= \frac{P_{\vec{v}}(\vec{a})}{\|\vec{v}\|} \\ &= \frac{\vec{v} \cdot \vec{a}}{\|\vec{v}\|^2}. \end{aligned}$$

Argumento Algébrico

Se você não gosta de contas e manipulações algébricas, e se você leu o *argumento geométrico* e entendeu tudo, fique à vontade pra pular essa subseção com esse *argumento algébrico*.

Lembre-se que queremos encontrar $t \in \mathbb{R}$ tal que quando

$$\vec{a} = t\vec{v} + \vec{n},$$

então $\vec{n} \perp \vec{v}$.

Ou seja, queremos encontrar t tal que

$$\begin{aligned} 0 &= \vec{n} \cdot \vec{v} \\ &= (\vec{a} - t\vec{v}) \cdot \vec{v} \\ &= \vec{a} \cdot \vec{v} - t\vec{v} \cdot \vec{v} \\ &= \vec{a} \cdot \vec{v} - t\|\vec{v}\|^2. \end{aligned}$$

Portanto,

$$t = \frac{\vec{v} \cdot \vec{a}}{\|\vec{v}\|^2}.$$

Exemplo

Seja pelo método geométrico, ou pelo método algébrico, já sabemos que para que

$$\vec{a} = t\vec{v} + \vec{n}$$

com $\vec{n} \perp \vec{v}$, é necessário e suficiente que

$$t = \frac{\vec{v} \cdot \vec{a}}{\|\vec{v}\|^2}. \tag{10.1.2}$$

Exemplo 10.1 (Plano inclinado)**.** Considere um corpo de peso $\vec{P}$ em um plano inclinado com normal $\vec{n}$, como na figura 10.2.

Figura 10.2: A componente do peso $\vec{P}$ normal à superfície é

$$\vec{P}_N = \frac{\vec{P} \cdot \vec{n}}{\|\vec{n}\|^2}\vec{n}.$$

Se $\vec{P} = (0, 0, -5)$, e se $\vec{n} = (1, 2, 3)$, então a componente de $\vec{P}$ normal ao plano inclinado é

$$\begin{aligned} \vec{P}_N &= \frac{\vec{P} \cdot \vec{n}}{\|\vec{n}\|^2}\vec{n} \\ &= \frac{-15}{14}(1, 2, 3). \end{aligned}$$

10.2 Bases Ortonormais

Se ao invés da base canônica $\vec{e}_1, \ldots, \vec{e}_n \in \mathbb{R}^n$ usarmos uma outra base $\vec{b}_1, \ldots, \vec{b}_n \in \mathbb{R}^n$ formada por vetores de norma 1 e ortogonais entre si, podemos usar o produto interno para facilmente determinar como podemos escrever um vetor qualquer $\vec{a}$ na base $\vec{b}_1, \dots, \vec{b}_n \in \mathbb{R}^n$.

Definição 10.2 (Base Ortonormal)**.** *Seja V um espaço vetorial e $\vec{b}_1, \dots, \vec{b}_n \in V$ uma base de V é* ortonormal *quando é formada por vetores de norma 1 que são ortogonais entre si. Em outras palavras, é uma base que satisfaz*

$$\vec{b}_j \cdot \vec{b}_k = \begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases}.$$

Geometricamente é bastante intuitivo que se quisermos escrever $\vec{a}$ na base ortonormal $\vec{b}_1, \dots, \vec{b}_n \in \mathbb{R}^n$, basta fazer

$$\vec{a} = \left\langle \vec{a}, \vec{b}_1 \right\rangle \vec{b}_1 + \cdots + \left\langle \vec{a}, \vec{b}_n \right\rangle \vec{b}_n.$$

Mas deixando a geometria de lado, podemos usar as propriedades do produto interno para verificar isso. De fato, como $\vec{b}_1, \dots, \vec{b}_n \in \mathbb{R}^n$ é uma base, podemos escrever $\vec{a}$ como combinação linear desses vetores:

$$\vec{a} = \alpha_1 \vec{b}_1 + \cdots + \alpha_n \vec{b}_n.$$

Basta então encontrar $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. Para tanto, vamos fazer o produto interno de $\vec{a}$ com $\vec{b}_j$:

$$\begin{aligned} \vec{a} \cdot \vec{b}_j &= \alpha_1 \vec{b}_1 + \cdots + \alpha_n \vec{b}_n \cdot \vec{b}_j \\ &= \alpha_1 \vec{b}_1 \cdot \vec{b}_j + \cdots + \alpha_n \vec{b}_n \cdot \vec{b}_j \\ &= \alpha_j \vec{b}_j \cdot \vec{b}_j \\ &= \alpha_j, \end{aligned}$$

onde as últimas igualdades são consequência do fato de $\vec{b}_k \cdot \vec{b}_j$ ser igual a 0 quando $k \neq j$ e igual a 1 quando $k = j$. Portanto, para encontrarmos α_j, basta calcularmos $\vec{a} \cdot \vec{b}_j$.

Observação 10.3. Sem geometria alguma, a conclusão do último parágrafo decorre apenas das propriedades básicas do produto interno. Por isso, na definição 9.1, chamamos nosso produto interno de *produto interno **canônico***. Existem outros produtos internos que possuem as mesmas propriedades e podem ser usados de modo similar.

Aula 11

Processo de Ortonormalização de Gram–Schmidt

Em $\mathbb{R}^n$, temos a base ortonormal $\vec{e}_1, \ldots, \vec{e}_n$. Mas se estivermos em um espaço vetorial que não seja $\mathbb{R}^n$? Será que conseguimos uma base ortonormal para esse espaço?

Suponha que $\vec{b}_1, \ldots, \vec{b}_n \in V$ seja uma base para V. Vamos usar o produto interno para construir uma base ortonormal $\vec{a}_1, \ldots, \vec{a}_n \in V$.

Primeiro, faça

$$\vec{a}_1 = \frac{\vec{b}_1}{\left\|\vec{b}_1\right\|}.$$

Escolhido $\vec{a}_1$, podemos tomar $\vec{b}_2$ e remover a componente $\vec{a}_1$ para ter um vetor ortogonal a $\vec{a}_1$

$$\vec{n} = \vec{b}_2 - (\vec{b}_2 \cdot \vec{a}_1)\vec{a}_1.$$

Agora, é só normalizá-lo para obter

$$\begin{aligned}\vec{a}_2 &= \frac{\vec{b}_2 - (\vec{b}_2 \cdot \vec{a}_1)\vec{a}_1}{\left\|\vec{b}_2 - (\vec{b}_2 \cdot \vec{a}_1)\vec{a}_1\right\|} \\ &= \frac{\vec{b}_2 - (\vec{b}_2 \cdot \vec{a}_1)\vec{a}_1}{\sqrt{\left\|\vec{b}_2\right\|^2 - (\vec{b}_2 \cdot \vec{a}_1)^2}}.\end{aligned}$$

Podemos então continuar o processo e fazer

$$\begin{aligned}\vec{a}_3 &= \frac{\vec{b}_3 - (\vec{b}_3 \cdot \vec{a}_1)\vec{a}_1 - (\vec{b}_3 \cdot \vec{a}_2)\vec{a}_2}{\left\| \vec{b}_3 - (\vec{b}_3 \cdot \vec{a}_1)\vec{a}_1 - (\vec{b}_3 \cdot \vec{a}_2)\vec{a}_2 \right\|} \\ &= \frac{\vec{b}_3 - (\vec{b}_3 \cdot \vec{a}_1)\vec{a}_1 - (\vec{b}_3 \cdot \vec{a}_2)\vec{a}_2}{\sqrt{\left\|\vec{b}_3\right\|^2 - (\vec{b}_3 \cdot \vec{a}_1)^2 - (\vec{b}_3 \cdot \vec{a}_2)^2}}.\end{aligned}$$

De um modo geral, se fizermos

$$\begin{aligned}\vec{a}_j &= \frac{\vec{b}_j - \sum_{k=1}^{j-1}(\vec{b}_j \cdot \vec{a}_k)\vec{a}_k}{\left\| \vec{b}_j - \sum_{k=1}^{j-1}(\vec{b}_j \cdot \vec{a}_k)\vec{a}_k \right\|} \\ &= \frac{\vec{b}_j - \sum_{k=1}^{j-1}(\vec{b}_j \cdot \vec{a}_k)\vec{a}_k}{\sqrt{\left\|\vec{b}_j\right\|^2 - \sum_{k=1}^{j-1}(\vec{b}_j \cdot \vec{a}_k)^2}},\end{aligned}$$

obtemos uma base ortonormal $\vec{a}_1, \ldots, \vec{a}_n$. Esse é o *processo de ortonormalização de Gram–Schmidt*.

Exemplo 11.1. Se fizermos o processo de ortonormalização para a base de $\mathbb{R}^4$

$$\begin{aligned}\vec{b}_1 &= (1,0,0,0) & \vec{b}_2 &= (1,1,0,0) \\ \vec{b}_3 &= (1,1,1,0) & \vec{b}_4 &= (1,1,1,1),\end{aligned}$$

teremos a base ortonormal canônica

$$\begin{aligned}\vec{a}_1 &= (1,0,0,0) & \vec{a}_2 &= (0,1,0,0) \\ \vec{a}_3 &= (0,0,1,0) & \vec{a}_4 &= (0,0,0,1).\end{aligned}$$

Mas se tomarmos a mesma base na ordem inversa:

$$\vec{b}_1 = (1,1,1,1) \qquad \vec{b}_2 = (1,1,1,0)$$
$$\vec{b}_3 = (1,1,0,0) \qquad \vec{b}_4 = (1,0,0,0),$$

obteremos

$$\vec{a}_1 = \frac{1}{2}(1,1,1,1) \qquad \vec{a}_2 = \frac{1}{2\sqrt{3}}(1,1,1,-3)$$
$$\vec{a}_3 = \frac{1}{\sqrt{6}}(1,1,-2,0) \qquad \vec{a}_4 = \frac{1}{\sqrt{2}}(1,-1,0,0).$$

11.1 Exercícios

xxx

Aula 12

Funcionais Lineares com Produto Interno

As aplicações lineares com contradomínio em $\mathbb{R}$ são também chamadas de *funcionais (lineares)*. Dado um vetor $\vec{a} \in \mathbb{R}^n$, podemos definir o functional

$$f_{\vec{a}} : \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R} \\ \vec{v} & \mapsto & \vec{a} \cdot \vec{v} \end{array} .$$

Por outro lado, os funcionais lineares de $\mathbb{R}^n$ em $\mathbb{R}$ são todos da forma

$$f : \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R} \\ (v_1, \ldots, v_n) & \mapsto & a_1 v_1 + \cdots + a_n v_n \end{array}$$

para $a_1, \ldots, a_n \in \mathbb{R}$ adequados. Portanto, podemos associar a f o vetor $\vec{a} = (a_1, \ldots, a_n)$, de modo que $f = f_{\vec{a}}$. Ou seja,

$$f(\vec{v}) = \vec{a} \cdot \vec{v}.$$

Assim, aplicar a transformação f é o mesmo que fazer o produto interno com o vetor $\vec{a}$. Neste caso, a repesentação matricial de f é a matriz linha

$$\left[f\right] = \left[a_1 \quad \ldots \quad a_n\right].$$

E o núcleo de f é o conjunto dos vetores ortogonais a $\vec{a}$:

$$\ker(f) = \left\{ \vec{v} \in \mathbb{R}^n \;\middle|\; \vec{v} \perp \vec{a} \right\}.$$

12.1 Linhas das matrizes

Uma *transformação linear* $T : \mathbb{R}^m \to \mathbb{R}^n$ pode ser escrita como

$$\begin{array}{rccl} T: & \mathbb{R}^m & \to & \mathbb{R}^n \\ & \vec{v} & \mapsto & (T_1\vec{v}, \dots, T_n\vec{v}) \end{array},$$

onde cada $T_j : \mathbb{R}^m \to \mathbb{R}$ é um *funcional linear*. Então, para cada $j = 1, \dots, n$ existe $\vec{a}_j = (a_{j1}, \dots, a_{jm})$ tal que

$$T_j\vec{v} = \vec{a}_j \cdot \vec{v}.$$

O vetor $\vec{a}_j$ é a j-ésima linha da matriz

$$\big[T\big] = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} = \begin{bmatrix} — & \vec{a}_1 & — \\ & \vdots & \\ — & \vec{a}_n & — \end{bmatrix}$$

que representa a transformação T.

12.2 Inversão de Matrizes

Suponha que M é uma matriz $n \times n$ inversível. Ou seja, M representa uma transformação linear inversível $T : \mathbb{R}^n \to \mathbb{R}^n$. Queremos encontrar a inversa de M.

Vamos escrever as linhas de M como vetores:

$$M = \begin{bmatrix} — & \vec{a}_1 & — \\ & \vdots & \\ — & \vec{a}_n & — \end{bmatrix}$$

Ou seja,

$$M = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix},$$

Queremos encontrar

$$N = \begin{bmatrix} | & & | \\ \vec{b}_1 & \dots & \vec{b}_n \\ | & & | \end{bmatrix}$$

tal que $MN = I$. Olhando para $\vec{b}_j$, a j-ésima coluna de N, isso é o mesmo que dizer que

1. $\vec{b}_j$ é ortogonal a $\vec{a}_k$ para todo $k \neq j$.

2. $\vec{b}_j \cdot \vec{a}_j = 1$.

Vamos nos concentrar no item 1. Suponha que tenhamos encontrado um vetor **não nulo** $\vec{v}_j$ que seja ortogonal a $\vec{a}_k$ para todo $k \neq j$. Então, basta fazer

$$\vec{b}_j = \frac{1}{\vec{v}_j \cdot \vec{a}_j} \vec{v}_j,$$

para que além do item 1 ($\vec{b}_j \perp \vec{a}_k$ para $k \neq j$), termos também o item 2:

$$\begin{aligned} \vec{b}_j \cdot \vec{a}_j &= \frac{1}{\vec{v}_j \cdot \vec{a}_j} \vec{v}_j \cdot \vec{a}_j \\ &= 1. \end{aligned}$$

Sendo assim, nosso problema de inverter a matriz M pode ser reduzido ao problema

> Dados $n-1$ vetores $\vec{a}_2, \dots, \vec{a}_n \in \mathbb{R}^n$, encontre um vetor **não nulo** $\vec{b} \in \mathbb{R}^n$ que seja ortogonal a todos eles.

Esse problema será resolvido com o *determinante* (capítulos 13 e 15).

Aula 13

Determinante

13.1 Definição de Determinante

xxx

13.2 Determinante 2×2

Nesta seção, vamos tratar apenas do espaço $\mathbb{R}^2$.

Dois vetores determinam um paralelogramo. Assim como o *tamanho da projeção* de um vetor em outro tem propriedades interessantes que nos levaram a definir o *produto interno*, a *área do paralelogramo* determinado por dois vetores também tem propriedades semelhantes, e nos levarão a definir o conceito de *determinante*. Apenas nesta seção, vamos denotar por $A(\vec{v}, \vec{w})$ a área do paralelogramo de lados $\vec{v}$ e $\vec{w}$.

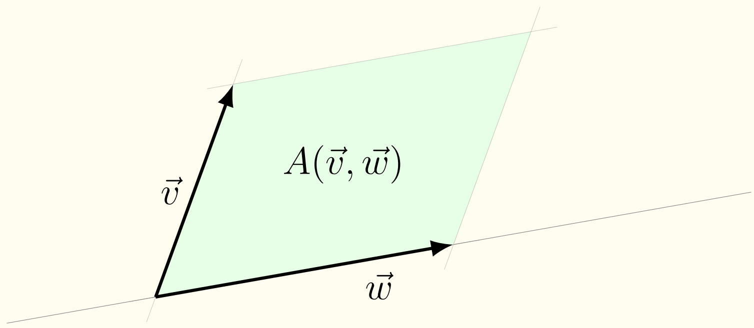


Figura 13.1: Os vetores $\vec{v}$ e $\vec{w}$ determinam um paralelogramo. Vamos denotar por $A(\vec{v}, \vec{w})$ a área desse paralelogramo.

Se $\alpha \geq 0$, então

$$A(\alpha\vec{v}, \vec{w}) = \alpha A(\vec{v}, \vec{w}) = A(\vec{v}, \alpha\vec{w}).$$

Consequentemente,

$$A(\alpha\vec{v}, \beta\vec{w}) = \alpha\beta A(\vec{v}, \vec{w}).$$

E se $\vec{v}$ e $\vec{w}$ são paralelos,

$$A(\vec{v}, \vec{w}) = 0.$$

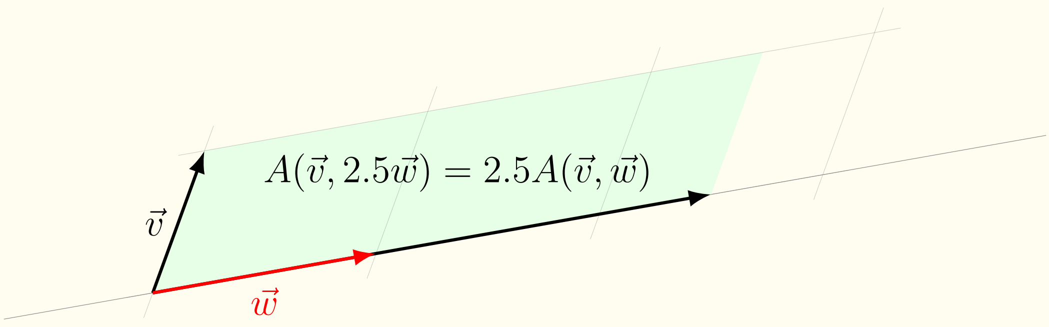


Figura 13.2: Assim como no caso da projeção, fixando o vetor $\vec{v}$, e dilatando ou encolhendo o vetor $\vec{w}$, a área $A(\vec{v}, \vec{w})$ aumenta ou diminui na mesma proporção.

Quanto à *linearidade*, fixado o vetor $\vec{w}$, a aplicação $A(\vec{v}, \vec{w})$ é *"praticamente" linear*. Lembre-se que os desenhos das figuras 13.3 e 13.4 **NÃO** são tridimensionais. Como mostrado na figura 13.3, existem casos tais que

$$A(\vec{u} + \vec{v}, \vec{w}) = A(\vec{u}, \vec{w}) + A(\vec{v}, \vec{w}).$$

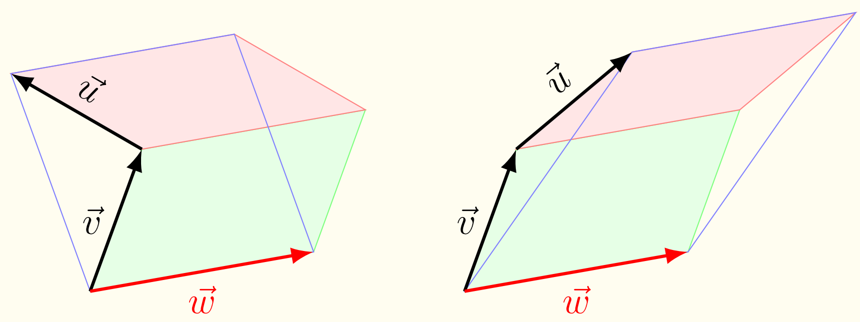


Figura 13.3: O valor de $A(\vec{u} + \vec{v}, \vec{w})$ corresponde à área delimitada pelo paralelogramo azul. Em rosa e verde, temos $A(\vec{u}, \vec{w})$ e $A(\vec{v}, \vec{w})$. Em alguns casos, $A(\vec{u} + \vec{v}, \vec{w})$ é a soma das áreas em rosa e verde.

Mas, assim como no caso da projeção, existem casos tais que

$$A(\vec{u} + \vec{v}, \vec{w}) = |A(\vec{u}, \vec{w}) - A(\vec{v}, \vec{w})|.$$

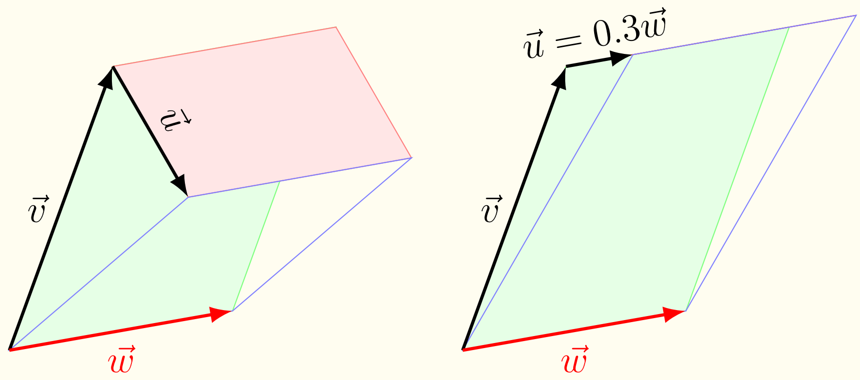


Figura 13.4: O valor de $A(\vec{u}+\vec{v},\vec{w})$ corresponde à área delimitada pelo paralelogramo azul. Em rosa e verde, temos $A(\vec{u},\vec{w})$ e $A(\vec{v},\vec{w})$. À direita, $A(\vec{u}+\vec{v},\vec{w}) = A(\vec{v},\vec{w})$. À esquerda, $A(\vec{u}+\vec{v},\vec{w}) = A(\vec{v},\vec{w}) - A(\vec{u},\vec{w})$.

Em analogia com a projeção ortogonal, uma pergunta que precisa ser respondida é:

> Quando é que $A(\vec{u}+\vec{v},\vec{w})$ é igual a $A(\vec{u},\vec{w})+A(\vec{v},\vec{w})$, e quando é que é igual a $|A(\vec{u},\vec{w}) - A(\vec{v},\vec{w})|$?

Ou seja, se fossemos atribuir um sinal a $A(\vec{v},\vec{w})$, qual seria o critério para escolher entre *positivo* e *negativo*?

A área é

$$\text{comprimento da base} \times \text{altura}.$$

O *comprimento da base* é $\|\vec{w}\|$. E a *altura*, com sinal, é

$$P_{R_{\frac{\pi}{2}}\vec{w}}(\vec{v}),$$

onde $R_{\frac{\pi}{2}}\vec{w} = (-w_2, w_1)$ é a rotação por um ângulo $\frac{\pi}{2}$ no sentido anti-horário. Assim,

$$\begin{aligned} A(\vec{w},\vec{v}) &= \|\vec{w}\| \left| P_{R_{\frac{\pi}{2}}\vec{w}}(\vec{v}) \right| \\ &= \|\vec{w}\| \left| \frac{\vec{v} \cdot R_{\frac{\pi}{2}}\vec{w}}{\left\| R_{\frac{\pi}{2}}\vec{w} \right\|} \right| \\ &= \left| \vec{v} \cdot R_{\frac{\pi}{2}}\vec{w} \right|. \end{aligned}$$

Vamos então, definir a *"área com sinal"* (o *determinante*) como sendo

$$\begin{aligned} \det(\vec{w},\vec{v}) &= \vec{v} \cdot R_{\frac{\pi}{2}}\vec{w} \\ &= w_1 v_2 - w_2 v_1. \end{aligned}$$

Se o leitor já ouviu falar do determinante, provavelmente aprendeu essa fórmula.

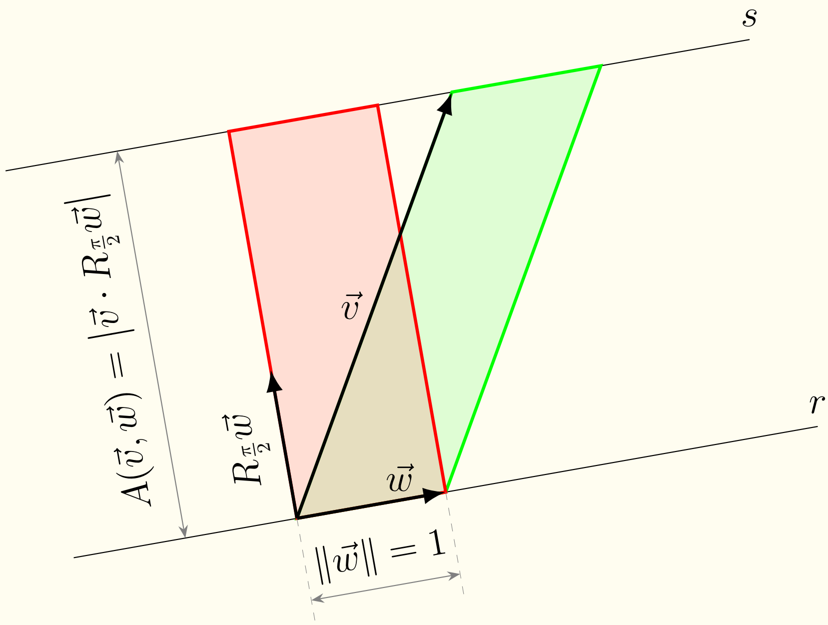


Figura 13.5: Note que se $\|\vec{w}\| = 1$, então a área, já com o sinal correto é exatamente $P_{R_{\frac{\pi}{2}}\vec{w}}(\vec{v})$. E mesmo que $\vec{w}$ não seja unitário,

$$A(\vec{v}, \vec{w}) = \|\vec{w}\| \left| P_{R_{\frac{\pi}{2}}\vec{w}}(\vec{v}) \right|.$$

Definido o *determinante*, vamos entender de modo mais geométrico o significado de seu sinal. No cálculo de $A(\vec{u} + \vec{v}, \vec{w})$, quando é que as áreas se somam, e quando é que elas se cancelam?

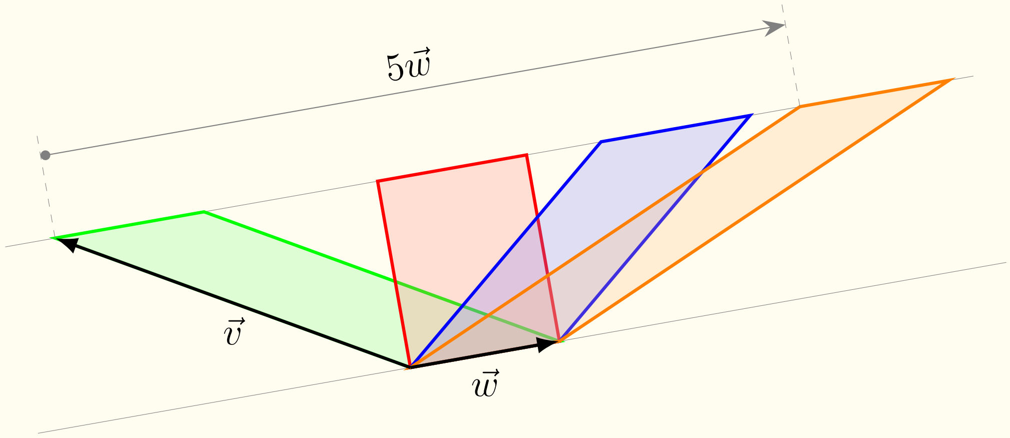


Figura 13.6: As áreas de todos esses paralelogramos são iguais. A área em verde corresponde a $A(\vec{v}, \vec{w})$, e é igual à área em laranja, que corresponde a $A(\vec{v} + 5\vec{w}, \vec{w})$.

A resposta é simples. Se considerarmos a reta s da figura 13.7, vemos que quando o vetor $\vec{u}$ está *"acima"* de s, então as áreas se somam. Quando está *"abaixo"*, as áreas se cancelam.

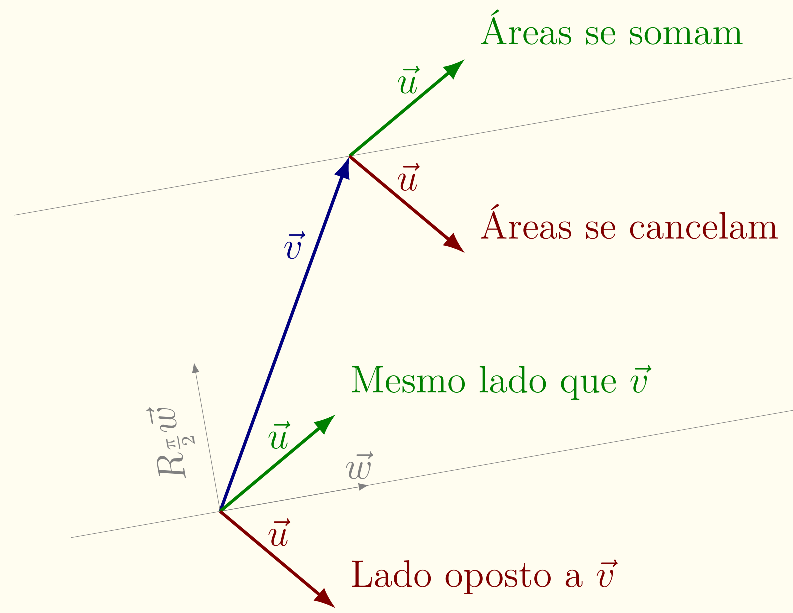


Figura 13.7:

E se ao invés de olharmos para a reta s, olharmos para a reta r, e se transportarmos o vetor $\vec{u}$ para o mesmo ponto base que $\vec{v}$, vemos que quando $\vec{u}$ e $\vec{v}$ estão do mesmo lado de r, as áreas se somam. Quando estão de lados opostos, as áreas se cancelam.

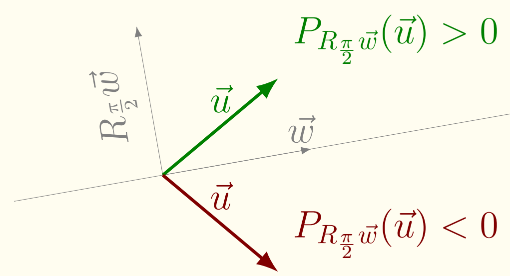


Figura 13.8:

Na notação matricial, quando $\vec{v} = (v_1, v_2)$ e $\vec{w} = (w_1, w_2)$, vamos usar a notação matricial para determinante e escrever

$$\det(\vec{w}, \vec{v}) = \begin{vmatrix} w_1 & v_1 \\ w_2 & v_2 \end{vmatrix}.$$

Quanto ao sinal, $\det(\vec{w}, \vec{v})$ é positivo quando $\vec{v} \cdot R_{\frac{\pi}{2}}\vec{w}$ é positivo. Ou seja, quando $\vec{v}$ estiver no semiplano à *"esquerda"* de $\vec{w}$. Ou então, quando a rotação mais curta de $\vec{w}$ a $\vec{v}$ for no sentido anti-horário.

Com toda essa discussão, sabemos que $\det(\vec{w}, \vec{v})$, é linear em $\vec{v}$ (quando fixamos $\vec{w}$). Pela fórmula encontrada,

$$\det(\vec{w}, \vec{v}) = -\det(\vec{v}, \vec{w}).$$

Se $\vec{v}$ está *"à esquerda"* de $\vec{w}$, então, $\vec{w}$ está *"à direita"* de $\vec{v}$. Isso mostra que $\det(\vec{w}, \vec{v})$ é linear tanto em $\vec{v}$ quanto em $\vec{w}$. Ou seja, além das propriedades já mostradas, também vale que

$$\det(\vec{w} + \vec{z}, \vec{v}) = \det(\vec{w}, \vec{v}) + \det(\vec{z}, \vec{v}).$$

Em outras palavras, det é *bilinear*. Repare também, que

$$\det(\vec{e}_1, \vec{e}_2) = 1.$$

Pra finalizar essa discussão, vamos (re)encontrar a fórmula para $\det(\vec{w}, \vec{v})$ usando apenas as equações

$$\begin{aligned} \det(\vec{e}_1, \vec{e}_2) &= 1 \\ \det(\vec{e}_2, \vec{e}_1) &= -1 \\ \det(\vec{e}_1, \vec{e}_1) &= \det(\vec{e}_2, \vec{e}_2) = 0, \end{aligned}$$

e mais o fato de det ser *bilinear*.

Exemplo 13.1 (fórmula com manipulações algébricas)**.** Note que $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2$, e $\vec{w} = w_1\vec{e}_1 + w_2\vec{e}_2$. Assim,

$$\begin{aligned}
\det(\vec{w}, \vec{v}) &= w_1 \det(\vec{e}_1, \vec{v}) + w_2 \det(\vec{e}_2, \vec{v}) \\
&= w_1 \left(v_1 \det(\vec{e}_1, \vec{e}_1) + v_2 \det(\vec{e}_1, \vec{e}_2)\right) \\
&\quad + w_2 \left(v_1 \det(\vec{e}_2, \vec{e}_1) + v_2 \det(\vec{e}_2, \vec{e}_2)\right) \\
&= w_1 \left(0 + v_2\right) + w_2 \left(-v_1 + 0\right) \\
&= w_1 v_2 - w_2 v_1.
\end{aligned}$$

Definição 13.2 (Determinante 2×2)**.** *Se* $\vec{v} = (v_1, v_2)$ *e* $\vec{w} = (w_1, w_2)$ *são vetores de* $\mathbb{R}^2$*, então a quantidade*

$$\det(\vec{w}, \vec{v}) = w_1 v_2 - w_2 v_1$$

é chamada de determinante *da matriz*

$$\begin{bmatrix} w_1 & v_1 \\ w_2 & v_2 \end{bmatrix}.$$

O determinante *também é denotado por*

$$\begin{vmatrix} w_1 & v_1 \\ w_2 & v_2 \end{vmatrix}.$$

13.3 Determinante $n \times n$

Na seção 13.2, substituímos a função $A(\vec{v}, \vec{w})$ pela função $\det(\vec{w}, \vec{v})$, que além de ser linear em cada entrada (*bilinear*), também é tal que a área de uma reta é 0. Ou seja,

$$\det(\vec{w}, \vec{w}) = 0.$$

Se fosse em $\mathbb{R}^3$, e estivéssemos falando do volume $V(\vec{u}, \vec{v}, \vec{w})$, talvez fosse mais interessante trabalhar, no lugar de V, com uma função

$\det : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, *trilinear*, e tal que se $\vec{u}$, $\vec{v}$ e $\vec{w}$ estiverem em um mesmo plano, ou seja, se forem linearmente dependentes, então,

$$\det(\vec{u}, \vec{v}, \vec{w}) = 0.$$

Figura 13.9: Escolhendo dois vetores para serem a *"base"*, o volume do prisma é *"base"* × *"altura"*.

No caso $\mathbb{R}^n$, estaríamos definindo uma espécie de *volume n-dimensional*, com sinal. Essa aplicação — que também vamos denotar por det — atribui a cada sequência de n vetores $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^n$, o *volume n-dimensional* sinalado do *"paralelepípedo"* associado a $\vec{v}_1, \dots, \vec{v}_n$. Esperamos que a aplicação $\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ tenha três propriedades elementares:

1. Que $\det(\vec{e}_1, \dots, \vec{e}_n) = 1$. Esse é o volume n-dimensional do n-cubo de lados de tamanho 1.

2. Que seja n-linear. Se fixarmos $n - 1$ vetores e os tratarmos como *"base"* do paralelepípedo, o volume é "base × altura". Assim como no caso 2×2, a "altura" é a projeção do n-ésimo vetor (o que não faz parte da base), na direção ortogonal à base. E é por isso que o n-volume é n-linear.

3. Para $\vec{v}_1, \dots, \vec{v}_n$, se $\vec{v}_i = \vec{v}_j$ para algum $i \neq j$, então

$$\det(\vec{v}_1, \dots, \vec{v}_n) = 0.$$

Será que tal aplicação existe? Uma maneira de mostrarmos que existe é exibindo uma fórmula. A proposição 13.3 nos permitirá encontrar uma fórmula para o *determinante*. A proposição é uma consequência do itens 2 e 3, que implica que o sinal do *determinante* se inverte quando permutamos um vetor $\vec{v}_j$ com um outro vetor $\vec{v}_k$.

Proposição 13.3. *Seja*

$$T : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$$

uma aplicação n-linear tal que se $\vec{v}_1, \ldots, \vec{v}_n \in \mathbb{R}^n$, tem alguma entrada repetida, então

$$T(\vec{v}_1, \ldots, \vec{v}_n) = 0.$$

Nesse caso, T é alternada. *Ou seja, se $\sigma \in P_n$ é uma permutação de $\{1, \ldots, n\}$, então*

$$T(\vec{v}_{\sigma(1)}, \ldots, \vec{v}_{\sigma(n)}) = \operatorname{sgn}(\sigma) T(\vec{v}_1, \ldots, \vec{v}_n).$$

Demonstração. Suponha que $\sigma = (j\, k)$, com $j < k$. Se, no lugar de v_j e v_k colocarmos $v_j + v_k$, o *determinante* será zero, pois as entradas j e k serão iguais. Assim,

$$\begin{aligned}
0 &= T(\vec{v}_1, \ldots, \vec{v}_j + \vec{v}_k, \ldots, \vec{v}_j + \vec{v}_k, \ldots, \vec{v}_n) \\
&= T(\vec{v}_1, \ldots, \cancel{\vec{v}_j}, \ldots, \cancel{\vec{v}_j}, \ldots, \vec{v}_n) \\
&\quad + T(\vec{v}_1, \ldots, \vec{v}_j, \ldots, \vec{v}_k, \ldots, \vec{v}_n) \\
&\quad + T(\vec{v}_1, \ldots, \vec{v}_k, \ldots, \vec{v}_j, \ldots, \vec{v}_n) \\
&\quad + T(\vec{v}_1, \ldots, \cancel{\vec{v}_k}, \ldots, \cancel{\vec{v}_k}, \ldots, \vec{v}_n) \\
&= T(\vec{v}_1, \ldots, \vec{v}_j, \ldots, \vec{v}_k, \ldots, \vec{v}_n) \\
&\quad + T(\vec{v}_1, \ldots, \vec{v}_k, \ldots, \vec{v}_j, \ldots, \vec{v}_n).
\end{aligned}$$

Mas isso é o mesmo que dizer que

$$T(\ldots, \vec{v}_j, \ldots, \vec{v}_k, \ldots) = -T(\ldots, \vec{v}_k, \ldots, \vec{v}_j, \ldots).$$

No caso de uma permutação σ qualquer, podemos escrevê-la como uma sequência de transposições. A cada transposição o sinal se inverte, de modo que o sinal do resultado final será exatamente $\operatorname{sgn}(\sigma)$.

Se

$$\begin{aligned}
\vec{v}_1 &= (v_{11}, v_{21}, \dots, v_{n1}) \\
\vec{v}_2 &= (v_{12}, v_{22}, \dots, v_{n2}) \\
&\vdots \\
\vec{v}_n &= (v_{1n}, v_{2n}, \dots, v_{nn}),
\end{aligned}$$

então, as propriedades de $\det : (\mathbb{R}^n)^n \to \mathbb{R}$ que foram discutidas implicam que

$$\begin{aligned}
\det(\vec{v}_1, \dots, \vec{v}_n) &= \sum_{\sigma \in \{1,\dots,n\}^n} v_{\sigma(1),1} \cdots v_{\sigma(n),n} \det(\vec{e}_{\sigma(1)}, \dots, \vec{e}_{\sigma(n)}) \\
&= \sum_{\sigma \in P_n} v_{\sigma(1),1} \cdots v_{\sigma(n),n} \det(\vec{e}_{\sigma(1)}, \dots, \vec{e}_{\sigma(n)}) \\
&= \sum_{\sigma \in P_n} \operatorname{sgn}(\sigma) v_{\sigma(1),1} \cdots v_{\sigma(n),n}.
\end{aligned}$$

Esta fórmula está *bem definida* porque, pela **??**, a aplicação $\operatorname{sgn} : P_n \to \{+1, -1\}$ está bem definida. Assim, chegamos a uma fórmula que utilizaremos para definir o *determinante*.

Definição 13.4 (Determinante)**.** *Para cada* $n = 1, 2, \dots,$ *o* determinante *é a aplicação*

$$\begin{array}{rrcl}
\det : & (\mathbb{R}^n)^n & \to & \mathbb{R} \\
& (\vec{v}_1, \dots, \vec{v}_n) & \mapsto & \sum_{\sigma \in P_n} \operatorname{sgn}(\sigma) v_{\sigma(1),1} \cdots v_{\sigma(n),n}
\end{array}.$$

Se $T : \mathbb{R}^n \to \mathbb{R}^n$ *é uma transformação linear,*

$$\det(T) = \det(T\vec{e}_1, \dots, T\vec{e}_n).$$

E o determinante *de uma matriz é o* determinante *da transformação linear correspondente.*

A notação da definição 13.4 usando permutações é bastante útil se soubermos trabalhar com ela. A melhor maneira de entender direitinho o significado dessa notação, é fazendo alguns casos manualmente.

Exemplo 13.5 (caso 2×2)**.** Seja

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Na notação da definição 13.4, $a_{11} = a$, $a_{12} = b$, $a_{21} = c$ e $a_{22} = d$.

As permutações de $\{1, 2\}$ são $(1, 2)$ e $(2, 1)$. O sinal da primeira é 1 e o da segunda é -1. Então,

$$\begin{aligned} \det(A) &= a_{11}a_{22} + (-1)a_{21}a_{12} \\ &= ad - cb. \end{aligned}$$

Vejamos o caso 3×3.

Exemplo 13.6. Considere a matriz

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}^t.$$

As permutações de $\{1, 2, 3\}$ que tem sinal positivo (permutações pares) são as permutações circulares: $\rho^0 = (1, 2, 3)$, $\rho^1 = (2, 3, 1)$ e $\rho^2 = (3, 1, 2)$. A permutação ρ^0 não faz nada, e equivale a somar 0. Assim, ρ^1 corresponde a somar 1, e ρ^2 corresponde a somar 2.

Ao fazer o produto dos números nas entradas das matrizes, ρ^0 corresponde à *diagonal principal*. E de modo geral, aplicar ρ significa, na matriz **transposta**, deslocar a diagonal para a direita. Aplicar duas vezes significa deslocar para a direita duas vezes. Se não estivéssemos tratando da transposta, deveríamos deslocar para baixo, já que em nossa fórmula a permutação é aplicada no primeiro índice, que se refere à linha da matriz. Na proposição 13.9, veremos que o determinante de uma matriz é igual ao determinante de sua transposta.

Figura 13.10: Regra de *Sarrus*. A diagonal principal corresponde à permutação identidade, que não permuta nada. Cada diagonal corresponde a uma permutação. As diagonais que apontam pra baixo tem sinal $+1$. As que apontam pra cima tem sinal -1.

As permutações ímpares são da forma $\gamma = (3, 2, 1)$, $\rho^1\gamma = (1, 3, 2)$ e $\rho^2\gamma = (2, 1, 3)$. A primeira, γ, corresponde à diagonal não principal. A segunda e a terceira correspondem a deslocar a diagonal principal para a direita 1 ou 2 vezes.

Para ver que o *determinante* é realmente a função que procurávamos, precisamos verificar os três itens enumerados antes da proposição 13.3. Isso fica como dever de casa! :-)

Proposição 13.7. *Vetores* $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^n$ *são uma base de* $\mathbb{R}^n$ *se, e somente se,*

$$\det(\vec{v}_1, \dots, \vec{v}_n) = 0.$$

Demonstração. Já sabemos da **??**, que um conjunto de vetores de $\mathbb{R}^n$ é uma base se, e somente se, for linearmente independente. E isso acontece se, se somente se, o conjunto for gerador.

Se não for uma base, ou seja, se são linearmente dependentes, então um dos vetores é combinação linear dos demais. Se, por exemplo,

$$\vec{v}_n = \alpha_1\vec{v}_1 + \cdots + \alpha_{n-1}\vec{v}_{n-1},$$

então

$$\begin{aligned}
\det(\vec{v}_1, \ldots, \vec{v}_n) &= \det(\vec{v}_1, \ldots, \vec{v}_{n-1}, \alpha_1\vec{v}_1 + \cdots + \alpha_{n-1}\vec{v}_{n-1}) \\
&= \alpha_1 \det(\vec{v}_1, \ldots, \vec{v}_{n-1}, \vec{v}_1) + \cdots \\
&\qquad \cdots + \alpha_{n-1} \det(\vec{v}_1, \ldots, \vec{v}_{n-1}, \vec{v}_{n-1}) \\
&= \alpha_1 0 + \cdots + \alpha_{n-1} 0 \\
&= 0.
\end{aligned}$$

Por outro lado, se os vetores formam uma base, então podemos escrever

$$\begin{aligned}
\vec{e}_1 &= a_{11}\vec{v}_1 + \cdots + a_{1n}\vec{v}_n \\
&\vdots \\
\vec{e}_n &= a_{n1}\vec{v}_1 + \cdots + a_{nn}\vec{v}_n.
\end{aligned}$$

Nesse caso,

$$\begin{aligned}
1 &= \det(\vec{e}_1, \ldots, \vec{e}_n) \\
&= \sum_{\sigma \in \{1,\ldots,n\}^n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \det(\vec{v}_{\sigma(1)}, \ldots, \vec{v}_{\sigma(n)}) \\
&= \sum_{\sigma \in P_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \det(\vec{v}_{\sigma(1)}, \ldots, \vec{v}_{\sigma(n)}) \\
&= \left(\sum_{\sigma \in P_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \right) \det(\vec{v}_1, \ldots, \vec{v}_n).
\end{aligned}$$

E portanto, $\det(\vec{v}_1, \ldots, \vec{v}_n)$ não pode ser nulo.

Temos então um critério simples pra saber se uma matriz ou equivalentemente, uma transformação linear, é inversível.

Corolário 13.8. *Uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ é inversível se, e somente se,* $\det(T) \neq 0$.

Demonstração. Basta juntar a proposição 13.7 com o **??**, que diz que T é inversível se, e somente se, $T\vec{e}_1, \ldots, T\vec{e}_n$ for uma base.

Por vezes, o ponto de vista geométrico é mais adequado pra entender uma propriedade das matrizes. Por vezes, o ponto de vista algébrico é melhor. Agora que temos uma fórmula, podemos relacionar o determinante de uma matriz com o determinante de sua transposta.

Proposição 13.9. *Seja A uma matriz $n \times n$. Então*

$$\det\left(A^t\right) = \det(A).$$

Demonstração. As permutações tem três propriedades interessantes. O leitor é convidado a demonstrá-las.

1. A operação
$$\begin{array}{rccc} \iota : & P_n & \to & P_n \\ & \sigma & \mapsto & \sigma^{-1} \end{array}$$
é uma bijeção.

2. $\operatorname{sgn}(\sigma) = \operatorname{sgn}\left(\sigma^{-1}\right)$.

3. Se trocarmos a ordem do produto $a_{\sigma(1),1} \cdots a_{\sigma(n),n}$ usando a permutação γ, obtemos o mesmo resultado, pois o produto de números reais é comutativo:
$$a_{\sigma(\gamma(1)),\gamma(1)} \cdots a_{\sigma(\gamma(n)),\gamma(n)} = a_{\sigma(1),1} \cdots a_{\sigma(n),n}.$$

Usando essas propriedades,

$$\begin{aligned}
\det(A) &= \sum_{\sigma\in P_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1}\cdots a_{\sigma(n),n} \\
&= \sum_{\sigma\in P_n} \operatorname{sgn}(\sigma) a_{\sigma(\sigma^{-1}(1)),\sigma^{-1}(1)}\cdots a_{\sigma(\sigma^{-1}(n)),\sigma^{-1}(n)} \\
&= \sum_{\sigma\in P_n} \operatorname{sgn}(\sigma) a_{1,\sigma^{-1}(1)}\cdots a_{n,\sigma^{-1}(n)} \\
&= \sum_{\sigma\in P_n} \operatorname{sgn}\left(\sigma^{-1}\right) a_{1,\sigma^{-1}(1)}\cdots a_{n,\sigma^{-1}(n)} \\
&= \sum_{\sigma\in P_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)}\cdots a_{n,\sigma(n)} \\
&= \det\left(A^t\right).
\end{aligned}$$

13.4 Propriedades

xxx

Aula 14

Produto Vetorial

Geometricamente, imaginamos o determinante 3×3 como uma espécie de *volume com sinal*. Do ponto de vista algébrico, o determinante nada mais é do que uma aplicação trilinear alternada tal que $\det(\vec{e}_1, \vec{e}_2, \vec{e}_3) = 1$. Vamos usar as propriedades algébricas para definir o chamado *produto vetorial*.

14.1 Definição

O determinante é multilinear. Isso significa que se escolhermos $\vec{a}, \vec{b} \in \mathbb{R}^3$,

$$\begin{array}{rrcl} f: & \mathbb{R}^3 & \to & \mathbb{R} \\ & \vec{v} & \mapsto & \det(\vec{v}, \vec{a}, \vec{b}) \end{array}$$

é um funcional linear. Como todo funcional linear de $\mathbb{R}^n$, existe $\vec{w} \in \mathbb{R}^3$ tal que

$$f(\vec{v}) = \vec{w} \cdot \vec{v}.$$

Ou seja, para todo $\vec{v} \in \mathbb{R}^3$,

$$\det(\vec{v}, \vec{a}, \vec{b}) = \vec{w} \cdot \vec{v}.$$

Definição 14.1 (Produto vetorial)**.** *Dados $\vec{a}, \vec{b} \in \mathbb{R}^3$, o* produto vetorial *de* $\vec{a}$ *por* $\vec{b}$*, denotado por* $\vec{a} \times \vec{b}$ *é o vetor de* $\mathbb{R}^3$ *tal que*

$$\det(\vec{v}, \vec{a}, \vec{b}) = \left(\vec{a} \times \vec{b}\right) \cdot \vec{v}.$$

14.2 Interpretação Geométrica

Dados $\vec{a}, \vec{b} \in \mathbb{R}^3$, podemos pensar no determinante $\det(\vec{v}, \vec{a}, \vec{b})$ como o "volume com sinal" de um prisma com base formada pelo paralelogramo de lados $\vec{a}$ e $\vec{b}$. Nesse caso, o volume pode ser calculado pela fórmula

$$\text{volume} = \text{área da base} \times \text{altura},$$

onde a altura é a projeção de $\vec{v}$ na direção de um vetor $\vec{n}$ ortogonal aos vetores $\vec{a}$ e $\vec{b}$. Se permitirmos uma "altura com sinal", temos o "volume com sinal"

$$\det(\vec{v}, \vec{a}, \vec{b}) = AP_{\vec{n}}(\vec{v}),$$

onde A é a áread do paralelogramo de lados $\vec{a}$ e $\vec{b}$, e $\vec{n}$ pode ser tomado unitário e com sentido de modo que o sinal do determinante coincida com o sinal da projeção. Em outras palavras,

$$\det(\vec{v}, \vec{a}, \vec{b}) = A\vec{n} \cdot \vec{v}.$$

Ou seja,

$$\vec{a} \times \vec{b} = A\vec{n}$$

é exatamente o vetor com norma A, direção ortogonal a $\vec{a}$ e $\vec{b}$, e sentido tal que o sinal do produto interno coincida com o sinal do determinante.

Apesar da interpretação geométrica, podemos chegar à conclusão de que $\vec{a} \perp (\vec{a} \times \vec{b})$ e $\vec{b} \perp (\vec{a} \times \vec{b})$, apenas usando as propriedades algébricas do determinante. De fato, por exemplo,

$$\begin{aligned} \vec{a} \cdot (\vec{a} \times \vec{b}) &= \det(\vec{a}, \vec{a}, \vec{b}) \\ &= 0. \end{aligned}$$

14.3 Propriedades Elementares

As propriedades elementares do produto vetorial são consequência direta das propriedades elementares do determinante 3×3.

Já vimos que $\vec{a}$ e $\vec{b}$ são ortogonais a$\vec{a} \times \vec{b}$. Além disso,

$$\vec{a} \times \vec{b} = -\vec{b} \times \vec{a},$$

pois

$$\det(\vec{v}, \vec{a}, \vec{b}) = \det(\vec{v}, \vec{b}, \vec{a}).$$

E o produto vetorial também é bilinear:

$$(\alpha\vec{a} + \beta\vec{b}) \times \vec{c} = \alpha(\vec{a} \times \vec{c}) + \beta(\vec{b} \times \vec{c}),$$

pois

$$\begin{aligned}
\det(\vec{v}, \alpha\vec{a} + \beta\vec{b}, \vec{c}) &= \alpha \det(\vec{v}, \vec{a}, \vec{c}) + \beta \det(\vec{v}, \vec{b}, \vec{c}) \\
&= \alpha\vec{v} \cdot \vec{a} \times \vec{c} + \beta\vec{v} \cdot \vec{b} \times \vec{c} \\
&= \vec{v} \cdot (\alpha(\vec{a} \times \vec{c}) + \beta(\vec{a} \times \vec{c})).
\end{aligned}$$

Por fim, temos a propriedade que foi usada para definir o produto vetorial. Para quaisquer $\vec{v}, \vec{a}, \vec{b} \in \mathbb{R}^3$,

$$\vec{v} \cdot (\vec{a} \times \vec{b}) = \det(\vec{v}, \vec{a}, \vec{b}).$$

14.4 Fórmula

Uma maneira de determinar as coordenadas de um vetor $\vec{v} \in \mathbb{R}^n$, é fazendo o produto interno com a base canônica $\vec{e}_1, \ldots, \vec{e}_n$. Sendo assim, a j-ésima coordenada de $\vec{c} = \vec{a} \times \vec{b}$ é dada por

$$\begin{aligned}
c_j &= \vec{c} \cdot \vec{e}_j \\
&= \det(\vec{e}_j, \vec{a}, \vec{b}).
\end{aligned}$$

Assim, pela **??**,

$$c_1 = \det \begin{bmatrix} a_2 & b_2 \\ a_3 & b_3 \end{bmatrix} \qquad c_2 = -\det \begin{bmatrix} a_1 & b_1 \\ a_3 & b_3 \end{bmatrix} \qquad c_3 = \det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}.$$

Perceba que para fazer o produto interno de $\vec{v} = (v_1, v_2, v_3)$ por $\vec{c} = c_1\vec{e}_1 + c_2\vec{e}_2 + c_3\vec{e}_3$, basta substituir $\vec{e}_1$, $\vec{e}_2$ e $\vec{e}_3$ por v_1, v_2 e v_3 na expressão de $\vec{c}$ na base canônica:

$$\vec{v} \cdot \vec{c} = c_1 v_1 + c_2 v_2 + c_3.$$

Sendo assim, uma maneira mais fácil de lembrar essa fórmula para encontrar $\vec{c}$, é calculando

$$\vec{v} \cdot \vec{c} = \det(\vec{v}, \vec{a}, \vec{b}),$$

mas substituindo v_1, v_2 e v_3 por $\vec{e}_1$, $\vec{e}_2$ e $\vec{e}_3$. Ou seja, "formalmente",

$$\vec{c} = \det \begin{bmatrix} \vec{e}_1 & a_1 & b_1 \\ \vec{e}_2 & a_2 & b_2 \\ \vec{e}_3 & a_3 & b_3 \end{bmatrix}.$$

Em geral, por algum motivo, as pessoas preferem usar a matriz transposta — já que o determinante de uma matriz é igual ao determinante da transposta — e substituir $\vec{e}_1$, $\vec{e}_2$ e $\vec{e}_3$ por $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ e $\hat{\mathbf{k}}$:

$$\vec{c} = \det \begin{bmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}.$$

Parte IV

Técnicas Algébricas

Aula 15

Transformações Multilineares Alternadas

O determinante é uma transformação multilinear alternada. Nessa lição, vamos tratar de transformações que tem propriedades semelhantes ao determinante.

15.1 Definição

Fixado $\vec{a} \in \mathbb{R}^n$,

$$\begin{array}{rrcl} \langle \vec{a}, \cdot \rangle : & \mathbb{R}^n & \to & \mathbb{R} \\ & \vec{v} & \mapsto & \vec{a} \cdot \vec{v} \end{array} \qquad \text{e} \qquad \begin{array}{rrcl} \langle \cdot, \vec{a} \rangle : & \mathbb{R}^n & \to & \mathbb{R} \\ & \vec{v} & \mapsto & \vec{v} \cdot \vec{a} \end{array}$$

são transformações lineares. Por isso, dizemos que o produto interno

$$\begin{array}{rrcl} \langle \cdot, \cdot \rangle : & \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ & (\vec{v}, \vec{w}) & \mapsto & \vec{v} \cdot \vec{w} \end{array}$$

é *bilinear*.

O determinante $n \times n$

$$\begin{array}{rrcl} \det : & \overbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}^{n\text{-vezes}} & \to & \mathbb{R} \\ & (\vec{v}, \vec{w}) & \mapsto & \vec{v} \cdot \vec{w} \end{array}$$

é uma aplicação n-linear. Por exemplo, fixados $\vec{a}_2, \cdots, \vec{a}_n \in \mathbb{R}^n$, a aplicação

$$\begin{array}{rccl} T: & \mathbb{R}^n & \to & \mathbb{R} \\ & \vec{v} & \mapsto & \det(\vec{v}, \vec{a}_2, \cdots, \vec{a}_n) \end{array}$$

Definição 15.1 (Transformação Multilinear)**.** *Dizemos que uma transformação*

$$T : V_1 \times \cdots \times V_n \to W,$$

onde $V_1, \ldots, V_n$ *e* W *são espaços vetoriais é* n-linear *quando fixados* $\vec{a}_1 \in V_1, \ldots, \vec{a}_n \in V_n$*, cada uma das transformações*

$$\begin{array}{rccl} T_j: & V_j & \to & W \\ & \vec{v}_j & \mapsto & T(\vec{a}_1, \ldots, \vec{a}_{j-1}, \vec{a}_j, \ldots, \vec{a}_n) \end{array}$$

é linear. Quando não queremos enfatizar o número n*, dizemos apenas que* T *é* multilinear.

Exemplo 15.2. Se $T : \mathbb{R}^p \to \mathbb{R}^n$ é linear, então

$$\begin{array}{rccl} J: & \mathbb{R}^n \times \mathbb{R}^p & \to & \mathbb{R} \\ & (\vec{v}, \vec{w}) & \mapsto & \vec{v} \cdot (T\vec{w}) \end{array}$$

é bilinear.

Exemplo 15.3. Vamos identificar as matrizes 1×1 com o conjunto dos números reais. Seja V o conjunto das matrizes $1 \times n$, e W o das matrizes $p \times 1$. Se M é uma matriz $n \times p$, então

$$\begin{array}{rccl} J: & V \times W & \to & \mathbb{R} \\ & (A, B) & \mapsto & AMB \end{array}$$

é bilinear. De fato, essa transformação é essencialmente a mesma do exemplo 15.2.

O determinanter também tem a propriedade de ser uma transformação linear *alternada*.

Definição 15.4 (Transformação multilinear alternada)**.** *Se $T : V^n \to W$ é uma transformação n-linear, dizemos que T é* alternada *quando para qualquer permutação $\sigma \in P_n$,*

$$T(\vec{a}_{\sigma(1)}, \vec{a}_{\sigma(2)}, \dots, \vec{a}_{\sigma(n)}) = \operatorname{sgn} \sigma T(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n).$$

A definição 15.4 é mais fácilmente compreendida se dissermos simplesmente que

quando trocamos dois vetores de lugar, o sinal de T muda.

Exemplo 15.5 (O determinante é alternado)**.** Sejam $\vec{a}_j = (a_{j1}, \dots, a_{jn}) \in \mathbb{R}^n$ e $\sigma \in P_n$. Então,

$$\begin{aligned}
\det(\vec{a}_{\sigma(1)}, \dots, \vec{a}_{\sigma(n)}) &= \sum_{\gamma \in P_n} \operatorname{sgn} \gamma a_{\gamma(\sigma(1))} \cdots a_{\gamma(\sigma(n))} \\
&= \operatorname{sgn} \sigma^{-1} \sum_{\gamma \in P_n} \operatorname{sgn} \gamma \operatorname{sgn} \sigma a_{\gamma(\sigma(1))} \cdots a_{\gamma(\sigma(n))} \\
&= \operatorname{sgn} \sigma^{-1} \sum_{\gamma \in P_n} \operatorname{sgn} \gamma \circ \sigma a_{(\gamma\circ\sigma)(1)} \cdots a_{(\gamma\circ\sigma)(n)} \\
&= \operatorname{sgn} \sigma^{-1} \sum_{\beta \in P_n} \operatorname{sgn} \beta a_{\beta(1)} \cdots a_{\beta(n)} \\
&= \operatorname{sgn} \sigma^{-1} \det(\vec{a}_1, \dots, \vec{a}_n) \\
&= \operatorname{sgn} \sigma \det(\vec{a}_1, \dots, \vec{a}_n).
\end{aligned}$$

Ao falar de transformações multilineares alternadas, tem gente que gosta de dizer que

> "quando trocamos dois vetores, o sinal muda".

Tem gente que prefere dizer que

> "quando permutamos os vetores, o valor da transformação é multiplicado pelo sinal da permutação".

Existem maneira diferentes de caracterizar o significa uma transformação linear ser alternada. Qualquer uma dessas caracterizações poderia ter sido usada no lugar da definição 15.4.

Proposição 15.6. *Dada uma transformação n-linear $T : V^n \to W$. São equivalentes*

1. *Se $\sigma \in P_n$,*
$$T(\vec{a}_{\sigma(1)}, \vec{a}_{\sigma(2)}, \dots, \vec{a}_{\sigma(n)}) = \operatorname{sgn} \sigma T(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n).$$
2. *Ao alternarmos a posição de dois vetores em*
$$T(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n),$$
o "resultado" é multiplicado por -1.
3. *Se os vetores $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ são linearmente dependentes, então*
$$T(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n) = 0.$$
4. *Se dois vetores em $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ são iguais, então*
$$T(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n) = 0.$$

Demonstração.

15.2 Inexistência

Talvez a propriedade mais importante do determinante $n \times n$ seja sua própria existência!

> Para todo $n \in \mathbb{N}^*$, existe uma transformação multilinear alternada não nula
> $$D : \overbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}^{n\text{-vezes}} \to \mathbb{R}.$$

Proposição 15.7. *Sejam $p, q \in \mathbb{N}^*$ tais que $p > q$. Então **não existe** uma transformação multilinear alternada **não nula***

$$D : \overbrace{\mathbb{R}^q \times \cdots \times \mathbb{R}^q}^{p\text{-vezes}} \to \mathbb{R}.$$

Demonstração. Seja D uma transformação multilinear alternada como no enunciado. Vamos mostrar que D é nula. Usando a multilinearidade, quais quer que sejam $\vec{a}_1, \ldots, \vec{a}_p \in \mathbb{R}^q$,

$$D(\vec{a}_1, \ldots, \vec{a}_p) = \sum_{f:[p]\to[q]} a_{1f(1)} \cdots a_{pf(p)} D(\vec{e}_{f(1)}, \ldots, \vec{e}_{f(p)}).$$

No entanto, como $q < p$, sempre tem ao menos um vetor repetido em $\vec{e}_{f(1)}, \ldots, \vec{e}_{f(p)}$.

Assim, pela alternância de D (item 4 da proposição 15.6), todos os termos da soma são iguais a 0.

15.3 Determinante Composto com Transformação Linear

Se V é um espaço vetorial e $T : V \to \mathbb{R}^n$ é uma transformação linear, então podemos construir a transformação multilinear alternada

$$\begin{array}{rccl} D_T : & \overbrace{V \times \cdots \times V}^{n\text{-vezes}} & \to & \mathbb{R} \\ & (\vec{v}_1, \ldots, \vec{v}_n) & \mapsto & \det(T\vec{v}_1, \ldots, T\vec{v}_n) \end{array}.$$

Proposição 15.8. *Se $p > q$, então* ***não existe*** *uma tranformação linear sobrejetiva $T : \mathbb{R}^q \to \mathbb{R}^p$.*

Em particular, se $p \neq q$, não existe uma bijeção linear entre $\mathbb{R}^p$ e $\mathbb{R}^q$.

Demonstração. Se existir uma tal T, então

$$\begin{array}{rccl} D_T : & \overbrace{\mathbb{R}^q \times \cdots \times \mathbb{R}^q}^{p\text{-vezes}} & \to & \mathbb{R} \\ & (\vec{v}_1, \ldots, \vec{v}_n) & \mapsto & \det(T\vec{v}_1, \ldots, T\vec{v}_n) \end{array}.$$

é uma transformação multilinear alternada não nula. Mas isso é impossível pela proposição 15.7.

15.4 Cardinalidade da Base

Se $\mathscr{B} = (\vec{b}_1, \ldots, \vec{b}_n)$ é uma base ordenada do espaço vetorial V, então podemos construir

$$\begin{array}{rccl} T_{\mathscr{B}} : & \mathbb{R}^n & \to & V \\ & (\alpha_1, \ldots, \alpha_n) & \mapsto & \alpha_1\vec{b}_1 + \cdots + \alpha_n\vec{b}_n \end{array}.$$

Como $\mathscr{B}$ é uma base, a aplicação $T_{\mathscr{B}}$ é uma bijeção.

Será que todo espaço vetorial possui uma base finita? Isso será respondido na capítulo 16.

Proposição 15.9. *Se V é um espaço vetorial com bases*

$$\mathscr{B} = (\vec{b}_1, \dots, \vec{b}_n)$$
$$\mathscr{C} = (\vec{c}_1, \dots, \vec{c}_m),$$

então $m = n$.

Demonstração. As aplicações

$$\begin{array}{rrcl} T_{\mathscr{B}}: & \mathbb{R}^n & \to & V \\ & (\alpha_1, \dots, \alpha_n) & \mapsto & \alpha_1 \vec{b}_1 + \cdots + \alpha_n \vec{b}_n \\ S_{\mathscr{C}}: & \mathbb{R}^m & \to & V \\ & \alpha_1 \vec{c}_1 + \cdots + \alpha_m \vec{c}_m & \mapsto & (\alpha_1, \dots, \alpha_m) \end{array}$$

são bijeções lineares. Portanto,

$$S_{\mathscr{C}} \circ T_{\mathscr{B}} : \mathbb{R}^n \to \mathbb{R}^m$$

é uma bijeção linear. Mas pela proposição 15.8, isso só pode acontecer quando $m = n$.

Observação 15.10. Na proposição 15.9, mostramos que se V tem duas bases finitas, as cardinalidades são as mesmas. Mas ainda não sabemos responder:

1. Será que todo espaço vetorial tem uma base?
2. Será que toda base é finita?
3. Será que todo conjunto gerador contém uma base?

4. Será que todo conjunto linearmente independente pode ser auemntado para se tornar uma base?

15.5 Unicidade do Determinante

Se $T : V \to W$ é uma transformação linear,

$$\begin{array}{rcl} 5T: \; V & \to & W \\ \vec{v} & \mapsto & 5T\vec{v} \end{array}$$

também é linear. Da mesma forma, se $M : V^n \to W$ é n-linear (alternada), então, para qualquer $\alpha \in \mathbb{R}$,

$$\begin{array}{rrcl} \alpha M: & V^n & \to & W \\ & (\vec{v}_1, \ldots, \vec{v}_n) & \mapsto & \alpha T(\vec{v}_1, \ldots, \vec{v}_n) \end{array}$$

também é.

A proposição a seguir mostra que toda transformação linear alternada

$$D : \overbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}^{n\text{-vezes}} \to \mathbb{R}.$$

é da forma $\alpha \det$ para algum $\alpha \in \mathbb{R}$.

Proposição 15.11. *Seja*

$$D : \overbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}^{n\text{-vezes}} \to \mathbb{R}$$

n-linear alternada. Então $D = \alpha \det$, onde

$$\alpha = D(\vec{e}_1, \ldots, \vec{e}_n).$$

Demonstração. Usando a *"distributividade"* de D e também sua alternância,

$$\begin{aligned} D(\vec{a}_1, \dots, \vec{a}_n) &= \sum_\sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)} D(\vec{e}_{\sigma(1)}, \dots, \vec{e}_{\sigma(n)}) \\ &= \sum_\sigma \operatorname{sgn}(\sigma) \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)} D(\vec{e}_1, \dots, \vec{e}_n) \\ &= \det(\vec{a}_1, \dots, \vec{a}_n) D(\vec{e}_1, \dots, \vec{e}_n). \end{aligned}$$

15.6 Determinante da Composta

Considere as transformação linear

$$T : \mathbb{R}^n \to \mathbb{R}^n.$$

Usando o determinante, podemos construir a transformação multilinear alternada

$$\begin{array}{rrcl} D_T : & \mathbb{R}^n \times \cdots \times \mathbb{R}^n & \to & \mathbb{R} \\ & (\vec{v}_1, \dots, \vec{v}_n) & \mapsto & \det(T\vec{v}_1, \dots, T\vec{v}_n) \end{array}.$$

Pela proposição 15.11, sabemos que D_T é um múltiplo de det.

Lema 15.12. *A constante $\alpha \in \mathbb{R}$ que satisfaz $D_T = \alpha \det$ é*

$$\alpha = \det(T).$$

Dito de outra forma,

$$\det(T\vec{v}_1, \dots, T\vec{v}_n) = \det(T) \det(\vec{v}_1, \dots, \vec{v}_n).$$

Demonstração. A primeira parte é uma reformulação da proposição 15.11. Ou seja,

$$\begin{aligned}\alpha &= D_T(\vec{e}_1, \dots, \vec{e}_n) \\ &= \det(T\vec{e}_1, \dots, T\vec{e}_n) \\ &= \det(T).\end{aligned}$$

Onde a última igualdade nada mais é do que nossa definição de $\det(T)$. Veja a definição 13.4.

A segunda parte é só uma questão de substituir a definição de α e D_T.

$$\begin{aligned}\det(T\vec{v}_1, \dots, T\vec{v}_n) &= D_T(\vec{v}_1, \dots, \vec{v}_n) \\ &= \det(T)\det(\vec{v}_1, \dots, \vec{v}_n).\end{aligned}$$

O lema 15.12 mostra uma outra maneira de definir o conceito de determinante de uma transformação linear. O determinante da transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ é o número $\det(T)$ que satisfaz

$$\det(T\vec{v}_1, \dots, T\vec{v}_n) = \det(T)\det(\vec{v}_1, \dots, \vec{v}_n).$$

Em particular, se T é sobrejetiva, então $\det(T) \neq 0$. Para verificar isso, basta tomar $\vec{v}_1, \dots, \vec{v}_n$ tais que $T\vec{v}_1 = \vec{e}_1, \dots, T\vec{v}_n = \vec{e}_n$, para obter

$$1 = \det(T\vec{v}_1, \dots, T\vec{v}_n) = \det(T)\det(\vec{v}_1, \dots, \vec{v}_n).$$

Agora, com duas transformações lineares

$$S : \mathbb{R}^n \to \mathbb{R}^n \quad \text{e} \quad T : \mathbb{R}^n \to \mathbb{R}^n,$$

podemos falar em $\det(S)$, $\det(T)$, $\det(S \circ T)$ e $\det(T \circ S)$.

Proposição 15.13. *Sejam*

$$S : \mathbb{R}^n \to \mathbb{R}^n \quad e \quad T : \mathbb{R}^n \to \mathbb{R}^n$$

transformações lineares. Então

$$\det(S \circ T) = \det(S)\det(T).$$

Demonstração. Basta calcular $\det(S \circ T)$. Para facilitar o entendimento, vamos escrever $\vec{v}_1 = T\vec{e}_1, \ldots, \vec{v}_n = T\vec{e}_n$ e utilizar a proposição 15.11.

$$\begin{aligned}
\det(S \circ T) &= \det(S(T\vec{e}_1), \ldots, S(T\vec{e}_n)) \\
&= \det(S\vec{v}_1, \ldots, S\vec{v}_n) \\
&= \det(S)\det(\vec{v}_1, \ldots, \vec{v}_n) \\
&= \det(S)\det(T\vec{e}_1, \ldots, \vec{e}_n) \\
&= \det(S)\det(T).
\end{aligned}$$

Aula 16

Dimensão

O espaço $\mathbb{R}^n$ tem dimensão n. A base $\vec{e}_1, \ldots, \vec{e}_n$ é formada por n vetores. Já sabemos, por tanto, que todas as bases **finitas** de $\mathbb{R}^n$ tem exatamente n vetores.

O objetivo deste capítulo é mostrar que

1. Todo espaço vetorial tem uma base.
2. Todas as bases de $V \subset \mathbb{R}^n$ tem a mesma cardinalidade $m \leq n$. Dizemos que m é a dimensão de V.

Na verdade, ao invés do item 1, vamos mostrar que

> Todo conjunto subconjunto linearmente independente de um espaço vetorial $V \subset \mathbb{R}^n$ pode ser estendido a uma base. Ou seja, se $S \subset V$ é linearmente independente, então existe uma base $B \subset V$ tal que $S \subset B$.

16.1 Excluindo Vetores Linearmente Dependentes

Se V é um espaço vetorial, todo subconjunto $S \subset V$ gera um subespaço de V. Se S é linearmente independente, então é base do espaço que gera.

Dizer que S é linearmente independente, é o mesmo que dizer que se retirarmos um vetor qualquer de S, o subespaço gerado será menor. E dizer que S **não é** linearmente independente, é o mesmo que dizer que existe um vetor em S que pode ser retirado de S sem modificar o subespaço gerado. Ou seja, S é *linearmente dependente* quando existe $\vec{v} \in S$ tal que

$$\text{gen}\,\{S \setminus \{\vec{v}\}\} = \text{gen}\,\{S\}.$$

Veja a **??**.

Proposição 16.1. *Seja V um espaço vetorial, $L \subset V$ um subconjunto linearmente independente e $G \subset V$ um conjunto gerador* ***finito*** *que contém L. Então existe uma base B de V tal que*

$$L \subset B \subset G.$$

Demonstração. Enquanto G for linearmente dependente, podemos retirar um elemento de G e construir um conjunto $G_1 \subsetneq G$ que gera o mesmo subespaço que G. Podemos ainda fazer mais, e garantir que o elemento retirado de G não é um elemento de L. Assim, garantiremos que o gerador G_1 é tal que

$$L \subset G_1 \subsetneq G.$$

Como G é finito, esse processo de retirada de vetores só pode ser repetido um número finito de vezes. Podemos então escolher $G_n \subset \cdots \subset G_1 \subset G$ por esse processo, de modo que para todo j,

$$\begin{aligned} L &\subset G_j \\ \text{gen}\,\{G_j\} &= V, \end{aligned}$$

e de modo que não seja possível retirar mais nenhum elemento de G_n sem deixar de gerar V.

Se não podemos retirar nenhum elemento de G_n sem reduzir o espaço gerado, então G_n é linearmente independente. Como é também gerador, concluímos que G_n é base de V que contém L.

Só ficou faltando mostrar que de fato, na construção de G_j podemos retirar um elemento que não está em L. Isso fica como exercício: **??**.

O argumento da proposição 16.1 pode ser feito de modo mais elegante. Faremos isso na seção 16.2.

Observação 16.2. Perceba que um conjunto linearmente independente L pode ser vazio.

$$\operatorname{gen}\{\emptyset\} = \left\{\vec{0}\right\}.$$

16.2 Estendendo Conjuntos Linearmente Independentes

Queremos começar com um conjunto linearmente independente e aumentá-lo de modo a construir uma base.

Lema 16.3. *Se V é um espaço vetorial e $L \subset V$ é linearmente independente, então, para qualquer*

$$\vec{v} \in V \setminus \operatorname{gen}\{L\},$$

$L \cup \{\vec{v}\}$ é linearmente independente e estritamente maior que L.

Demonstração. Suponha que $M = L \cup \{\vec{v}\}$ seja linearmente dependente. Então, existem vetores distintos $\vec{v}_1, \dots, \vec{v}_k \in M$ e $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ não nulos tais que

$$\alpha_1\vec{v}_1 + \cdots + \alpha_k\vec{v}_k = \vec{0}.$$

Como L é linearmente independente, algum dos $\vec{v}_j$ tem que ser igual a $\vec{v}$. Reordenando os índices, vamos assumir que $\vec{v}_1 = \vec{v}$. Mas isso significaria que

$$\vec{v} = \frac{1}{\alpha_1} (\alpha_2 \vec{v}_2 + \cdots + \alpha_k \vec{v}_k) \in \operatorname{gen}\{L\}.$$

Mas como $\vec{v}$ foi escolhido de modo que $\vec{v} \notin \operatorname{gen}\{L\}$, podemos concluir que M não é linearmente dependente.

Note que o vetor $\vec{v}$ do lema 16.3 só existe quando L não é gerador. Ou seja, quando L não é uma base. O lema pode ser entendido como:

> Todo conjunto linearmente independente que não é uma base pode ser "aumentado" de modo que o conjunto aumentado continue sendo linearmente independente.

Ou então:

> Se um conjunto linearmente independente B não pode ser "aumentado" presevando a independência linear, então B é uma base.

Agora, a prometida demonstração mais elegante da proposição 16.1.

(demonstração mais elegante da proposição 16.1). Seja $B \subset G$ um subconjunto de G que contém L, que é linearmente independente, e que não pode ser aumentado sem perder essas propriedades. Note que B existe porque G é finito e L possuí as propriedades listadas: L contém L e é linearmente independente.

Ora, se B não pode ser aumentado, então, do lema 16.3,

$$G \subset \operatorname{gen}\{B\}.$$

Ou seja, B é gerador. Como é também linearmente indepentende, então B é uma base.

Observação 16.4. Na demonstração anterior, um matemático diria que B é um subconjunto de G que contém L e que é maximal. Não dizemos que B é o maior porque existem vários que satisfazem essas condições. Por exemplo,

$$B_1 = \{(1,0),(0,1)\}$$
$$B_2 = \{(1,0),(1,1)\}$$
$$B_3 = \{(1,0),(5,7)\}$$

são todos linearmente independentes que contém $L = \{(1,0)\}$, todos estão contidos em $G = \{(1,0),(0,1),(1,1),(5,7\}$ e nenhum deles pode ser "aumentado" sem deixar de ser linearmente independente.

Mesmo com esses resultados anteriores, será que não é possível ir "aumentando" indefinidamente o conjunto linearmente independente? Será que não podemos construir uma sequência infinita de conjuntos lineamente independentes

$$L_1 \subsetneq L_2 \subsetneq \cdots \subsetneq L_k \subsetneq \cdots ?$$

Proposição 16.5. *Todo subconjunto $L \subset \mathbb{R}^n$ que é linearmente independente tem no máximo n elementos.*

Demonstração. Se L tiver mais que n elementos, podemos escolher $S \subset L$ com exatamente $n+1$ elementos. Então, pela proposição 16.1, existe uma base $B \subset \mathbb{R}^n$ tal que

$$S \subset B \subset S \cup \{\vec{e}_1, \dots, \vec{e}_n\}.$$

Mas sabemos da proposição 15.9 que todas as bases finitas de $\mathbb{R}^n$ possuem exatamente n elementos. Ou seja, B tem n elementos e portanto S não pode ter mais que n elementos.

Corolário 16.6. *Se $V \subset \mathbb{R}^n$ é um espaço vetorial, então todas as bases de V são finitas e tem a mesma cardinalidade (quantidade de elementos).*

Demonstração. Sejam B_1 e B_2 bases de V. Então, pela proposição 16.5, ambos os conjuntos são finitos. Agora, da proposição 15.9, sabemos que B_1 e B_2 tem a mesma cardinalidade.

Agora que sabemos que todo espaço vetorial $V \subset \mathbb{R}^n$ possuem uma base, e que todas são finitas com a mesma cardinalidade, podemos definir o que é a *dimensão* do espaço vetorial V.

Definição 16.7. *Se $V \subset \mathbb{R}^n$ é um espaço vetorial, então sua* dimensão *é a cardinalidade de qualquer uma de suas bases.*

16.3 Organizando as Ideias

De tudo o que falamos até agora nesta lição, o que é importante mesmo são as ideias que desenvolvemos. Agora que já temos alguma experiência, podemos refazer tudo de uma forma muito mais prática! Vamos precisar apenas do lema 16.3 e da proposição 16.5. A proposição 16.5, por sua vez, utiliza a proposição 16.1, que tem uma "demonstração elegante" na seção 16.2.

Teorema 16.8. *Seja $V \subset \mathbb{R}^n$ um espaço vetorial, $G \subset V$ um gerador e $L \subset G$ um conjunto linearmente independente. Então existe $B \subset G$*

que é base de V e que contém L.

Demonstração. Seja $B \subset G$ um subconjunto de G maximal que contenha L e que seja linearmente independente. Tal B existe por causa da proposição 16.5.

Então, o lema 16.3 implica que B, além de linearmente independente, é também um gerador.

Agora que sabemos que todo espaço vetorial $V \subset \mathbb{R}^n$ possui uma base, e que todas são finitas com a mesma cardinalidade, podemos definir o que é a *dimensão* do espaço vetorial V.

Definição 16.9. *A* dimensão *do espaço vetorial $V \subset \mathbb{R}^n$ é a cardinalidade de qualquer uma de suas bases.*

Aula 17

Mudança de Base

Aula 18

Autovalores e Autovetores

xxx

Aula 19

Diagonalização

Matrizes diagonais.

Diagonalização.

Coelhos.

Dicas e Respostas dos Exercícios

Dicas

Respostas

Referências Bibliográficas

[Ban13] Pedro Bandeira, *O poeta e o cavaleiro*, 3ª ed., Editora Moderna, 2013.