Notas de Aula de GAAL

Ricardo Souza

6 de Maio de 2019

Conteúdo

Introdução

Este texto foi escrito como material auxiliar para um curso de GAAL ministrado em 2019/1.

Capítulo 1

Matrizes

1.1 Definições e Propriedades Básicas

Matrizes são simplesmente o nome matemático dado a tabelas de valores. Por exemplos, podemos ter matrizes numéricas

$$\begin{pmatrix} 1 & 2 & 9 \\ -8 & 5 & \pi \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -16 \end{pmatrix}, \cdots$$

ou de qualquer outra natureza, realmente

$$\begin{pmatrix} \text{Geometria} & \text{Analítica} \\ \text{Álgebra} & \text{Linear} \end{pmatrix}, \begin{pmatrix} \sqcup & \triangle & \triangledown \\ \odot & \nabla & \sum \end{pmatrix}, \cdots .$$

A princípio pode parecer meio arbitrário estudar esses objetos - e mais ainda definir operações e fazer matemática com eles. A motivação para isso ficará em um capítulo posterior. Neste momento, vamos dar uma motivação bem "rasa":

EXEMPLO(S):

Suponha que cinco amigos, Ana, Bernardo, Carlos, Diogo e Eliza resolveram sair para comemorar o aniversário de Carlos no boliche. Eles acabaram jogando três partidas e obtiveram os seguintes resultados:

$$\begin{array}{r c} & \begin{array}{ccccc} \text{Ana} & \text{Bernando} & \text{Carlos} & \text{Diogo} & \text{Eliza} \end{array} \\ \begin{array}{r} \text{Primeiro jogo} \\ \text{Segundo jogo} \\ \text{Terceiro jogo} \end{array} & \begin{pmatrix} 101 & 96 & 99 & 87 & 123 \\ 95 & 100 & 110 & 80 & 102 \\ 90 & 103 & 80 & 86 & 110 \end{pmatrix} \end{array}.$$

Inconformado com o resultado, Diogo praticou bastante. Alguns meses depois, em seu aniversário, convidou os amigos para repetirem a jogatina, e obtiveram os seguintes resultados:

$$\begin{array}{r c} & \begin{array}{ccccc} \text{Ana} & \text{Bernando} & \text{Carlos} & \text{Diogo} & \text{Eliza} \end{array} \\ \begin{array}{r} \text{Primeiro jogo} \\ \text{Segundo jogo} \\ \text{Terceiro jogo} \end{array} & \begin{pmatrix} 97 & 87 & 90 & 150 & 103 \\ 80 & 105 & 100 & 170 & 98 \\ 120 & 110 & 80 & 127 & 115 \end{pmatrix} \end{array}.$$

Surpresos com o resultados, eles resolveram computar quem tinha o maior total de pontos, combinando os dois aniversários. Com isso, eles obtiveram

$$\begin{array}{r c} & \begin{array}{ccccc} \text{Ana} & \text{Bernando} & \text{Carlos} & \text{Diogo} & \text{Eliza} \end{array} \\ \begin{array}{r} \text{Primeiro jogo} \\ \text{Segundo jogo} \\ \text{Terceiro jogo} \end{array} & \begin{pmatrix} 101+97 & 96+87 & 99+90 & 87+150 & 123+103 \\ 95+80 & 100+105 & 110+100 & 80+170 & 102+98 \\ 90+120 & 103+110 & 80+80 & 86+127 & 110+115 \end{pmatrix} \end{array}$$

$$
\begin{array}{c}
 \\
=
\end{array}
\begin{array}{ccccc}
\text{Ana} & \text{Bernando} & \text{Carlos} & \text{Diogo} & \text{Eliza} \\
\end{array}
$$

$$
= \begin{pmatrix} 198 & 183 & 189 & 237 & 226 \\ 175 & 205 & 210 & 250 & 200 \\ 210 & 213 & 160 & 213 & 225 \end{pmatrix}
$$

1.1.1 Somas e Produtos por Números

Definição 1.1.1. *Denotaremos por $M_{n\times m}(\mathbb{R})$ o* ***conjunto das matrizes com*** *n* ***linhas e*** *m* ***colunas, e com entradas em*** *$\mathbb{R}$. Caso $n = m$ diremos que nossas matrizes são* ***quadradas*** *e notaremos simplesmente $M_n(\mathbb{R})$.*

De maneira análoga, denotaremos o elemento na linha i, coluna j de uma matriz M por $M_{i,j}$.

EXEMPLO(S):

A matriz

$$M = \begin{pmatrix} 1 & 2 & 9 \\ -8 & 5 & \pi \end{pmatrix}$$

claramente pertence a $M_{2\times 3}(\mathbb{R})$. Além disso, temos: $M_{1,1} = 1, M_{1,2} = 2, M_{1,3} = 9, M_{2,1} = -8, M_{2,2} = 5$ e $M_{2,3} = \pi$.

Reciprocamente, se dissermos que uma matriz $N \in M_{3\times 2}(\mathbb{R})$ tem como entradas $N_{1,1} = 0 = N_{3,2}, N_{1,2} = 1, N_{2,1} = -1, N_{2,2} = \pi$ e $N_{3,1} = -\pi$, podemos recuperar a matriz N:

$$N = \begin{pmatrix} 0 & 1 \\ -1 & \pi \\ -\pi & 0 \end{pmatrix}$$

Desse exemplo tiramos uma informação muito importante: **uma matriz está unicamente determinada por seus elementos -** isto é, dada uma matriz M, então existe uma única coleção de elementos $M_{i,j}$; e dada qualquer coleção $a_{i,j}$ existe uma única matriz A cujos elementos são exatamente $A_{i,j} = a_{i,j}$.

Isso pode parecer trivial, mas nos permite, por exemplo, criar a seguinte definição:

Definição 1.1.2. *Dadas duas matrizes $M, N \in M_{n\times m}(\mathbb{R})$, definimos a* ***soma de*** *M* ***e*** *N como sendo uma matriz $M + N$ dada por*

$$(M + N)_{i,j} := M_{i,j} + N_{i,j}.$$

Observação 1.1.3. *Aqui, o símbolo "$x := y$" significa "estamos definindo x como sendo igual a y" ou "$x = y$ por definição".*

Observação 1.1.4. *Note que a definição acima faz todo sentido: para cada par de índices i, j, $M_{i,j}$ e $N_{i,j}$ são números reais (que nós já sabemos somar!) e portanto $M_{i,j} + N_{i,j}$ também é um número real. Nós estamos, então, coletando todas as somas, variando i e j, e chamando essa matriz, cujos elementos são somas dos elementos de M e N, de $M + N$.*

Exercício (1.1.1)

Do jeito que definimos, só sabemos somar matrizes que têm o mesmo número de linhas e colunas. Tente criar uma definição de soma de matrizes que funcione para quaisquer duas matrizes. Por que não usamos uma definição desse tipo?

Similarmente a como fizemos com a soma, definindo elemento a elemento, podemos também definir multiplicação por números:

Definição 1.1.5. *Seja $a \in \mathbb{R}$ um número real qualquer e $M \in M_{n\times m}(\mathbb{R})$ uma matriz $n \times m$ real. Definimos o* ***produto de** a **cópias de** M* *como sendo a matriz aM dada por*

$$(aM)_{i,j} := a \cdot M_{i,j}.$$

Novamente, pelo mesmo argumento acima, isso faz sentido, porque como a e cada $M_{i,j}$ são números reais (que nós sabemos multiplicar!), então $aM_{i,j}$ também é um número real.

EXEMPLO(S):

Sejam $M = \begin{pmatrix} 1 & -8 \\ 2 & 5 \\ 9 & \pi \end{pmatrix}$ e $N = \begin{pmatrix} 0 & 1 \\ -1 & \pi \\ -\pi & 0 \end{pmatrix}$ matrizes reais, e $a = \pi$. Então, podemos computar:

$$\begin{aligned} (M+N)_{1,1} &= M_{1,1} + N_{1,1} = 1 + 0 = 1 & (M+N)_{1,2} &= M_{1,2} + N_{1,2} = -8 + 1 = -7 \\ (M+N)_{2,1} &= M_{2,1} + N_{2,1} = 2 + (-1) = 1 & (M+N)_{2,2} &= M_{2,2} + N_{2,2} = 5 + \pi \\ (M+N)_{3,1} &= M_{3,1} + N_{3,1} = 9 + (-\pi) = 9 - \pi & (M+N)_{3,2} &= M_{3,2} + N_{3,2} = \pi + 0 = \pi \end{aligned}$$

e escrever

$$M + N = \begin{pmatrix} 1 & -7 \\ 1 & 5+\pi \\ 9-\pi & \pi \end{pmatrix}.$$

Similarmente, podemos computar

$$\begin{aligned} (\pi M)_{1,1} &= \pi \cdot 1 = \pi & (\pi M)_{1,2} &= \pi \cdot (-8) = -8\pi \\ (\pi M)_{2,1} &= \pi \cdot 2 = 2\pi & (\pi M)_{2,2} &= \pi \cdot 5 = 5\pi \\ (\pi M)_{3,1} &= \pi \cdot 9 = 9\pi & (\pi M)_{3,2} &= \pi \cdot \pi = \pi^2 \end{aligned}$$

e escrever

$$\pi M = \begin{pmatrix} \pi & -8\pi \\ 2\pi & 5\pi \\ 9\pi & \pi^2 \end{pmatrix}.$$

Exercício (1.1.2)

Calcule, usando os dados do exemplo acima, aN e $a(M+N)$. Em seguida, calcule $(M+M)+N$ e $2M+N$. O que podemos dizer dessas duas últimas matrizes?

1.1.2 Produtos(?)

Agora que sabemos somar matrizes e multiplicar matrizes por números, o mais natural seria definirmos uma multiplicação de matrizes. Poderíamos, intuitivamente, nos inspirar nas construções acima e definir que dadas duas matrizes $M, N \in M_{n\times m}(\mathbb{R})$, o produto delas será uma matriz $M \times N$ dada por

$$(M \times N)_{i,j} := M_{i,j} \cdot N_{i,j}$$

o que faria sentido, já que para cada par de índices i, j, ambos $M_{i,j}$ e $N_{i,j}$ são números reais - o que já sabemos multiplicar. Além disso, essa operação de produto teria excelentes propriedades: Seria comutativa, associativa, teria inverso, teria elemento neutro...

Por que então não definimos a multiplicação de matrizes assim?

A resposta simples é que o produto que vamos definir, apesar de não parecer intuitivo, é o que mais faz sentido quando consideramos as aplicações de matrizes que veremos mais à frente.

Definição 1.1.6 (Multiplicação Clássica)**.** *Dadas duas matrizes $M \in M_{n\times m}(\mathbb{R})$ e $N \in M_{m\times l}(\mathbb{R})$, definimos o* ***produto de** M **com** N* *como sendo uma matriz MN dada por*

$$(MN)_{i,j} := M_{i,1}N_{1,j} + M_{i,2}N_{2,j} + \cdots + M_{i,m}N_{m,j}.$$

Observação 1.1.7. *Novamente, convém notar que essa definição faz todo sentido, porque para cada trio de índices i, j e k, $M_{i,k}$ e $N_{k,j}$ são números reais (que nós sabemos multiplicar!) e portanto $M_{i,k}N_{k,j}$ também é um número real; e como $(MN)_{i,j}$ é uma soma de números reais, $(MN)_{i,j}$ é, em si, um número real. Além disso, essa definição explica a necessidade de o número de colunas de M ser o número de linhas de N: esse número é exatamente o número de termos da soma.*

Essa é a multiplicação de matrizes que já conhecemos desde sempre. Contudo, como comentamos previamente, ela "não faz sentido" - parece que surge do nada, e é desnecessariamente complicada.

Visando "facilitar" esse processo, vamos gastar um tempo tentando criar uma intuição de onde surge essa multiplicação de matrizes. Mas não se assuste - o resultado que vamos chegar é "o mesmo". A única diferença é que vamos explicar cada passo e tentar justificar essa definição.

Sendo assim, continuem calculando produtos de matrizes como sempre fizeram, mas atentem para os raciocínios que virão a seguir para entender de onde esses produtos vêm.

Definição 1.1.8. *Definimos por $e^n_{i,j} \in M_n(\mathbb{R})$ a matriz dada por:*

$$(e^n_{i,j})_{r,s} := \begin{cases} 1 \text{ se } r = i \text{ e } s = j \\ 0 \text{ caso contrário} \end{cases}.$$

EXEMPLO(S):

A matriz $e^2_{1,2} \in M_2(\mathbb{R})$ é a matriz 2×2 que tem 1 na linha 1, coluna 2, e 0 em todo o resto - ou seja,

$$e^2_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Já a matriz $e^3_{1,2} \in M_3(\mathbb{R})$ é a matriz 3×3 que tem 1 na linha 1, coluna 2, e 0 em todo o resto, ou seja,

$$e^3_{1,2} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Mais à frente, vamos estender essa definição para matrizes de tamanho arbitrário - não necessariamente quadradas. Por enquanto, vamos nos abster às matrizes quadradas para facilitar a vida.

Definição 1.1.9. *Definimos por $E^n_{i,j} : M_{n\times m}(\mathbb{R}) \to M_{n\times m}(\mathbb{R})$ a **função** dada por:*

$$\begin{array}{r} \\ \\ \\ j\text{-ésima linha} \\ \\ \\ \end{array} \begin{pmatrix} M_{1,1} & M_{1,2} & \cdots & M_{1,m} \\ M_{2,1} & M_{2,2} & \cdots & M_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ M_{j,1} & M_{j,2} & \cdots & M_{j,m} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n,1} & M_{n,2} & \cdots & M_{n,m} \end{pmatrix} \mapsto \begin{array}{r} \\ \\ \\ i\text{-ésima linha} \\ \\ \\ \end{array} \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ M_{j,1} & M_{j,2} & \cdots & M_{j,m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Ou seja, $E^n_{i,j}$ é a função que leva matrizes de n linhas (e qualquer quantidade de colunas) em matrizes de n linhas (e a mesma quantidade de colunas) simplesmente colando uma cópia da j-ésima linha da matriz original na i-ésima linha da imagem, e preenchendo o resto com 0s.

EXEMPLO(S):

A função $E^3_{1,2}$ aplicada na matriz $M = \begin{pmatrix} 1 & -8 \\ 2 & 5 \\ 9 & \pi \end{pmatrix}$ nos dá a matriz $E^3_{1,2}(M) = \begin{pmatrix} 2 & 5 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ em que nós simplesmente copiamos a segunda linha de M e colocamos na primeira linha da matriz nova, preenchendo o resto com 0s.

Analogamente, se $N = \begin{pmatrix} 0 & 1 \\ -1 & \pi \\ -\pi & 0 \end{pmatrix}$, então $E^3_{1,2}(N) = \begin{pmatrix} -1 & \pi \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ em que nós simplesmente copiamos a segunda linha de N e colocamos na primeira linha da matriz nova, preenchendo o resto com 0s.

Observação 1.1.10. *Aqui, o "expoente" da função simplesmente serve para indicar a quantidade de linhas das matrizes que você quer computar. Nos exemplos acima, o 3 no expoente de $E^3_{1,2}$ simplesmente indica que a função $E^3_{1,2}$ só pode ser aplicada em matrizes com três linhas.*

Exercício (1.1.3)

Usando os dados do exemplo acima, calcule $E^3_{1,1}, E^3_{1,3}, E^3_{2,1}, E^3_{2,2}, E^3_{2,3}, E^3_{3,1}, E^3_{3,2}$ e $E^3_{3,3}$ aplicadas tanto em M quanto em N.

O que podemos dizer sobre $E^3_{i,j}$ quando $i = j$? Ou seja, calcule $E^3_{1,1}, E^3_{2,2}$ e $E^3_{3,3}$ de M e N. O que essas matrizes têm de especial?

Exercício (1.1.4)

Usando o exemplo acima, existe algum par de índices i, j tal que $E^3_{i,j}(E^3_{i,j}(M)) = E^3_{i,j}(M)$? E para N?

Com isso temos nossa primeira noção de multiplicação de matrizes:

Definição 1.1.11. *Dadas uma matriz $M \in M_{n\times m}(\mathbb{R})$ qualquer e $e^n_{i,j} \in M_n(\mathbb{R})$, definimos o **produto de** $e^n_{i,j}$ **e** M como sendo a matriz $e^n_{i,j}M$ dada por*

$$e^n_{i,j}M := E^n_{i,j}(M).$$

EXEMPLO(S):

Ainda com os dados do exemplo anterior, podemos agora computar o produto

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & -8 \\ 2 & 5 \\ 9 & \pi \end{pmatrix} = e^3_{1,2}M := E^3_{1,2}(M) = \begin{pmatrix} 2 & 5 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

e

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -1 & \pi \\ -\pi & 0 \end{pmatrix} = e^3_{1,2}N := E^3_{1,2}(N) = \begin{pmatrix} -1 & \pi \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Observação 1.1.12. *Vamos mostrar, com alguns exemplos, que nossa definição até agora coincide com a definição clássica 1.1.6:*

$$\begin{pmatrix}0&1&0\\0&0&0\\0&0&0\end{pmatrix}\cdot\begin{pmatrix}1&-8\\2&5\\9&\pi\end{pmatrix}=\begin{pmatrix}0\cdot1+1\cdot2+0\cdot9&0\cdot(-8)+1\cdot5+0\cdot\pi\\0\cdot1+0\cdot2+0\cdot9&0\cdot(-8)+0\cdot5+0\cdot\pi\\0\cdot1+0\cdot2+0\cdot9&0\cdot(-8)+0\cdot5+0\cdot\pi\end{pmatrix}=\begin{pmatrix}2&5\\0&0\\0&0\end{pmatrix}$$

que é exatamente o resultado que obtivemos aplicando $E^3_{1,2}$ em M.

Exercício (1.1.5)

Calcule $E^4_{3,2}$ de uma matriz $M=\begin{pmatrix}a&b\\c&d\\e&f\\g&h\end{pmatrix}\in M_{4\times2}(\mathbb{R})$ genérica. Em seguida, calcule $e^4_{3,2}M$ usando a multiplicação clássica de matrizes e compare os resultados.

O próximo passo é considerar matrizes que são *quase* $e^n_{i,j}$. Por exemplo, como vamos multiplicar $\begin{pmatrix}0&8&0\\0&0&0\\0&0&0\end{pmatrix}$ por $M=\begin{pmatrix}1&-8\\2&5\\9&\pi\end{pmatrix}$? Bom, podemos perceber que $\begin{pmatrix}0&8&0\\0&0&0\\0&0&0\end{pmatrix}=8\cdot\begin{pmatrix}0&1&0\\0&0&0\\0&0&0\end{pmatrix}=8\cdot e^3_{1,2}$ e que nós já sabemos multiplicar números por matrizes e $e^3_{1,2}$ por M.

Dito de outra maneira, queremos calcular $(8\cdot e^3_{1,2})\cdot M$ sendo que já sabemos calcular $e^3_{1,2}\cdot M$. Como a multiplicação de números reais é associativa ($(ab)c=a(bc)$) faz sentido que desejemos que essa operação que vamos definir seja associativa. Assim, podemos simplesmente definir

$$(8\cdot e^3_{1,2})\cdot M:=8\cdot(e^3_{1,2}\cdot M)$$

o que faz sentido, já que $8\in\mathbb{R}$, $e^3_{1,2}\cdot M\in M_{n\times m}(\mathbb{R})$ e **já sabemos multiplicar matrizes por números**.

De maneira geral, podemos definir:

Definição 1.1.13. *Dado um número real $a\in\ R$, uma matriz real $M\in M_{n\times m}(\mathbb{R})$ e uma matriz quadrada $e^n_{i,j}\in M_n(\mathbb{R})$, definimos o* ***produto de** ae^n_{ij} **com*** *M como sendo a matriz $(ae^n_{i,j})M$ dada por*

$$(ae^n_{i,j})M:=a(e^n_{i,j}M).$$

Exercício (1.1.6)

Usando M e N do exemplo anterior, calcule $\begin{pmatrix}0&5&0\\0&0&0\\0&0&0\end{pmatrix}\cdot M$ e $\begin{pmatrix}0&0&0\\0&0&-\pi\\0&0&0\end{pmatrix}\cdot N$ usando nossa nova definição e compare com a multiplicação clássica de matrizes.

Outro tipo de matrizes que são *quase* $e^n_{i,j}$ são matrizes que são *somas* de matrizes $e^n_{i,j}$. Por exemplo, a matriz $A=\begin{pmatrix}0&1&1\\1&0&0\\0&1&0\end{pmatrix}$ pode ser escrita como

$$A=\begin{pmatrix}0&1&1\\1&0&0\\0&1&0\end{pmatrix}=\begin{pmatrix}0&1&0\\0&0&0\\0&0&0\end{pmatrix}+\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix}+\begin{pmatrix}0&0&0\\1&0&0\\0&0&0\end{pmatrix}+\begin{pmatrix}0&0&0\\0&0&0\\0&1&0\end{pmatrix}=e^3_{1,2}+e^3_{1,3}+e^3_{2,1}+e^3_{3,2}.$$

Agora, dada uma matriz $M = \begin{pmatrix} 1 & -8 \\ 2 & 5 \\ 9 & \pi \end{pmatrix}$ queremos definir a multiplicação $A \cdot M$. Novamente, vamos nos lembrar que **já sabemos somar matrizes**, **já sabemos multiplicar matrizes $e^3_{i,j}$ por** M e **o produto de números reais distribui sobre a soma**. Assim, podemos definir, inspirados por essas três propriedades,

$$AM = (e^3_{1,2} + e^3_{1,3} + e^3_{2,1} + e^3_{3,2})M := e^3_{1,2}M + e^3_{1,3}M + e^3_{2,1}M + e^3_{3,2}M,$$

e calculando

$$e^3_{1,2}M = \begin{pmatrix} 2 & 5 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, e^3_{1,3}M = \begin{pmatrix} 9 & \pi \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, e^3_{2,1}M = \begin{pmatrix} 0 & 0 \\ 1 & -8 \\ 0 & 0 \end{pmatrix}, e^3_{3,2}M = \begin{pmatrix} 0 & 0 \\ 9 & \pi \\ 0 & 0 \end{pmatrix}$$

vemos que

$$AM = \begin{pmatrix} 2 & 5 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 9 & \pi \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 1 & -8 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 9 & \pi \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 11 & 5+\pi \\ 10 & \pi - 8 \\ 0 & 0 \end{pmatrix}$$

Exercício (1.1.7)

Compare o resultado obtido acima com a definição clássica de multiplicação de matrizes.

Definição 1.1.14. *Dados um conjunto de pares de índices* $\{i_1, j_1\}, \{i_2, j_2\}, \cdots, \{i_k, j_k\}$ *e uma matriz* $M \in M_{n\times m}(\mathbb{R})$, *definimos o* ***produto da soma de todos os*** $e^n_{i_r,j_r}$ ***com*** M *como sendo a matriz* $(e^n_{i_1,j_1} + e^n_{i_2,j_2} + \cdots + e^n_{i_k,j_k})M$ *dada por*

$$(e^n_{i_1,j_1} + e^n_{i_2,j_2} + \cdots + e^n_{i_k,j_k})M := e^n_{i_1,j_1}M + e^n_{i_2,j_2}M + \cdots + e^n_{i_k,j_k}M.$$

Agora podemos combinar os dois resultados: Considere a matriz $A = \begin{pmatrix} 0 & 5 & 2 \\ -7 & 0 & 0 \\ 0 & 10 & 0 \end{pmatrix}$ e a matriz $M = \begin{pmatrix} 1 & -8 \\ 2 & 5 \\ 9 & \pi \end{pmatrix}$. Como vamos definir o produto AM?

Primeiro vamos decompor A como soma de matrizes *quase* $e^n_{i,j}$

$$A = \begin{pmatrix} 0 & 5 & 2 \\ -7 & 0 & 0 \\ 0 & 10 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 5 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ -7 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 10 & 0 \end{pmatrix}$$

e em seguida escrever cada uma dessas matrizes como produto de um número por uma matriz $e^n_{i,j}$

$$\begin{aligned} A = \begin{pmatrix} 0 & 5 & 2 \\ -7 & 0 & 0 \\ 0 & 10 & 0 \end{pmatrix} &= \begin{pmatrix} 0 & 5 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ -7 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 10 & 0 \end{pmatrix} \\ &= 5e^3_{1,2} + 2e^3_{1,3} + (-7)e^3_{2,1} + 10e^3_{3,2} \end{aligned}$$

e finalmente, como já sabemos multiplicar cada pedaço por M, vamos definir

$$AM = (5e^3_{1,2} + 2e^3_{1,3} + (-7)e^3_{2,1} + 10e^3_{3,2})M := 5e^3_{1,2}M + 2e^3_{1,3}M + (-7)e^3_{2,1}M + 10e^3_{3,2}M.$$

Efetuando, vamos obter:

$$\begin{aligned}
AM &= 5\begin{pmatrix}2 & 5\\ 0 & 0\\ 0 & 0\end{pmatrix} + 2\begin{pmatrix}9 & \pi\\ 0 & 0\\ 0 & 0\end{pmatrix} + (-7)\begin{pmatrix}0 & 0\\ 1 & -8\\ 0 & 0\end{pmatrix} + 10\begin{pmatrix}0 & 0\\ 9 & \pi\\ 0 & 0\end{pmatrix}\\
&= \begin{pmatrix}10 & 25\\ 0 & 0\\ 0 & 0\end{pmatrix} + \begin{pmatrix}18 & 2\pi\\ 0 & 0\\ 0 & 0\end{pmatrix} + \begin{pmatrix}0 & 0\\ -7 & 56\\ 0 & 0\end{pmatrix} + \begin{pmatrix}0 & 0\\ 90 & 10\pi\\ 0 & 0\end{pmatrix} = \begin{pmatrix}28 & 25+2\pi\\ 83 & 56+10\pi\\ 0 & 0\end{pmatrix}.
\end{aligned}$$

Observação 1.1.15. *Note que qualquer matriz quadrada pode ser escrita como acima - isto é, somando vários produtos de números com matrizes $e^n_{i,j}$. Isso nos permite finalmente definir produtos quase totalmente gerais:*

Definição 1.1.16. *Seja $A \in M_n(\mathbb{R})$ uma matriz quadrada e $M \in M_{n\times m}(\mathbb{R})$ qualquer. Definimos o* ***produto de*** *A* ***com*** *M como sendo a matriz AM dada por*

$$\begin{aligned}
AM &= (A_{1,1}e^n_{1,1} + A_{1,2}e^n_{1,2} + \cdots + A_{1,n}e^n_{1,n} + A_{2,1}e^n_{2,1} + \cdots A_{n,n}e^n_{n,n})M\\
&:= A_{1,1}e^n_{1,1}M + A_{1,2}e^n_{1,2}M + \cdots + A_{1,n}e^n_{1,n}M + A_{2,1}e^n_{2,1}M + \cdots A_{n,n}e^n_{n,n}M
\end{aligned}$$

EXEMPLO(S):

Dada a matriz quadrada $A = \begin{pmatrix}a & b\\ c & d\end{pmatrix}$ e uma matriz qualquer $M = \begin{pmatrix}a' & b' & c'\\ d' & e' & f'\end{pmatrix}$, temos que AM é simplesmente

$$\begin{aligned}
&\begin{pmatrix}a & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}a' & b' & c'\\ d' & e' & f'\end{pmatrix} + \begin{pmatrix}0 & b\\ 0 & 0\end{pmatrix}\begin{pmatrix}a' & b' & c'\\ d' & e' & f'\end{pmatrix} + \begin{pmatrix}0 & 0\\ c & 0\end{pmatrix}\begin{pmatrix}a' & b' & c'\\ d' & e' & f'\end{pmatrix} + \begin{pmatrix}0 & 0\\ 0 & d\end{pmatrix}\begin{pmatrix}a' & b' & c'\\ d' & e' & f'\end{pmatrix}\\
&= \begin{pmatrix}aa' & ab' & ac'\\ 0 & 0 & 0\end{pmatrix} + \begin{pmatrix}bd' & be' & bf'\\ 0 & 0 & 0\end{pmatrix} + \begin{pmatrix}0 & 0 & 0\\ ca' & cb' & cc'\end{pmatrix} + \begin{pmatrix}0 & 0 & 0\\ dd' & de' & df'\end{pmatrix}\\
&= \begin{pmatrix}aa'+bd' & ba'+be' & ac'+bf'\\ ca'+dd' & cb'+de' & cc'+df'\end{pmatrix}
\end{aligned}$$

Observação 1.1.17. *Em matemática, quando vamos escrever uma soma muito grande, ou com muitos termos, costumamos usar um símbolo especial para isso - $\sum$ - o* ***somatório****. Ele funciona da seguinte maneira: Ao invés de escrever $x_1 + x_2 + \cdots + x_n$ vamos escrever*

$$\sum_{i=1}^{n} x_i.$$

A ideia é a seguinte: i denota um índice variável, e os números 1 e n que aparecem abaixo e acima, respectivamente, do somatório indicam qual o valor mínimo e máximo que i pode assumir, sempre variando de 1 em 1.

EXEMPLO(S):

Podemos escrever a soma de todos os números naturais entre 1 e n usando $\sum_{i=1}^{n} i$, por exemplo,

$$1+2+3+4 = \sum_{i=1}^{4} i.$$

Podemos escrever a soma dos quadrados dos números naturais entre 1 e n como $\sum_{i=1}^{n} i^2$, por exemplo,

$$1+4+9+16+25 = 1^2+2^2+3^2+4^2+5^2 = \sum_{i=1}^{5} i^2.$$

Podemos escrever a soma dos n primeiros números ímpares e pares, respectivamente, como $\sum_{i=1}^{n} 2i-1$ e $\sum_{i=1}^{n} 2i$, por exemplo,

$$1+3+5+7+9+11 = (2\cdot1-1)+(2\cdot2-1)+(2\cdot3-1)+(2\cdot3-1)+(2\cdot4-1)+(2\cdot5-1)+(2\cdot6-1) = \sum_{i=1}^{6} 2i-1$$

$$2+4+6+8+10+12+14 = (2\cdot1)+(2\cdot2)+(2\cdot3)+(2\cdot4)+(2\cdot5)+(2\cdot6)+(2\cdot7) = \sum_{i=1}^{7} 2i$$

Podemos escrever o produto clássico (1.1.6) das matrizes $M \in M_{n,m}(\mathbb{R})$ e $N \in M_{m,l}(\mathbb{R})$ como

$$(MN)_{i,j} = \sum_{k=1}^{m} M_{i,k}N_{k,j}.$$

Com isso, por exemplo, podemos tornar a definição de produto mais compacta:

Definição 1.1.18. *Seja $A \in M_n(\mathbb{R})$ uma matriz quadrada e $M \in M_{n\times m}(\mathbb{R})$ qualquer. Definimos o **produto de** A **com** M como sendo a matriz AM dada por*

$$\begin{aligned} AM &= \left(\sum_{j=1}^{n}\sum_{i=1}^{n} A_{i,j}e^n_{i,j}\right) M \\ &:= \sum_{j=1}^{n}\sum_{i=1}^{n} \left(A_{i,j}e^n_{i,j}M\right) \end{aligned}$$

Finalmente, para encerrar esta seção, vamos aprender a multiplicar matrizes de tamanho qualquer, usando o raciocínio acima.

Definição 1.1.19. *Vamos denotar por $e^{n,m}_{i,j}$ a matriz $n\times m$ dada por*

$$(e^{n,m}_{i,j})_{r,s} := \begin{cases} 1, & \text{se } r=i \text{ e } s=j \\ 0, & \text{caso contrário.} \end{cases}$$

Ou seja, é apenas uma generalização das nossas já familiares matrizes $e^n_{i,j}$ para matrizes não-quadradas com a mesma propriedade - elas têm 1 na entrada i,j e 0 em todas as outras entradas.

Novamente, podemos definir agora funções que vão realizar nossas multiplicações:

Definição 1.1.20. *Definimos por $E^{l,n}_{i,j} : M_{n\times m}(\mathbb{R}) \to M_{l\times m}(\mathbb{R})$ a **função** dada por:*

$$\begin{array}{c} \\ \\ \\ j\text{-ésima linha} \\ \\ \\ \end{array}\begin{pmatrix} M_{1,1} & M_{1,2} & \cdots & M_{1,m} \\ M_{2,1} & M_{2,2} & \cdots & M_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ M_{j,1} & M_{j,2} & \cdots & M_{j,m} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n,1} & M_{n,2} & \cdots & M_{n,m} \end{pmatrix} \mapsto \begin{array}{c} \\ \\ \\ i\text{-ésima linha} \\ \\ \\ \end{array}\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ M_{j,1} & M_{j,2} & \cdots & M_{j,m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Observação 1.1.21. *Assim como antes, o expoente em $E^{l,n}_{i,j}$ significa apenas que estamos transformando matrizes de n linhas em matrizes de l linhas.*

EXEMPLO(S):

Dada uma matriz $M = \begin{pmatrix} 5 & -8 & \sqrt{2} \\ 4 & 4 & 0 \end{pmatrix} \in M_{2\times 3}(\mathbb{R})$, temos que

$$E^{3,2}_{1,1}(M) = \begin{pmatrix} 5 & -8 & \sqrt{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad E^{3,2}_{1,2}(M) = \begin{pmatrix} 4 & 4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$E^{3,2}_{2,1}(M) = \begin{pmatrix} 0 & 0 & 0 \\ 5 & -8 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \qquad E^{3,2}_{2,2}(M) = \begin{pmatrix} 0 & 0 & 0 \\ 4 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$E^{3,2}_{3,1}(M) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 5 & -8 & \sqrt{2} \end{pmatrix} \qquad E^{3,2}_{3,2}(M) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 4 & 4 & 0 \end{pmatrix}$$

Com isso, vamos simplesmente imitar o procedimento anterior:

Definição 1.1.22. *Dada uma matriz $M \in M_{n\times m}(\mathbb{R})$, $l \in \mathbb{N}$ e qualquer par de índices i, j com $i \leq l$ e $j \leq n$, definimos o produto de $e^{l,n}_{i,j}$ com M como sendo a matriz $e^{l,n}_{i,j}M$ dada por*

$$e^{l,n}_{i,j}M := E^{l,n}_{i,j}(M).$$

EXEMPLO(S):

Continuando o exemplo acima, em que $M = \begin{pmatrix} 5 & -8 & \sqrt{2} \\ 4 & 4 & 0 \end{pmatrix} \in M_{2\times 3}(\mathbb{R})$, temos que

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 5 & -8 & \sqrt{2} \\ 4 & 4 & 0 \end{pmatrix} = e^{3,2}_{1,1}M := E^{3,2}_{1,1}(M) = \begin{pmatrix} 5 & -8 & \sqrt{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 5 & -8 & \sqrt{2} \\ 4 & 4 & 0 \end{pmatrix} = e^{3,2}_{1,2}M := E^{3,2}_{1,2}(M) = \begin{pmatrix} 4 & 4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 5 & -8 & \sqrt{2} \\ 4 & 4 & 0 \end{pmatrix} = e^{3,2}_{2,1}M := E^{3,2}_{2,1}(M) = \begin{pmatrix} 0 & 0 & 0 \\ 5 & -8 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 5 & -8 & \sqrt{2} \\ 4 & 4 & 0 \end{pmatrix} = e^{3,2}_{1,1}M := E^{3,2}_{2,2}(M) = \begin{pmatrix} 0 & 0 & 0 \\ 4 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 5 & -8 & \sqrt{2} \\ 4 & 4 & 0 \end{pmatrix} = e^{3,2}_{3,1}M := E^{3,2}_{3,1}(M) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 5 & -8 & \sqrt{2} \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 5 & -8 & \sqrt{2} \\ 4 & 4 & 0 \end{pmatrix} = e^{3,2}_{3,2}M := E^{3,2}_{3,2}(M) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 4 & 4 & 0 \end{pmatrix}$$

Exercício (1.1.8)

Essa multiplicação coincide com a multiplicação clássica de matrizes? Se sim, você consegue mostrar que isso sempre é verdade? Se não, você consegue mostrar um caso onde falha?

Feito isso, vamos agora simplesmente repetir o que fizemos antes acima para ensinar a multiplicar matrizes de tamanho “qualquer”.

EXEMPLO(S):

Considere as matrizes $M = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \in M_{2\times 3}(\mathbb{R})$ e $N = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \\ \epsilon & \eta \end{pmatrix} \in M_{3\times 2}(\mathbb{R})$. Para calcular MN podemos fazer como fizemos com as matrizes quadradas:
Primeiro, vamos decompor M em matrizes $e_{i,j}^{l,n}$:

$$\begin{aligned} M &= \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \\ &= \begin{pmatrix} a & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & b & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ d & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & e & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & f \end{pmatrix} \\ &= a\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + e\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + f\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

e cada uma dessas matrizes $e_{i,j}^{l,n}$ sabemos multiplicar por N:

$$\begin{aligned} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \\ \epsilon & \eta \end{pmatrix} &:= \begin{pmatrix} \alpha & \beta \\ 0 & 0 \end{pmatrix} & \qquad \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \\ \epsilon & \eta \end{pmatrix} &:= \begin{pmatrix} 0 & 0 \\ \alpha & \beta \end{pmatrix} \\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \\ \epsilon & \eta \end{pmatrix} &:= \begin{pmatrix} \gamma & \delta \\ 0 & 0 \end{pmatrix} & \qquad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \\ \epsilon & \eta \end{pmatrix} &:= \begin{pmatrix} 0 & 0 \\ \gamma & \delta \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \\ \epsilon & \eta \end{pmatrix} &:= \begin{pmatrix} \epsilon & \eta \\ 0 & 0 \end{pmatrix} & \qquad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \\ \epsilon & \eta \end{pmatrix} &:= \begin{pmatrix} 0 & 0 \\ \epsilon & \eta \end{pmatrix} \end{aligned}$$

e cada uma dessas matrizes sabemos multiplicar por números:

$$\begin{aligned} a\begin{pmatrix} \alpha & \beta \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} a\alpha & a\beta \\ 0 & 0 \end{pmatrix} & \qquad d\begin{pmatrix} 0 & 0 \\ \alpha & \beta \end{pmatrix} &= \begin{pmatrix} 0 & 0 \\ d\alpha & d\beta \end{pmatrix} \\ b\begin{pmatrix} \gamma & \delta \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} b\gamma & b\delta \\ 0 & 0 \end{pmatrix} & \qquad e\begin{pmatrix} 0 & 0 \\ \gamma & \delta \end{pmatrix} &= \begin{pmatrix} 0 & 0 \\ e\gamma & e\delta \end{pmatrix} \\ c\begin{pmatrix} \epsilon & \eta \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} c\epsilon & c\eta \\ 0 & 0 \end{pmatrix} & \qquad f\begin{pmatrix} 0 & 0 \\ \epsilon & \eta \end{pmatrix} &= \begin{pmatrix} 0 & 0 \\ f\epsilon & f\eta \end{pmatrix} \end{aligned}$$

e como todas têm o mesmo tamanho, nós sabemos somar todas elas:

$$\begin{gathered} \begin{pmatrix} a\alpha & a\beta \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} b\gamma & b\delta \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} c\epsilon & c\eta \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ d\alpha & d\beta \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ e\gamma & e\delta \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ f\epsilon & f\eta \end{pmatrix} \\ = \begin{pmatrix} a\alpha + b\gamma + c\epsilon & a\beta + b\delta + c\eta \\ d\alpha + e\gamma + f\epsilon & d\beta + e\delta + f\eta \end{pmatrix} \end{gathered}$$

e finalmente, vemos chamar esse resultado de MN.

Finalmente podemos definir o produto de matrizes compatíveis:

Definição 1.1.23. *Sejam $M \in M_{n\times m}(\mathbb{R})$ e $N \in M_{m\times l}(\mathbb{R})$. Definimos o produto de M com N como sendo a matriz MN dada por*

$$MN := \sum_{j=1}^{m}\sum_{i=1}^{n}\left(M_{i,j}e_{i,j}^{n,m}N\right)$$

Exercício (1.1.9)

Calcule, usando os dados do exemplo acima, NM, usando o método que preferir (isto é, o método clássico, ou o que estamos definindo) e compare o resultado que você obtiver com o MN calculado acima.

Exercício (1.1.10)

Sejam A e B matrizes quaisquer, tais que ambos os produtos AB e BA existem. Isso significa que $AB = BA$? Se sim, tente provar por que isso é verdade. Se não, dê um contra-exemplo.

Exercício (1.1.11)

Tente provar as seguintes propriedades:

(Comutatividade da soma) Para quaisquer duas matrizes $M, N \in M_{n\times m}(\mathbb{R})$ temos que $M + N = N + M$;

(Elemento neutro da soma) Para qualquer matriz $M \in M_{n\times m}(\mathbb{R})$ existe uma única matriz $Z \in M_{n\times m}(\mathbb{R})$ tal que $M + Z = M$. Vamos chamar essa matriz de **zero** e notar por 0;

(Inverso da soma) Para qualquer matriz $M \in M_{n\times m}(\mathbb{R})$ existe uma única matriz $N \in M_{n\times m}(\mathbb{R})$ tal que $M + N = 0$. Vamos chamar essa matriz de **inversa aditiva de** M e notar por $-M$;

(Associatividade da soma) Para quaisquer três matrizes $L, M, N \in M_{n\times m}(\mathbb{R})$ temos que $L + (M + N) = (L + M) + N$;

(Comutatividade do produto por número) Para qualquer número real $a \in \mathbb{R}$ e qualquer matriz $M \in M_{n\times m}(\mathbb{R})$ temos que $aM = Ma$;

(Elemento neutro do produto por número) Para qualquer matriz $M \in M_{n\times m}(\mathbb{R})$ existe um único número $u \in \mathbb{R}$ tal que $uM = M$;

(Associatividade do produto por número) Para quaisquer dois números $a, b \in \mathbb{R}$ e qualquer matriz $M \in M_{n\times m}(\mathbb{R})$, temos que $a(bM) = (ab)M$;

(Distributividade do produto por número sobre a soma) Para quaisquer duas matrizes $M, N \in M_{n\times m}(\mathbb{R})$ e qualquer número real $a \in \mathbb{R}$ temos que $a(M + N) = aM + aN$;

(Distributividade do produto por número sobre a soma de números) Para quaisquer dois números reais $a, b \in \mathbb{R}$ e qualquer matriz $M \in M_{n\times m}(\mathbb{R})$ temos que $(a+b)M = aM+bM$;

(Elemento neutro do produto) Para qualquer matriz $M \in M_{n\times m}(\mathbb{R})$ existe uma única matriz $I \in M_m(\mathbb{R})$ tal que $MI = M = IM$. Vamos chamar essa matriz de **identidade**;

(Associatividade do produto) Para quaisquer três matrizes $L \in M_{n\times m}(\mathbb{R})$, $M \in M_{m\times l}(\mathbb{R})$ e $N \in M_{l\times k}(\mathbb{R})$ temos que $(LM)N = L(MN)$;

(Associatividade do produto com produto por números) Para quaisquer duas matrizes $M \in M_{n\times m}(\mathbb{R})$ e $N \in M_{m\times l}(\mathbb{R})$ e qualquer número real $a \in \mathbb{R}$ temos que $a(MN) = (aM)N$;

(Distributividade do produto sobre a soma) Para quaisquer duas matrizes $L, M \in M_{n\times m}(\mathbb{R})$ qualquer matriz $N \in M_{m\times l}(\mathbb{R})$ e qualquer matriz $K \in M_{o\times n}(\mathbb{R})$ temos que $K(L+M) = KL + KM$ e $(L + M)N = LN + MN$.

1.2 Sistemas Lineares

Vamos agora tentar dar uma primeira justificativa para o estudo de matrizes reais: A resolução de sistemas lineares.

Definição 1.2.1. *Um **sistema linear de** n **equações em** m **variáveis** consiste em uma coleção de n equações, em que a i-ésima equação é da forma $a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,m}x_m = b_i$, em que $\{a_{i,j}\}_{i,j\in\mathbb{N}}$ e $\{b_i\}_{i\in\mathbb{N}}$ são números reais. Vamos denotar sistemas dessa forma por*

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + & \cdots & + a_{1,m}x_m = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + & \cdots & + a_{2,m}x_m = b_2 \\ & \vdots & \\ a_{n,1}x_1 + a_{n,2}x_2 + & \cdots & + a_{n,m}x_m = b_n \end{cases}$$

Observação 1.2.2. *A motivação por trás de chamarmos tais sistemas de **lineares** ficará para um capítulo posterior.*

EXEMPLO(S):

Considere o sistema de equações

$$\begin{cases} 5x + 4y - 7z = 0 \\ 3x + y - z = 9. \end{cases}$$

Ele é composto de duas equações ($5x + 4y - 7z = 0$ e $3x + y - z = 9$), ambas são lineares e ambas têm três variáveis (x, y e z). Assim, o sistema acima é um sistema linear de duas equações e três incógnitas.

Contudo, o sistema

$$\begin{cases} 8x + y + z = 2 \\ -4x + y + z^2 = 4 \end{cases}$$

não é linear. Ele é composto de duas equações e três variáveis - contudo, a segunda equação **não é linear**.

Em outras palavras, sistemas lineares são sistemas de equações onde só aparecem somas de números multiplicados a variáveis - nada de funções trigonométricas, nada de potências, nada de funções exponenciais ou qualquer outro tipo de funções. Apenas multiplicação por números.

Se voltarmos para a definição de sistemas lineares de n equações e m incógnitas, vamos ver que cada equação tem exatamente m números (chamados coeficientes) multiplicando as variáveis. Assim, como temos n equações, temos $n \times m$ coeficientes no total. Isso sugere que podemos pegar um sistema

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + & \cdots & + a_{1,m}x_m = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + & \cdots & + a_{2,m}x_m = b_2 \\ & \vdots & \\ a_{n,1}x_1 + a_{n,2}x_2 + & \cdots & + a_{n,m}x_m = b_n \end{cases}$$

e representá-lo por uma matriz, da seguinte forma:

Primeiro, temos uma matriz

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix}$$

chamada de **matriz de coeficientes** que faz exatamente o que o nome diz - coleta todos os coeficientes do sistema.

Em seguida, temos uma matriz

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

chamada de **matriz de variáveis** que, novamente, é auto-descritiva: ela contém todas as variáveis do sistema.

Por fim, temos uma matriz

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

chamada de **matriz resultante** que, mais uma vez, é simplesmente a matriz que contém todos os resultados do sistema.

Por que isso é legal? Bom, primeiro vamos notar que $A \in M_{n\times m}(\mathbb{R})$, X é uma matriz $m \times 1$ com entradas que são variáveis, e que $B \in M_{n,1}(\mathbb{R})$. Agora, lembrando da multiplicação de matrizes, não é difícil ver que podemos multiplicar AX. Mas o que seria, por exemplo, o elemento na posição i, j de AX? Bom, por definição,

$$(AX)_{i,j} = \sum_{l=1}^{m} A_{i,l}X_{l,j} = \sum_{l=1}^{m} a_{i,l}X_{l,j}$$

e como $AX \in M_{n,1}(\mathbb{R})$, $j = 1$, de forma que podemos continuar:

$$(AX)_{i,j} = (AX)_{i,1} = \sum_{l=1}^{m} a_{i,l}X_{l,j} = \sum_{l=1}^{m} a_{i,l}X_{l,1} = \sum_{l=1}^{m} a_{i,l}x_l,$$

e expandindo temos

$$(AX)_{i,1} = a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,m}x_m$$

que é exatamente b_i. Mas $b_i = B_{i,1}$. Disso temos que $(AX)_{i,1} = B_{i,1}$ - ou seja, as matrizes AX e B são iguais em **todas** as entradas e, portanto, são **iguais**.

Em outras palavras, ao escrever um sistema usando as matrizes A, X e B, vemos que o sistema pode ser descrito como a igualdade de matrizes $AX = B$. Ou, dito de outra maneira - se $A' \in M_{n\times m}(\mathbb{R})$, X' é uma matriz $m \times 1$ com entradas que são variáveis, e $B' \in M_{n,1}(\mathbb{R})$ são matrizes quaisquer, então uma equação matricial $A'X' = B'$ determina um único sistema linear.

1.2.1 Escalonamentos

Contudo, resolver sistemas lineares não é tarefa fácil. Vamos aqui apresentar uma técnica - chamada **escalonamento** - para resolver (se possível!) um sistema linear. Antes disso, porém, precisamos entender o que significa *resolver um sistema linear.*

Definição 1.2.3. *Dado um sistema linear $AX = B$, dizemos que C* ***é uma solução para o sistema*** *se $AC = B$.*

EXEMPLO(S):

Dado o sistema linear

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \\ 1 \end{pmatrix},$$

temos que uma solução do sistema é a matriz

$$\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

De fato,

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3+2+1 \\ 2+1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \\ 1 \end{pmatrix}.$$

Afirmamos ainda que essa solução é única.

De fato, suponha que temos alguma matrix $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ tal que

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \\ 1 \end{pmatrix}.$$

Então teríamos três equações: $a+b+c=6$, $b+c=3$ e $c=1$ Como $c=1$ e $b+c=3$, segue que $b=2$. Similarmente, como $b+c=3$ e $a+b+c=6$, segue que $a=3$ - ou seja, a matriz dada é, de fato, a única solução do sistema.

Exercício (1.2.1)

Encontre uma solução do sistema

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 13 \\ 21 \end{pmatrix}.$$

Essa solução é única? Se não, encontre outra. Se sim, prove.

Vamos agora lembrar de como resolvíamos sistemas de equações lineares antes:

EXEMPLO(S):

Considere o sistema

$$\begin{cases} 4x+2y=0 \\ -4x+9y=22 \end{cases}.$$

Nós podemos perceber que o coeficiente de x em ambas as equações é o mesmo, com o sinal invertido. Isso nos diz que se somarmos as equações, obteremos uma nova equação com o coeficiente de x sendo 0

$$\begin{array}{rr} & 4x+2y=0 \\ + & -4x+9y=22 \\ \hline & 0x+11y=22 \end{array}$$

com isso, conseguimos resolver a equação $11y = 22$, que tem como única solução $y = 2$. Agora, com o valor de y em mãos, podemos escolher qualquer equação original e resolvê-la - por exemplo a primeira

$$\begin{aligned} 4x + 2y &= 0 \\ 4x + 2 \cdot 2 &= 0 \quad \text{(usando } y = 2\text{)} \\ 4x + 4 &= 0 \\ 4(x+1) &= 0 \quad \text{(colocando 4 em evidência)} \\ x + 1 &= 0 \\ x &= -1, \end{aligned}$$

e agora afirmamos que a solução do sistema original é $(-1, 2)$.
O que fizemos, em suma, foi trocar o sistema

$$\begin{cases} 4x + 2y = 0 \\ -4x + 9y = 22 \end{cases}$$

pelo sistema

$$\begin{cases} 4x + 2y = 0 \\ 11y = 22, \end{cases}$$

resolver esse segundo sistema e afirmar que essa solução obtida também é solução do sistema original.

Pensando do ponto de vista de matrizes, vamos chamar

$$A = \begin{pmatrix} 4 & 2 \\ -4 & 9 \end{pmatrix}$$

a matriz de coeficientes do sistema original e

$$A' = \begin{pmatrix} 4 & 2 \\ 0 & 11 \end{pmatrix}$$

a matriz de coeficientes do sistema que obtemos somando as duas equações do sistema original.
Mas como obtivemos a matriz A'? Vamos analisar ela por partes:

- A linha 1 de A' é simplesmente a linha 1 de A;
- A linha 2 de A' é simplesmente a soma das linhas 1 e 2 de A.

Ou seja, se escrevermos

$$A' = \begin{pmatrix} 4 & 2 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 11 \end{pmatrix}$$

podemos ver que $\left(\begin{smallmatrix} 4 & 2 \\ 0 & 0 \end{smallmatrix}\right) = e^2_{1,1}A$. Além disso, como já dissemos,

$$\begin{pmatrix} 0 & 0 \\ 0 & 11 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 4 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -4 & 9 \end{pmatrix} = e^2_{2,1}A + e^2_{2,2}A,$$

ou seja,

$$A' = \begin{pmatrix} 4 & 2 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 11 \end{pmatrix} = e^2_{1,1}A + e^2_{2,1}A + e^2_{2,2}A = (e^2_{1,1} + e^2_{2,1} + e^2_{2,2})A$$

que é simplesmente

$$A' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} A.$$

Em outras palavras, nós multiplicamos A por uma soma de matrizes $e^n_{i,j}$ e obtivemos uma matriz A' com as mesmas soluções de A.

EXEMPLO(S):

Similarmente ao exemplo anterior, considere o sistema linear

$$\begin{cases} -7x + 24y = 8 \\ x - 3y = 1 \end{cases}.$$

Para resolvê-lo, geralmente somaríamos a linha 1 com sete vezes a linha 2:

$$\begin{array}{rr} & -7x + 24y = 8 \\ + & 7(x - 3y = 1) \\ \hline & 0x + 3y = 15 \end{array}$$

e repetindo acima, podemos encontrar $y = 5$. Agora, com y em mãos, podemos escolher alguma das equações originais (por exemplo, a segunda) e resolvê-la, obtendo $x = 16$, e, portanto, a solução é o par $(16, 5)$.

Novamente, temos duas matrizes: A matriz do sistema original

$$A = \begin{pmatrix} -7 & 24 \\ 1 & -3 \end{pmatrix}$$

e a matriz que obtivemos somando a primeira linha com sete vezes a segunda linha

$$A' = \begin{pmatrix} 1 & -3 \\ 0 & 3 \end{pmatrix}.$$

Será que conseguimos expressar A' como um produto de alguma matriz por A - como fizemos antes?
Vamos começar desmembrando A':

$$A' = \begin{pmatrix} 1 & -3 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix}$$

em que o primeiro pedaço não é nada mais que uma cópia da segunda linha de A - ou seja, $e^2_{1,2}A$ - e o segundo pedaço é a soma da primeira linha de A com sete vezes a segunda linha de A - ou seja,

$$\begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -7 & 24 \end{pmatrix} + 7\begin{pmatrix} 0 & 0 \\ 1 & -3 \end{pmatrix} = e^2_{2,1}A + 7e^2_{2,2}A$$

e, portanto,

$$A' = e^2_{1,2}A + e^2_{2,1}A + 7e^2_{2,2}A = (e^2_{1,2} + e^2_{2,1} + 7e^2_{2,2})A,$$

ou seja,

$$A' = \begin{pmatrix} 0 & 1 \\ 1 & 7 \end{pmatrix} A,$$

como queríamos.
Neste caso, multiplicamos A por uma matriz que era soma de múltiplos de matrizes $e^n_{i,j}$ e ainda obtivemos uma matriz que tem as mesmas soluções.

Observação 1.2.4. *Cuidado! Nem toda multiplicação de A por matrizes $e^n_{i,j}$ preserva soluções. Por exemplo, fazendo $\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$ vezes $A = \begin{pmatrix} -7 & 24 \\ 1 & -3 \end{pmatrix}$ obtemos a matriz $\begin{pmatrix} 1 & -3 \\ 1 & -3 \end{pmatrix}$ que claramente não tem as mesmas soluções que A.*

Em breve veremos condições para que isso não aconteça.

Mais à frente veremos exatamente *porquê* isso é verdade. Por agora, nos basta focar nessas operações.

Definição 1.2.5. *Um **escalonamento** de uma matriz $A \in M_{n\times m}(\mathbb{R})$ qualquer, é uma matriz A' que pode ser obtida de A por composições das seguintes operações:*

- *Somar uma linha de A a um múltiplo de outra linha de A;*
- *Trocar duas linhas de A;*
- *Multiplicar uma linha inteira de A por um mesmo número.*

Pelos exemplos acima, podemos ver que sempre podemos realizar escalonamentos de A por multiplicações XA, em que X é uma soma de múltiplos de matrizes $e^n_{i,j}$.

Exercício (1.2.2)

Construa uma matriz $X \in M_2(\mathbb{R})$ diferente de 0 que seja soma de matrizes $e^2_{i,j}$ tal que XA **não** seja um escalonamento de A (ou seja, não tenha as mesmas soluções de A) para qualquer matriz $A \in M_{2\times m}(\mathbb{R})$.

Finalmente, podemos enunciar o teorema mais forte desta seção:

Teorema 1.2.6. *Uma matriz A possui solução se, e somente se, todo escalonamento de A possui solução.*

Não vamos demonstrar esse teorema agora, mas vamos chamar a atenção para o seguinte corolário:

Corolário 1.2.7. *Se, a matriz identidade é um escalonamento de A, então A possui solução única.*

Observação 1.2.8. *Em matemática, um **corolário** é um resultado que segue imediatamente de um resultado anterior. Então estamos afirmando que o Corolário 1.2.7 segue imediatamente do Teorema 1.2.6.*

Exercício (1.2.3)

Prove o Corolário 1.2.7 assumindo que o Teorema 1.2.6 seja verdade.

1.2.2 Resolvendo sistemas lineares

Vamos finalmente aprender a obter soluções usando escalonamentos. Para isso, considere o exemplo abaixo:

EXEMPLO(S):

Sejam

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \\ 1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

e $AX = B$ um sistema linear. Para escalonar A vamos fazer:

$$\begin{pmatrix}1 & 1\\ 1 & 1\\ 1 & -1\\ 1 & -1\end{pmatrix} \rightsquigarrow \begin{array}{c} \\ l_2 - l_1\\ l_3 - l_1\\ l_4 - l_1\end{array}\begin{pmatrix}1 & 1\\ 0 & 0\\ 0 & -2\\ 0 & -2\end{pmatrix} \rightsquigarrow \begin{array}{c} \\ \\ \frac{l_3}{-2}\\ \\ \end{array}\begin{pmatrix}1 & 1\\ 0 & 0\\ 0 & 1\\ 0 & -2\end{pmatrix} \rightsquigarrow \begin{array}{c} l_1 - l_3\\ \\ \\ l_4 + 2l_3\end{array}\begin{pmatrix}1 & 0\\ 0 & 0\\ 0 & 1\\ 0 & 0\end{pmatrix}$$

no primeiro passo multiplicamos A por $P_1 = \begin{pmatrix}1 & 0 & 0 & 0\\ -1 & 1 & 0 & 0\\ -1 & 0 & 1 & 0\\ -1 & 0 & 0 & 1\end{pmatrix}$, no segundo por $P_2 = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & \frac{1}{-2} & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$ e no terceiro por $P_3 = \begin{pmatrix}1 & 0 & -1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 2 & 1\end{pmatrix}$, obtendo a matrix escalonada $P_3P_2P_1A$.
Se fizermos, agora, $(P_3P_2P_1A)X$ e lembrarmos que o produto de matrizes é associativo, teremos

$$(P_3P_2P_1A)X = P_3P_2P_1(AX) = P_3P_2P_1B$$

onde do lado direito aparece B escalonada pelas mesmas matrizes que A foi escalonada. Além disso, a equação $(P_3P_2P_1A)X = P_3P_2P_1B$ nos diz que a solução (X) do sistema original continua sendo solução do sistema escalonado. Assim, podemos calcular $P_3P_2P_1B$:

$$P_1B = \begin{array}{l} l_1\\ l_2 - l_1\\ l_3 - l_1\\ l_4 - l_1\end{array}\begin{pmatrix}1\\ -2\\ 0\\ -2\end{pmatrix}$$

$$P_2(P_1B) = \begin{array}{l} l_1\\ l_2\\ \frac{l_3}{-2}\\ l_4\end{array}\begin{pmatrix}1\\ -2\\ 0\\ -2\end{pmatrix}$$

$$P_3(P_2P_1B) = \begin{array}{l} l_1 - l_3\\ l_2\\ l_3\\ l_4 + 2l_3\end{array}\begin{pmatrix}1\\ -2\\ 0\\ -2\end{pmatrix},$$

ou seja,

$$\begin{pmatrix}1 & 0\\ 0 & 0\\ 0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}1\\ -2\\ 0\\ -2\end{pmatrix} \quad \text{e} \quad \begin{pmatrix}1 & 1\\ 1 & 1\\ 1 & -1\\ 1 & -1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}1\\ -1\\ 1\\ -1\end{pmatrix}$$

têm as mesmas soluções.
Contudo, o sistema escalonado nos diz que (pela última linha) $0x + 0y = -2$, e nós sabemos que $0a = 0$ para qualquer $a \in \mathbb{R}$. Em particular, não existe $(x, y) \in \mathbb{R}^2$ tal que $0x + 0y = -2$ - ou seja, o sistema escalonado **não admite solução** e, portanto, o sistema original também não.

Observação 1.2.9. *De maneira geral, sempre que, ao escalonar uma matriz, obtivermos uma linha de zeros sendo igual a algo diferente de zero, podemos parar o escalonamento e concluir, imediatamente, que o sistema não possui soluções.*

Definição 1.2.10. *Seja $AX = B$ um sistema linear. Denotamos por $\left(\begin{array}{c:c} A & B \end{array}\right)$ a* ***matrix aumentada do sistema*** *obtida de A simplesmente colando uma cópia de B à direita.*

EXEMPLO(S):

No exemplo anterior, em que $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \\ 1 & -1 \end{pmatrix}$ e $B = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$, a **matrix aumentada** é

$$A \mid B = \left(\begin{array}{cc:c} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{array}\right).$$

A vantagem de se trabalhar com a matriz aumentada é a seguinte: Ao invés de escalonar A e depois repetir o procedimento em B, nós escalonamos ambas as matrizes ao mesmo tempo:

$$\left(\begin{array}{cc:c} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 1 & 1 & 1 \\ 0 & 0 & -2 \\ 0 & -2 & 0 \\ 0 & -2 & -2 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 1 & 1 & 1 \\ 0 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & -2 & -2 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 1 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{array}\right).$$

Por exemplo, para o sistema linear

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

temos a matriz aumentada $\left(\begin{array}{cc:c} 1 & 1 & 1 \\ 1 & -1 & 1 \end{array}\right)$ que podemos escalonar e obter

$$\left(\begin{array}{cc:c} 1 & 1 & 1 \\ 1 & -1 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 1 & 1 & 1 \\ 0 & -2 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 1 & 1 & 1 \\ 0 & 1 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 1 & 0 & 1 \\ 0 & 1 & 0 \end{array}\right)$$

e podemos dizer imediatamente que as equações são $1x = 1$ e $1y = 0$, donde vemos que a única solução do sistema é $(1, 0) \in \mathbb{R}^2$.

EXEMPLO(S):

Vamos resolver um último sistema antes de avançar:

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

cuja matriz aumentada é $\left(\begin{array}{cccc:c} 1 & 2 & 1 & 1 & 1 \\ 1 & 3 & -1 & 2 & 3 \end{array}\right)$. Escalonando obtemos

$$\left(\begin{array}{cccc:c} 1 & 2 & 1 & 1 & 1 \\ 1 & 3 & -1 & 2 & 3 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc:c} 1 & 2 & 1 & 1 & 1 \\ 0 & 1 & -2 & 1 & 2 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc:c} 1 & 0 & 5 & -1 & -3 \\ 0 & 1 & -2 & 1 & 2 \end{array}\right)$$

que nos dá as equações $x + 5z - w = -3$ e $y - 2z + w = 2$. Não temos mais restrições no sistema, então todos os pontos que satisfazem a essas equações são soluções.
Para denotar, então, essa solução, note que podemos isolar x e y acima, obtendo $x = -5z + w - 3$ e $y = 2z - w + 2$. Assim, vemos que *para cada valor distinto de z e w temos uma solução diferente do sistema*. Por exemplo, se $w = z = 0$, temos que $x = -3$ e $y = 2$. Substituindo, então, o ponto $(-3, 2, 0, 0)$ no sistema original vemos que

$$(-3) + 2(2) + (0) + (0) = 4 - 3 = 1$$

$$(-3) + 3(2) - 1(0) + 2(0) = 6 - 3 = 3$$

e, de fato, o ponto $(-3, 2, 0, 0)$ é solução.
De maneira geral, como para cada valor possível de z e w temos uma solução, vamos notar o conjunto de todas as soluções por

$$S = \{(-5z + w - 3, 2z - w + 2, z, w) \in \mathbb{R}^4 \mid z \in \mathbb{R}, w \in \mathbb{R}\}.$$

Observação 1.2.11. *No caso do exemplo acima dizemos que z e w são **variáveis livres**. A ideia é que elas não têm um valor fixo, mas são "livres" para assumir o valor que quiserem e o sistema continua tendo solução.*

Exercício (1.2.4)

Use os exemplos acima para responder: Seja $AX = B$ um sistema $n \times m$, com $n \neq m$. O que podemos dizer sobre a existência de soluções do sistema?
E no caso das matrizes quadradas, i.e., $n = m$?

1.2.3 Sistemas homogêneos

Definição 1.2.12. *Um sistema linear $AX = B$ é dito **homogêneo** se $B = 0$.*

Poderíamos nos perguntar a importância de destacar sistemas homogêneos dentre todos os sistemas lineares. Contudo, se nos lembrarmos do que já fizemos, vamos ver que um sistema não possui solução se, e somente se, sua forma escalonada possui uma linha de zeros, com a entrada correspondente na matriz resultante diferente de zero - por exemplo, algo da forma

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

Mas ao lidar com sistemas homogêneos, todas as entradas da matriz resultante são nulas - ou seja, mesmo que durante o processo de escalonamento alguma linha se anule, ainda assim o sistema continua tendo solução.

Proposição 1.2.13. *Todo sistema linear homogêneo possui solução.*

Um jeito óbvio de ver isso é:

EXEMPLO(S):

Considere o sistema linear homogêneo

$$\begin{pmatrix} 1 & 7 \\ 8 & 9 \\ 17 & -10 \\ 25 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Ao invés de tentar escalonar o sistema, note que estamos multiplicando a matrix $\begin{pmatrix} 1 & 7 \\ 8 & 9 \\ 17 & -10 \\ 25 & 3 \end{pmatrix}$ por uma outra matriz, e queremos que o resultado seja 0. Mas já sabemos que $A0 = 0$ para qualquer matriz A - ou seja, tomando $X_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ vemos que X_0 é solução do sistema.

De fato, todo sistema homogêneo possui a solução $(0, 0)$ - por isso ela é chamada de **solução trivial**.

EXEMPLO(S):

Vamos voltar a um exemplo que fizemos acima:

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Já vimos que o conjunto de soluções desse sistema, S, é da forma

$$S = \{(x, y, z, w) \in \mathbb{R}^4 \mid x = -5z + w - 3, y = 2z - w + 2\}.$$

Vamos, agora, resolver o sistema homogêneo associado ao sistema acima, em que nós simplesmente trocamos a matriz resultante pela matriz de zeros:

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Escalonando obtemos

$$\left(\begin{array}{cccc:c} 1 & 2 & 1 & 1 & 0 \\ 1 & 3 & -1 & 2 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc:c} 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & -2 & 1 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc:c} 1 & 0 & 5 & -1 & 0 \\ 0 & 1 & -2 & 1 & 0 \end{array}\right)$$

que nos dá as equações $x + 5z - w = 0$ e $y - 2z + w = 0$. Não temos mais restrições no sistema, então todos os pontos que satisfazem a essas equações são soluções - ou seja, o conjunto de soluções do sistema homogêneo, S_0, é da forma

$$S_0 = \{(x, y, z, w) \in \mathbb{R}^4 \mid x = -5z + w, y = 2z - w\}.$$

Comparando os dois conjuntos, não é difícil ver que se X_0 é uma solução do sistema homogêneo, então $X_0 + \begin{pmatrix} -3 \\ 2 \\ 0 \\ 0 \end{pmatrix}$ é solução do sistema original.

Por exemplo, escolhendo $X_0 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}$ vamos primeiro conferir que X_0 é, de fato, solução do sistema homogêneo:

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot (-1) + 1 \cdot 0 + 1 \cdot 1 \\ 1 \cdot 1 + 3 \cdot (-1) + (-1) \cdot 0 + 2 \cdot 1 \end{pmatrix} = \begin{pmatrix} 1 - 2 + 0 + 1 \\ 1 - 3 + 0 + 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Agora vamos somar $\begin{pmatrix} -3 \\ 2 \\ 0 \\ 0 \end{pmatrix}$ a X_0, obtendo uma nova matriz

$$X = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -3 \\ 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

Afirmamos que X é solução do sistema original. Vamos conferir:

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot (-2) + 2 \cdot 1 + 1 \cdot 0 + 1 \cdot 1 \\ 1 \cdot (-2) + 3 \cdot 1 + (-1) \cdot 0 + 2 \cdot 1 \end{pmatrix} = \begin{pmatrix} -2 + 2 + 0 + 1 \\ -2 + 3 + 0 + 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

ou seja, X é de fato uma solução do sistema original.
Será que isso acontece por acaso?

Proposição 1.2.14. *Dado um sistema linear $AX = B$, uma solução X_0 do sistema homogêneo $AX = 0$, e X_1 uma solução qualquer do sistema $AX = B$, então $X_0 + X_1$ é solução de $AX = B$.*

Demonstração. Se X_0 é solução do sistema homogêneo, temos que $AX_0 = 0$. Similarmente, se X_1 é solução do sistema $AX = B$, temos que $AX_1 = B$. Agora, como o produto de matrizes distribui sobre a soma, temos que

$$A(X_0 + X_1) = AX_0 + AX_1 = 0 + B = B,$$

ou seja, $X_0 + X_1$ também é solução do sistema $AX = B$. □

Mas será que qualquer solução do sistema pode ser calculada assim - sabendo uma solução do sistema homogêneo e uma solução do sistema original? A resposta é sim, e vamos mostrar em seguida, mas antes, vamos fazer um exemplo:

EXEMPLO(S):

Ainda no exemplo anterior, considere o sistema

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Já vimos que $X_0 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}$ é solução do sistema homogêneo. Vamos escolher qualquer outra solução do sistema original, por exemplo, $X_1 = \begin{pmatrix} -13 \\ 6 \\ 2 \\ 0 \end{pmatrix}$ (verifique que isso é uma solução!). Será que existe alguma solução X_2 tal que $X_1 = X_0 + X_2$? Ora, resolver isso é o mesmo que resolver $X_2 = X_1 - X_0$ - o que nós já sabemos fazer:

$$X_2 = \begin{pmatrix} -13 \\ 6 \\ 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -14 \\ 7 \\ 2 \\ -1 \end{pmatrix}.$$

Nos resta mostrar que X_2 é solução:

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} -14 \\ 7 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \cdot (-14) + 2 \cdot 7 + 1 \cdot 2 + 1 \cdot (-1) \\ 1 \cdot (-14) + 3 \cdot 7 + (-1) \cdot 2 + 2 \cdot (-1) \end{pmatrix} = \begin{pmatrix} -14 + 14 + 2 - 1 \\ -14 + 21 - 2 - 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

como queríamos.

Proposição 1.2.15. *Dado um sistema linear $AX = B$, uma solução do sistema homogêneo X_0 e uma solução do sistema original X_1, então existe uma solução do sistema original X_2 tal que $X_1 = X_0 + X_2$.*

Finalmente, vamos caminhar para uma caracterização geral desse tipo de problema:

EXEMPLO(S):

Mais uma vez, vamos voltar ao sistema

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Já vimos que $S_0 = \{(x, y, z, w) \in \mathbb{R}^4 \mid x = -5z + w, y = 2z - w\}$ - ou seja, uma matrix X_0 está em S_0 se, e somente se, X_0 é da forma

$$X_0 = \begin{pmatrix} -5z + w \\ 2z - w \\ z \\ w \end{pmatrix}.$$

Mas podemos reescrever isso como

$$X_0 = \begin{pmatrix} -5z \\ 2z \\ z \\ 0 \end{pmatrix} + \begin{pmatrix} w \\ -w \\ 0 \\ w \end{pmatrix},$$

que finalmente se torna

$$X_0 = z\begin{pmatrix}-5\\2\\1\\0\end{pmatrix} + w\begin{pmatrix}1\\-1\\0\\1\end{pmatrix}.$$

Ou seja, para quaisquer valores de $z, w \in \mathbb{R}$ temos uma solução do sistema homogêneo.
Também já vimos que $S = \{(x, y, z, w) \in \mathbb{R}^4 \mid x = -5z + w - 3, y = 2z - w + 2\}$, e podemos fazer a mesma análise: X_1 é solução se, e somente se, é da forma

$$X_1 = \begin{pmatrix}-5z+w-3\\2z-w+2\\z\\w\end{pmatrix} = z\begin{pmatrix}-5\\2\\1\\0\end{pmatrix} + w\begin{pmatrix}1\\-1\\0\\1\end{pmatrix} + \begin{pmatrix}-3\\2\\0\\0\end{pmatrix} = X_0 + \begin{pmatrix}-3\\2\\0\\0\end{pmatrix}.$$

Isso nos mostra que qualquer solução X_1 do sistema original é da forma $X_0 + \begin{pmatrix}-3\\2\\0\\0\end{pmatrix}$.

Mas o que é a matriz $\begin{pmatrix}-3\\2\\0\\0\end{pmatrix}$? Se voltarmos a quando resolvemos o sistema pela primeira vez, vamos ver que -3 e 2 é exatamente como fica a matriz resultante após o escalonamento. Os zeros nas linhas de baixo, então, simbolizam o fato de que o sistema não tem informação suficiente para determinar todas as quatro variáveis.
Por fim, note que se X_0 e X_0' são soluções do sistema homogêneo, então, pelo que fizemos acima,

$$X_0 = z\begin{pmatrix}-5\\2\\1\\0\end{pmatrix} + w\begin{pmatrix}1\\-1\\0\\1\end{pmatrix}$$

e

$$X_0' = z'\begin{pmatrix}-5\\2\\1\\0\end{pmatrix} + w'\begin{pmatrix}1\\-1\\0\\1\end{pmatrix},$$

e como $z, z', w.w'$ são números reais, podemos escrever $w' = \frac{ww'}{w} = w\frac{w'}{w}$ e $z' = \frac{zz'}{z} = z\frac{z'}{z}$, ou seja,

$$X_0' = z'\begin{pmatrix}-5\\2\\1\\0\end{pmatrix} + w'\begin{pmatrix}1\\-1\\0\\1\end{pmatrix} = \frac{z'}{z}\left(z\begin{pmatrix}-5\\2\\1\\0\end{pmatrix}\right) + \frac{w'}{w}\left(w\begin{pmatrix}1\\-1\\0\\1\end{pmatrix}\right),$$

e vemos que duas soluções do sistema homogêneo diferem apenas por multiplicações de números.

Para obter um grande resultado conclusivo para esta seção, precisaremos avançar mais no curso.

1.3 Matrizes Inversas

Já vimos que as matrizes possuem o que costumamos chama de *identidade multiplicativa* - ou seja, uma matriz tal que toda matriz vezes ela é a própria matriz.

Definição 1.3.1. *A matriz $I_n \in M_n(\mathbb{R})$ dada por*

$$I_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

*é chamada de **matriz identidade** $n \times n$.*

EXEMPLO(S):

Dada a matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$, é fácil ver que

$$AI_2 = I_2A = A.$$

De fato:

$$AI_2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a \cdot 1 + b \cdot 0 & a \cdot 0 + b \cdot 1 \\ c \cdot 1 + d \cdot 0 & c \cdot 0 + d \cdot 0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A$$

e

$$I_2A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 \cdot a + 0 \cdot c & 1 \cdot b + 0 \cdot d \\ 0 \cdot a + 1 \cdot c & 0 \cdot b + 1 \cdot d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A.$$

Os números reais também têm essa propriedade: Existe um número real (1) tal que para qualquer número real a temos $a \cdot 1 = 1 \cdot a = a$.

Além disso, os números reais têm outra propriedade: Para qualquer número real a existe um (único) número real a^{-1} tal que $a \cdot a^{-1} = a^{-1} \cdot a = 1$. Nós chamamos esse número de *inverso de a*.

Surge então a pergunta natural: Será que matrizes reais têm inversas?

Observação 1.3.2. *Note que se toda matriz quadrada $A \in M_n(\mathbb{R})$ possui inversa, então qualquer sistema da forma $AX = B$ pode ser resolvido multiplicando ambos os lados por A^{-1}:*

$$AX = B \Leftrightarrow A^{-1}(AX) = A^{-1}B \Leftrightarrow (A^{-1}A)X = A^{-1}B \Leftrightarrow X = A^{-1}B.$$

Dito de outra forma, se toda matriz A for inversível, todo sistema $AX = B$ teria solução dada por $X = A^{-1}B$.

Contudo, como já vimos acima, *nem todo sistema possui solução*. Isso nos diz imediatamente que *nem toda matriz possui inversa.*

Para ver quais são as matrizes que possuem inversa, vamos precisar de dar uma definição formal para elas.

Definição 1.3.3. *Dada uma matriz $A \in M_n(\mathbb{R})$, dizemos que A **é inversível** se existe uma matriz $B \in M_n(\mathbb{R})$ tal que $AB = BA = I_n$.*

Exercício (1.3.1)

Prove que se B e C são duas inversas para A, então $B = C$ (dica: comece assim: "como I_n vezes qualquer matriz é a própria matriz, $B = BI_n$ e $C = I_nC$. Além disso, como B e C são inversos de A, temos que $I_n = AC = BA$").

Vamos agora abordar um método para calcular inversos, quando estes existirem:

EXEMPLO(S):

Considere a matriz
$$A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}.$$

Queremos achar uma matriz B tal que $AB = I_2$. Mas isso é um sistema linear! Nós temos uma matriz de dados iniciais (A) que multiplicada por uma matriz que queremos determinar (B) dá uma matriz de resultados (I_2). E nós já sabemos resolver sistemas lineares - via escalonamento! Então, vamos lá:

$$\left(\begin{array}{cc:cc} 2 & 5 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:cc} 1 & \frac{5}{2} & \frac{1}{2} & 0 \\ 1 & 3 & 0 & 1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{cc:cc} 1 & \frac{5}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{cc:cc} 1 & \frac{5}{2} & \frac{1}{2} & 0 \\ 0 & 1 & -1 & 2 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:cc} 1 & 0 & 3 & -5 \\ 0 & 1 & -1 & 2 \end{array}\right)$$

Então estamos dizendo que a matriz inversa de A é a matriz

$$A^{-1} = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}.$$

Vamos testar:

$$AA^{-1} = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 2\cdot 3 + 5\cdot(-1) & 2\cdot(-5) + 5\cdot 2 \\ 1\cdot 3 + 3\cdot(-1) & 1\cdot(-5) + 3\cdot 2 \end{pmatrix} = \begin{pmatrix} 6-5 & -10+10 \\ 3-3 & -5+6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

e

$$A^{-1}A = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3\cdot 2 + (-5)\cdot 1 & 3\cdot 5 + (-5)\cdot 3 \\ (-1)\cdot 2 + 2\cdot 1 & (-1)\cdot 5 + 2\cdot 3 \end{pmatrix} = \begin{pmatrix} 6-5 & -5+5 \\ -2+2 & -5+6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Proposição 1.3.4. *Seja* $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$ *uma matriz. Então a inversa, se existir, é da forma*

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

Essa é a primeira vez que vemos o determinante de uma matriz aparecendo. Mais pra frente vamos definir o determinante de outra maneira e ver para que ele serve.

Outro resultado que vamos usar agora, mesmo que não sejamos capazes de mostrar ainda é o seguinte:

Lema 1.3.5. *Seja $A \in M_n(\mathbb{R})$ matriz quadrada. Se existe matriz $B \in M_n(\mathbb{R})$ tal que $AB = I_n$, então B é inversa de A. Similarmente, se existe matriz $C \in M_n(\mathbb{R})$ tal que $CA = I_n$, então C é inversa de A.*

Em outras palavras, para matrizes, basta checar se o produto em uma ordem dá a identidade, que isso é suficiente para concluir que o produto na outra ordem também dará.

Agora sim, vamos ver um resultado que conseguimos provar:

Lema 1.3.6. *Uma matriz $A \in M_n(\mathbb{R})$ é inversível se, e somente se, A pode ser escalonada em I_n.*

Demonstração. Por um lado, é óbvio que a matriz identidade pode ser escalonada em si mesma (fazendo nada) e que a matriz identidade é inversível: $I_n I_n = I_n$ - de fato, ela é seu próprio inverso. Além disso, já vimos que se A' pode ser obtida de A via escalonamento, então o sistema $AX = B$ tem as mesmas soluções do sistema $A'X = B'$, em que B' é obtida de B pelo mesmo escalonamento que leva A em A'.

Então, se I_n pode ser obtida de A por escalonamento, o sistema $AX = I_n$ tem as mesmas soluções do sistema $I_nX = C$, em que C é obtida de I_n pelo mesmo escalonamento que leva A em I_n. Mas $I_nX = C$ nos diz que $C = X$ e, logo, $AC = I_n$. O lema acima agora nos garante que $CA = I_n$ e portanto A é inversível.

Analogamente, suponha que A é inversível - ou seja, o sistema $AX = I_n$ tem solução B - em outras palavras, $X = B$. Mas podemos re-escrever isso como $I_nX = B$ e ver que esse sistema tem a mesma solução de $AX = I_n$, donde podemos concluir que I_n pode ser obtido de A por escalonamentos. □

Finalmente, vamos encerrar essa seção retomando as matrizes de escalonamento.
Como dissemos anteriormente, nem toda soma de matrizes $e^n_{i,j}$ é um escalonamento.

Proposição 1.3.7. *Toda matriz invertível é um escalonamento.*

Demonstração. Dado um sistema $AX = B$, se E é inversível com inversa E^{-1}, tome Z solução de $(EA)X = EB$. Vamos mostrar que Z também é solução de $AX = B$ - e portanto, E é escalonamento.
De fato:

$$AZ = I_n(AZ) = (E^{-1}E)(AZ) = E^{-1}((EA)Z) = E^{-1}(EB) = (E^{-1}E)B = I_nB = B,$$

logo Z é solução de $AX = B$ e, portanto, E é um escalonamento. □

Gostaríamos de mostrar mais - gostaríamos de mostrar que todo escalonamento é invertível, mas não temos ferramentas para isso ainda. Contudo, para matrizes 2×2 é fácil:

EXEMPLO(S):

Os possíveis escalonamentos são combinações de

- Troca de linhas;
- Multiplicação de uma linha por um número;
- Soma de uma linha a outra linha.

Vamos exibir as matrizes que realizam cada operação:

- A matriz que troca as duas linhas de uma matriz 2×2 é dada por $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, de fato:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 \cdot a + 1 \cdot c & 0 \cdot b + 1 \cdot d \\ 1 \cdot a + 0 \cdot c & 1 \cdot b + 0 \cdot d \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}.$$

- A matriz que multiplica a linha 1 por $\lambda \in \mathbb{R}$ é $\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}$ e a matriz que multiplica a linha 2 por $\mu \in \mathbb{R}$ é $\begin{pmatrix} 1 & 0 \\ 0 & \mu \end{pmatrix}$. De fato,

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda \cdot a + 0 \cdot c & \lambda \cdot b + 0 \cdot d \\ 0 \cdot a + \mu \cdot c & 0 \cdot b + \mu \cdot d \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b \\ \mu c & \mu d \end{pmatrix}.$$

- A matrix que soma a linha 1 na linha 2 é $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ e a matriz que soma a linha 2 na linha 1 é $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. De fato,

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 \cdot a + 0 \cdot c & 1 \cdot b + 0 \cdot d \\ 1 \cdot a + 1 \cdot c & 1 \cdot b + 1 \cdot d \end{pmatrix} = \begin{pmatrix} a & b \\ a + c & b + d \end{pmatrix}$$

e
$$\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}a & b\\c & d\end{pmatrix} = \begin{pmatrix}1\cdot a+1\cdot c & 1\cdot b+1\cdot d\\0\cdot a+1\cdot c & 0\cdot b+1\cdot d\end{pmatrix} = \begin{pmatrix}a+c & b+d\\c & d\end{pmatrix}.$$

Mas todas essas matrizes são inversíveis:

- O inverso da matriz $\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}$ é ela mesma, já que trocar as duas linhas duas vezes é a mesma coisa de não fazer nada.
- Os inversos das matrizes $\begin{pmatrix}\lambda & 0\\0 & 1\end{pmatrix}$ e $\begin{pmatrix}1 & 0\\0 & \mu\end{pmatrix}$ são as matrizes $\begin{pmatrix}\lambda^{-1} & 0\\0 & 1\end{pmatrix}$ e $\begin{pmatrix}1 & 0\\0 & \mu^{-1}\end{pmatrix}$, respectivamente, já que o inverso de "multiplicar uma linha por x" é "dividir uma linha por x".
- Os inversos das matrizes $\begin{pmatrix}1 & 0\\1 & 1\end{pmatrix}$ e $\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}$ são, respectivamente, as matrizes $\begin{pmatrix}1 & 0\\-1 & 1\end{pmatrix}$ e $\begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}$, já que o inverso de "somar a linha i na linha j" é "subtrair a linha i da linha j".

Finalmente, note que se A e B são inversíveis, então o produto AB também é, pois $(AB)(B^{-1}A^{-1}) = A((BB^{-1})A^{-1}) = A(I_nA^{-1}) = AA^{-1} = I_n$. Assim, como qualquer escalonamento é produto das matrizes acima, segue que qualquer escalonamento é invertível, já que cada uma delas é.

Exercício (1.3.2)

Prove que os inversos que apresentamos acima são, de fato, inversos.

Capítulo 2

Vetores em $\mathbb{R}^2$

2.1 Definições e Propriedades Básicas

Em matemática, dados dois conjuntos A e B, nós denotamos por $A \times B$ o conjunto de todos os pares ordenados de elementos de A e B.

EXEMPLO(S):

Se $A = \{a, b, c, d\}$ e $B = \{1, 2, 3\}$, então

$$A \times B = \{(a, 1), (a, 2), (a, 3), (b, 1), (b, 2), (b, 3), (c, 1), (c, 2), (c, 3), (d, 1), (d, 2), (d, 3)\}.$$

Se $C = \{\ltimes, \rtimes\}$, então

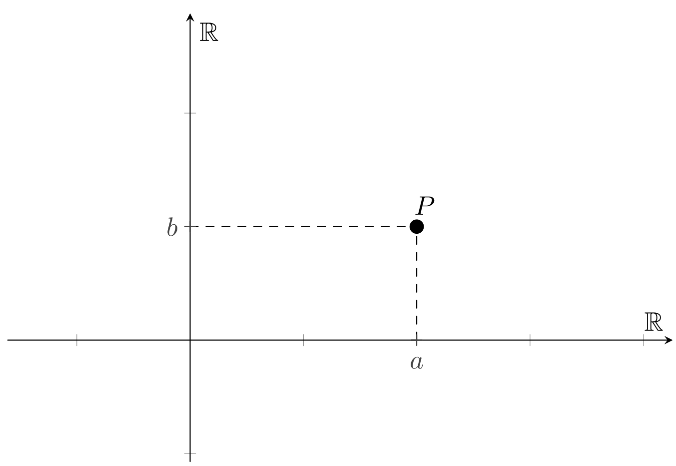
$$C \times B = \{(\ltimes, 1), (\ltimes, 2), (\ltimes 3), (\rtimes, 1), (\rtimes, 2), (\rtimes 3)\}.$$

Se $D = \{$calça, camiseta$\}$ e $E = \{$azul, verde, amarela$\}$, então

$D \times E = \{$calça azul, calça verde, calça amarela, camiseta azul, camiseta verde, camiseta amarela$\}$.

Assim, o conjunto de todos os pares ordenados de números reais é o conjunto $\mathbb{R} \times \mathbb{R}$. Como o símbolo $\times$ remete a um produto, nós chamamos esse conjunto de $\mathbb{R}^2$.

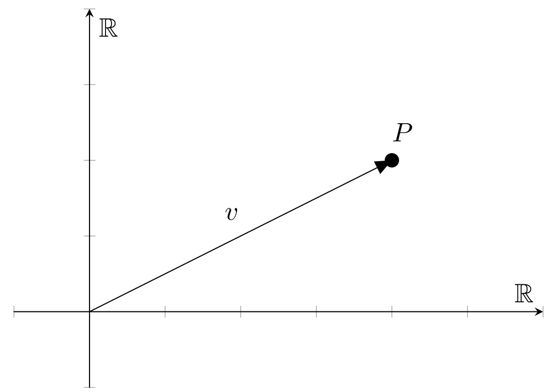
Vamos representar os elementos de $\mathbb{R}^2$ como pontos no plano da seguinte maneira: Correspondemos o elemento $(a, b) \in \mathbb{R}^2$ ao ponto P dado pelas coordenadas a e b, como abaixo:



Claramente essa correspondência é biunívoca - pois dado um ponto no plano, ele é unicamente determinado por duas coordenadas.

Assim estabelecemos um paralelo entre **pontos no plano** e **elementos de** $\mathbb{R}^2$. Por esse motivo, elementos de $\mathbb{R}^2$ são muitas vezes chamados de *pontos* de $\mathbb{R}^2$.

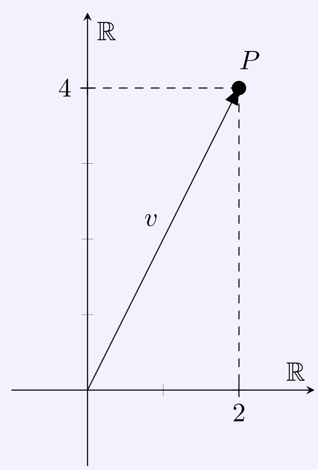
Além disso, podemos fazer outra correspondência:



Vamos chamar de **vetor em** $\mathbb{R}^2$ uma seta partindo da origem e terminando em qualquer ponto do plano. A figura acima estabelece uma correspondência biunívoca entre vetores no plano e pontos no plano: Para cada vetor v existe um único ponto final P. Similarmente, para cada ponto P existe um único vetor v que termina em P.

Juntando tudo isso, nós vemos que **o conjunto de pontos, pares ordenados e vetores em** $\mathbb{R}^2$ **são "a mesma coisa"**. Com isso em mente, daqui para frente vamos usar todos esses significados, usando sempre o significado mais conveniente naquele momento.

EXEMPLO(S):



A figura acima em $\mathbb{R}^2$ representa simultaneamente os três conceitos: O vetor v tem como ponto final o ponto P cujas coordenadas são $(2, 4)$.

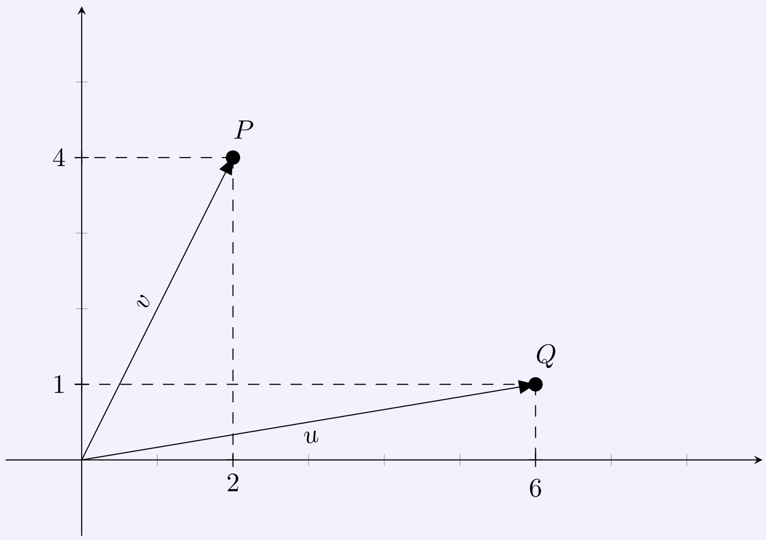
2.1.1 Somas e produtos por números

Dados dois pares ordenados (a, b) e (c, d) em $\mathbb{R}^2$, nós podemos notar que como a, b, c, d são números reais, nós podemos somar $a + c$ e $b + d$, e ambos serão números reais. Assim, faria sentido definir uma operação de soma em $\mathbb{R}^2$ dada por

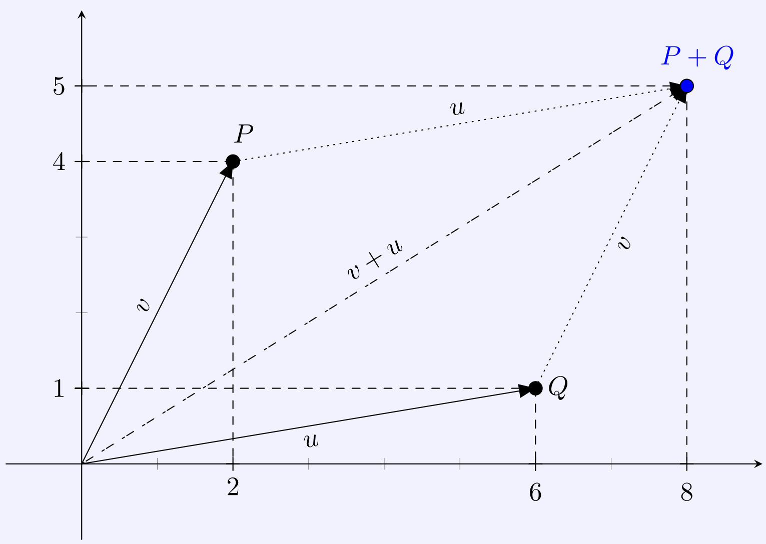
$$(a, b) + (c, d) := (a + c, b + d).$$

Contudo, pelo que já vimos acima, pares, pontos e vetores são "a mesma coisa". Será que essa soma tem alguma interpretação via pontos e vetores?

EXEMPLO(S):



Na figura acima temos os pares $(2, 4)$ e $(6, 1)$ representados pelos pontos P e Q e pelos vetores v e u, respectivamente.



Então, somando os pares $(2, 4)$ e $(6, 1)$ obtemos o par $(2 + 6, 4 + 1) = (8, 5)$. Geometricamente, temos a figura acima: Dados os vetores v e u, com coordenadas $(2, 4)$ e $(6, 1)$, respectivamente, o vetor $v + u$ que é obtido simplesmente pela concatenação do vetor u ao vetor v.

Definição 2.1.1. *Dados dois vetores $u, v \in \mathbb{R}^2$ com coordenadas $u = (a, b)$ e $v = (c, d)$, definimos a* ***soma de*** *u* ***com*** *v como sendo o vetor $u + v$ dado por*

$$u + v := (a + c, b + d).$$

Exercício (2.1.1)

Mostre que essa soma satisfaz as seguintes propriedades:

- (Comutatividade) Para quaisquer vetores $u, v \in R^2$, $u + v = v + u$;
- (Associatividade) Para quaisquer três vetores $u, v, w \in \mathbb{R}^2$, $u + (v + w) = (u + v) + w$;
- (Existência e unicidade de elemento neutro) Existe um (único) vetor $\overrightarrow{0} \in \mathbb{R}^2$ tal que $v + \overrightarrow{0} = v$ para qualquer $v \in \mathbb{R}^2$;
- (Existência e unicidade de inversos) Para qualquer vetor $v \in \mathbb{R}^2$ existe um (único) vetor $-v \in \mathbb{R}^2$ tal que $v + (-v) = \overrightarrow{0}$.

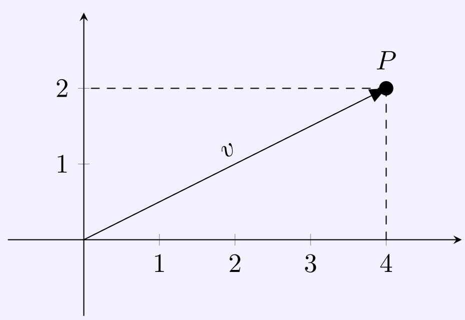
Por outro lado, é intuitivamente óbvio que se $(a, b) \in \mathbb{R}^2$ e $\lambda \in \mathbb{R}$, certamente o par $(\lambda a, \lambda b) \in \mathbb{R}^2$. Então parece natural definir

$$\lambda(a, b) := (\lambda a, \lambda b).$$

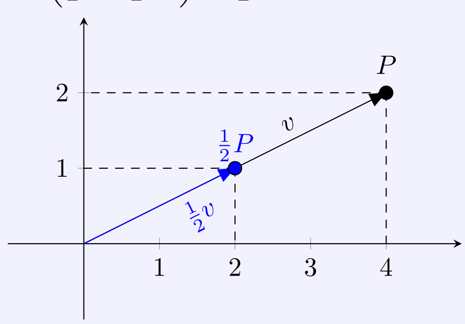
Contudo, será que temos uma boa interpretação geométrica pra isso, como para a soma?

EXEMPLO(S):

Fixe o número $\lambda = \dfrac{1}{2} \in \mathbb{R}$ e o vetor $v = (4, 2) \in \mathbb{R}^2$.



O que seria o vetor $u = (2, 1) = \left(\dfrac{1}{2} \cdot 4, \dfrac{1}{2} \cdot 2\right) = \dfrac{1}{2} \cdot v$?



Então o vetor $\dfrac{1}{2} \cdot v$ é o vetor **na mesma direção de** v, mas **com tamanho igual a** $\dfrac{1}{2}$ **vezes o tamanho de** v.

Exercício (2.1.2)

O que seria, então multiplicar por um número negativo, por exemplo $\lambda = -1 \in \mathbb{R}$? Faça um desenho do que seria uma interpretação geométrica para $(-1) \cdot (4, 2)$.

Exercício (2.1.3)

Mostre que a multiplicação de vetores por números satisfaz:

- (Comutatividade) $(\lambda\mu)v = (\mu\lambda)v$ para quaisquer números $\lambda, \mu \in \mathbb{R}$ e vetor $v \in \mathbb{R}^2$;
- (Associatividade) $\lambda(\mu v) = (\lambda\mu)v$ e para quaisquer números $\lambda, \mu \in \mathbb{R}$ e vetor $v \in \mathbb{R}^2$;
- (Existência e unicidade de elemento neutro) Existe um (único) numero real $u \in \mathbb{R}$ tal que $uv = v$ para qualquer vetor $v \in \mathbb{R}^2$;
- (Distributividades) Para quaisquer $\lambda, \mu \in \mathbb{R}$ e $v, u \in \mathbb{R}^2$, temos que $\lambda(v + u) = \lambda v = \lambda u$ e $(\lambda + \mu)v = \lambda v + \mu v$.

2.2 Vetores e Matrizes

Antes de estudar mais a fundo os vetores, contudo, precisamos estabelecer um paralelo entre o estudo de matrizes e o estudo de vetores:

Considere a função $[-] : \mathbb{R}^2 \to M_{2\times1}(\mathbb{R})$ dada por

$$[(x, y)] := \begin{pmatrix} x \\ y \end{pmatrix} \in M_{2\times1}(\mathbb{R}).$$

Essa função "não faz nada": Ela só nos permite enxergar vetores como matrizes coluna. Similarmente, podemos considerar a função $V : M_{2\times1}(\mathbb{R}) \to \mathbb{R}^2$ dada por

$$V \begin{pmatrix} x \\ y \end{pmatrix} := (x, y) \in \mathbb{R}^2.$$

Essa função também "não faz nada": Ela só nos permite enxergar matrizes coluna como vetores. Contudo, note que:

$$V([(x, y)]) = V \begin{pmatrix} x \\ y \end{pmatrix} = (x, y)$$

$$\left[V \begin{pmatrix} x' \\ y' \end{pmatrix} \right] = [(x', y')] = \begin{pmatrix} x' \\ y' \end{pmatrix}$$

ou seja, *essas funções são inversas*! Isso nos diz que os conjuntos $\mathbb{R}^2$ e $M_{2\times1}(\mathbb{R})$ são "a mesma coisa", apenas vistos com outros olhos. Por isso, vamos adicionar mais um significado à palavra vetor: Não só pontos no plano, ou pares ordenados em $\mathbb{R}^2$ ou setas partindo da origem, mas também matrizes coluna com duas entradas.

2.2.1 Funções lineares

Nesta seção, então, vamos interpretar vários dos resultados do capítulo anterior, mas agora com o auxílio dos diversos significados de vetor.

Por exemplo, dada uma matriz quadrada $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, a multiplicação AX, em que X é alguma matriz coluna com duas entradas, é da seguinte forma:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

Então para cada valor de x e y, o resultado do produto AX é diferente - ora, isso nos permite definir:

Definição 2.2.1. *Para cada matriz $A \in M_2(\mathbb{R})$, definimos $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ a função dada por*

$$T_A(x, y) := V(A[(x, y)]).$$

Ou seja, para calcular $T_A(x, y)$, a gente escreve (x, y) como uma matriz coluna $\begin{pmatrix} x \\ y \end{pmatrix}$, multiplica ele por A e escreve o resultado como um vetor.

EXEMPLO(S):

Dada a matriz $A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$, a função T_A é tal que

$$T_A(x, y) = V(A[(x, y)]) = V\left(\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}\right) = V\left(\begin{pmatrix} 2x + 5y \\ x + 3y \end{pmatrix}\right) = (2x + 5y, x + 3y)$$

em outras palavras, a função T_A é tal que $T_A(x, y)$ é simplesmente

$$\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + 5y \\ x + 3y \end{pmatrix}$$

escrito em forma de vetor, ou seja $(2x + 5y, x + 3y)$.
Similarmente,

$$T_{I_2}(x, y) = V(I_2[(x, y)]) = V\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}\right) = V\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = (x, y),$$

e, analogamente a como fizemos acima, ou seja, $T_{I_2}(x, y)$ é simplesmente

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

escrito em forma de vetor, ou seja, (x, y) - em outras palavras, a função T_{I_2} é a função identidade que leva todo vetor nele mesmo.

Como essas funções são dadas por matrizes, será que elas têm as propriedades de matrizes?

Proposição 2.2.2. *Dados uma matriz $A \in M_2(\mathbb{R})$, qualquer número $\lambda \in \mathbb{R}$ e dois vetores $v, u \in \mathbb{R}^2$, temos:*

$$T_A(\lambda v) = \lambda T_A(v)$$

$$T_A(v + u) = T_A(v) + T_A(u).$$

Exercício (2.2.1)

Prova a proposição acima (dica: calcule explicitamente $T_A(\lambda v)$ e $\lambda T_A(v)$ e veja que são a mesma coisa. Idem para as somas.).

Essas funções são bastante interessantes. Elas nos dão um “dicionário” que nos permite ver o mundo dos vetores como se fosse o mundo das matrizes, e vice-e-versa. Seria, então, interessante se a gente conseguisse um método de determinar quando uma função $f : \mathbb{R}^2 \to \mathbb{R}^2$ é dessa forma.

EXEMPLO(S):

A função $f : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $f(x, y) = (x, x^2)$ **não** é dada por uma matriz. Podemos ver isso usando a proposição acima: Se f fosse igual a T_A para alguma matriz $A \in M_2(\mathbb{R})$, $f(v + u)$ seria igual a $f(v) + f(u)$. Mas

$$f(0, 0) = (0, 0^2) = (0, 0),$$

enquanto

$$f(1, 0) + f(-1, 0) = (1, 1^2) + (-1, 1^2) = (0, 2).$$

Agora, note que $(0, 0) = (1, 0) + (-1, 0)$. Então nós acabamos de mostrar que $f((1, 0) + (-1, 0)) \neq f(1, 0) + f(-1, 0)$, ou seja, f **não pode ser dada por** T_A.

Proposição 2.2.3. *Uma função $f : \mathbb{R}^2 \to \mathbb{R}^2$ é dada por uma matriz se, e somente se, existem números reais $a, b, c, d \in \mathbb{R}$ tais que $f(x, y) = (ax + by, cx + dy)$.*

Demonstração. Suponha que f é dada por uma matriz, ou seja, $f = T_A$ para alguma matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$. Então, para qualquer vetor $(x, y) \in \mathbb{R}^2$ temos:

$$f(x, y) = T_A(x, y) = V(A([x, y])) = V\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\right) = V\left(\begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}\right) = (ax + by, cx + dy),$$

como queríamos mostrar.

Por outro lado, suponha que $f(x, y) = (ax + by, cx + dy)$. Afirmamos que a matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ é tal que $f = T_A$. De fato,

$$T_A(x, y) = (ax + by, cx + dy) = f(x, y),$$

o que encerra a prova. □

Finalmente, agora vamos resolver um problema que será central nesse curso: Será que toda função $f : \mathbb{R}^2 \to \mathbb{R}^2$ que preserva somas e produtos por números é dada por matrizes?

Teorema 2.2.4. *Seja $f : \mathbb{R}^2 \to \mathbb{R}^2$ uma função. Então $f(v + \lambda u) = f(v) + \lambda f(u)$ para todo $\lambda \in \mathbb{R}$ e todos $v, u \in \mathbb{R}^2$ se, e somente se, $f = T_A$ para alguma matriz $A \in M_2(\mathbb{R})$.*

Demonstração. Já mostramos que se $f = T_A$, então $f(v + \lambda u) = f(v) + \lambda f(u)$.

Suponha, então, que $f(v + \lambda u) = f(v) + \lambda f(u)$. Vamos mostrar que $f = T_A$ para alguma matriz A.

Calculando $f(x, y)$ temos:

$$\begin{aligned} f(x, y) &= f(x, 0) + f(0, y) && \text{(já que } f \text{ preserva somas)} \\ &= xf(1, 0) + yf(0, 1) && \text{(já que } f \text{ preserva multipli-} \\ & && \text{cação por números).} \end{aligned}$$

Sejam então $f(1, 0) = (a, b)$ e $f(0, 1) = (c, d)$. Assim,

$$\begin{aligned} f(x, y) &= xf(1, 0) + yf(0, 1) \\ &= x(a, b) + y(c, d) \\ &= (xa, ab) + (yc, yd) = (xa + yc, xb + yd) \end{aligned}$$

e, pela proposição anterior, sabemos que $f = T_A$, onde $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$, como queríamos mostrar. □

Definição 2.2.5. *Uma função $f : \mathbb{R}^2 \to \mathbb{R}^2$ será dita **linear** se $f(v + \lambda u) = f(v) + \lambda f(u)$ para quaisquer $\lambda \in \mathbb{R}$ e $v, u \in \mathbb{R}^2$.*

Então o teorema acima nos diz que as funções lineares são exatamente as funções "multiplique o seu vetor por uma matriz".

EXEMPLO(S):

Sejam $f, g : \mathbb{R}^2 \to R^2$ funções lineares. Será que $g \circ f$ e $f \circ g$ também são lineares?
Por exemplo, se $f = T_A$ e $g = T_B$, com

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} \text{ e } B = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix},$$

podemos lembrar que $(f \circ g)(x, y)$ é simplesmente

$$\begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} \left(\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right)$$

escrito em forma de vetor. Contudo, o produto de matrizes é associativo! Então esse produto é a mesma coisa que

$$\left(\begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \right) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 & 14 \\ 9 & 25 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

ou seja, a função $f \circ g$ é dada pela matriz $AB = \begin{pmatrix} 5 & 14 \\ 9 & 25 \end{pmatrix}$.
Similarmente, a função $g \circ f$ leva um vetor (x, y) no vetor dado pelo produto de matrizes

$$\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \left(\begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right)$$

escrito em forma de vetor. Contudo, como notamos acima, o produto de matrizes é associativo, então esse produto pode ser escrito como

$$\left(\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} \right) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 12 & 31 \\ 7 & 18 \end{pmatrix},$$

ou seja, a função $g \circ f$ é dada pela matriz BA.
Em suma, isso nos diz o seguinte: A definição de multiplicação de matrizes é feita de forma que a composição de função lineares seja a função dada pelo produto das matrizes correspondentes.
Ou, em outras palavras, poderíamos definir AB como sendo a matriz tal que $T_A \circ T_B = T_{AB}$.

Exercício (2.2.2)

Mostre, sem utilizar exemplos concretos, que se f e g são lineares, então $f \circ g$ e $g \circ f$ também são (dica: use o exemplo acima).

2.3 Subespaços

2.3.1 Núcleo e imagem

Agora vamos introduzir dois conceitos que vão nos ajudar a entender ainda melhor o espaço de vetores e transformações lineares.

Definição 2.3.1. *Dada uma função linear* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *definimos o* ***núcleo de*** f *como sendo o conjunto* $\operatorname{Ker} f$ *dado por*

$$\operatorname{Ker} f := \{v \in \mathbb{R}^2 \mid f(v) = (0,0)\},$$

ou seja, são os pontos de $\mathbb{R}^2$ *que vão na origem quando aplicamos* f*.*

Observação 2.3.2. *O símbolo* Ker *vem da palavra inglesa* kernel *que significa "a parte macia de um grão" ou "a parte central".*

EXEMPLO(S):

Considere função $f : \mathbb{R}^2 \to \mathbb{R}^2$ que leva (x, y) em $(x, 0)$, ou seja, $f(x, y) = (x, 0)$. Claramente, f é linear (verifique!) e é dada pela matriz $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$:

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

Quem é o núcleo de f? São todos os pontos $(x, y) \in \mathbb{R}^2$ tais que $f(x, y) = (0, 0)$. Mas $f(x, y) = (x, 0)$. Então $(x, 0) = f(x, y) = (0, 0)$ se, e somente se, $x = 0$ - ou seja,

$$\operatorname{Ker} f = \{(x, y) \in \mathbb{R}^2 \mid x = 0\}.$$

Será que isso nos diz alguma coisa sobre a matriz A? Bom, resolvendo o sistema $AX = 0$ temos:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ou seja, $x = 0$ e $0y = 0$. Isso nos diz que

$$S_0 := \{(x, y) \in \mathbb{R}^2 \mid x = 0\}.$$

Mas isso é a mesma coisa que $\operatorname{Ker} f$!
Em outras palavras, o núcleo de uma função linear **nada mais é do que o conjunto solução do sistema homogêneo associado àquela função**.

Antes de avançarmos, vamos ver o que acontece quando temos uma matriz com uma linha nula:

$$A = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}.$$

Nesse caso, o sistema homogêneo

$$\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

nos dá a equação $ax + by = 0$... E só. Então se ambos $a = b = 0$, isso não teria graça - estaríamos dizendo simplesmente que $0 + 0 = 0$.
Vamos supor, então, que ou a ou b é diferente de 0 - por exemplo, a: Nesse caso, podemos isolar x, fazendo $x = \dfrac{-b}{a} y$.
Assim, vemos que para cada possível valor de y podemos encontrar um único valor de x correspondente. Ora, isso nos diz que se uma matriz tem uma linha nula, então certamente o sistema homogêneo tem várias soluções além da trivial - basta dar valores para y na fórmula acima!

Proposição 2.3.3. *Para qualquer função linear* $f : \mathbb{R}^2 \to \mathbb{R}^2$*,* $\operatorname{Ker} f \neq \varnothing$*.*

Demonstração. Como f é linear, existe matriz A tal que $f = T_A$. Mas, pelo exemplo acima, vimos que $\operatorname{Ker} f = S_0$, em que S_0 é o conjunto de soluções de $AX = 0$. Contudo, já vimos que $(0,0) \in S_0$. Segue que $(0,0) \in \operatorname{Ker} f$ e, portanto, $\operatorname{Ker} f \neq \varnothing$.

Alternativamente, podemos calcular explicitamente:

$$f(0,0) = f(1,0) + f(-1,0) = f(1,0) - f(1,0) = (0,0)$$

e chegar à mesma conclusão. □

Agora vamos usar o núcleo para inferir alguns resultados importantes da álgebra vetorial:

Proposição 2.3.4. *Seja $f : \mathbb{R}^2 \to \mathbb{R}^2$ linear. Então f é injetiva se, e somente se,* $\operatorname{Ker} f = \{(0,0)\}$.

Demonstração. Se f é injetiva, então $f(v) = f(u)$ implica $v = u$. Escolha, então, qualquer $v \in \operatorname{Ker} f$ - ou seja, $f(v) = (0,0)$. Mas, pela proposição anterior, sabemos que $f(0,0) = (0,0)$. Agora, como f é injetiva e v e $(0,0)$ têm a mesma imagem, isso implica que $v = (0,0)$. Ora, nós mostramos que qualquer elemento v no núcleo **tem** que ser $(0,0)$ - segue que o único elemento do núcleo é $(0,0)$.

Por outro lado, suponha que $\operatorname{Ker} f = \{(0,0)\}$. Tome, então, $v, u \in \mathbb{R}^2$ tais que $f(v) = f(u)$. Se mostrarmos que $v = u$, teremos mostrado que f é injetiva. Mas f é linear, então dizer que $f(v) = f(u)$ é a mesma coisa que dizer que $f(v - u) = (0,0)$. Mas estamos supondo que $\operatorname{Ker} f = \{(0,0)\}$ - ou seja, se $f(w) = (0,0)$, então $w = (0,0)$. Como nós temos que $f(v - u) = (0,0)$, nossa hipótese de $\operatorname{Ker} f = \{(0,0)\}$ nos garante que $v - u = (0,0)$. Finalmente, isso é a mesma coisa que $v = u$. Logo, nós concluímos que f é, de fato, injetiva, o que encerra a demonstração. □

EXEMPLO(S):

Vamos voltar a comparar núcleos e soluções do sistema homogêneo. O resultado que mostramos acima nos diz que uma função linear é injetiva se, e somente se, o núcleo é "trivial" - ou seja, o sistema homogêneo só possui a solução trivial $S_0 = \{(0,0)\}$. Como podemos interpretar, do ponto de vista de matrizes, a injetividade de uma função linear?
Pela proposição acima, uma função linear é injetiva se, e somente se, o único ponto que vai no zero é o zero. Em termos de matrizes, então, isso significa "uma matriz A é *injetiva* se $AX = 0$ implica $X = 0$". Em outras palavras, uma matriz é injetiva se o sistema homogêneo *possui solução única* (a solução trivial).
Mas já vimos acima que um sistema tem solução única se, e somente se, a forma escalonada da matriz A não possui linhas de zeros.
Juntando tudo isso, vemos que *uma matriz corresponde a uma função linear injetiva se, e somente se, sua forma escalonada* ***não possui*** *linhas de zeros.*

Por exemplo, considere as matrizes

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix}.$$

A já está escalonada, e não possui linhas de zeros, então certamente T_A é injetiva.
B ainda não está escalonada:

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Então B escalonada é A, que já vimos que não possui linhas de zeros. Segue que T_B é injetiva.
C também não está escalonada ainda:

$$\begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 3/2 \\ 4 & 6 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 3/2 \\ 0 & 0 \end{pmatrix}.$$

Então C escalonada tem uma linha de zeros, ou seja, T_C **não** é injetiva.

A discussão e os exemplos acima sugerem que toda função linear injetiva possui inverso. Isso pode soar estranho a princípio: Por exemplo, esquecendo o adjetivo "linear", sabemos que uma função f entre dois conjuntos quaisquer possui inversa se, e somente se, f é injetiva e sobrejetiva. Dito de outra maneira, em geral *não basta uma função ser injetiva para que ela seja inversível.*

Essa é uma das grandes vantagens de funções lineares: Com o acréscimo desse adjetivo "linear", podemos mostrar o seguinte resultado:

Teorema 2.3.5. *Seja $f : \mathbb{R}^2 \to \mathbb{R}^2$ uma função linear. Então f é inversível se, e somente se, f é injetiva, o que acontece se, e somente se, f é sobrejetiva.*

Em outras palavras, no mundo dos vetores e funções lineares, *ser injetivo, sobrejetivo e inversível é* ***a mesma coisa***.

Vamos agora caminhar em direção a demonstra esse resultado. Para isso vamos precisar de alguns conceitos.

Definição 2.3.6. *Seja $S \subset \mathbb{R}^2$ um subconjunto de $\mathbb{R}^2$ tal que para quaisquer $v, u \in S$ e qualquer $\lambda \in \mathbb{R}$ temos que $v + \lambda u \in S$. Então diremos que S é um* ***subespaço de*** *$\mathbb{R}^2$.*

Observação 2.3.7. *Note que, por definição, se $v \in S$ e S é subespaço de $\mathbb{R}^2$, então λv também está em S, para qualquer $\lambda \in \mathbb{R}$. Disso segue que se $S \neq \varnothing$, então S tem infinitos elementos.*

Exercício (2.3.1)

Mostre que $\{(0,0)\}$, $\{(x,y) \in \mathbb{R}^2 \mid y = \lambda x \text{ para algum } \lambda \in \mathbb{R}\}$ e $\mathbb{R}^2$ são subespaços de $\mathbb{R}^2$.

Com isso, podemos finalmente ter um bom resultado:

Proposição 2.3.8. *Para qualquer função linear $f : \mathbb{R}^2 \to \mathbb{R}^2$, temos que $\operatorname{Ker} f$ é um subespaço de $\mathbb{R}^2$.*

Demonstração. Tome $v, u \in \operatorname{Ker} f$ e $\lambda \in \mathbb{R}$. Vamos mostrar que $f(v + \lambda u) = (0,0)$ e, portanto, $v + \lambda u \in \operatorname{Ker} f$. Calculando:

$$f(v + \lambda u) = f(v) + f(\lambda u) = f(v) + \lambda f(u),$$

já que f é linear. Agora, como $v, u \in \operatorname{Ker} f$, temos que $f(v) = f(u) = (0,0)$, logo

$$f(v + \lambda u) = f(v) + \lambda f(u) = (0,0) + \lambda(0,0) = (0,0) + (0,0) = (0,0),$$

ou seja, $v + \lambda u \in \operatorname{Ker} f$.

Segue que $\operatorname{Ker} f$ é subespaço de $\mathbb{R}^2$, como queríamos mostrar. □

Corolário 2.3.9. *Seja $f : \mathbb{R}^2 \to \mathbb{R}^2$ linear. Então $\operatorname{Ker} f$ ou tem um único ponto ($(0,0)$) ou tem infinitos pontos.*

Observação 2.3.10. *O resultado acima é uma justificativa para a afirmação de que um sistema linear ou não tem solução, ou tem uma solução ou tem infinitas soluções.*

Exercício (2.3.2)

Prove a observação acima - ou seja, tome qualquer sistema linear $AX = B$ com $A \in M_2(\mathbb{R})$ e mostre que o conjunto de soluções ou é vazio, ou tem um único ponto, ou tem infinitos pontos.

Temos, por outro lado, outro subespaço associado a uma função linear - a imagem desta:

Definição 2.3.11. *Seja $f : X \to Y$ uma função entre os conjuntos X e Y. Denotamos por $\operatorname{Im} f$ a* ***imagem de*** *f o conjunto dado por*

$$\operatorname{Im} f := \{v \in Y \mid \exists u \in X \textit{ tal que } f(u) = v\}.$$

Ou seja, a imagem de uma função são todos os pontos que você pode obter aplicando aquela função.

EXEMPLO(S):

Considere a função $f : \mathbb{N} \to \mathbb{N}$ dada por $f(n) = n + 1$. Qual a imagem de f? São todos os números naturais que podem ser escritos como "algum número natural + 1". Por exemplo, 2 está na imagem, pois $2 = 1 + 1$. Similarmente, 7 está na imagem, pois $7 = 6 + 1$. O único número natural que não está na imagem é o 0 - de fato, não existe nenhum número natural n tal que $0 = n + 1$. Então

$$\operatorname{Im} f = \mathbb{N} - \{0\} = \{1, 2, 3, 4, 5, \cdots\}.$$

Considere agora a função $\|-\| : \mathbb{R}^2 \to \mathbb{R}$ dada por $\|(x, y)\| := \sqrt{x^2 + y^2}$. Essa função é chamada de **norma** - ela mede o tamanho de um vetor em $\mathbb{R}^2$. Qual a imagem dessa função, ou seja, quais são todos os possíveis tamanhos de vetores?

Olhando para a fórmula da função, vemos que a norma é a raiz quadrada (que é sempre não-negativa) da soma de dois números positivos (que é sempre não-negativa), ou seja, a imagem de $\|-\|$ são todos os números reais não-negativos.

De fato, para qualquer número não-negativo $\lambda \in \mathbb{R}$ podemos exibir um vetor com essa norma: $(\lambda, 0)$. Computando a norma de $(\lambda, 0)$ temos:

$$\|(\lambda, 0)\| = \sqrt{\lambda^2 + 0^2} = \sqrt{\lambda^2}$$

e como $\lambda \geq 0$, temos finalmente $\sqrt{\lambda^2} = \lambda$, ou seja, existe (pelo menos) um vetor v tal que $\|v\| = \lambda$ e, portanto, $\lambda \in \operatorname{Im}\|-\|$.

Proposição 2.3.12. *Para qualquer função linear $f : \mathbb{R}^2 \to \mathbb{R}^2$, temos que $\operatorname{Im} f$ é um subespaço de $\mathbb{R}^2$.*

Demonstração. Tome v, u em $\operatorname{Im} f$ e $\lambda \in \mathbb{R}$ qualquer. Vamos mostrar que $v + \lambda u$ também está na imagem.

Por definição, se $v \in \operatorname{Im} f$, então v é imagem de alguém, ou seja, existe $v' \in \mathbb{R}^2$ tal que $f(v') = v$. Similarmente, como $u \in \operatorname{Im} f$ ele é imagem de alguém, ou seja, existe $u' \in \mathbb{R}^2$ tal que $f(u') = u$.

Afirmamos que $v + \lambda u$ é imagem de $v' + \lambda u'$ por f: De fato, $f(v' + \lambda u') = f(v') + \lambda f(u')$, já que f é linear. Agora, usando que $v = f(v')$ e $u = f(u')$ temos finalmente que $f(v' + \lambda u') = v + \lambda u$, donde podemos concluir que $v + \lambda u \in \operatorname{Im} f$ e, portanto, $\operatorname{Im} f$ é subespaço. □

Corolário 2.3.13. *Para qualquer função linear $f : \mathbb{R}^2 \to \mathbb{R}^2$, $(0, 0) \in \operatorname{Im} f$.*

2.3.2 Geradores, dependência e independência linear

Na seção anterior introduzimos o conceito de subespaços de $\mathbb{R}^2$ que são subconjuntos que se comportam "como $\mathbb{R}^2$" - ou seja, são fechados para soma e multiplicação por escalar. Será que subespaços herdam alguma outra propriedade interessante de $\mathbb{R}^2$?

EXEMPLO(S):

Considere qualquer vetor $(x, y) \in \mathbb{R}^2$. Esse vetor pode ser escrito de maneira **única** como $(x, y) = x(1, 0) + y(0, 1)$.

Por outro lado, também podemos escrever de maneira única qualquer vetor (x, y) de $\mathbb{R}^2$ como $(x, y) = \frac{x + y}{2}(1, 1) + \frac{x - y}{2}(1, -1)$ - por exemplo, $(5, 3)$ pode ser escrito tanto como $5(1, 0) + 3(0, 1)$ quanto como $4(1, 1) + (1, -1)$: De fato, $5(1, 0) + 3(0, 1) = (5, 0) + (0, 3) = (5, 3)$ e $4(1, 1) + (1, -1) = (4, 4) + (1, -1) = (5, 3)$.

De maneira geométrica, isso significa que para chegar em qualquer ponto no plano partindo da origem podemos dar passos horizontais e verticais ($(1, 0)$ e $(0, 1)$) ou passos nas direções nordeste e noroeste ($(1, 1)$ e $(1, -1)$) (sempre lembrando que podemos dar passos para frente e para trás).

Definição 2.3.14. *Dados dois vetores $v, u \in \mathbb{R}^2$, dizemos que u **é gerado por** v se existe algum número real λ tal que $u = \lambda v$.*

*Similarmente, dado um vetor v, o conjunto de todos os vetores gerados por v será chamado de **conjunto gerado por** v e denotado por $\mathrm{Gen}(v)$. Nesse caso diremos que v é um **gerador de** $\mathrm{Gen}(v)$.*

Definição 2.3.15. *Dados um vetor $u \in \mathbb{R}^2$ e uma coleção finita de vetores $V = \{v_1, v_2, \cdots, v_n\} \subseteq \mathbb{R}^2$, dizemos que u **é gerado por** V se existe uma coleção de números reais $\{\lambda_1, \lambda_2, \cdots, \lambda_n\} \subseteq \mathbb{R}$ tal que*

$$u = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n.$$

*Similarmente, dada uma coleção de vetores V, o conjunto de todos os vetores gerados por V será chamado de **conjunto gerado por** V e denorado por $\mathrm{Gen}(V)$. Nesse caso diremos que V é um **conjunto de geradores de** $\mathrm{Gen}(V)$ ou, equivalentemente, um **gerador de** $\mathrm{Gen}(V)$.*

Observação 2.3.16. *A terminologia* Gen *para se referir ao conjunto gerado vem do termo inglês* generated*, "gerado" em português.*

EXEMPLO(S):

Continuando o exemplo acima, vimos que qualquer vetor $(x, y) \in \mathbb{R}^2$ pode ser escrito tanto como $x(1,0) + y(0,1)$ quanto como $\frac{x+y}{2}(1,1) + \frac{x-y}{2}(1,-1)$. Ou seja, tanto $E = \{(1,0),(0,1)\}$ quanto $V = \{(1,1),(1,-1)\}$ são geradores de $\mathbb{R}^2$ - ou, em outras palavras, $\mathbb{R}^2 = \mathrm{Gen}(E) = \mathrm{Gen}(V)$.
Por outro lado, dado o vetor $v = (2,3)$, qual o conjunto gerado por v - ou seja, quem é $\mathrm{Gen}(v)$?
Por definição, temos:

$$\mathrm{Gen}(v) = \{u \in \mathbb{R}^2 \mid \exists \lambda \in \mathbb{R} \text{ tal que } u = \lambda v\},$$

ou seja, $\mathrm{Gen}(v)$ é o conjunto de todos os múltiplos de v. Por exemplo, $(4,6)$, $(-2,-3)$, $(2\pi, 3\pi)$ e $(2000, 3000)$ são gerados por v.

Proposição 2.3.17. *Para qualquer coleção finita de vetores V, temos que $\mathrm{Gen}(V)$ é um subespaço de $\mathbb{R}^2$.*

Demonstração. Vamos denotar $V = \{v_1, v_2, \cdots, v_n\}$ os elementos de V (o que faz sentido já que, por hipótese, V tem uma quantidade finita de elementos). Tome então u e w em $\mathrm{Gen}(V)$, e qualquer número real α.

Como u é gerado por V, por definição existem $\{\lambda_1, \lambda_2, \cdots, \lambda_n\} \subseteq \mathbb{R}$ tais que

$$u = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n.$$

Similarmente, como w é gerado por V, por definição existem $\{\mu_1, \mu_2, \cdots, \mu_n\} \subseteq \mathbb{R}$ tais que

$$w = \mu_1 v_1 + \mu_2 v_2 + \cdots \mu_n v_n.$$

Isso nos diz que multiplicando w por α obtemos

$$\alpha w = \alpha(\mu_1 v_1 + \mu_2 v_2 + \cdots \mu_n v_n) = (\alpha\mu_1) v_1 + (\alpha\mu_2) v_2 + \cdots (\alpha\mu_n) v_n$$

o que nos diz que $\alpha w \in \mathrm{Gen}(V)$, ou seja, $\mathrm{Gen}(V)$ é fechado por multiplicação por escalares.

Por outro lado, somando u e w obtemos

$$u+w = (\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) + (\mu_1 v_1 + \mu_2 v_2 + \cdots \mu_n v_n) = (\lambda_1 + \mu_1) v_1 + (\lambda_2 + \mu_2) v_2 + \cdots + (\lambda_n + \mu_n) v_n$$

o que nos diz que $u + w \in \mathrm{Gen}(V)$, ou seja, $\mathrm{Gen}(V)$ é fechado por somas.

Juntando as duas conclusões, vemos que $\mathrm{Gen}(V)$ é um subespaço, como queríamos mostrar. ☐

Com isso, então, vamos passar a classificar os subespaços de $\mathbb{R}^2$:

Proposição 2.3.18. *O conjunto* $\{(0,0)\}$ *é um subespaço de* $\mathbb{R}^2$.

Demonstração. Claramente $(0,0)+(0,0)=(0,0)$ e $\lambda(0,0)=(0,0)$ para qualquer $\lambda \in \mathbb{R}$. Segue que $\{(0,0)\}$ é fechado por somas e produtos por escalares. $\square$

Corolário 2.3.19. $\{(0,0)\}$ *é o único subespaço de* $\mathbb{R}^2$ *com apenas um ponto.*

Demonstração. Já vimos que todo subespaço contém a origem. Então se um subespaço contém apenas um ponto, esse ponto é a origem, como queríamos mostrar. $\square$

Observação 2.3.20. *Muitas vezes vamos denotar o conjunto* $\{(0,0)\}$ *e o vetor* $(0,0)$ *simplesmente por* 0*, quando não houver ambiguidade.*

EXEMPLO(S):

Com o resultado acima, então, vemos que nossos subespaços sempre têm pelo menos um ponto. Será que é possível que um subespaço tenha exatamente dois pontos?
Bom, se o subespaço S tem dois pontos v e u, ou v ou u é diferente de $(0,0)$. Suponha (sem perda de generalidade) que $v \neq (0,0)$. Nesse caso, como S é subespaço, $2v \in S$. Mas como S só tem dois pontos, ou $2v = u$, ou $2v = v$.
Se $2v = v$, então subtraindo v dos dois lados obtemos $v = (0,0)$, o que é absurdo, porque nós escolhemos $v \neq (0,0)$ acima. Então $2v$ tem que ser u.
Mas como $v \neq (0,0)$, necessariamente $u = (0,0)$ e, portanto, $2v = (0,0)$. Mas isso nos diz, novamente, que $v = (0,0)$ que, como já vimos, contraria o fato de $v \neq (0,0)$.
Ou seja, se nós fôssemos capazes de construir um subespaço com exatamente dois pontos, nós seríamos capazes de mostrar que $0 \neq 0$, o que é um absurdo! A única conclusão possível, então, é que *não somos capazes de construir um subespaço com exatamente dois pontos.*

Na verdade, o que fizemos acima foi uma prova do seguinte resultado:

Proposição 2.3.21. *Seja* $S \subseteq \mathbb{R}^2$. *Se* $S \neq 0$*, então* S *tem infinitos pontos.*

Demonstração. Se $S \neq 0$, existe $v \in S$ diferente de $(0,0)$. Como S é subespaço, $\lambda v \in S$ para todo $\lambda \in \mathbb{R}$. Como $\lambda v \neq \mu v$ se $\lambda \neq \mu$, temos que para cada número real temos um vetor diferente, e como existem infinitos números reais temos infinitos vetores em S. $\square$

Isso nos dá uma intuição do próximo resultado:

Proposição 2.3.22. *Toda reta passando pela origem é um subespaço gerado por um único vetor não-nulo, e todo subespaço gerado por um único vetor não-nulo é uma reta passando pela origem.*

Demonstração. Tome r uma reta passando pela origem. Então $r = \{(x,y) \in \mathbb{R}^2 \mid ax+by=0\}$ para algum par de números reais $a, b \in \mathbb{R}$ não ambos nulos. Afirmamos que $r = \mathrm{Gen}(v)$, em que $v = (-b, a)$.

Para mostrar isso, vamos mostrar que todo ponto da reta r é gerado por v, e que todo ponto gerado por v está na reta r.

Tome $u = \lambda v$ para algum $\lambda \in \mathbb{R}$, ou seja, $u = (-\lambda b, \lambda a)$. Então computando

$$a(-\lambda b) + b(\lambda a) = -\lambda ab + \lambda ab = 0$$

vemos que $u \in r$.

Por outro lado, tome $P \in r$, ou seja, $P = (p,q)$ com $ap + bq = 0$. Como r é reta, $a \neq 0$ ou $b \neq 0$.

Se $a \neq 0$, então podemos reescrever $ap + bq = 0$ como $p = \dfrac{-b}{a}q$ e, portanto, $P = (p,q) = \left(\dfrac{-b}{a}q, q\right)$.
Disso temos que $P = q\left(\dfrac{-b}{a}, 1\right)$ e, finalmente, $aP = q(-b,a) = qv$. Novamente, como $a \neq 0$, isso nos diz que $P = \dfrac{q}{a}v$, donde vemos que $P \in \mathrm{Gen}(v)$.

Similarmente se $b \neq 0$ podemos mostrar que $P = \dfrac{p}{b} v$, e chegar na mesma conclusão.
Segue que $r = \operatorname{Gen} v$, como queríamos mostrar. $\square$

O próximo passo, a essa altura, seria mostrar que $\mathbb{R}^2$ é gerado por dois vetores não-nulos e que qualquer subespaço gerado por dois vetores não-nulos é $\mathbb{R}^2$. A primeira dessas afirmações é fácil: $\mathbb{R}^2$ é gerado por $\{(1,0),(0,1)\}$ que são dois vetores não-nulos. Contudo, a segunda afirmação é *falsa*:

EXEMPLO(S):

Considere o conjunto $V = \{(1,1),(2,2)\} \subseteq \mathbb{R}^2$. Quem é $\operatorname{Gen}(V)$?
Por definição, são os vetores $v \in \mathbb{R}^2$ tais que existem números reais $\lambda_1, \lambda_2 \in \mathbb{R}$ tais que

$$v = \lambda_1(1,1) + \lambda_2(2,2).$$

Primeira afirmação: $\mathbb{R}^2 \neq \operatorname{Gen}(V)$. Para ver isso, basta notar que o ponto $(1,0)$ não é gerado por V. Dito de outra maneira, suponha que existem números reais x, y tais que $(1,0) = x(1,1)+y(2,2)$. Em particular, comparando as coordenadas, teríamos que $1 = x+2y$ e $0 = x+2y$. Disso, poderíamos concluir que $1 = x + 2y = 0$, ou seja, $1 = 0$, um absurdo! Portanto não podem existir x, y tais que $(1,0) = x(1,1) + y(2,2)$, ou seja, $(1,0) \notin \operatorname{Gen}(V)$ e, portanto, $\mathbb{R}^2 \neq \operatorname{Gen}(V)$.
Segunda afirmação: $\operatorname{Gen}(V) = \operatorname{Gen}((1,1))$ (e, portanto, uma reta, pois $(1,1) \neq 0$). Isso é mais simples: Tome qualquer $v \in \operatorname{Gen}(V)$ e escreva $v = \lambda_1(1,1) + \lambda_2(2,2)$. Agora note que $(2,2) = 2(1,1)$, e podemos reescrever a expressão anterior de v como

$$v = \lambda_1(1,1) + \lambda_2 \cdot 2(1,1) = (\lambda_1 + 2\lambda_2)(1,1),$$

ou seja, $v \in \operatorname{Gen}((1,1))$.
Por outro lado, dado qualquer $u \in \operatorname{Gen}((1,1))$, escreva $u = \lambda(1,1) = \lambda(1,1) + 0(2,2)$, logo $u \in \operatorname{Gen}(V)$.
Disso concluímos que $\operatorname{Gen}(V)$ é uma reta, e certamente não pode ser o plano todo.

Precisamos então de criar um critério para saber quando um conjunto com dois vetores gera ou não o plano todo.

Definição 2.3.23. *Dada uma coleção finita de vetores não-nulos $V = \{v_1, v_2, \cdots, v_n\}$ diremos que ela é **linearmente dependente** se $v_i \in \operatorname{Gen}(V - \{v_i\})$ para algum $i \in \{1, 2, \cdots, n\}$, ou seja V é linearmente dependente se algum dos vetores v_i for gerado pelos outros.*

*Similarmente, se a coleção V não for linearmente dependente, diremos que V é **linearmente independete**, ou seja, nenhum dos vetores v_i é gerado pelos outros.*

Observação 2.3.24. *Vamos usar as abreviações **l.d.** e **l.i.** para linearmente dependente e linearmente independente, respectivamente.*

EXEMPLO(S):

O exemplo que demos acima, com $V = \{(1,1),(2,2)\}$ é um exemplo de uma coleção l.d.: O vetor $(2,2)$ é gerado pelo vetor $(1,1)$.

A coleção $\{(1,0),(0,1),(2,3)\}$ é l.d. pois o vetor $(2,3)$ é gerado pela coleção $\{(1,0),(0,1)\}$.
Por outro lado, a coleção $\{(1,0),(0,1)\}$ é l.i. pois $(1,0)$ não gera nem é gerado por $(0,1)$.

Observação 2.3.25. *Qualquer coleção unitária (i.e. com um só vetor não-nulo) é l.i. Por definição, considere a coleção $V = \{v\}$. Para essa coleção ser l.d. deveria haver algum vetor (diferente de v) em V que gerasse v. Mas V não tem nenhum vetor diferente de v, então não tem nenhum vetor diferente de v que gere v e, portanto, V é l.i.*

Lema 2.3.26. *Toda coleção finita de vetores l.d. pode ser reduzida em uma coleção l.i.*

Demonstração. Se a coleção V é finita, ela tem uma quantidade de elementos, n. Como a coleção é l.d., existe algum vetor v_1 que é gerado pelos outros.

Considere a coleção $V - \{v_1\}$. Ela tem $n-1$ elementos. Se for l.i., acabamos. Se não, existe algum vetor v_2 que é gerado pelos outros.

Considere a coleção $V - \{v_1, v_2\}$. Ela tem $n-2$ elementos. Se for l.i., acabamos. Se não, existe algum vetor v_3 que é gerado pelos outros.

E assim por diante.

Contudo, já vimos que conjuntos unitários são l.i., então, no pior dos casos, vamos ter que remover $n-1$ vetores de V, ficando com o conjunto $V - \{v_1, v_2, \cdots, v_{n-1}\} = \{v_n\}$ que é l.i., como queríamos mostrar. □

Lema 2.3.27. *Seja V conjunto finito de vetores l.i. e $u \in \mathbb{R}^2$ vetor tal que $V \cup \{u\}$ é l.d. Então* $\mathrm{Gen}(V) = \mathrm{Gen}(V \cup \{u\})$.

Demonstração. Tome qualquer $w \in \mathrm{Gen}(V \cup \{u\})$. Vamos mostrar que $w \in \mathrm{Gen}(V)$.

Se V é l.i. com n elementos e $V \cup \{u\}$ é l.d. temos duas possibilidades:

1) u é gerado por V, ou;

2) algum elemento de V é gerado pelos outros elementos de V e $\{u\}$.

   Nesse caso, vamos chamar esse elemento de v_1. Então, existem números reais $\{\lambda_2, \lambda_3, \cdots, \lambda_n, \mu\}$ tais que
$$v_1 = \lambda_2 v_2 + \lambda_3 v_2 + \cdots + \lambda_n v_n + \mu u.$$
   Se $\mu = 0$, estaríamos dizendo que v_1 é gerado pelos elementos de V, o que contrariaria o fato de V ser l.i. Então $\mu \neq 0$. Assim, podemos reescrever a expressão acima isolando u
$$u = \frac{1}{\mu} v_1 + \frac{-\lambda_2}{\mu} v_2 + \cdots + \frac{-\lambda_n}{\mu} v_n$$
   e ver que $u \in \mathrm{Gen}(V)$, ou seja, os dois casos são a mesma coisa.

   Agora, $w \in \mathrm{Gen}(V \cup \{u\})$ implica que existem números reais $\{w_1, w_2, \cdots, w_n, w_u\}$ tais que
$$w = w_1 v_1 + w_2 v_2 + \cdots + w_n v_n + w_u u.$$
   Mas como u é gerado por V, existem $\{u_1, u_2, \cdots, u_n\}$ números reais tais que
$$u = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$
   e, substituindo acima, temos:
$$\begin{aligned} w &= w_1 v_1 + w_2 v_2 + \cdots + w_n v_n + w_u u \\ &\ w_1 v_1 + w_2 v_2 + \cdots + w_n v_n + w_u(u_1 v_1 + u_2 v_2 + \cdots + u_n v_n) \\ &\ (w_1 + w_u u_1) v_1 + (w_2 + w_u u_2) v_2 + \cdots + (w_n + w_u u_n) v_n, \end{aligned}$$
   ou seja $w \in \mathrm{Gen}(V)$, como queríamos mostrar.

□

Lema 2.3.28. *Se V é finito e l.i., então $V - \{v\}$ também é l.i. para qualquer $v \in V$.*

Demonstração. Seja V finito e l.i. $v \in V$. Em particular, $V - \{v\}$ também é finito. Seja $V - \{v\} = \{v_1, v_2, \cdots, v_n\}$. Suponha que v_1 é gerado pelos outros v_i, ou seja, existem números reais $\{\lambda_2, \lambda_3, \cdots, \lambda_n\}$ tais que $v_1 = \lambda_2 v_2 + \lambda_3 v_3 + \cdots + \lambda_n v_n$. Em particular, poderíamos escrever $v_1 = \lambda_2 v_2 + \lambda_3 v_3 + \cdots + \lambda_n v_n + 0v$ e ver que v_1 já era gerado em V, o que contraria o fato de V ser l.i. Segue que v_1 não pode ser gerado pelos outros v_i.

Como temos um número finito de v_i, podemos repetir o mesmo raciocínio para eles e concluir que cada v_j não podem ser gerados pelos outros v_i, ou seja, o conjunto $V - \{v\}$ é l.i., como queríamos mostrar. □

Finalmente, vamos enunciar o resultado que precisamos:

AVISO!

A demonstração abaixo é bastante técnica.
Não se preocupe em entendê-la por agora. Leia o enunciado do teorema e certifique-se de que entendeu.
Se tiver interesse, por favor, leia e tente acompanhar a prova, mas não tenha vergonha em passar para frente caso tenha dificuldade.

Teorema 2.3.29. *Sejam V e W dois conjuntos finitos de geradores l.i. para algum subespaço S. Então $\#V = \#W$.*

Demonstração. Como V e W são finitos, vamos dizer que $\#V = n$ e $\#W = m$. Escreva então

$$V = \{v_1, v_2, \cdots, v_n\}, \quad W = \{w_1, w_2, \cdots, w_m\}$$

os elementos de V e W. Como V gera S e $W \subseteq S$, podemos escrever cada $w_i \in W$ como

$$w_i = \lambda_{i,1}v_1 + \lambda_{i,2}v_2 + \cdots + \lambda_{i,n}v_n.$$

Dito de outra maneira, temos uma matriz $m \times n$

$$M = \begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \cdots & \lambda_{1,n} \\ \lambda_{2,1} & \lambda_{2,2} & \cdots & \lambda_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{m,1} & \lambda_{m,2} & \cdots & \lambda_{m,n} \end{pmatrix}$$

tal que

$$\begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \cdots & \lambda_{1,n} \\ \lambda_{2,1} & \lambda_{2,2} & \cdots & \lambda_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{m,1} & \lambda_{m,2} & \cdots & \lambda_{m,n} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix}.$$

Analogamente, como W gera S e $V \subseteq S$, podemos escrever cada $v_i \in V$ como

$$v_i = \mu_{i,1}w_1 + \mu_{i,2}w_2 + \cdots + \mu_{i,m}w_m.$$

Dito de outra maneira, temos uma matriz $n \times m$

$$N = \begin{pmatrix} \mu_{1,1} & \mu_{1,2} & \cdots & \mu_{1,m} \\ \mu_{2,1} & \mu_{2,2} & \cdots & \mu_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n,1} & \mu_{n,2} & \cdots & \mu_{n,m} \end{pmatrix}$$

tal que

$$\begin{pmatrix} \mu_{1,1} & \mu_{1,2} & \cdots & \mu_{1,m} \\ \mu_{2,1} & \mu_{2,2} & \cdots & \mu_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n,1} & \mu_{n,2} & \cdots & \mu_{n,m} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Substituindo a expressão obtida acima para a matriz $\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ na equação anterior, temos:

$$\begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \cdots & \lambda_{1,n} \\ \lambda_{2,1} & \lambda_{2,2} & \cdots & \lambda_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{m,1} & \lambda_{m,2} & \cdots & \lambda_{m,n} \end{pmatrix} \begin{pmatrix} \mu_{1,1} & \mu_{1,2} & \cdots & \mu_{1,m} \\ \mu_{2,1} & \mu_{2,2} & \cdots & \mu_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n,1} & \mu_{n,2} & \cdots & \mu_{n,m} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix}.$$

Vamos reescrever a matriz MN como

$$MN = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,m} \end{pmatrix}.$$

Então podemos reescrever a equação matricial anterior como

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,m} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix}$$

e obter as equações $w_i = a_{i,1}w_1 + a_{i,2}w_2 + \cdots + a_{i,m}w_m$, ou seja, $(1 - a_{i,i})w_i = a_{i,1}w_1 + a_{i,2}w_2 + \cdots + a_{i,i-1}w_{i-1} + a_{i,i+1}w_{i+1} + \cdots + a_{i,m}w_m$. Se $1 - a_{i,i} \neq 0$, podemos dividir ambos os lados por $1 - a_{i,i}$ e concluir que w_i é combinação linear dos outros elementos de W.

Mas W é l.i., o que nos levaria a um absurdo. Então $1 - a_{i,i} = 0$ e, portanto, $a_{i,i} = 1$ para todos os valores de i, ou seja, a diagonal de MN é composta de 1s.

Contudo, as equações acima nos garantem que $(1 - a_{i,i})w_i = a_{i,1}w_1 + a_{i,2}w_2 + \cdots + a_{i,i-1}w_{i-1} + a_{i,i+1}w_{i+1} + \cdots + a_{i,m}w_m$, e como já sabemos que $a_{i,i} = 1$, temos que $a_{i,1}w_1 + a_{i,2}w_2 + \cdots + a_{i,i-1}w_{i-1} + a_{i,i+1}w_{i+1} + \cdots + a_{i,m}w_m = 0$.

Se algum dos $a_{i,j}$ for diferente de 0, seríamos capazes de isolar w_j e escrever ele como combinação linear dos outros elementos de W. Mas W é l.i., o que nos diz que isso é impossível - ou seja, $a_{i,j} = 0$ sempre que $i \neq j$. Isso nos diz que nossa matriz MN tem 0 em todo elemento fora da diagonal.

Juntando essas duas conclusões, vemos que $MN = I_m$, a matriz identidade $m \times m$.

Um raciocínio análogo nos permite concluir que $NM = I_n$, a matriz identidade $n \times n$.

Mas uma matriz possui inverso bilateral se, e somente se, a matriz é quadrada - segue que tanto M quanto N são quadradas, ou seja, $n = m$, como queríamos mostrar. $\square$

O que esse teorema nos garante é que sempre que nós escolhermos uma coleção l.i. de geradores de algum subespaço, seja ela qual for, ela vai ter sempre a mesma quantidade de elementos. Isso nos permite fazer a seguinte definição:

Definição 2.3.30. *Seja S um subespaço de $\mathbb{R}^2$. Definimos como* ***base*** *qualquer conjunto de vetores l.i. que geram S.*

Analogamente, definimos a ***dimensão*** *de S como sendo o número natural $\dim S$ definido por*

$$\dim S := \#B,$$

em que B é qualquer base de S.

EXEMPLO(S):

O subespaço $\{(0,0)\}$ só tem um vetor - $(0,0)$. Por definição, esse conjunto não é l.i. apesar de ser gerador. Assim, esse subespaço não possui base e, portanto, possui dimensão 0.

Por outro lado, considere o subespaço gerado pelo vetor $(2,3)$, que já vimos ser uma reta que passa pela origem e por $(2,3)$. Claramente o conjunto $\{(2,3)\}$ é l.i. (pois é unitário) e gera esse subespaço (por definição) - ou seja, é uma base. Segue que a reta que passa pela origem e pelo ponto $(2,3)$ tem dimensão 1.
Não é difícil ver que qualquer reta que passa pela origem é um subespaço de dimensão 1.

Finalmente, considere $\mathbb{R}^2$. Já sabemos que $\mathbb{R}^2$ é gerado por $\{(1,0),(0,1)\}$ e que esse conjunto é l.i.

- ou seja, uma base de $\mathbb{R}^2$. Segue que $\dim \mathbb{R}^2 = 2$.

Olha que interessante: Nós mostramos que subespaços gerados ou têm dimensão 0 (o subespaço 0), ou têm dimensão 1 (as retas pela origem), ou têm dimensão 2 (o plano todo).

Para finalizar, mais alguns resultados técnicos.

Lema 2.3.31. *Dada uma coleção finita de conjuntos não-nulos V e um subespaço S de dimensão n, as seguintes afirmações são equivalentes:*

(a) V é uma base para S;

(b) V é l.i. e tem n elementos;

(c) V gera S e tem n elementos.

Demonstração. Claramente (a) implica tanto (b) quanto (c).

Suponha (b), ou seja, que V é l.i. e tem n elementos. Se V não gera, existe algum elemento $v \in S$ tal que $v \notin \text{Gen}(V)$ - em particular, o conjunto $V \cup \{v\}$ seria l.i. e teria $n+1$ elementos.

Isso é um absurdo, pois se $V \cup \{v\}$ gera S, temos então um conjunto de geradores l.i. com mais de n elementos. Por outro lado, se $V \cap \{v\}$ não gera, isso é pior ainda - significa que precisamos adicionar ainda mais gente para gerar S, e no final vamos ficar com muito mais do que n elementos.

Como isso não pode acontecer, V tem que já gerar S, ou seja, supondo (b) concluímos (a) e (c).

Analogamente, suponha (c), ou seja, V gera e tem n elementos. Se V não é l.i., ou $V = \{0, 0\}$ ou V é l.d. Como estamos assumindo, por hipótese, que V contém apenas vetores não nulos, podemos descartar o primeiro caso.

Como V é l.d., podemos reduzir V até obter um conjunto V' l.i. tal que $\text{Gen}(V) = \text{Gen}(V')$ (pelos lemas 2.3.26 e 2.3.27). Como $\text{Gen}(V) = S$, por hipótese, temos que V' é um conjunto gerador l.i. com menos elementos do que n - o que contraria o Teorema 2.3.29, um absurdo.

Segue que V é l.i., ou seja, (c) implica (a) e (b), o que encerra a demonstração. $\square$

Finalmente, vamos mostrar que todo subespaço é gerado:

Proposição 2.3.32. *Seja S subespaço de $\mathbb{R}^2$. Então existe $V \subseteq S$ conjunto de geradores de S - ou seja, $\text{Gen}(V) = S$.*

Demonstração. Certamente $(0, 0) \in S$. Se $S - \{(0, 0)\} = \varnothing$, acabou: $S = 0$ e portanto é gerado por $(0, 0)$.

Caso contrário, existe algum $v \in S$ diferente de $(0, 0)$. Então $\text{Gen}(v) \subseteq S$, já que S é subespaço. Se $S - \text{Gen}(v) = \varnothing$, acabou: mostramos que $S = \text{Gen}(v)$.

Caso contrário, existe algum $w \in S$ diferente de $(0, 0)$ e que não é gerado por v. Então $\text{Gen}(v, w) \subseteq S$, já que S é subespaço.

Contudo, como $\{v, w\}$ foi construído l.i., o lema acima nos garante que $\{v, w\}$ é base de $\mathbb{R}^2$, donde $\text{Gen}(\{v, w\}) = \mathbb{R}^2$. Então temos $\mathbb{R}^2 \subseteq S \subseteq \mathbb{R}^2$, ou seja, $S = \mathbb{R}^2$, e S é, portanto, gerado por $\{v, w\}$, como queríamos mostrar. $\square$

Juntando isso tudo, temos um grande corolário que encerra esta seção:

Proposição 2.3.33. *Todos os subespaços de $\mathbb{R}^2$ são: 0, retas que passam pela origem e $\mathbb{R}^2$.*

Demonstração. Já vimos, no exemplo, que todo subespaço gerado tem dimensão 0, 1 ou 2. Contudo, a proposição acima nos diz que todo subespaço é gerado. O resultado se segue. $\square$

Exercício (2.3.3)

Mostre que a dimensão do núcleo de uma transformação linear é exatamente o número de linhas nulas na forma escalonada da matriz dessa transformação linear.

Exercício (2.3.4)

Mostre que a dimensão da imagem de uma transformação linear é exatamente o número de linhas não-nulas na forma escalonada da matriz dessa transformação.

Exercício (2.3.5)

Junte os dois exercícios acima para concluir que dada uma transformação linear $f : \mathbb{R}^2 \to \mathbb{R}^2$, temos que $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim \mathbb{R}^2 = 2$.

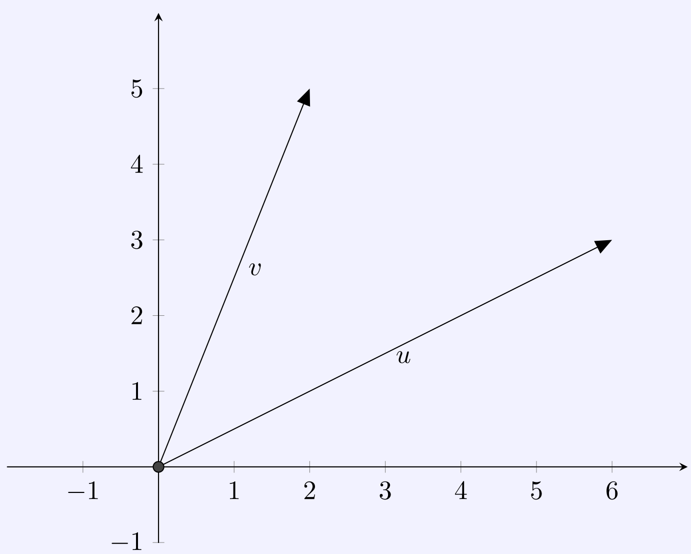
2.4 Produtos interno e vetorial

2.4.1 Projeção ortogonal

Em muitos problemas de física é interessante decompor um vetor não como tendo uma coordenada X e uma coordenada Y, mas em termos de outras coordenadas arbitrárias.

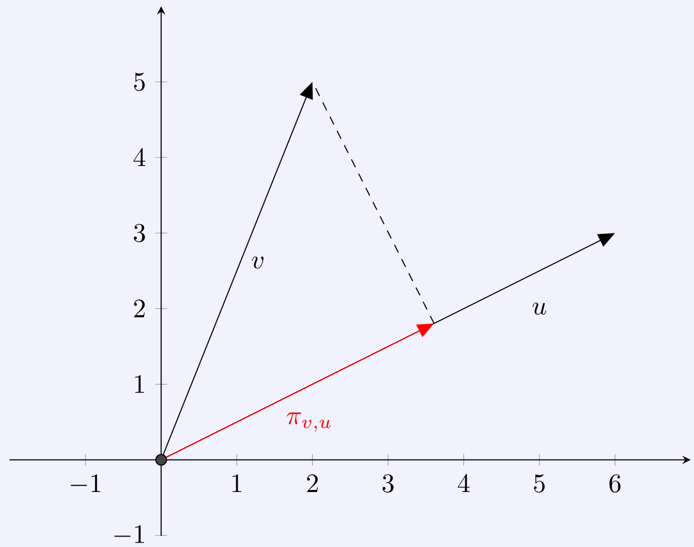
EXEMPLO(S):

Considere os vetores abaixo:

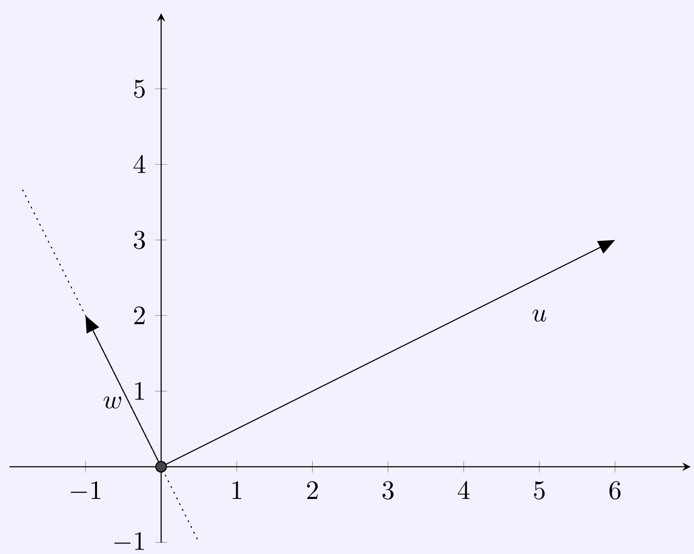


em que $u = (6, 3)$ e $v = (2, 5)$.

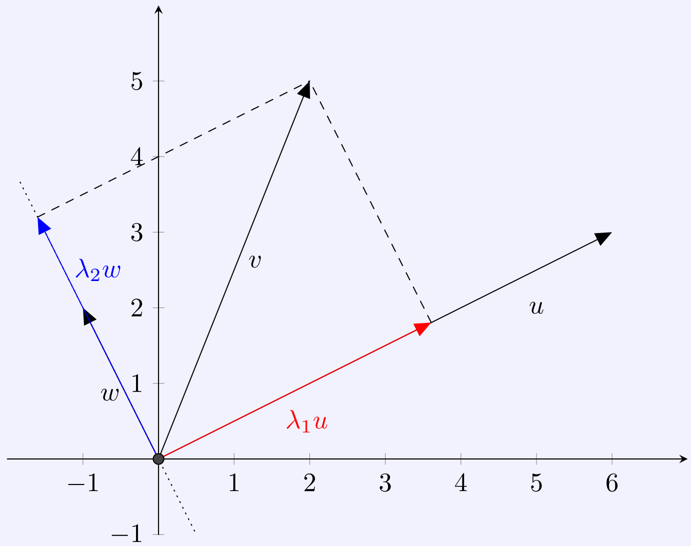
A **projeção ortogonal de v em u** é o vetor $\pi_{v,u}$ representado abaixo:



Esse vetor tem a seguinte propriedade especial: Considere a reta ortogonal a u que passa pela origem e qualquer vetor w nela:



nesse caso, $w = (-1, 2)$. Como $u \perp w$, claramente o conjunto $\{u, w\}$ é l.i. e, portanto, gera $\mathbb{R}^2$ - em particular, existem λ_1, λ_2 em $\mathbb{R}$ tais que $v = \lambda_1 u + \lambda_2 w$, como vemos abaixo:



Assim, a projeção ortogonal de v em u é definida como sendo $\pi_{v,u} := \lambda_1 u$, e a projeção ortogonal de v em w é definida como sendo $\pi_{v,w} := \lambda_2 w$.

Para calcular essas projeções, então, precisamos encontrar λ_1 e λ_2, ou seja, resolver a equação

$$(2, 5) = \lambda_1(6, 3) + \lambda_2(-1, 2)$$

que se traduz nas equações $2 = 6\lambda_1 - \lambda_2$ e $5 = 3\lambda_1 + 2\lambda_2$. Ora, já sabemos que podemos representar isso por uma equação matricial

$$\begin{pmatrix} 6 & -1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix},$$

ou seja, encontrar as projeções equivale a resolver o sistema linear acima.
Vamos lá!

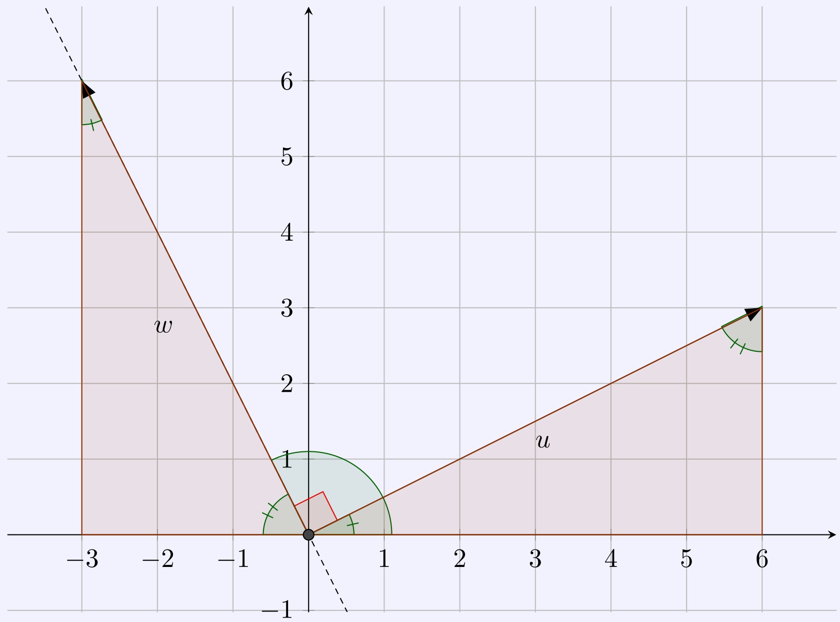
$$\left(\begin{array}{cc:c} 6 & -1 & 2 \\ 3 & 2 & 5 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 6 & -1 & 2 \\ 0 & 5 & 8 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 30 & 0 & 18 \\ 0 & 5 & 8 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 5 & 0 & 3 \\ 0 & 5 & 8 \end{array}\right)$$

então nosso sistema tem solução única $5\lambda_1 = 3$ e $5\lambda_2 = 8$, ou seja, $\lambda_1 = 3/5$ e $\lambda_2 = 8/5$ (de fato, $6(3/5) - 8/5 = 18/5 - 8/5 = 10/5 = 2$ e $3(3/5) + 2(8/5) = 9/5 + 16/5 = 25/5 = 5$). Com isso, então, podemos ver que $\pi_{v,u} = \lambda_1 u = 3/5(6, 3)$ e $\pi_{v,w} = \lambda_2 w = 8/5(-1, 2)$.

2.4.2 Produto interno

EXEMPLO(S):

Continuando o exemplo acima, vamos calcular a projeção ortogonal de um vetor arbitrário $v = (v_1, v_2)$ em um vetor arbitrário $u = (u_1, u_2)$. Nesse caso, vamos usar $w = u^\perp = (-u_2, u_1)$ - isso pode ser visto notando que $\tan\theta = -\tan(\theta + 90°)$ - como podemos ver no desenho abaixo:



Assim, vamos calcular $\pi_{v,u}$ como sendo igual ao $\lambda_1 \in \mathbb{R}$ tal que $v = \lambda_1 u + \lambda_2 w$, ou seja,

$$\left(\begin{array}{cc:c} u_1 & -u_2 & v_1 \\ u_2 & u_1 & v_2 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 1 & -u_2/u_1 & v_1/u_1 \\ u_2 & u_1 & v_2 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 1 & -u_2/u_1 & v_1/u_1 \\ 0 & u_1 - u_2(-u_2/u_1) & v_2 - u_2(v_1/u_1) \end{array}\right) \rightsquigarrow$$
$$\rightsquigarrow \left(\begin{array}{cc:c} 1 & -u_2/u_1 & v_1/u_1 \\ 0 & (u_1^1 + u_2^2)/u_1 & (u_1v_2 - u_2v_1)/u_1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 1 & -u_2/u_1 & v_1/u_1 \\ 0 & u_1^1 + u_2^2 & u_1v_2 - u_2v_1 \end{array}\right) \rightsquigarrow$$
$$\rightsquigarrow \left(\begin{array}{cc:c} 1 & -u_2/u_1 & v_1/u_1 \\ 0 & 1 & (u_1v_2 - u_2v_1)/(u_1^2 + u_2^2) \end{array}\right) \rightsquigarrow \left(\begin{array}{cc:c} 1 & 0 & v_1/u_1 + (u_2/u_1)(u_1v_2 - u_2v_1)/(u_1^2 + u_2^2) \\ 0 & 1 & (u_1v_2 - u_2v_1)/(u_1^2 + u_2^2) \end{array}\right)$$

então nosso λ_1 vale... $v_1/u_1 + (u_2/u_1)(u_1v_2 - u_2v_1)/(u_1^2 + u_2^2)$?! Ok, sem pânico, vamos resolver isso:

$$\begin{aligned} v_1/u_1 + (u_2/u_1)(u_1v_2 - u_2v_1)/(u_1^2 + u_2^2) &= ((v_1/u_1)(u_1^2 + u_2^2) + (u_2/u_1)(u_1v_2 - u_2v_1))/(u_1^2 + u_2^2) \\ &= (v_1u_1 + v_1u_2^2/u_1 + u_2v_2 - u_2^2v_1/u_1)/(u_1^2 + u_2^2) \\ &= (v_1u_1 + v_2u_2)/(u_1^2 + u_2^2) \end{aligned}$$

Pronto! Agora sabemos que $\lambda_1 = (v_1u_1 + v_2u_2)/(u_1^2 + u_2^2)$. Se notarmos que o denominador é simplesmente $\|u\|^2$, podemos escrever uma expressão fechada para $\pi_{v,u}$:

$$\pi_{v,u} = (v_1u_1 + v_2u_2)/\|u\|^2.$$

O que significa o termo $v_1u_1 + v_2u_2$ que apareceu acima?

Se pensarmos novamente em vetores como matrizes coluna, não é difícil ver que $v_1u_1 + v_2u_2 = [v]^t[u]$, onde ^{t} representa a matriz transposta, ou seja:

$$\begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} v_1u_1 + v_2u_2 \end{pmatrix}.$$

Ou seja, usando o conceito de vetores como matrizes nós fomos capazes de inventar uma multiplicação de vetores cujo resultado é um escalar!

Definição 2.4.1. *Dados dois vetores $v, u \in \mathbb{R}^2$, definimos o* **produto interno (ou escalar) de** v **e** u *como sendo o número real $\langle v, u\rangle \in \mathbb{R}$ dado por*

$$\langle v, u\rangle := [v]^t[u].$$

Proposição 2.4.2. *Para quaisquer vetores $v, u, w \in \mathbb{R}^2$ e qualquer número real $\lambda \in \mathbb{R}$ valem as seguintes propriedes:*

- *(Linearidade à esquerda) $\langle v + \lambda u, w\rangle = \langle v, w\rangle + \lambda\langle u, w\rangle$;*
- *(Linearidade à direita) $\langle v, u + \lambda w\rangle = \langle v, u\rangle + \lambda\langle v, w\rangle$;*
- *(Comutatividade) $\langle v, u\rangle = \langle u, v\rangle$;*
- *(Não-degenerado) $\langle v, v\rangle \geq 0$ e a igualdade acontece apenas no caso $v = 0$.*

Lema 2.4.3. *Para qualquer vetor $v \in \mathbb{R}^2$, temos que $\langle v, v\rangle = \|v\|^2$.*

EXEMPLO(S):

Vamos voltar ao exemplo motivador do produto interno: Projeções ortogonais.
Agora, com o produto interno, dados dois vetores $v, u \in \mathbb{R}^2$ podemos escrever a projeção de v em u como sendo $\pi_{v,u} = \langle v, u\rangle/\langle u, u\rangle$ - dito de outra maneira, $\langle v, u\rangle = \pi_{v,u}\langle u, u\rangle$.
O que acontece então se escolhermos v ortogonal a u, ou seja, $v = \lambda u^\perp$, para algum $\lambda \in \mathbb{R}$? Nesse caso, temos:

$$\langle v, u\rangle = \langle \lambda u^\perp, u\rangle = \lambda\langle u^\perp, u\rangle = \lambda(-u_2u_1 + u_1u_2) = \lambda(0) = 0$$

ou seja, se $v \perp u$, temos que $\langle v, u\rangle = 0$ e, portanto, $\pi_{v,u} = 0$ - o que condiz com nossa expectativa geométrica!
Por outro lado, se temos algum vetor $w \in \mathbb{R}^2$ tal que $\langle w, u\rangle = 0$, então temos que $w_1u_1 + w_2u_2 = 0$, ou seja, $w_1/w_2 = -u_2/u_1$ e, portanto $w \perp u$ - geométricamente isso é óbvio: Se algum vetor tem projeção de tamanho 0 em u, esse vetor só pode ser ortogonal a u!

Lema 2.4.4. *Dados dois vetores $v, u \in \mathbb{R}^2$, temos que $v \perp u$ se, e somente se, $\langle v, u\rangle = 0$.*

Isso nos permite, então, verificar ortogonalidade sem fazer desenhos, usando apenas coordenadas!
Para encerrar esta subseção, vamos dar um aplicação interessante:

EXEMPLO(S):

Já vimos que uma reta em $\mathbb{R}^2$ pode ser descrita de duas maneiras distintas: como todas as soluções de uma equação do tipo $ax + by = 0$ ou como todos os múltiplos de algum vetor u fixado. Como essas definições se relacionam?
Dito de outra maneira, como encontrar os coeficientes a, b da equação conhecendo apenas o vetor u que gera a reta; e como descobrir qual o vetor gerador de uma reta sabendo apenas os coeficientes da equação?
Simples! Vamos mostrar que uma reta é dada pelo conjunto de soluções reais da equação $ax+by = 0$ se, e somente se, é gerada pelo vetor $(a, b)^\perp$.
Um lado é trivial: Considere $r = \{v \in \mathbb{R}^2 \mid \exists\lambda \in \mathbb{R}$ tal que $v = \lambda(a, b)^\perp\}$ e lembre que $(a, b)^\perp = (-b, a)$. Ou seja, todo elemento de r é da forma $\lambda(-b, a)$. Claramente elementos dessa forma são soluções da equação $ax + by = 0$.
Por outro lado, se v é solução da equação $ax + by = 0$, isso significa que $av_1 + bv_2 = 0$, ou seja, $\langle (a, b), v\rangle = 0$ - ou seja, $(a, b) \perp v$ e, portanto, v é gerado por $(a, b)^\perp$, como queríamos mostrar.

Resumindo, quando escrevemos uma reta usando sua *equação* estamos, implicitamente, definindo uma reta como “o conjunto de todos os vetores ortogonais a um vetor fixado”, e quando escrevemos uma reta com um *subespaço gerado por um vetor* estamos definindo a reta como “o conjunto de todos os vetores múltiplos de um vetor fixado”.