Álgebra Linear

Melissa Weber Mendonça

2019.2

Conteúdo

1 Matrizes e Sistemas **2**
1.1 Introdução 2
1.2 Matrizes 4
1.2.1 Tipos especiais de matrizes 4
1.3 Operações com matrizes 7
1.3.1 Adição de matrizes 7
1.3.2 Multiplicação por escalar 7
1.3.3 Transposição 8
1.3.4 Produto de matrizes 9
1.3.5 4 diferentes formas de se fazer um produto de matrizes . . 11
1.3.6 Matriz Inversa 12
1.4 Fatoração $PA = LU$ de uma matriz A. 13
1.4.1 Sistemas Lineares 13
1.4.2 Forma matricial: Forma escalonada do sistema 14
1.4.3 Eliminação Gaussiana e a Forma Escalonada 15
1.4.4 Operações elementares: Matrizes elementares 16
1.4.5 Fatoração $PA = LU$. 22

2 Espaços Vetoriais **24**
2.1 Introdução aos espaços vetoriais 24
2.2 Subespaços vetoriais 28
2.3 Interseção e soma de subespaços vetoriais 30
2.4 Dependência linear entre vetores 33
2.4.1 Base e dimensão de um espaço vetorial 34
2.4.2 Coordenadas e Mudança de Base 41

3 Transformações Lineares **47**
3.1 Introdução 47
3.2 Matriz de uma transformação linear 50
3.2.1 Matriz de uma transformação linear em relação a uma base do domínio e a uma base do contradomínio. 52
3.3 Rotações, projeções e reflexões. 54
3.4 Produto de transformações lineares 57
3.5 Os Espaços Fundamentais 58
3.6 Sistemas Retangulares 62
3.7 Existência de Inversas 67
3.8 Transformações lineares inversíveis. 69
3.8.1 Teorema do núcleo e da imagem de uma transformação linear. 71
3.9 Mudança de base 73
3.10 Soma direta 75

CAPÍTULO 1

Matrizes e Sistemas

Nesta unidade, discutiremos tipos especiais de matrizes, operações com matrizes e o escalonamento, também conhecido como eliminação Gaussiana. Discutiremos a fatoração $PA = LU$ e sua implementação, através de uma introdução à linguagem Python.

1.1 Introdução

A álgebra linear se preocupa principalmente com a solução de sistemas lineares. Vamos considerar primeiro o caso mais simples, em que o número de equações é igual ao número de variáveis. Considere o exemplo abaixo:

$$\begin{cases} 2x - y = 1 \\ x + y = 5. \end{cases} \tag{1.1}$$

É comum aprendermos duas maneiras de resolver esse tipo de problemas. A primeira é a ideia de eliminar variáveis de uma equação para resolver a outra. No exemplo, poderíamos fazer

$$\begin{cases} 2x - y = 1 \\ x + y = 5 \end{cases} \Rightarrow \begin{cases} 2x - y = 1 \\ x = 5 - y \end{cases}$$

Substituindo o valor de x na primeira equação, teremos que

$$\Rightarrow 2(5 - y) - y = 1 \Rightarrow 10 - 3y = 1 \Rightarrow y = 3.$$

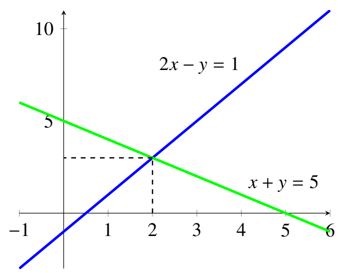


Figura 1.1: Resolução gráfica do sistema linear (1.1). A solução é $(x, y) = (2, 3)$.

Agora, substituindo este valor de y na primeira equação, deduzimos que

$$x = 5 - y = 5 - 3 = 2.$$

Este método é bastante simples em um problema com poucas equações e poucas variáveis, mas não é fácil imaginar a resolução de um problema de grande porte desta forma. Da mesma forma, o segundo método, que envolve determinantes e é normalmente conhecido como *regra de Cramer*, é bastante complicado e, na prática, pouco eficiente.

Felizmente, existe uma maneira inteligente de se aplicar o primeiro método a um sistema de grande porte. O algoritmo que permite a resolução de problemas desta forma e que vamos estudar é conhecido como *eliminação Gaussiana.*

O que é um algoritmo?

Já sabemos (da Geometria Analítica) que um sistema de equações do tipo

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

pode ser reescrito em uma forma matricial, como sendo

$$Ax = b \tag{1.2}$$

e que temos algumas possibilidades para a solução deste sistema:

- Se a matriz A for quadrada e inversível, então $x = A^{-1}b$, ou seja, é possível encontrarmos solução única do sistema para qualquer lado direito b. Em outras palavras, o sistema é possível e determinado.

- Se a matriz A não é inversível (se não for quadrada, ou se for quadrada mas não tiver inversa), então a existência e unicidade de soluções dependem do lado direito b do sistema. Para algumas escolhas de b o sistema terá infinitas soluções (sistema possível e indeterminado) e para outras o sistema não terá soluções (sistema impossível).

Para identificarmos cada um destes casos, estudaremos alguns fatos sobre as matrizes.

1.2 Matrizes

Definição 1. *Uma matriz é uma tabela de elementos dispostos em linhas e colunas.*

Exemplo 2.1. *Alguns exemplos de matrizes de números reais:*

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \begin{pmatrix} 2 & -1 \\ 2 & 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 & 1 & -3 \end{pmatrix}, \begin{pmatrix} 2 \end{pmatrix}.$$

Os elementos de uma matriz podem ser números (reais ou complexos), funções, ou ainda outras matrizes. Representamos uma matriz de m linhas e n colunas por

$$A_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{m\times n}.$$

1.2.1 Tipos especiais de matrizes

Nos exemplos abaixo, as matrizes só contém números reais. No entanto, estas definições tem equivalentes para matrizes contendo outros tipos de elemento (números complexos, funções etc).

Matriz Quadrada é aquela cujo número de linhas é igual ao número de colunas. $(m = n)$.

Exemplo 2.2.

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

Neste caso, dizemos que A é uma matriz de ordem m (ou n).

Matriz Nula é aquela em que $a_{ij} = 0$ para todo i e j.[1]

Exemplo 2.3.

$$A_{2\times 2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, B_{3\times 5} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Matriz Coluna é aquela que possui uma única coluna ($n = 1$).

Matriz Linha é aquela que possui uma única linha ($m = 1$).

Matriz Simétrica é uma matriz quadrada A em que $a_{ij} = a_{ji}$, para $i = 1, \ldots, m$, $j = 1, \ldots, n$ ($m = n$).

Exemplo 2.4.

$$\begin{pmatrix} 4 & 3 & -1 \\ 3 & 2 & 0 \\ -1 & 0 & 5 \end{pmatrix}$$

Matriz Diagonal é uma matriz A em que os únicos elementos (possivelmente) não-nulos são os elementos a_{ij} com $i = j$, $i = 1, \ldots, m$, $j = 1, \ldots, n$.

Exemplo 2.5.

$$\begin{pmatrix} 7 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

Matriz triangular é uma matriz quadrada em que todos os elementos abaixo (ou acima) da diagonal são nulos.

Exemplo 2.6 (Matriz Triangular Superior)**.**

$$\begin{pmatrix} 2 & -1 & 0 \\ 0 & -1 & 4 \\ 0 & 0 & 3 \end{pmatrix}$$

Neste caso, $a_{ij} = 0$ *para todo* $i > j$.

[1] Aqui, o correto seria dizer que a_{ij} é igual ao elemento neutro da soma do conjunto onde vivem os elementos da matriz.

Exemplo 2.7 (Matriz Triangular Inferior)**.**

$$\begin{pmatrix} 2 & 0 & 0 \\ -1 & -1 & 0 \\ 3 & 2 & 1 \end{pmatrix}$$

Neste caso, $a_{ij} = 0$ *para todo* $i < j$*.*

Matriz esparsa é uma matriz que contem muitos elementos nulos.

Não existe uma definição precisa de matriz esparsa. Quando falamos de matrizes esparsas, em geral estamos falando de matrizes que possuem menos da metade de seus elementos não-nulos, ou que possuem algum tipo de estrutura especial que concentra seus elementos não-nulos.

Exemplo 2.8. *Matrizes diagonais são esparsas.*

Exemplo 2.9.

$$\begin{pmatrix} 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Matriz de banda é uma matriz esparsa cujos elementos não-nulos estão concentrados próximos à sua diagonal principal.

Exemplo 2.10.

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

A largura da banda *desta matriz é 3.*

Exemplo 2.11.

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

Esta matriz é uma matriz de banda, de largura 2. Também é uma matriz triangular inferior.

1.3 Operações com matrizes

1.3.1 Adição de matrizes

A soma de duas matrizes de mesma ordem $A_{m\times n} = (a_{ij})$ e $B_{m\times n} = (b_{ij})$ é uma matriz $m \times n$ denotada por $A + B$ cujos elementos são somas dos elementos correspondentes de A e de B, ou seja

$$A + B = (a_{ij} + b_{ij})_{m\times n}.$$

Exemplo 3.1.

$$\begin{pmatrix} 1 & -1 \\ 4 & 0 \\ 2 & 5 \end{pmatrix} + \begin{pmatrix} 0 & 4 \\ -2 & 5 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 3 & 5 \end{pmatrix}$$

Propriedades A maneira como a adição de matrizes foi definida garante que ela possuirá as mesmas propriedades da adição de números reais. Dadas as matrizes A, B e C de mesma ordem $m \times n$, temos:

(i) $A + B = B + A$ (comutatividade)

(ii) $A + (B + C) = (A + B) + C$ (associatividade)

(iii) $A + 0 = A$, onde 0 denota a matriz nula $m \times n$.

1.3.2 Multiplicação por escalar

Seja $A = (a_{ij})_{m\times n}$ e k um escalar (número). Definimos o produto da matriz A pelo escalar k como uma nova matriz

$$kA = (ka_{ij})_{m\times n}.$$

Exemplo 3.2.

$$-2\begin{pmatrix} 2 & 10 \\ 1 & -3 \end{pmatrix} = \begin{pmatrix} -4 & 20 \\ 2 & -6 \end{pmatrix}$$

Propriedades Dadas as matrizes A, B de mesma ordem $m \times n$ e números k, k_1, k_2 escalares, temos:

(i) $k(A+B) = kA + kB$

(ii) $(k_1 + k_2)A = k_1 A + k_2 A$

(iii) $0A = 0$, ou seja, se multiplicarmos o número 0 pela matriz $A_{m\times n}$ obteremos a matriz nula $0_{m\times n}$.[2]

(iv) $k_1(k_2 A) = (k_1 k_2)A$.

1.3.3 Transposição

Dada uma matriz $A = (a_{ij})_{m\times n}$, podemos obter uma outra matriz $A^T = (b_{ij})_{n\times m}$, cujas linhas são as colunas de A, ou seja, $b_{ij} = a_{ji}$. A^T é denominada a *transposta* de A.

Exemplo 3.3.

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 3 \\ -1 & 4 \end{pmatrix}_{3\times 2} \qquad A^T = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 3 & 4 \end{pmatrix}_{2\times 3}$$

Exemplo 3.4.

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix}^T = (1\,2)\,.$$

Exemplo 3.5.

$$B = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix} \qquad B^T = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$$

Propriedades

(i) Uma matriz é *simétrica* se e somente se ela é igual à sua transposta.

(ii) $\left(A^T\right)^T = A$.

(iii) $(A+B)^T = A^T + B^T$.

(iv) $(kA)^T = kA^T$, para todo escalar k.

[2]Novamente, aqui poderíamos substituir o escalar (número) 0 pelo elemento neutro do conjunto que estamos considerando.

1.3.4 Produto de matrizes

Vamos supor que temos dois alimentos, com certas quantidades de vitaminas A, B e C, definidas na seguinte tabela.

	A	B	C
Alimento 1	4	3	0
Alimento 2	5	0	1

Se ingerirmos 5 unidades do Alimento 1 e 2 unidades do Alimento 2, quanto teremos consumido de cada tipo de vitamina?

Podemos representar esse problema em termos de matrizes da seguinte forma. As quantidades de vitamina de cada alimento podem ser representadas pela matriz

$$\begin{pmatrix} 4 & 3 & 0 \\ 5 & 0 & 1 \end{pmatrix}$$

e as unidades do Alimento 1 e do Alimento 2 consumidas podem ser representadas pelo (ou matriz-linha)

$$\begin{pmatrix} 5 & 2 \end{pmatrix}.$$

Desta forma, a operação que nos fornecerá a quantidade total de cada vitamina ingerida é o produto definido por

$$\begin{aligned} &\begin{pmatrix} 5 & 2 \end{pmatrix} \begin{pmatrix} 4 & 3 & 0 \\ 5 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 5 \cdot 4 + 2 \cdot 5 & 5 \cdot 3 + 2 \cdot 0 & 5 \cdot 0 + 2 \cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} 30 & 15 & 2 \end{pmatrix} \end{aligned}$$

Ou seja, serão ingeridas 30 unidades de vitamina A, 15 de vitamina B e 2 de C.

Esta operação é definida como o produto de duas matrizes, e pode ser expresso formalmente da seguinte maneira: sejam $A = (a_{ij})_{m \times n}$ e $B = (b_{ij})_{n \times p}$. Definimos $AB = (c_{ij})_{m \times p}$, com

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + \ldots + a_{in} b_{nj}.$$

Só podemos efetuar o produto de duas matrizes se o número de colunas da primeira for igual ao número de linhas da segunda.

Exemplo 3.6.

$$\begin{pmatrix} 2 & 1 \\ 4 & 2 \\ 5 & 3 \end{pmatrix}_{3\times 2} \cdot \begin{pmatrix} 1 & -1 \\ 0 & 4 \end{pmatrix}_{2\times 2} = \begin{pmatrix} 2\cdot 1 + 1\cdot 0 & 2(-1) + 1\cdot 4 \\ 4\cdot 1 + 2\cdot 0 & 4(-1) + 2\cdot 4 \\ 5\cdot 1 + 3\cdot 0 & 5(-1) + 3\cdot 4 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 4 & 4 \\ 5 & 7 \end{pmatrix}$$

Exemplo 3.7.

$$\begin{pmatrix} 1 & -1 \\ 0 & 4 \end{pmatrix}_{2\times 2} \cdot \begin{pmatrix} 2 & 1 \\ 4 & 2 \\ 5 & 3 \end{pmatrix}_{3\times 2}$$

Não é possível efetuar este produto.

Propriedades

(i) Em geral, $AB \neq BA$.

Exemplo 3.8.

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

$$AB = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ BA = \begin{pmatrix} -11 & 6 & 1 \\ -22 & 12 & -2 \\ -11 & 6 & -1 \end{pmatrix}$$

Além disso, AB pode ser nula sem que A ou B sejam nulas.

(ii) $AI = IA = A$

(iii) $0A = 0$ e $A0 = 0$.

(iv) Sejam $A_{m\times p}$, $B_{p\times n}$ e $C_{p\times n}$. Então, $A(B + C)$ é uma matriz com elementos dados por

$$\begin{aligned}
(A(B+C))_{ij} &= \sum_{k=1}^{p} a_{ik}(b_{kj} + c_{kj}) \\
&= \sum_{k=1}^{p} a_{ik}b_{kj} + a_{ik}c_{kj} \\
&= \sum_{k=1}^{p} a_{ik}b_{kj} + \sum_{k=1}^{p} a_{ik}c_{kj} \\
&= (AB)_{ij} + (AC)_{ij}.
\end{aligned}$$

Logo, $A(B + C) = AB + AC$ (distributividade à esquerda da multiplicação)

(v) $(A + B)C = AC + BC$ (distributividade à direita da multiplicação) (análogo ao caso anterior)

(vi) Sejam $A_{m\times p}$, $B_{p\times q}$, $C_{q\times n}$. Então, $(AB)C = A(BC)$ (associatividade da multiplicação).

Precisamos mostrar que $((AB)C)_{ij} = (A(BC))_{ij}$. Mas, note que:

$$\begin{aligned}
(A(BC))_{ij} &= \sum_{k=1}^{p} A_{ik}(BC)_{kj} \\
&= \sum_{k=1}^{p} A_{ik} \sum_{\ell=1}^{q} B_{k\ell}C_{\ell j} \\
&= \sum_{k=1}^{p} \sum_{\ell=1}^{q} A_{ik}B_{k\ell}C_{\ell j} \\
&= \sum_{\ell=1}^{q} \sum_{k=1}^{p} A_{ik}B_{k\ell}C_{\ell j} \\
&= \sum_{\ell=1}^{q} \sum_{k=1}^{p} (A_{ik}B_{k\ell})C_{\ell j} \\
&= \sum_{\ell=1}^{q} (AB)_{i\ell}C_{\ell j} \\
&= ((AB)C)_{ij}.
\end{aligned}$$

(vii) $(AB)^T = B^T A^T$ (Exercício!)

1.3.5 4 diferentes formas de se fazer um produto de matrizes

Podemos encarar o produto de duas matrizes como vimos anteriormente, de forma monolítica, mas existem três outros jeitos de interpretarmos esta operação.

(i) Cada entrada da matriz-produto AB é o produto entre uma linha e uma coluna:

$$(AB)_{ij} = \text{ linha } i \text{ de } A \text{ vezes a coluna } j \text{ de } B$$

(ii) Cada coluna de AB é o produto entre uma matriz e uma coluna:

$$\text{coluna } j \text{ de } AB = A \text{ vezes coluna } j \text{ de } B$$

(iii) Cada linha de AB é o produto entre uma linha e uma matriz:

$$\text{linha } i \text{ de } AB = \text{ linha } i \text{ de } A \text{ vezes } B.$$

Exemplo.

$$\begin{aligned}
\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} &= \begin{bmatrix} 1\cdot 2 + 0\cdot 3 & 1\cdot 1 + 0\cdot 4 \\ 2\cdot 2 + 3\cdot 3 & 2\cdot 1 + 3\cdot 4 \end{bmatrix}\left(\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix}\right) \\
&= \begin{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \\ \begin{bmatrix} 2 & 3 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \end{bmatrix}
\end{aligned}$$

1.3.6 Matriz Inversa

Matriz Identidade é uma matriz $I_{n\times n}$ tal que para toda matriz $A_{n\times n}$,

$$AI = IA = A.$$

Dada $A_{n\times n}$, se pudermos encontrar uma outra matriz $B_{n\times n}$ tal que

$$AB = BA = I,$$

então diremos que A é inversível e que $B = A^{-1}$. Senão, diremos que A é singular.

Teorema 1. *A inversa de $A_{n\times n}$ é única.*

Demonstração. Suponha que existem duas inversas de A, B e C. Então,

$$BA = I \Leftrightarrow (BA)C = IC = C.$$

No entanto,

$$AC = I \Leftrightarrow B(AC) = BI = B.$$

Logo, como $(BA)C = B(AC)$ pela propriedade associativa do produto de matrizes, temos que $B = C$. □

Teorema 2. *Se A, B são matrizes $n \times n$ inversíveis, então $(AB)^{-1} = B^{-1}A^{-1}$.*

Demonstração. Basta verificarmos que $(AB)B^{-1}A^{-1} = B^{-1}A^{-1}AB$, já que a inversa é única. □

Teorema 3. *Se A é inversível, A^T também o é, e $(A^T)^{-1} = (A^{-1})^T$.*

Demonstração. Basta verificarmos que $(A^T)(A^{-1})^T = (A^{-1})^T A^T = I$.

Mas $A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I$. Além disso, $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$. □

1.4 Fatoração $PA = LU$ de uma matriz A.

1.4.1 Sistemas Lineares

Considere o sistema linear

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

em que $A_{m\times n}$ é uma matriz de números reais, x é um vetor com n componentes reais e b é um vetor com m componentes reais. Se $b = 0$, chamamos este sistema de *homogêneo*.

Ideia: transformar a matriz (quadrada) geral em uma matriz triangular para resolver o sistema linear por substituição. Vamos considerar primeiramente sistemas quadrados.

Exemplo.

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

1.4.2 Forma matricial: Forma escalonada do sistema

Poderíamos resolver um sistema $Ax = b$ encontrando a inversa de A, pois se $Ax = b$, então

$$x = A^{-1}b$$

No entanto, encontrar a inversa de uma matriz não é tarefa fácil; assim, vamos procurar resolver o sistema usando outras estratégias.

Exemplo.

$$\begin{cases} x + y + 2z = 9 \\ 2x + 4y - 3z = 1 \\ 3x + 6y - 5z = 0 \end{cases}$$

Inicialmente, vamos tentar eliminar x na segunda e terceira equações:

$$\begin{cases} x & +y & +2z & = 9 \\ & y & -\frac{7z}{2} & = -\frac{17}{2} \\ & -y & +\frac{11z}{3} & = 9 \end{cases}$$

Agora, vamos eliminar y na terceira equação:

$$\begin{cases} x & +y & +2z & = 9 \\ & y & -(7/2)z & = -17/2 \\ & & (1/6)z & = 1/2 \end{cases}$$

Resolvendo para z, temos que

$$z = 3$$

Além disso,

$$2y - 7z = -17 \Leftrightarrow 2y = -17 + 21 = 4 \Leftrightarrow y = 2$$

Por último,

$$x + y + 2z = 9 \Leftrightarrow x = 9 - 2(3) - 2 = 9 - 8 = 1.$$

Exemplo. *Considere o seguinte exemplo: (homogêneo)*

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \\ x_1 + x_3 = 0 \end{cases}$$

Podemos representar o processo, agora numa forma matricial:

$$\begin{pmatrix} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ x_1 & + & x_2 & + & x_3 & = & 0 \\ x_1 & & & + & x_3 & = & 0 \end{pmatrix} \rightarrow \begin{pmatrix} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & - & 3x_2 & & & = & 0 \\ & - & 2x_2 & & & = & 0 \end{pmatrix}$$

Aqui, podemos ver que, pelas equações acima, $x_2 = 0$*,* $x_1 + x_3 = 0$*, ou seja* x_1 *é* livre, $x_2 = 0$ *e* $x_3 = -x_1$*. Neste caso, o sistema tem infinitas soluções.*

Se, por outro lado, o lado direito for diferente de nulo, teremos:

$$\begin{pmatrix} x_1 & - & 2x_2 & + & x_3 & = & 1 \\ x_1 & + & x_2 & + & x_3 & = & 2 \\ x_1 & & & + & x_3 & = & 3 \end{pmatrix} \rightarrow \begin{pmatrix} x_1 & - & 2x_2 & + & x_3 & = & 1 \\ & - & 3x_2 & & & = & 1 \\ & - & 2x_2 & & & = & 2 \end{pmatrix}$$

De onde vemos que o valor de x_2 *não pode ser encontrado. Portanto, esse sistema é inconsistente.*

No entanto, se tivéssemos

$$\begin{pmatrix} x_1 & - & 2x_2 & + & x_3 & = & 6 \\ x_1 & + & x_2 & + & x_3 & = & -3 \\ x_1 & & & + & x_3 & = & 0 \end{pmatrix} \rightarrow \begin{pmatrix} x_1 & - & 2x_2 & + & x_3 & = & 6 \\ & - & 3x_2 & & & = & 9 \\ & - & 2x_2 & & & = & 6 \end{pmatrix}$$

então o sistema seria consistente, mas teria infinitas soluções.

Qual é o problema que podemos encontrar aqui? **Pivôs nulos.** Isso indica problemas no sistema. Assim, esse procedimento pode ser usado para identificar classes de sistemas lineares; se o procedimento puder ser realizado até o fim, sabemos que teremos uma solução definida para o sistema.

1.4.3 Eliminação Gaussiana e a Forma Escalonada

O procedimento que descrevemos acima é chamado de *Eliminação Gaussiana*. O objetivo da eliminação gaussiana é obter a chamada *forma escalonada da matriz*.

Definição 2. *Uma matriz A está na forma escalonada quando:*

- *Se uma linha não for inteiramente nula, o primeiro número não-nulo da linha (indo da esquerda para a direita) é 1 (pivô);*
- *Se uma linha é nula, ela está abaixo de todas as linhas que contém elementos não-nulos;*

- *Para cada duas linhas não-nulas, o pivô da de baixo está mais à direita que o da de cima;*

- *Cada coluna que tem pivô é nula no resto. (se utilizarmos o procedimento de Gauss-Jordan até o final)*

Esta definição é válida também para sistemas não quadrados, mas por enquanto vamos nos concentrar em sistemas quadrados.

Para estes sistemas, temos algumas situações que podem ocorrer.

Para sistemas não-homogêneos:

- Se pudermos realizar o escalonamento até o fim sem encontrarmos pivôs nulos, temos dois casos:
  - se o lado direito for consistente com as equações, então temos um sistema consistente e determinado (caso quadrado)
  - se o lado direito for inconsistente, não poderemos encontrar solução para o sistema.
- Se não pudermos realizar o escalonamento até o fim, também temos duas situações:
  - Se o lado direito for consistente, teremos um sistema consistente indeterminado (infinitas soluções)
  - caso contrário, teremos um sistema inconsistente, sem soluções.

Para o caso homogêneo,

- Se pudermos realizar o escalonamento até o fim sem encontrarmos pivôs nulos, então a solução trivial é solução do sistema.
- Caso contrário, o sistema tem infinitas soluções (indeterminado).

1.4.4 Operações elementares: Matrizes elementares

Nestes exemplos, resolvemos o sistema ao realizarmos as chamadas operações elementares nas linhas nestas matrizes. As operações elementares são:

- Multiplicar uma linha por um escalar não-nulo;

- Trocar uma linha r do sistema por $r + cs$, onde c é um escalar e s é outra linha do sistema ($r \neq s$);
- Trocar duas linhas de lugar.

Estas operações são importantes pois podem ser revertidas! Vamos analisar essas operações na forma matricial.

Definição 3. *Uma matriz $n \times n$ que pode ser obtida da matriz I_n executando-se uma única operação elementar sobre linhas é uma matriz elementar.*

Se E resulta de uma operação elementar em $I_{m\times m}$ e $A_{m\times n}$, então EA realiza as mesmas operações em A.

Exemplo.

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} = EI.$$

$$A = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix}; \qquad EA = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 4 & 4 & 10 & 9 \end{pmatrix}$$

Teorema 4. *Toda E elementar é inversível, e E^{-1} também é elementar.*

Demonstração. Seja E_0 a matriz que resulta da aplicação da operação elementar inversa de E em I (garantida pois estamos em um corpo!!). Então,

$$E_0E = I = EE_0$$

E_0 também é elementar pois representa também uma operação elementar. □

Exemplo.

$$\begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix} \qquad b = \begin{pmatrix} 5 \\ -2 \\ 9 \end{pmatrix}$$

Temos então, na sequência:

$$\begin{aligned} E^{(1)}A &= A^{(1)} \\ \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix} &= \begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ -2 & 7 & 2 \end{pmatrix} \end{aligned}$$

Além disso,

$$
\begin{aligned}
E^{(1)}b &= b^{(1)} \\
\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 5 \\ -2 \\ 9 \end{pmatrix} &= \begin{pmatrix} 5 \\ -12 \\ 9 \end{pmatrix}
\end{aligned}
$$

Continuando o processo,

$$
\begin{aligned}
E^{(2)}A^{(1)} &= A^{(2)} \\
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ -2 & 7 & 2 \end{pmatrix} &= \begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 8 & 3 \end{pmatrix}
\end{aligned}
$$

Além disso,

$$
\begin{aligned}
E^{(2)}b^{(1)} &= b^{(2)} \\
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 5 \\ -12 \\ 9 \end{pmatrix} &= \begin{pmatrix} 5 \\ -12 \\ 14 \end{pmatrix}
\end{aligned}
$$

Por último,

$$
\begin{aligned}
E^{(3)}A^{(2)} &= A^{(3)} \\
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 8 & 3 \end{pmatrix} &= \begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}
$$

Além disso,

$$
\begin{aligned}
E^{(3)}b^{(2)} &= b^{(3)} \\
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 5 \\ -12 \\ 14 \end{pmatrix} &= \begin{pmatrix} 5 \\ -12 \\ 2 \end{pmatrix}
\end{aligned}
$$

Desta forma,

$$
E = E^{(3)}E^{(2)}E^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}
$$

Existem algumas maneiras de representar a matriz depois das operações elementares:

- Se quisermos obter a matriz na forma *escalonada*, realizamos as operações por linhas, de forma que a matriz resultante seja triangular superior com pivôs 1;

- Se quisermos obter a matriz na forma *reduzida por linhas*, realizamos as operações por linhas, de forma que a matriz resultante tenha pivôs 1, e zero no resto de cada coluna com pivô;

- Se quisermos apenas resolver o sistema, realizamos as operações por linhas, de forma que a matriz resultante seja triangular superior com pivôs não-nulos (não necessariamente 1). Neste caso, obtemos a decomposição (ou fatoração) LU da matriz.

As operações elementares realizadas no momento do escalonamento do sistema podem ser acumuladas em uma matriz E, que é o produto das matrizes elementares utilizadas. Assim,

$$EA = U$$

e ainda,

$$A = E^{-1}U = LU.$$

Como isso se relaciona com o sistema?

Podemos agora fazer

$$Ax = b \Leftrightarrow LUx = b$$

ou seja,

Resolvemos $Ly = b$ para y;

Resolvemos $Ux = y$ para x.

Felizmente, não precisamos montar as matrizes E! Basta fazermos o seguinte procedimento:

- Reduzir A à forma escalonada U, guardando os multiplicadores;

- Na diagonal de L colocar 1's;

- Abaixo da diagonal de L, colocar o negativo do multiplicador usado para zerar o respectivo elemento de A.

Exemplo.

$$\begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix}$$

Temos então, na sequência:

$$\begin{aligned} A &= LU \\ \begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Definição 4. *Se A e B são matrizes $m \times n$, dizemos que A é equivalente por linhas a B se B puder ser obtida através de um número finito de operações elementares efetuadas em A.*

Desta forma, é fácil ver que, como cada operação elementar pode ser representada por uma matriz elementar, e como a aplicação sucessiva destas matrizes elementares é representada pelo produto de todas estas matrizes, então

A e B são equivalentes por linhas se e somente se $A = PB$, onde P é um produto de matrizes elementares.

Teorema 5. *Se A e B são equivalentes por linhas, então os sistemas $Ax = 0$ e $Bx = 0$ tem exatamente as mesmas soluções.*

Demonstração. Basta mostrarmos que uma operação elementar não altera o conjunto de soluções. Mas, note que

$$EAx = 0 \Leftrightarrow Ax = E^{-1}0 = 0.$$

Portanto, $EAx = 0$ tem o mesmo conjunto de soluções que $Ax = 0$. Logo, se cada operação elementar não altera o conjunto solução, e aplicamos uma sequência finita de operações elementares em A para chegarmos a B, então $Ax = 0$ e $Bx = 0$ tem o mesmo conjunto de soluções. □

Teorema 6. *Se A é $n \times n$, então sua forma reduzida por linhas escalonada (escalonada, em que cada coluna que contém um pivô é nula em todas as outras entradas) terá ao menos uma linha toda nula, ou será a matriz identidade.*

Demonstração. Suponha que B é a forma reduzida por linhas escalonada da matriz. Se B tiver pelo menos uma linha toda nula então não há nada a provar. Suponha desta forma que B não tem uma linha toda nula. Então, toda linha deve ter um 1.

Agora, sabemos que o 1 de uma linha deve estar à esquerda do 1 da linha de baixo. Suponha então que o 1 da primeira linha não está em b_{11}. Assim, suponha que ele está em b_{12}. Se este for o caso, então vamos supor que o melhor cenário acontece, ou seja, cada 1 da linha 2 em diante está exatamente uma posição à direita do 1 da linha de cima. Então, na linha 2 o primeiro 1 deve estar pelo menos em b_{23}, e o de baixo em b_{34} e assim por diante, até que, na linha $n-1$, teremos $b_{n-1,n} = 1$. Desta forma, na última linha não teremos a posição $b_{n,n+1}$ para colocar o próximo 1, e esta linha deve ficar nula, o que contradiz nossa hipótese.

Se tomarmos um caso pior ainda, em que o 1 de uma linha está mais de uma casa deslocado para a direita, acabaremos com mais linhas nulas, o que também confirma nosso teorema. Desta forma, b_{11} deve ser 1, e assim podemos concluir para todos os outros elementos não nulos da forma reduzida por linhas escalonada de A. $\square$

Corolário 1. *Toda $A_{n\times n}$ pode ser reduzida à forma escalonada (reduzida por linhas, ou reduzida por linhas escalonada).*

Demonstração. Segue direto da definição de matriz reduzida por linhas escalonada, já que podemos ter inclusive linhas nulas. Todas as outras operações estão bem definidas para quaisquer entradas da matriz A. $\square$

Teorema 7. *Se A é uma matriz $n \times n$ então as seguintes afirmações são equivalentes:*

(a) A é inversível;

(b) A única solução para o sistema homogêneo $Ax = 0$ é a solução trivial;

(c) A é equivalente por linhas a I;

(d) A pode ser expressa como um produto de matrizes elementares.

Demonstração.

(a) $\Rightarrow$ (b) Suponha que A é inversível, vamos mostrar então que $Ax = 0$ só possui a solução trivial. Suponha que x_0 é uma solução deste sistema. Como A é inversível, basta vermos que $A^{-1}Ax_0 = A^{-1}0 = 0$, ou seja $x_0 = 0$.

(b) $\Rightarrow$ (c) Primeiramente, escrevemos a matriz aumentada deste sistema, $[A|0]$. Agora, como sabemos que a única solução deste sistema é a solução trivial, então podemos reduzir A à forma reduzida por linhas escalonada que nos dá essa solução, ou seja, A deve ser equivalente por linhas a I.

(c) $\Rightarrow$ (d) Como A é equivalente por linhas a I, sabemos que pode-se escrever

$$E_k E_{k-1} \dots E_2 E_1 A = I$$

Como já mostramos que as matrizes elementares são inversíveis, isto nos dá

$$A = E_1^{-1} E_2^{-1} \dots E_{k-1}^{-1} E_k^{-1}.$$

(d) $\Rightarrow$ (a) Se A é um produto de matrizes inversíveis, e como mostramos que tal produto é inversível também, então A deve ser inversível.

$\square$

1.4.5 Fatoração $PA = LU$.

Vamos considerar uma situação que não tínhamos considerado antes: a troca de linhas.

Exemplo.

$$\begin{pmatrix} 0 & 2 \\ 3 & 4 \end{pmatrix}$$

Aqui, o primeiro pivô é zero! Não podemos multiplicar nenhum número pela primeira equação para zerar a segunda. Podemos então usar a operação elementar que não havíamos usado até agora: troca de linhas. Assim:

$$PA = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 0 & 2 \end{pmatrix}$$

que, inclusive, já está na forma escalonada! Portanto, podemos reescrever esta situação como

$$PA = LU.$$

Exemplo.

$$\begin{pmatrix} 1 & 1 & 1 & \_ \\ 2 & 2 & 5 & \_ \\ 4 & 6 & 8 & \_ \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 1 & \_ \\ 0 & 0 & 3 & \_ \\ 0 & 2 & 4 & \_ \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 1 & \_ \\ 0 & 2 & 4 & \_ \\ 0 & 0 & 3 & \_ \end{pmatrix}$$

Exemplo.

$$\begin{pmatrix} 1 & 1 & 1 & \_ \\ 2 & 2 & 5 & \_ \\ 4 & 4 & 8 & \_ \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 1 & \_ \\ 0 & 0 & 3 & \_ \\ 0 & 0 & 4 & \_ \end{pmatrix}$$

Não há permutação que permita encontrar um pivô.

A matriz de permutação é sempre a identidade com linhas trocadas. Desta forma, uma sequência de permutações consecutivas pode ser encontrada através do produto das permutações, assim como havíamos feito com as operações elementares.

CAPÍTULO 2

Espaços Vetoriais

Nesta unidade, estudaremos as propriedades dos espaços vetoriais e sua relação com os sistemas lineares.

2.1 Introdução aos espaços vetoriais

A Álgebra Linear é o estudo dos espaços vetoriais sobre corpos arbitrários e das transformações lineares entre esses espaços.

Definição 5. *Um conjunto não-vazio $\mathbb{K}$ é um* corpo *se em $\mathbb{K}$ pudermos definir duas operações, denotadas por + (soma) e $\cdot$ (multiplicação), satisfazendo as seguintes propriedades:*

(i) $a + b = b + a$, $\forall a, b \in \mathbb{K}$ *(comutativa)*

(ii) $a + (b + c) = (a + b) + c$, $\forall a, b, c \in \mathbb{K}$ *(associativa)*

(iii) Existe um elemento em $\mathbb{K}$, denotado por 0 *e chamado de* elemento neutro da soma*, que satisfaz* $0 + a = a + 0 = a$, $\forall a \in \mathbb{K}$.

(iv) Para cada $a \in \mathbb{K}$, existe um elemento em $\mathbb{K}$ denotado por $-a$ e chamado de oposto de *a (ou* inverso aditivo de *a) tal que* $a + (-a) = (-a) + a = 0$.

(v) $a \cdot b = b \cdot a$, $\forall a, b \in \mathbb{K}$ *(comutativa)*

(vi) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$, $\forall a, b, c \in \mathbb{K}$

(vii) *Existe um elemento em* $\mathbb{K}$ *denotado por* 1 *e chamado de* elemento neutro da multiplicação, *tal que* $1 \cdot a = a \cdot 1 = a$, $\forall a \in \mathbb{K}$.

(viii) *Para cada elemento não-nulo* $a \in \mathbb{K}$, *existe um elemento em* $\mathbb{K}$, *denotado por* a^{-1} *e chamado de* inverso multiplicativo de a, *tal que* $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

(ix) $(a + b) \cdot c = a \cdot c + b \cdot c$, $\forall a, b, c \in \mathbb{K}$ *(distributiva).*

Exemplo. *São corpos:* $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.

Definição 6. *Um conjunto não vazio E é um* espaço vetorial *sobre um corpo* $\mathbb{K}$ *se em seus elementos, denominados* vetores, *estiverem definidas duas operações:*

- soma*: A cada* $u, v \in E$, *associa* $u + v \in E$
- multiplicação por um escalar*: a cada escalar* $\alpha \in \mathbb{K}$ *e a cada vetor* $v \in E$, *associa* $\alpha v \in E$.

Estas operações devem satisfazer as condições abaixo, denominadas axiomas *de espaço vetorial:*

(i) **Comutatividade:** $u + v = v + u$, $\forall u, v \in E$

(ii) **Associatividade:** $(u + v) + w = u + (v + w)$ *e* $(\alpha\beta)v = \alpha(\beta v)$, $\forall u, v, w \in E$ *e* $\alpha, \beta \in \mathbb{K}$

(iii) **Existência do vetor nulo:** *existe um vetor* $0 \in E$, *chamado* vetor nulo, *tal que* $v + 0 = 0 + v = v$ *para todo* $v \in E$.

(iv) **Existência do inverso aditivo:** *para cada vetor* $v \in E$ *existe um vetor* $-v \in E$ *chamado* inverso aditivo *tal que* $-v + v = v + (-v) = 0 \in E$.

(v) **Distributividade:** $(\alpha + \beta)v = \alpha v + \beta v$ *e* $\alpha(u + v) = \alpha u + \alpha v$, $\forall u, v \in E$ *e* $\forall \alpha, \beta \in \mathbb{K}$

(vi) **Multiplicação por 1:** $1 \cdot v = v$, *em que* 1 *é o elemento neutro da multiplicação em* $\mathbb{K}$.

Exemplo 1.1. *Todo corpo é um espaço vetorial sobre si mesmo. De fato, se* $\mathbb{K}$ *é um corpo, então as duas operações internas em* $\mathbb{K}$ *podem ser vistas como a soma de vetores e a multiplicação por escalares.*

Exemplo 1.2. *Para todo número natural n, o conjunto* $\mathbb{K}^n$*, definido como*

$$\mathbb{K}^n = \underbrace{\mathbb{K} \times \cdots \times \mathbb{K}}_{n \text{ vezes}} = \{(u_1, \ldots, u_n) : u_i \in \mathbb{K}, \forall i = 1, \ldots, n\}$$

é um espaço vetorial sobre $\mathbb{K}$*.*

Por definição, a igualdade vetorial $u = v$ *significa as n igualdades numéricas*

$$\begin{aligned} \alpha_1 &= \beta_1 \\ \alpha_2 &= \beta_2 \\ &\vdots \\ \alpha_n &= \beta_n. \end{aligned}$$

Os números $\alpha_1, \ldots, \alpha_n$ *são chamados de* coordenadas *do vetor u. As operações do espaço vetorial* $\mathbb{K}^n$ *são definidas naturalmente por*

$$\begin{aligned} u + v &= (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n) \\ \rho u &= (\rho\alpha_1, \rho\alpha_2, \ldots, \rho\alpha_n). \end{aligned}$$

O vetor zero é, por definição, $(0, 0, \ldots, 0)$*, em que* 0 *é o elemento neutro da soma em* $\mathbb{K}$*.*

O inverso aditivo de $u = (\alpha_1, \ldots, \alpha_n)$ *é* $-u = (-\alpha_1, \ldots, -\alpha_n)$*.*

Exemplo 1.3. *Os elementos do espaço vetorial* $\mathbb{R}^\infty$ *são as sequências infinitas* $u = (\alpha_1, \ldots, \alpha_n, \ldots)$*,* $v = (\beta_1, \ldots, \beta_n, \ldots)$ *de números reais. O elemento zero é a sequência formada por infinitos zeros* $0 = (0, \ldots, 0, \ldots)$ *e o inverso aditivo da sequência u é* $-u = (-\alpha_1, \ldots, -\alpha_n, \ldots)$*. As operações de adição e multiplicação por escalar são definidas por*

$$\begin{aligned} u + v &= (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n, \ldots) \\ \rho u &= (\rho\alpha_1, \ldots, \rho\alpha_n, \ldots). \end{aligned}$$

Exemplo 1.4. *Uma matriz* $m \times n$*, definida no corpo* $\mathbb{K}$ *e denotada por* $A = [a_{ij}]$*, é uma lista de elementos* $a_{ij} \in \mathbb{K}$ *com índices duplos, onde* $1 \leq i \leq m$ *e* $1 \leq j \leq n$*. Costuma-se representar a matriz na forma*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

O vetor

$$(a_{i1}, a_{i2}, \ldots, a_{in}) \in \mathbb{K}^n$$

é o i-ésimo vetor linha da matriz A e o vetor

$$(a_{1j}, a_{2j}, \ldots, a_{mj}) \in \mathbb{K}^m$$

é o j-ésimo vetor coluna de A. Quando $m = n$, dizemos que A é uma matriz quadrada. O conjunto $\mathcal{M}_{m\times n}(\mathbb{K})$ de todas as matrizes $m \times n$ com elementos em $\mathbb{K}$ torna-se um espaço vetorial quando nele se define a soma das matrizes $A = [a_{ij}]$ e $B = [b_{ij}]$ como

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

e o produto de uma matriz A pelo escalar $\rho \in \mathbb{K}$ como

$$\rho A = \begin{pmatrix} \rho a_{11} & \rho a_{12} & \cdots & \rho a_{1n} \\ \rho a_{21} & \rho a_{22} & \cdots & \rho a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho a_{m1} & \rho a_{m2} & \cdots & \rho a_{mn} \end{pmatrix}$$

A matriz nula $0 \in \mathcal{M}_{m\times n}$ é aquela formada por zeros e o inverso aditivo da matriz $A = [a_{ij}]$ é a matriz $-A = [-a_{ij}]$.

Exemplo 1.5. *O conjunto de polinômios*

$$\mathcal{P}(\mathbb{K}) = \{p(x) = a_n x^n + \ldots + a_1 x + a_0 : a_i \in \mathbb{K} \text{ e } n \geq 0\}$$

é um espaço vetorial com as operações usuais de soma de polinômios e multiplicação por escalar.

Exemplo 1.6. *Seja X um conjunto não-vazio qualquer. O símbolo $\mathcal{F}(X; \mathbb{R})$ representa o conjunto de todas as funções reais $f : X \to \mathbb{R}$. Ele se torna um espaço vetorial quando se define a soma $f + g$ de duas funções e o produto αf do número α pela função f da maneira natural:*

$$(f + g)(x) = f(x) + g(x) \qquad e \qquad (\alpha f)(x) = \alpha f(x).$$

Variando o conjunto X, obtemos diversos exemplos de espaços vetoriais da forma $\mathcal{F}(X;\mathbb{R})$. Por exemplo, se $X = \{1,\ldots,n\}$, então $\mathcal{F}(X;\mathbb{R}) = \mathbb{R}^n$, pois a cada número em X associamos um número real α, gerando assim uma lista de n valores reais para cada elemento do conjunto. Similarmente, se $X = \mathbb{N}$, então $\mathcal{F}(X;\mathbb{R}) = \mathbb{R}^\infty$. Se X é o produto cartesiano dos conjuntos $\{1,\ldots,m\}$ e $\{1,\ldots,n\}$ então $\mathcal{F}(X;\mathbb{R}) = \mathcal{M}_{m\times n}$.

Como consequência dos axiomas, valem num espaço vetorial as regras operacionais habitualmente usadas nas manipulações numéricas:

1. Para todos $u, v, w \in E$, temos que $w + u = w + v \Rightarrow u = v$. Em particular, $w + u = w \Rightarrow u = 0$ e $w + u = 0 \Rightarrow u = -w$.

2. Dados $0 \in \mathbb{K}$ e $v \in E$, temos que $0v = 0 \in E$. Analogamente, dados $\alpha \in \mathbb{K}$ e $0 \in E$, temos que $\alpha 0 = 0$.

3. Se $\alpha \neq 0$ e $v \neq 0$ então $\alpha v \neq 0$.

4. $(-1)v = -v$.

Observação. Um espaço vetorial sobre um corpo $\mathbb{K}$ é um conjunto E de *vetores*, com uma operação de soma que é uma função $+ : E \to E$ e uma operação de produto por escalar, que é uma função $\cdot : \mathbb{K} \times E : E$, satisfazendo os axiomas listados acima. Note que os axiomas não involvem a propriedade de inverso multiplicativo do corpo, e podemos definir uma estrutura semelhante à de espaço vetorial sobre um anel, que chamamos de *módulo* sobre $\mathbb{K}$. No entanto, a maioria dos teoremas provados para espaços vetoriais não seria válida nos módulos; por exemplo, não podemos falar da dimensão de um módulo.

2.2 Subespaços vetoriais

Definição 7. *Um subespaço vetorial do espaço vetorial E é um subconjunto $F \subset E$ que, relativamente às operações de E, é ainda um espaço vetorial, ou seja, satisfaz*

(i) Para todo $u, v \in F$, $u + v \in F$

(ii) Para todo $u \in F$ e $\alpha \in \mathbb{K}$, $\alpha u \in F$.

Note que no caso de um subespaço, não é necessário verificar as seis propriedades listadas anteriormente pois elas já são satisfeitas para E, e $F \subset E$. No entanto, um subespaço deve ser *fechado* para a adição e a multiplicação por escalar. Mais geralmente, dados $v_1, \ldots, v_m \in F$ e $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$,

$$v = \alpha_1 v_1 + \ldots + \alpha_m v_m$$

deve pertencer a F.

Observação: O vetor nulo pertence a *todos* os subespaços. Para verificar isto, basta tomarmos $0 \in \mathbb{K}$: $0v$ deve pertencer a F para todo $v \in F$. O espaço inteiro E também é um exemplo trivial de subespaço de E. Todo subespaço é, em si mesmo, um espaço vetorial. O conjunto vazio não pode ser um subespaço vetorial.

Exemplo 2.1. *(i) Seja $v \in E$ um vetor não-nulo. O conjunto $F = \{\alpha v : \alpha \in \mathbb{K}\}$ de todos os múltiplos de v é um subespaço vetorial de E, chamado de* reta que passa pela origem e contém v.

(ii) Seja $E = \mathcal{F}(\mathbb{R};\mathbb{R})$ o espaço vetorial das funções reais de uma variável real $f : \mathbb{R} \to \mathbb{R}$. Para cada $k \in \mathbb{N}$, o conjunto $C^k(\mathbb{R})$ das funções k vezes continuamente diferenciáveis é um subespaço vetorial de E.

(iii) Sejam $a_1, \ldots, a_n$ números reais. O conjunto $\mathcal{H}$ de todos os vetores $v = (x_1, \ldots, x_n) \in \mathbb{R}^n$ tais que

$$a_1 x_1 + \ldots + a_n x_n = 0$$

é um subespaço vetorial de $\mathbb{R}^n$. No caso trivial em que $a_1 = \ldots = a_n = 0$, o subespaço $\mathcal{H}$ é todo o $\mathbb{R}^n$. Se, ao contrário, pelo menos um dos $a_i \neq 0$, $\mathcal{H}$ chama-se hiperplano *de $\mathbb{R}^n$ que passa pela origem.*

(iv) Seja E o espaço das matrizes 3×3: $E = \{A \in \mathbb{R}^{3\times 3}\}$. O conjunto das matrizes triangulares inferiores de dimensão 3 é um subespaço de E, assim como o conjunto das matrizes simétricas.

(v) Dentro do espaço $\mathbb{R}^3$, os subespaços possíveis são: o subespaço nulo, o espaço inteiro, as retas que passam pela origem, e os planos que passam pela origem. Qualquer reta que não passe pela origem não pode ser um subespaço (pois não contem o vetor nulo).

2.3 Interseção e soma de subespaços vetoriais

Teorema 8. *Dados um espaço vetorial E e subespaços $F_1, F_2 \subset E$, a interseção $F_1 \cap F_2$ ainda é um subespaço de E.*

Demonstração. Primeiramente, note que $F_1 \cap F_2$ nunca é vazio, pois $0 \in F_1$ e $0 \in F_2$. Precisamos então verificar as duas condições que definem um subespaço vetorial.

(i) Sejam $u, v \in F_1 \cap F_2$. Então, $u, v \in F_1$ e $u, v \in F_2$. Logo, como F_1 e F_2 são ambos subespaços de E, $u + v \in F_1$ e $u + v \in F_2$, portanto $u + v \in F_1 \cap F_2$.

(ii) Seja $u \in F_1 \cap F_2$ e $\alpha \in \mathbb{K}$. Então, $u \in F_1$ e $u \in F_2$. Como ambos F_1 e F_2 são subespaços de E, $\alpha u \in F_1$ e $\alpha u \in F_2$. Portanto, $\alpha u \in F_1 \cap F_2$.

Assim, provamos que a interseção dos dois subespaços é também um subespaço vetorial de E. □

A união de dois subespaços vetoriais *não* é (em geral) um subespaço vetorial.

Exemplo 3.1. *Considere o espaço E das matrizes reais $n \times n$. É fácil verificar que $F_1 = \{$matrizes triangulares superiores$\}$ e $F_2 = \{$matrizes triangulares inferiores$\}$ são ambos subespaços de E. Então, $F_1 \cap F_2 = \{$matrizes diagonais$\}$ também é um subespaço de E. Já a união de F_1 e F_2 é formada pelas matrizes triangulares (superiores e inferiores). Note que se tomarmos uma matriz triangular superior e outra matriz triangular inferior, a soma destas matrizes é uma matriz que não tem estrutura especial, e que portanto não está na união dos dois subespaços.*

Exemplo 3.2. *Considere $E = \mathbb{R}^3$. Sejam F_1, F_2 dois planos em $\mathbb{R}^3$ passando pela origem. Então, $F_1 \cap F_2$ é a reta de interseção de F_1 e F_2 passando pela origem. $F_1 \cup F_2$ é a união dos dois planos.*

Exemplo 3.3. *Considere $E = \mathbb{R}^3$, F_1 e F_2 duas retas que passam pela origem. Ambos F_1 e F_2 são subespaços, mas sua união, representada pelo feixe das duas retas, não o é, pois a soma de dois elementos nesta união pode se encontrar no plano definido pelas duas retas, que não está inteiramente contido no subespaço união.*

Como vimos no último exemplo, a união de dois subespaços vetoriais não é necessariamente um subespaço vetorial. No entanto, podemos construir um conjunto S que contém F_1 e F_2 e que é subespaço de E, como veremos no Teorema a seguir.

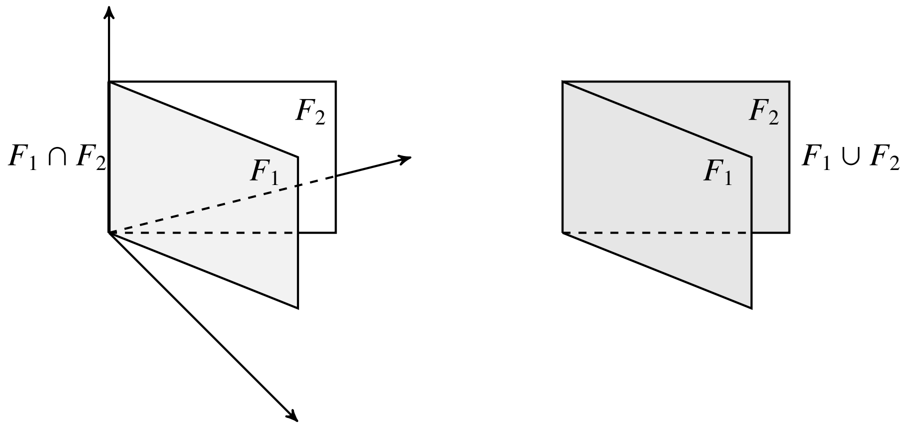


Figura 2.1: Interseção e união de subespaços vetoriais.

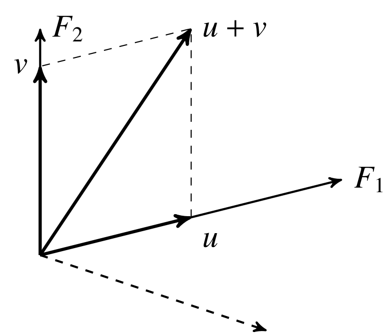


Figura 2.2: União de subespaços vetoriais.

Teorema 9. *Sejam F_1 e F_2 subespaços de um espaço vetorial E. Então o conjunto*

$$S = F_1 + F_2 = \{w \in E : w = w_1 + w_2, w_1 \in F_1, w_2 \in F_2\}$$

é um subespaço de E.

Demonstração. Vamos verificar as condições para que S seja um subespaço de E. Primeiramente, note que $0 \in S$ pois $0 \in F_1$ e $0 \in F_2$.

(i) Sejam $v, w \in S$. Então $v = v_1 + v_2$, $v_1 \in F_1$, $v_2 \in F_2$ e $w = w_1 + w_2$, $w_1 \in F_1$,

$w_2 \in F_2$. Assim

$$\begin{aligned} v + w &= (v_1 + v_2) + (w_1 + w_2) \\ &= (v_1 + w_1) + (v_2 + w_2) \in S, \end{aligned}$$

pois $v_1 + w_1 \in F_1$ e $v_2 + w_2 \in F_2$ já que ambos são subespaços de E e $v_1, w_1 \in F_1$ e $v_2, w_2 \in F_2$, e a última igualdade segue das propriedades da soma no espaço vetorial E.

(ii) Sejam $\alpha \in \mathbb{K}$ e $w \in S$. Então,

$$\alpha w = \alpha(w_1 + w_2) = \alpha w_1 + \alpha w_2 \in S$$

já que $\alpha w_1 \in F_1$ e $\alpha w_2 \in F_2$ pois ambos são subespaços de E.

□

Exemplo 3.4. *No Exemplo 3.3, $S = F_1 + F_2$ é o plano que contém as duas retas.*

Considere agora uma matriz $A \in \mathbb{R}^{m\times n}$. Chegamos então ao caso interessante de subespaços ligados à matriz A em um sistema linear $Ax = b$, com $x \in \mathbb{R}^n$ e $b \in \mathbb{R}^m$. Para relacionarmos o conceito de subespaços vetoriais à resolução de sistemas de equações lineares, precisamos do conceito de combinação linear.

Definição 8. *Seja E um espaço vetorial, e sejam $u_1, u_2, \ldots, u_n \in E$ vetores neste espaço. Então uma* combinação linear *destes vetores é um vetor u no espaço E dado por*

$$u = \alpha_1 u_1 + \ldots + \alpha_n u_n,$$

para $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$.

Definição 9. *Uma vez fixados os vetores $\{v_1, \ldots, v_n\}$ em V, o conjunto $W \subset V$ que contém todas as combinações lineares destes vetores é chamado de* espaço gerado *pelos vetores $v_1, \ldots, v_n$. Denotamos isto por*

$$\begin{aligned} W &= span\{v_1, \ldots, v_n\} \\ &= \{v \in V : v = a_1 v_1 + a_2 v_2 + \ldots + a_n v_n, a_i \in \mathbb{K}, 1 \le i \le n\}. \end{aligned}$$

Uma outra caracterização de subespaço gerado é a seguinte: W é o menor subespaço de V que contém o conjunto de vetores $\{v_1, \ldots, v_n\}$ no sentido que qualquer outro subespaço W' de V que contenha estes vetores satisfará $W' \supset W$, já que como $v_1, \ldots, v_n \in W'$ e W' é um subespaço vetorial, então qualquer combinação linear destes vetores também está incluida em W'; logo $W \subset W'$.

2.4 Dependência linear entre vetores

Definição 10. *Seja* $X = \{v_1, v_2, \ldots, v_n\} \subset E$ *um conjunto de vetores. Se nenhum dos vetores* v_i *puder ser escrito como combinação linear dos outros vetores, dizemos que este conjunto é linearmente independente (l.i.). Formalmente, o conjunto* X *é l.i. se e somente se a única combinação linear nula dos vetores de* X *for aquela cujos coeficientes são todos nulos, ou seja,*

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0 \Rightarrow \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0.$$

Evidentemente, todo subconjunto de um conjunto l.i. é também l.i.

Exemplo 4.1. *Em* $\mathbb{R}^2$*, quaisquer dois vetores que não sejam colineares são l.i.*

Exemplo 4.2. *Em* $\mathbb{R}^n$*, chamamos de vetores canônicos os vetores definidos como, para todos* $i, j = 1, \ldots, n$

$$(e_i)_j = \begin{cases} 1, & \textit{se } j = i \\ 0, & \textit{caso contrário.} \end{cases}$$

em que o subíndice j *denota a coordenada* j *do* i*-ésimo vetor canônico. Estes vetores são l.i.*

Exemplo 4.3. *Em* $\mathbb{R}^{2\times 2}$*, as matrizes*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \textit{ e } B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

são l.i.

Exemplo 4.4. *O conjunto* $\mathcal{P}(\mathbb{R})$ *dos polinômios*

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n$$

é um subespaço de $\mathcal{F}(\mathbb{R}; \mathbb{R})$*, assim como o conjunto* $\mathcal{P}_n$ *dos polinômios de grau* $\leq n$*. Note que o conjunto dos polinômios de grau* n *não é um subespaço, pois a soma de dois polinômios de grau* n *pode ter grau* $< n$*. Então, os monômios* $1, x, \ldots, n^n$ *em* $\mathcal{P}_n$ *são l.i., pois* $\alpha_0 + \alpha_1 x + \ldots + \alpha n x^n = p(x)$ *é o vetor nulo em* $\mathcal{P}_n$ *somente quando* $p(x)$ *é o polinômio identicamente nulo, ou seja,* $p(x) = 0$ *para todo* $x \in \mathbb{R}$*. Isto implica que* $\alpha_0 = \ldots = \alpha_n = 0$*, pois um polinômio não nulo de grau* k *tem no máximo* k *raízes reais. Podemos, além disso, concluir que o conjunto* $X = \{1, x, \ldots, x^n, \ldots\} \subset \mathcal{P}$ *é um conjunto infinito l.i.*

Teorema 10. *Se $v = \alpha_1 v_1 + \ldots + \alpha_m v_m = \beta_1 v_1 + \ldots + \beta_m v_m$ e os vetores $v_1, \ldots, v_m$ são l.i., então $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_m = \beta_m$.*

Se um conjunto X de vetores em um espaço vetorial E não é l.i., dizemos que ele é linearmente dependente (l.d.).

Variedades afins

Definição 11. *Um subconjunto $V \subset E$ chama-se uma* variedade afim *quando a reta que une dois pontos quaisquer de V está contida em V. Assim, $V \subset E$ é uma variedade afim se e somente se cumpre a seguinte condição:*

$$x, y \in V,\ t \in \mathbb{R} \Rightarrow (1-t)x + ty \in V.$$

Exemplo 4.5. *Todo subespaço é também uma variedade afim.*

Se $V_1, \ldots, V_m \subset E$ são variedades afins, então a intersecção $V_1 \cap V_2 \cap \ldots \cap V_m$ é ainda uma variedade afim. Todo ponto $p \in E$ é uma variedade afim.

Exemplo 4.6. *Sejam $a_1, \ldots, a_n, b$ números reais. O conjunto dos pontos $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ tais que*

$$a_1 x_1 + \ldots + a_n x_n = b$$

é uma variedade afim, que não contém a origem quando $b \neq 0$. Se os números a_i não forem todos nulos, chamamos esta veriedade $\mathcal{H}$ de hiperplano. *Se $a_1 = \ldots = a_n = 0$, então $\mathcal{H} = \emptyset$ quando $b \neq 0$ e $\mathcal{H} = \mathbb{R}^n$ quando $b = 0$. Mais geralmente, o conjunto das soluções de um sistema linear de m equações com n incógnitas é uma variedade afim, intersecção das m variedades afins definidas pelas equações do sistema.*

2.4.1 Base e dimensão de um espaço vetorial

Gostaríamos de encontrar, para um espaço W qualquer, um conjunto de vetores de forma que qualquer outro vetor em W possa ser escrito como combinação linear destes vetores (como i, j, k em $\mathbb{R}^3$, por exemplo)

Definição 12. *Uma* base *de um espaço vetorial E é um conjunto $\mathcal{B} \subset E$ linearmente independente que gera E, ou seja, todo vetor $v \in E$ se exprime, de modo único, como combinação linear $v = \alpha_1 v_1 + \ldots + \alpha_m v_m$ de elementos $v_1, \ldots, v_m$ da base $\mathcal{B}$. Se $\mathcal{B} = \{v_1, \ldots, v_m\}$ é uma base de E e $v = \alpha_1 v_1 + \ldots + \alpha_m v_m$, então os números $\alpha_1, \ldots, \alpha_m$ chamam-se as* coordenadas *do vetor v na base $\mathcal{B}$.*

Exemplo 4.7. *Base canônica no* $\mathbb{R}^n$.

Exemplo 4.8. *Os monômios* $1, x, \ldots, x^n$ *formam uma base para o espaço vetorial* $\mathcal{P}_n$ *dos polinômios de grau* $\leq n$. *O conjunto*

$$\{1, x, \ldots, x^n, \ldots\}$$

dos monômios de graus arbitrários constitui uma base (infinita) para o espaço vetorial $\mathcal{P}$ *de todos os polinômios reais.*

Resultados sobre bases

Lema 1. *Sejam* $v_1, \ldots, v_n \neq 0$ *que geram um e.v.* E. *Então dentre estes vetores podemos extrair uma base de* E.

Demonstração. Se $v_1, \ldots, v_n$ forem l.i., não há nada a fazer. Suponha então que eles sejam l.d. Então,

$$x_1 v_1 + \ldots + x_n v_n = 0$$

com pelo menos algum $x_i \neq 0$. Sem perda de generalidade, suponha que $x_n \neq 0$ (a ordem não importa). Então, escreva

$$v_n = -\frac{x_1}{x_n} v_i - \ldots - \frac{x_{n-1}}{x_n} v_{n-1}$$

Desta forma, $v_1, \ldots, v_{n-1}$ ainda geram E. Prossiga desta maneira até que todos os elementos l.d. tenham sido eliminados e teremos uma base de E. □

O Lema seguinte nos dá uma amostra da ligação entre os espaços vetoriais e as soluções dos sistemas lineares.

Lema 2. *Todo sistema linear homogêneo cujo número de incógnitas é maior do que o número de equações admite uma solução não-trivial.*

Demonstração. Consideremos o sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = 0 \end{cases}$$

de m equações e n incógnitas, onde $m < n$. Vamos provar o resultado por indução no número de equações do sistema.

Se tivermos apenas uma equação do tipo

$$a_{11}x_1 + \ldots + a_{1n}x_n = 0$$

com $n > 1$ incógnitas, devemos ter um dos dos coeficientes $a_{1i} \neq 0$ (caso contrário esta equação não faria sentido). Podemos supor então, sem perda de generalidade, que $a_{1n} \neq 0$. Isolando x_n na equação dada, temos

$$x_n = -\left(\frac{a_{11}}{a_{1n}}x_1 + \ldots + \frac{a_{1,n-1}}{a_{1n}}x_{n-1}\right).$$

Para obtermos uma solução não-trivial para a equação do sistema, basta escolhermos valores quaisquer para os $x_1, \ldots, x_{n-1}$ (que são variáveis livres) e obteremos x_n.

Para completar a indução, vamos supor que o lema seja verdadeiro para um sistema com $m - 1$ equações. Podemos primeiramente admitir que, no sistema original, temos $a_{mn} \neq 0$ (caso contrário, o sistema não teria m, mas $m-1$ equações). Então, a m-ésima equação pode ser reescrita como

$$x_n = -\left(\frac{a_{m1}}{a_{mn}}x_1 + \ldots + \frac{a_{m,n-1}}{a_{mn}}x_{n-1}\right).$$

Substituindo em cada uma das $m - 1$ primeiras equações a incógnita x_n por esta expressão, obtemos um sistema homogêneo de $m-1$ equações nas $n-1$ primeiras incógnitas. Pela hipótese de indução, este sistema admite uma solução não-trivial $(\alpha_1, \ldots, \alpha_{n-1})$, pois $n - 1 > m - 1$. Escrevendo então

$$\alpha_n = -\left(\frac{a_{m1}}{a_{mn}}\alpha_1 + \ldots + \frac{a_{m,n-1}}{a_{mn}}\alpha_{n-1}\right),$$

obtemos uma solução não-trivial $(\alpha_1, \ldots, \alpha_{n-1}, \alpha_n)$ do sistema proposto. □

Usando os lemas anteriores, podemos provar o seguinte resultado.

Teorema 11. *Se o conjunto finito de vetores* $\{v_1, \ldots, v_m\}$ *gera o espaço vetorial* E*, então qualquer conjunto com mais de* m *vetores em* E *é l.d.*

Demonstração. Dados os vetores $w_1, \ldots, w_n \in E$, com $n > m$, para cada $j = 1, \ldots, n$ podemos escrever

$$w_j = \alpha_{1j}v_1 + \ldots + \alpha_{mj}v_m,$$

pois os vetores v_j geram E. Para mostrar que os vetores w_j são l.d., devemos achar coeficientes $x_1, \dots, x_n$, com pelo menos um deles não-nulo, tais que $x_1 w_1 + \dots + x_n w_n = 0$. Substituindo os w_j por suas expressões em termos de v_j e reorganizando a soma, esta igualdade significa que

$$\left(\sum_{j=1}^{n} x_j \alpha_{1j}\right) v_1 + \left(\sum_{j=1}^{n} x_j \alpha_{2j}\right) v_2 + \dots + \left(\sum_{j=1}^{n} x_j \alpha_{mj}\right) v_m = 0.$$

Esta condição será satisfeita se todos os somatórios forem nulos, ou seja,

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1n}x_n & = 0 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \dots + \alpha_{2n}x_n & = 0 \\ \vdots & \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \dots + \alpha_{mn}x_n & = 0 \end{cases}$$

Tal solução existe pelo Lema 2, pois $n > m$. Portanto, w_j são l.d. e o teorema está provado. □

Corolário 2. *Se os vetores $v_1, \dots, v_m$ geram o espaço vetorial E e os vetores $u_1, \dots, u_n$ são l.i., então $n \leq m$.*

Corolário 3. *Se o espaço vetorial E admite uma base $\mathcal{B} = \{u_1, \dots, u_n\}$ com n elementos, então qualquer outra base de E possui também n elementos.*

Demonstração. Seja $\mathcal{B}'$ outra base de E com m elementos. Como $\mathcal{B}'$ gera E e $\mathcal{B}$ é l.i., temos que $n \leq m$, pelo Corolário 2. Como $\mathcal{B}$ gera E e $\mathcal{B}'$ é l.i., do mesmo Corolário segue-se que $m \leq n$. Logo, $m = n$. □

Definição 13. *Diz-se que o espaço vetorial E tem* dimensão finita *quando admite uma base $\mathcal{B}$ com um número finito n de elementos. Este número, que é o mesmo para todas as bases de E, chama-se* dimensão *do espaço vetorial E, $n = dim(E)$. Por extensão, diz-se que o espaço vetorial $E = \{0\}$ tem dimensão zero.*

Corolário 4. *Se a dimensão de E é n, um conjunto com n vetores gera E se e somente se é l.i.*

Demonstração. Se $X = \{v_1, \dots, v_n\}$ gera E e não é l.i., então um dos seus elementos é combinação dos $n - 1$ vetores restantes. Logo, estes $n - 1$ vetores restantes formariam também um conjunto de geradores de E, o que contradiz o Teorema 11, pois todas as bases de E devem ter n vetores linearmente independentes.

Reciprocamente, suponha que X seja l.i. Se X não gerasse E, existiria um vetor $v \in E$ que não seria combinação linear dos elementos de X. Então o conjunto $\{v_1, \ldots, v_n, v\}$ seria l.i., em contradição com o Teorema 11, pois uma base de E com n elementos gera todo o espaço E. □

Exemplo 4.9. *Em* $\mathbb{R}^2$*, duas possíveis bases são* $\{(1, 0), (0, 1)\}$ *e* $\{(1, 1), (0, 1)\}$*. Este espaço tem dimensão 2.*

Exemplo 4.10. *O espaço* $\mathbb{R}^n$ *tem dimensão n (basta pensar na base canônica de cada espaço destes, para* $n \in \mathbb{N}$*).*

Exemplo 4.11. *O espaço* $\mathcal{M}_{2\times 2}$ *tem dimensão 4.*

A seguir, demonstramos um resultado que garante que qualquer espaço vetorial (de dimensão finita) admite uma base.

Teorema 12. *Qualquer conjunto l.i. de um espaço vetorial E com dimensão finita pode ser completado para formar uma base.*

Demonstração. Seja $n =$dim(E) e $v_1, \ldots, v_r$ um conjunto l.i. Pelo Corolário 2, $r \leq n$. Se esse conjunto gera E, então ele é base e $n = r$. Suponha que não. Então existe um $v_{r+1} \in E$ tal que $v_{r+1} \notin \{v_1, \ldots, v_r\}$. Então v_{r+1} não pode ser combinação linear dos v_i, pois caso contrário os v_i seriam base para E. Logo, $\{v_1, \ldots, v_r, v_{r+1}\}$ é l.i. Se este conjunto gera E, terminamos. Senão, continuamos no mesmo procedimento até que uma base tenha sido encontrada. □

Corolário 5. *Seja E um espaço vetorial de dimensão finita. Se U é subespaço de E, então*

$$dim(U) \leq dim(E).$$

Demonstração. Para isto, basta pensarmos que uma base de U deve estar contida em E, e assim não pode ter mais elementos do que uma base de E. □

Corolário 6. *Se* $n =$*dim*(V)*, com* $V \subset W$ *e dim*$(W) = n$*, então* $V = W$*.*

Mais uma vez, o próximo Corolário é um resultado que pode parecer surpreendente, mas que mostra a ligação entre os espaços vetoriais e os sistemas lineares.

Corolário 7. *Seja* $A \in \mathbb{R}^{n\times n}$ *e suponha que as linhas de A formam um conjunto l.i. Então A é inversível.*

Demonstração. Seja para cada $i = 1, \ldots, n$ o vetor $\alpha_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ (a linha i da matriz A). Suponha que W é o subespaço gerado por estes vetores. Como as linhas são linearmente independentes por hipótese, a dimensão de W é n. Pelo Corolário 6, $W = \mathbb{R}^n$. Portanto, devem existir escalares $b_{ij} \in \mathbb{R}$, $1 \leq i, j \leq n$, tais que

$$\begin{aligned}
e_1 &= b_{11}\alpha_1 + b_{12}\alpha_2 + \ldots + b_{1n}\alpha_n \\
e_2 &= b_{21}\alpha_1 + b_{22}\alpha_2 + \ldots + b_{2n}\alpha_n \\
&\vdots \\
e_n &= b_{n1}\alpha_1 + b_{n2}\alpha_2 + \ldots + b_{nn}\alpha_n
\end{aligned}$$

em que cada e_i é o i-ésimo vetor canônico em $\mathbb{R}^n$. Portanto, se construirmos uma matriz com os b_{ij} teremos que

$$BA = I$$

ou seja, $B = A^{-1}$. □

Finalmente, temos o seguinte resultado.

Teorema 13. *Se U e W são subespaços de E (espaço vetorial com dimensão finita), então*

$$dim(U + W) = dim(U) + dim(W) - dim(U \cap W)$$

Demonstração. Note que $U \cap W$ e é subespaço de E. Portanto, deve ter dimensão finita e sua dimensão deve ser menor que a dimensão de E. Assim, sua base deve ser um subconjunto $\alpha_1, \ldots, \alpha_k$ da base de U e da base de W. Escreva isso então como

$$\begin{aligned}
U &= \text{span}[\alpha_1, \ldots, \alpha_k, \beta_{k+1}, \ldots, \beta_m] \\
W &= \text{span}[\alpha_1, \ldots, \alpha_k, \gamma_{k+1}, \ldots, \gamma_n]
\end{aligned}$$

Assim, o subespaço $U + W$ é gerado pelos vetores $\{\alpha_i, \beta_j, \gamma_p\}$, com $1 \leq i \leq k$, $k + 1 \leq j \leq m$, $k + 1 \leq p \leq n$. Agora, note que estes vetores formam um conjunto linearmente independente, pois caso contrário poderíamos escrever, por exemplo, $\sum x_i\alpha_i + \sum y_j\beta_j + \sum z_p\gamma_p = 0$ e assim teríamos

$$-\sum y_j\beta_j = \sum x_i\alpha_i + \sum z_p\gamma_p$$

o que implicaria que $\beta_j \in W$. Como os $\beta_j \in U$, isso implicaria que poderíamos escrever

$$\sum y_j\beta_j = \sum c_i\alpha_i$$

para escalares c_i. Mas, como a base de U é linearmente independente, cada um dos escalares y_j deve ser igual a zero; logo,

$$0 = -\sum y_j\beta_j = \sum x_i\alpha_i + \sum z_p\gamma_p$$

e como $\{\alpha_i, \gamma_q\}$ também é l.i., todos os x_i e todos os z_p devem ser iguais a zero. Portanto, $\{\alpha_i, \beta_j, \gamma_p\}$ formam uma base para $U + W$.

Além disso,

$$\dim(U) + \dim(W) = m + n = k + (m + n - k) = \dim(U \cap W) + \dim(U + W).$$

□

Exemplo. *Seja P o espaço dos polinômios em $\mathbb{R}$ (de qualquer grau). Então os vetores deste espaço tem a forma*

$$f(x) = a_0 + a_1x + \ldots + a_nx^n$$

Seja $f_k(x) = x^k$, $k = 0, 1, \ldots$. O conjunto infinito $\{f_0, f_1, \ldots\}$ é uma base para V. Claramente este conjunto gera V pois as funções f são, como acima,

$$f(x) = a_0f_0 + a_1f_1 + \ldots + a_nf_n$$

Precisamos mostrar então que as funções na base são l.i. Mas para isto basta mostrarmos que todo subconjunto finito deste conjunto infinito é l.i. (para qualquer coleção finita de vetores neste conjunto). Tome então os conjuntos $\{f_0, \ldots, f_n\}$ para cada $n \in \mathbb{N}$. Suponha então que

$$a_0f_0 + a_1f_1 + \ldots + a_nf_n = 0$$

Isto é o mesmo que dizer que

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n = 0$$

para cada $x \in \mathbb{R}$; em outras palavras, todo $x \in \mathbb{R}$ seria uma raiz deste polinômio. Mas sabe-se que um polinômio de grau n não pode ter mais do que n raizes distintas; portanto, os coeficientes $a_0 = a_1 = \ldots = a_n = 0$.

Mostramos então que existe uma base infinita para P. Isto implica que ele não tem dimensão finita, pelo teorema anterior que dizia que um conjunto l.i. não pode ter mais elementos do que a dimensão do espaço.

Uma última observação é que uma base infinita não requer combinações lineares infinitas; de fato, cada vetor no espaço pode ser obtido com uma combinação linear finita *dos elementos da base. Basta vermos que $\sum_{k=0}^{\infty} a_k x^k$ não está neste espaço.*

2.4.2 Coordenadas e Mudança de Base

Para determinarmos as coordenadas de um vetor (isto é, os coeficientes da combinação linear dos elementos da base que o definem) precisamos fazer isso numa certa ordem.

Exemplo. $(2,3) = 2(1,0) + 3(0,1)$. *Mas se* $\overline{e_1} = (0,1)$ *e* $\overline{e_2} = (1,0)$*, então as coordenadas mudam para*

$$(2,3) = 3\overline{e_1} + 2\overline{e_2} = (3,2)_{\textit{(nova base)}}$$

Definição 14. *Se E é um espaço vetorial de dimensão finita, uma base ordenada de E é uma* sequência *finita de vetores que é l.i. e que gera E.*

Desta forma, dada uma base ordenada $\mathcal{B} = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ de V, então dado $v \in V$, existe uma única n-tupla de escalares x_i tais que

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

Frequentemente, ao invés de trabalharmos com as coordenadas de um vetor, vamos precisar trabalhar com a matriz de v relativa à base ordenada $\mathcal{B}$:

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}$$

Isto é útil pois vamos tentar descrever o que acontece quando mudamos de base.

Teorema 14. *Seja E um espaço vetorial de dimensão n e sejam $\mathcal{B}$ e $\mathcal{B}'$ duas bases ordenadas de E. Então existe uma matriz única P, inversível e $n \times n$, com entradas tais que*

$$[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'} \quad e \quad [v]_{\mathcal{B}'} = P^{-1}[v]_{\mathcal{B}},$$

para todo $v \in E$. As colunas de P são dadas por

$$P_j = [\alpha'_j]_{\mathcal{B}}, \quad j = 1, \ldots, n.$$

Demonstração. Sejam as bases

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\} \text{ e } \mathcal{B}' = \{\alpha'_1, \ldots, \alpha'_n\}$$

Então, existe um conjunto único de escalares P_{ij} tais que

$$\alpha'_j = \sum_{i=1}^{n} P_{ij}\alpha_i, 1 \leq j \leq n.$$

Sejam agora $x_1, \ldots, x_n$ as coordenadas de um vetor v na base ordenada $\mathcal{B}$ e $x'_1, \ldots, x'_n$ as coordenadas do mesmo vetor v na base ordenada $\mathcal{B}'$. Então,

$$\begin{aligned} v = x'_1\alpha'_1 + \ldots + x'_n\alpha'_n &= \sum_{j=1}^{n} x'_j\alpha'_j \\ &= \sum_{j=1}^{n} x'_j \sum_{i=1}^{n} P_{ij}\alpha_i \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} (P_{ij}x'_j)\alpha_i \\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} P_{ij}x'_j \right) \alpha_i \end{aligned}$$

Como as coordenadas são unicamente determinadas para cada base, isso implica que

$$x_i = \sum_{j=1}^{n} P_{ij}x'_j, 1 \leq i \leq n$$

Seja então P a matriz formada pelos P_{ij} e X e X' as matrizes coordenadas do vetor v nas bases $\mathcal{B}$ e $\mathcal{B}'$, respectivamente. Então,

$$X = PX'.$$

Como as duas bases são linearmente independentes, $X = 0$ se e somente se $X' = 0$. Logo, segue de um teorema anterior que P é inversível; ou seja

$$X' = P^{-1}X.$$

Em outras palavras,
$$[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'} \text{ e } [v]_{\mathcal{B}'} = P^{-1}[v]_{\mathcal{B}}$$

□

Isto quer dizer que para construirmos P que leva um vetor descrito na base $\mathcal{B}'$ em sua descrição na base $\mathcal{B}$, devemos escrever cada vetor da base $\mathcal{B}'$ em suas coordenadas na base B. Podemos denotar também
$$P = I_{\mathcal{B}}^{\mathcal{B}'}.$$

Finalmente, podemos mostrar o seguinte:

Teorema 15. *Seja P uma matriz $n \times n$ inversível, e seja V um espaço n-dimensional definido no mesmo corpo; além disso, seja $\mathcal{B}$ uma base ordenada de V. Então existe uma única base ordenada $\mathcal{B}'$ de V tal que P é a matriz de mudança de base de $\mathcal{B}'$ para $\mathcal{B}$, ou seja,*
$$[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'} \textit{ e } [v]_{\mathcal{B}'} = P^{-1}[v]_{\mathcal{B}}$$
para qualquer vetor $v \in V$.

Demonstração. Seja $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$. Se $\mathcal{B}' = \{\alpha'_1, \ldots, \alpha'_n\}$ for uma base ordenada de V para a qual a primeira igualdade é válida, então devemos ter
$$\alpha'_j = \sum_{i=1}^{n} P_{ij}\alpha_i.$$
Logo, precisamos somente mostrar que estes vetores α'_j formam uma base de V. Mas:
$$\begin{aligned}
\sum_j P^{-1}_{jk}\alpha'_j &= \sum_j P^{-1}_{jk} \sum_i P_{ij}\alpha_i \\
&= \sum_j \sum_i P_{ij}P^{-1}_{jk}\alpha_i \\
&= \alpha_k
\end{aligned}$$
Logo, o subespaço gerado por $\mathcal{B}'$ contém $\mathcal{B}$ e é portanto igual a V. Logo, $\mathcal{B}'$ é base; logo é claro que as duas afirmações são verdadeiras. □

Exemplo. *Seja* $\mathcal{B} = \{(1,0),(0,1)\}$ *e* $\mathcal{B}' = \{(1,1),(1,0)\}$ *bases de* $\mathbb{R}^2$. *Então*

$$\begin{aligned}(1,0) &= 0(1,1) + 1(1,0)\\ (0,1) &= 1(1,1) - 1(1,0)\end{aligned}$$

Logo,

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}_{\mathcal{B}'} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}_{\mathcal{B}}$$

Assim,

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix}_{\mathcal{B}'} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix}_{\mathcal{B}} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}_{\mathcal{B}}.$$

Exemplo. *Ainda no* $\mathbb{R}^2$, *considere* $\mathcal{B} = \{(1,0),(0,1)\}$, $\mathcal{B}' = \{(2,3),(-1,2)\}$. *Para construirmos* $I_{\mathcal{B}}^{\mathcal{B}'}$, *escrevemos cada vetor da base* $\mathcal{B}'$ *na base* $\mathcal{B}$*:*

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

Logo,

$$I_{\mathcal{B}}^{\mathcal{B}'} = \begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix}$$

Note que, se $v = (1,1)_{\mathcal{B}'}$, *então*

$$v = 1(2,3) + 1(-1,2) = (1,5)_{\mathcal{B}}.$$

De fato:

$$I_{\mathcal{B}}^{\mathcal{B}'} v_{\mathcal{B}'} = \begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}.$$

Por outro lado,

$$\begin{aligned}(1,0) &= \frac{2}{7}(2,3) - \frac{3}{7}(-1,2)\\ (0,1) &= \frac{1}{7}(2,3) + \frac{2}{7}(-1,2)\end{aligned}$$

Assim,

$$I_{\mathcal{B}'}^{\mathcal{B}} = \begin{pmatrix} 2/7 & 1/7 \\ 3/7 & 2/7 \end{pmatrix}.$$

De fato,

$$I^{\mathcal{B}}_{\mathcal{B}'} v_{\mathcal{B}} = \begin{pmatrix} 2/7 & 1/7 \\ -3/7 & 2/7 \end{pmatrix} \begin{pmatrix} 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\mathcal{B}}.$$

Ainda:

$$I^{\mathcal{B}}_{\mathcal{B}'} I^{\mathcal{B}'}_{\mathcal{B}} = I.$$

Exemplo. *Se $\mathcal{B} = \{(1,2),(3,5)\}$ e $\mathcal{B}' = \{(1,-1),(1,-2)\}$, para encontrarmos $I^{\mathcal{B}'}_{\mathcal{B}}$ devemos escrever os elementos de $\mathcal{B}'$ na base $\mathcal{B}$. Mas isso pode ser difícil. Considere então a base canônica em $\mathbb{R}^2$. Então:*

$$I^{\mathcal{B}'}_{C} = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}.$$

Além disso,

$$\begin{aligned} (1,0) &= -5(1,2) + 2(3,5) \\ (0,1) &= 3(1,2) - 1(3,5) \end{aligned}$$

e assim

$$I^{C}_{\mathcal{B}} = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix}.$$

Então:

$$I^{\mathcal{B}'}_{\mathcal{B}} = I^{C}_{\mathcal{B}} I^{\mathcal{B}'}_{C} = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} -8 & -11 \\ 3 & 4 \end{pmatrix}.$$

Note que

$$I^{\mathcal{B}'}_{\mathcal{B}} v_{\mathcal{B}'} = \begin{pmatrix} -8 & -11 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\mathcal{B}'} = \begin{pmatrix} -19 \\ 7 \end{pmatrix}$$

De fato,

$$\begin{aligned} (1,1)_{\mathcal{B}'} &= (1,-1) + (1,-2) = (2,3)_C \\ (-19,7)_{\mathcal{B}} &= -19(1,2) + 7(3,5) = (2,-3)_C. \end{aligned}$$

Exemplo. *Em $\mathbb{R}^3$, se consideramos as bases*

$$\begin{aligned} E &= \{(1,0,0),(0,1,0),(0,0,1)\} \\ S &= \{(1,0,1),(2,1,2),(1,2,2)\} \end{aligned}$$

temos que

$$I^{S}_{E} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix}$$

e

$$I_S^E = \begin{pmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ 1 & 0 & 1 \end{pmatrix}$$

Assim, se $v = (1, 1, 1)_E$*, temos*

$$I_S^E v_E = \begin{pmatrix} -2 & -2 & 3 \\ 2 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}_E = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}_S .$$

Exemplo. *Seja* $\theta \in \mathbb{R}$*; a matriz*

$$P = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

é inversível com inversa

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

Logo, para cada θ*, o conjunto* $\mathcal{B}'$ *formado pelos vetores* $(\cos\theta, \sin\theta)$*,* $(-\sin\theta, \cos\theta)$ *é uma base de* $\mathbb{R}^2$*. Intuitivamente esta base é obtida ao rotacionarmos a base canônica num ângulo* θ*. Se* $\alpha = (x_1, x_2)$*, então*

$$[\alpha]_{\mathcal{B}'} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

ou ainda

$$\begin{aligned} x_1' &= x_1 \cos\theta + x_2 \sin\theta \\ x_2' &= -x_1 \sin\theta + x_2 \cos\theta \end{aligned}$$

CAPÍTULO 3

Transformações Lineares

3.1 Introdução

Considere uma matriz $A \in \mathbb{R}^{m \times n}$. Quando efetuamos um produto entre uma matriz A e um vetor x, estamos levando o espaço $\mathbb{R}^n$ no espaço $\mathbb{R}^m$:

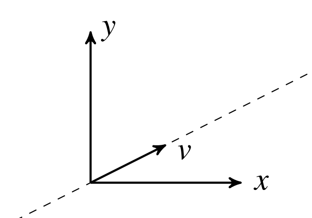
$$\begin{aligned} A : \mathbb{R}^n &\to \mathbb{R}^m \\ x &\mapsto Ax = b \end{aligned}$$

Portanto, podemos encarar uma matriz como uma transformação no espaço.

Exemplo 1.1.

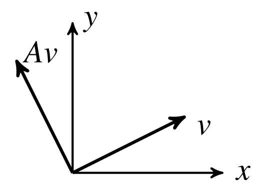
$$A = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}$$

Esta transformação (às vezes chamada de homotetia*) não movimenta um vetor no plano, mas altera seu tamanho. Se $c \in \mathbb{R}$, e $v = (x, y)$, $Av = (cx, cy)$ está sobre a reta que passa pelo vetor v.*



Exemplo 1.2.

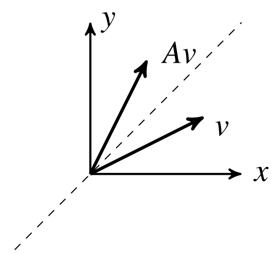
$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$



Esta matriz rotaciona qualquer vetor por 90 graus no sentido anti-horário: $A(1,0) = (0,1)$, $A(0,1) = (-1,0)$.

Exemplo 1.3.

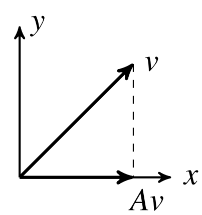
$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$



Esta matriz reflete todo vetor no eixo de simetria de 45 graus.

Exemplo 1.4.

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



Esta matriz toma qualquer vetor do espaço e projeta no subespaço definido por $x_2 = 0$*, ou seja, no eixo* x_1*. Este eixo é o espaço coluna de A, enquanto que seu espaço nulo é o eixo* $x_1 = 0$.

É importante notar, no entanto, que algumas transformações não podem ser realizadas através de matrizes:

(i) É impossível mover a origem, já que $A0 = 0$.

(ii) Se $Ax = x'$, então $A(2x) = 2x'$, ou seja, $A(cx) = cAx$.

(iii) Se $Ax = x'$ e $Ay = y'$, então $A(x+y) = x' + y'$, ou seja, $A(x+y) = Ax + Ay$.

Estas regras vem da definição da multiplicação entre matrizes, e definem o que chamamos de *transformação linear*.

Definição 15. *Sejam E, F espaços vetoriais sobre um mesmo corpo $\mathbb{K}$. Uma transformação linear T é uma função $T : E \to F$ que associa a cada $u \in E$ um vetor $v = T(u) \in F$ e que satisfaz a condição seguinte: para todos $u, w \in E$ e $\alpha \in \mathbb{K}$ temos que*

$$T(\alpha u + w) = \alpha T(u) + T(w).$$

Para todo número $c, d \in \mathbb{R}$ e vetores $x, y \in \mathbb{R}^n$, a multiplicação de matrizes satisfaz a regra da linearidade

$$A(cx + dy) = c(Ax) + d(Ay).$$

Toda transformação que satisfaz esta propriedade é uma transformação linear. Portanto, toda matriz define uma transformação linear. Mas, será que toda transformação linear leva a uma matriz? Veremos que, em espaços de dimensão finita, isso é verdadeiro.

Tome como exemplo o espaço $\mathcal{P}_n$, polinômios de grau $\leq n$. Este espaço tem dimensão $n + 1$.

Exemplo 1.5. *A diferenciação, d/dx, é uma transformação linear:*

$$\begin{aligned} Ap &= \frac{d}{dx} p(x) \\ &= \frac{d}{dx}(a_0 + a_1 x + \ldots + a_n x^n) \\ &= a_1 + 2a_2 x + \ldots + n a_n x^{n-1}. \end{aligned}$$

Exemplo 1.6. *A integração de 0 a x também é linear (leva $\mathcal{P}_n$ a $\mathcal{P}_{n+1}$):*

$$\begin{aligned} Ap &= \int_0^x p(x) dx \\ &= \int_0^x (a_0 + a_1 x + \ldots + a_n x^n) dx \\ &= a_0 x + \ldots + \frac{a_n}{n+1} x^{n+1}. \end{aligned}$$

A definição abaixo nos dá uma ideia de onde essas transformações lineares vivem e das relações entre elas.

Definição 16. *Seja $\mathcal{L}(E; F)$ o conjunto das transformações lineares de E em F. Então, $\mathcal{L}(E; F)$ é um espaço vetorial. As transformações lineares $T : E \to E$ são chamadas* operadores lineares *em E. Por sua vez, as transformações lineares $\varphi : E \to \mathbb{R}$, com valores numéricos, são chamadas* funcionais *lineares. Escreve-se E^* em vez de $\mathcal{L}(E; \mathbb{R})$ e o conjunto E^* dos funcionais lineares $\varphi : E \to \mathbb{R}$ chama-se* espaço vetorial dual *de E.*

3.2 Matriz de uma transformação linear

A linearidade é importante pois nos dá uma propriedade crucial: se conhecermos a ação de uma transformação em todos os vetores da base, conhecemos a ação da transformação em todos os vetores do espaço gerado por esta base, visto que cada vetor do espaço é apenas combinação linear de todos os vetores da base. Depois que sabemos a ação de uma transformação na base, não há mais graus de liberdade possíveis: a transformação fica inteiramente determinada.

Exemplo 2.1. *Que transformação linear leva*

$$x_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ em } Ax_1 = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \text{ e}$$

$$x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ em } Ax_2 = \begin{pmatrix} 4 \\ 6 \\ 8 \end{pmatrix}?$$

A resposta deve ser a multiplicação pela matriz

$$A = \begin{pmatrix} 2 & 4 \\ 3 & 6 \\ 4 & 8 \end{pmatrix}.$$

Agora, considere outro problema: encontrar a matriz que representa a diferenciação, e a matriz que representa a integração em um intervalo. Basta, para isto, definirmos uma base para o espaço onde as transformações serão aplicadas. Para os polinômios de grau ≤ 3, cuja dimensão é 4, existe uma base natural que é a base dos monômios,

$$p_1 = 1, \quad p_2 = x, \quad p_3 = x^2, \quad p_4 = x^3.$$

Esta base não é única, mas é bastante conveniente. Agora, vamos olhar para o efeito da diferenciação nesta base: sejam $q_1 = 1$, $q_2 = x$, $q_3 = x^2$ em $\mathcal{P}_2$. Então

$$\begin{aligned} &Ap_1 = 0, \quad Ap_2 = 1 = q_1, \\ &Ap_3 = 2x = 2q_2, \quad Ap_4 = 3x^2 = 3q_3. \end{aligned}$$

Note que

$$A : \mathcal{P}_3 \to \mathcal{P}_2, \text{ tal que } A_{\text{diff}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

A derivada de qualquer outra combinação de polinômios é uma combinação linear dos membros da base, e portanto aplicar a matriz da transformação a este polinômio é equivalente a aplicá-la aos vetores da base e combinar os resultados.

Suponha que os vetores $p_1, \ldots, p_n$ são base para o espaço V, e $q_1, \ldots, q_m$ formam uma base para o espaço W. Então, cada transformação linear A de V para W é representada por uma matriz. A j-ésima coluna é encontrada ao aplicarmos A ao j-ésimo vetor da base de W; o resultado Ap_j é combinação dos q e os coeficientes desta combinação vão na coluna j de A:

$$Ap_j = a_{1,j}q_1 + \ldots + a_{m,j}q_m.$$

Para a matriz de diferenciação, a coluna 1 veio de p_1: sua derivada era zero, então a primeira coluna era nula. A última coluna veio de x^3: a derivada era $3x^2$, e assim o coeficiente 3 está na linha correspondente a $x^2 = p_3$.

Fazendo a mesma coisa para a integração, que vai do espaço $\mathcal{P}_2$ no espaço $\mathcal{P}_3$. Devemos então escolher uma base para $\mathcal{P}_3$. A base mais natural para este espaço é justamente $q_1 = 1, q_2 = x, q_3 = x^2, q_4 = x^3$. Desta forma, teremos, para cada vetor da base do espaço $\mathcal{P}_3$:

$$\begin{aligned}
&Ap_1 = x = q_2, \quad Ap_2 = \frac{1}{2}x^2 = \frac{1}{2}q_3, \\
&Ap_3 = \frac{1}{3}x^3 = \frac{1}{3}q_4.
\end{aligned}$$

Assim, a matriz que representa esta transformação linear é

$$A_{\text{int}} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

Observação: Enxergamos a integração como a operação inversa da diferenciação, isto é, a integração seguida da diferenciação nos dão o resultado original de volta. Se tomarmos a matriz de diferenciação na base das cúbicas, que é uma matriz 3×4, teremos:

$$A_{\text{diff}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \text{ e } A_{\text{diff}}A_{\text{int}} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}.$$

Portanto, o operador de diferenciação nesta base é a inversa à esquerda da integração. Como para matrizes retangulares é impossível termos inversas dos dois lados, o produto $A_{\text{int}}A_{\text{diff}}$ não pode ser igual à identidade; mas isto não ocorre de qualquer forma, já que a derivada de uma constante é zero, e a integral de zero nunca pode trazer esta constante de volta (a primeira linha do produto citado é nula).

As transformações lineares aqui representadas (e todas as outras) tem representação matricial, mas uma transformação linear não é uma matriz: uma matriz representa a transformação em uma base dada. Portanto, a matriz usada para representarmos uma transformação linear varia de acordo com a base escolhida para o espaço.

3.2.1 Matriz de uma transformação linear em relação a uma base do domínio e a uma base do contradomínio.

Uma transformação linear $T : E \to F$ é uma função especial que é linear, e vai do espaço vetorial E no espaço vetorial F. Em geral, para definirmos uma função, precisamos definir o valor de $f(x)$ para todo x no domínio de f. No caso das transformações lineares, é bem mais fácil definirmos esta função, pois basta fazê-lo em cada elemento da base.

Sejam E, F espaços vetoriais de dimensão finita e $\beta = \{u_1, \ldots, u_n\}$ uma base de E. Todo vetor $v \in E$ se exprime, de modo único, como combinação linear

$$v = x_1u_1 + \ldots + x_nu_n$$

de elementos da base β. Para qualquer transformação linear $T : E \to F$ e para qualquer vetor $v \in E$, temos que

$$\begin{aligned} T(v) &= T(x_1u_1 + \ldots + x_nu_n) \\ &= x_1T(u_1) + \ldots + x_nT(u_n), \end{aligned}$$

ou seja, a definição de T depende apenas da aplicação da transformação linear nos vetores $u_i \in \beta$.

Como consequência, se quisermos definir uma transformação linear $T : E \to F$, basta escolhermos uma base γ para F, e para cada $j = 1, \ldots, n$ um vetor

$$v_j = (a_{1j}, a_{2j}, \ldots, a_{mj})_\gamma \in F$$

e dizer que $v_j = T(u_j)$ é a imagem do j-ésimo vetor da base β pela transformação linear T. A partir daí, fica determinada a imagem $T(v)$ de qualquer vetor $v \in E$, pois se

$$v = x_1u_1 + \ldots + x_nu_n = (x_1, \ldots, x_n)_\beta,$$

então

$$
\begin{aligned}
y = T(v) = T\left(\sum_{j=1}^{n} x_j u_j\right) \\
&= \sum_{j=1}^{n} x_j T(u_j) \\
&= \sum_{j=1}^{n} \left(a_{1j}x_j, a_{2j}x_j, \ldots, a_{mj}x_j\right)_\gamma \\
&= \left(\sum_{j=1}^{n} a_{1j}x_j, \sum a_{2j}x_j, \ldots, \sum a_{mj}x_j\right)_\gamma,
\end{aligned}
$$

ou seja, cada coordenada y_i do vetor y na base γ é calculada como

$$
\begin{aligned}
y_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\
y_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\
&\vdots \\
y_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n
\end{aligned}
$$

Portanto, toda transformação linear $T : E \to F$ fica inteiramente determinada por uma matriz $A \in \mathbb{K}^{m\times n}$. Os vetores-coluna desta matriz são as coordenadas dos vetores v_j na base γ, em que cada $v_j = T(u_j)$, ou seja, imagem do vetor u_j da base β de E. A imagem de $T(v)$ é o vetor $(y_1, \ldots, y_m)_\gamma \in F$ cujas coordenadas são dadas pelas equações acima.

O resultado a seguir será útil mais à frente.

Teorema 16. *Para toda matriz $A \in \mathcal{M}(m\times n)$, o número de linhas l.i. de A é igual ao número de colunas l.i. de A.*

Demonstração. Seja p o número de colunas l.i. da matriz A. Então existem p vetores

$$
w_k = \begin{pmatrix} w_{1k} \\ \vdots \\ w_{mk} \end{pmatrix} \in \mathbb{R}^m
$$

tais que cada uma das colunas

$$
a_{(:,j)} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}, 1 \le j \le n
$$

de A é combinação linear dos $w_1, \ldots, w_p$:

$$a_{(:,j)} = \sum_{k=1}^{p} c_{kj} w_k, \qquad 1 \leq j \leq n.$$

Tomando a i-ésima coordenada de cada um dos membros desta equação, vemos que

$$a_{ij} = \sum_{k=1}^{p} c_{kj} w_{ik} = \sum_{k=1}^{p} w_{ik} c_{kj}, \tag{3.1}$$

para quaisquer i, j, com $1 \leq i \leq m$ e $1 \leq j \leq n$. Considerando agora os vetores-linha $a_{(i,:)} = (a_{i1}, \ldots, a_{in})$ da matriz A, juntamente com os vetores $c_k = (c_{k1}, \ldots, c_{kn})$, $1 \leq k \leq p$, observamos que a igualdade entre o primeiro e o terceiro membros de (3.1) significa que, para todo $i = 1, \ldots, m$ tem-se

$$a_{(i,:)} = \sum_{k=1}^{p} w_{ik} c_k, \qquad 1 \leq i \leq m.$$

Assim, os vetores linha de A são combinações lineares de $c_1, \ldots, c_p$, portanto o número de linhas l.i. de A é $\leq p$. Aplicando este resultado à matriz A^T, que tem como linhas as colunas de A, concluimos que o o número de colunas l.i. de A é menor ou igual ao número de linhas l.i. de A e assim temos o resultado completo. □

3.3 Rotações, projeções e reflexões.

Já vimos que rotações de 90 graus, projeções no eixo x e reflexões com relação à linha de 45 graus em $\mathbb{R}^2$ eram representadas por matrizes simples da forma

$$Q = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

respectivamente. Mas é lógico pensar que rotações com outros ângulos, reflexões e projeções em outros eixos serão igualmente simples, já que estas são todas transformações lineares que passam pela origem ($A0 = 0$). Assim, vamos primeiramente considerar o plano e seus vetores básicos $(1, 0), (0, 1)$ para tentarmos visualizar estas transformações em geral.

Rotação. Na Figura 3.1, mostramos a rotação por um ângulo θ no sentido anti-horário. Podemos ver o efeito desta rotação nos dois vetores básicos, onde $c = \cos\theta$ e $s = \sin\theta$. A rotação aplicada ao primeiro vetor da base resulta em $(\cos\theta, \sin\theta)$, cujo comprimento é ainda 1. Se aplicada ao segundo vetor, a rotação resulta em $(-\sin\theta, \cos\theta)$. Assim, pela definição da transformação linear que vimos antes, estas são as colunas da matriz da transformação, e assim temos Q_θ. Com esta transformação, podemos exemplificar o comportamento das trans-

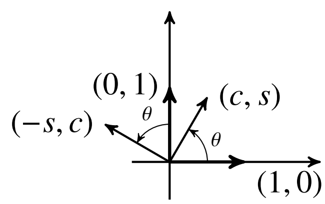


Figura 3.1: Rotação por um ângulo θ.

formações enquanto matrizes.

(i) A inversa de Q_θ equivale à rotação por θ no sentido contrário:

$$Q_\theta Q_{-\theta} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(ii) O quadrado de Q_θ (ou seja, a aplicação desta matriz duas vezes sobre o mesmo vetor) equivale à rotação por um ângulo 2θ:

$$\begin{aligned} Q_\theta^2 &= \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \\ &= \begin{pmatrix} c^2 - s^2 & -2cs \\ 2cs & c^2 - s^2 \end{pmatrix} \\ &= \begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}. \end{aligned}$$

(iii) O produto de Q_θ e Q_φ é equivalente à rotação por $\theta + \varphi$:

$$\begin{aligned}
Q_\theta Q_\varphi &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta\cos\varphi - \sin\theta\sin\varphi & -\cos\theta\sin\varphi - \sin\theta\cos\varphi \\ \cos\theta\sin\varphi + \sin\theta\cos\varphi & -\sin\theta\sin\varphi + \cos\theta\cos\varphi \end{pmatrix} \\
&= \begin{pmatrix} \cos(\theta+\varphi) & -\sin(\theta+\varphi) \\ \sin(\theta+\varphi) & \cos(\theta+\varphi) \end{pmatrix} = Q_{\theta+\varphi}.
\end{aligned}$$

Estas propriedades não acontecem por acidente: a multiplicação de matrizes é definida para que estas propriedades sejam verdadeiras para as transformações lineares: o produto das matrizes das transformações lineares é o produto das transformações, ou mais exatamente, a composição das transformações.

Projeção Na Figura 3.2 representamos a projeção dos vetores básicos na reta determinada pelo ângulo θ. Note que o comprimento da projeção é $c = \cos\theta$. As-

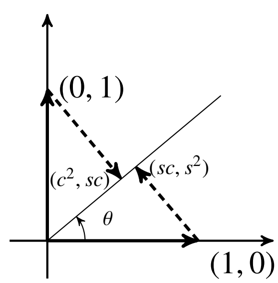


Figura 3.2: Projeção ortogonal na reta que faz ângulo θ com o eixo x.

sim, o ponto de projeção de $(1, 0)$ é exatamente $c(c, s)$. Similarmente, a projeção de $(0, 1)$ é $s(c, s)$.

A matriz desta transformação não pode possuir inversa, pois a projeção não pode ser desfeita. Pontos do tipo $(-s, c)$ são projetados na origem (voltaremos a este fato mais tarde). Ao mesmo tempo, todos os pontos na linha θ são projetados neles mesmos! Em outras palavras, projetar duas vezes é o mesmo que projetar uma vez, e então $P^2 = P$:

$$P^2 = \begin{pmatrix} c^2 & cs \\ cs & s^2 \end{pmatrix}^2 = \begin{pmatrix} c^2(c^2+s^2) & cs(c^2+s^2) \\ cs(c^2+s^2) & s^2(c^2+s^2) \end{pmatrix}.$$

Reflexão Na Figura 3.3, vemos a reflexão de $(1,0)$ na reta que faz ângulo θ com o eixo x. O comprimento da reflexão é igual ao comprimento do vetor original, assim como na rotação. No entanto, estas transformações são bem diferentes: aqui, a reta definida por θ continua a mesma, e os pontos são apenas refletidos, como que por um espelho.

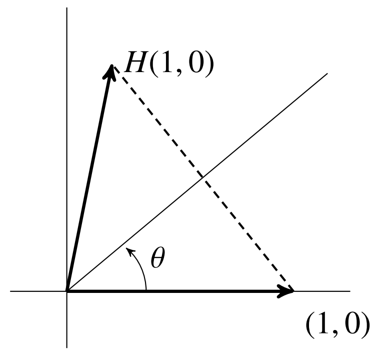


Figura 3.3: Reflexão em torno do eixo que faz ângulo θ com o eixo x.

A matriz desta transformação tem uma propriedade especial: $H^2 = I$, ou seja, duas reflexões trazem de volta o original. Assim, uma reflexão é sua própria inversa. Para vermos isso em termos matriciais, note primeiramente que $H = 2P - I$. Assim, $Hx + x = 2Px$, ou seja, o elemento refletido mais o original é igual a duas vezes sua projeção. Portanto,

$$H^2 = (2P - I)^2 = 4P^2 - 4P + I = I,$$

já que toda projeção satisfaz $P^2 = P$.

3.4 Produto de transformações lineares

Definição 17. *Suponha que T e S são transformações lineares tais que*

$$T : V \to W, \qquad S : U \to V.$$

Então,

$$\begin{aligned}(T \circ S) :\, & U \to V \to W \\ & u \mapsto S(u) \mapsto T(S(u)).\end{aligned}$$

Logo, a composição $(T \circ S) : U \to W$ também é uma transformação linear.

Observação Sejam $P, Q : \mathbb{R}^2 \to \mathbb{R}^2$ projeções ortogonais sobre duas retas do plano, uma das quais é perpendicular à outra. Todo vetor $v \in \mathbb{R}^2$ é a diagonal de um retângulo que tem Pv e Qv como lados. Segue-se então que $v = Pv + Qv$ para todo $v \in \mathbb{R}^2$, ou seja, $P+Q = I$. Logo, $Q = I-P$ e portanto $PQ = P(I-P) = P-P^2$. Como sabemos que a projeção satisfaz $P^2 = P$, temos que $PQ = 0$ mesmo com $P, Q \neq 0$.

Definição 18. *Um operador A chama-se* nilpotente *quando, para algum $n \in \mathbb{N}$, tem-se $A^n = 0$.*

Exemplo 4.1. *O operador de derivação que vimos anteriormente é nilpotente, pois se ele age de $\mathcal{P}_n$ em $\mathcal{P}_{n-1}$, então $D^{n+1}p = 0$ para todo p, ou seja, $D^{n+1} = 0$.*

3.5 Os Espaços Fundamentais

Imagem

Agora, vamos considerar o exemplo de um sistema de equações lineares de três equações e duas incógnitas:

$$\begin{pmatrix} 1 & 0 \\ 5 & 4 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

Como nesse caso temos mais equações que incógnitas, é bastante provável que o sistema não tenha solução. Um sistema com $m > n$ poderá ser resolvido se o lado direito do sistema (aqui, b) estiver contido em um subespaço vetorial especial: *o sistema $Ax = b$ poderá ser resolvido se e somente se o vetor b puder ser escrito como combinação linear das colunas de A.*

Para ver isto, note que o sistema pode ser reescrito da seguinte forma:

$$u \begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix} + v \begin{pmatrix} 0 \\ 4 \\ 4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

Assim, note que o subconjunto dos vetores que podem ser gerados como lado direito do sistema é exatamente o conjunto de todas as combinações lineares das colunas de A. Portanto, $Ax = b$ pode ser resolvido se e somente se b estiver contido no plano que é gerado pelos dois vetores coluna de A. Se b não estiver contido neste plano, então ele não pode ser obtido como combinação linear das colunas de b. Neste caso, o sistema não tem solução.

Se A representa uma transformação linear, este plano definido pelas colunas de A, é um subespaço importante chamado *espaço coluna de A*, ou *imagem* de A (ou ainda, imagem da transformação representada pela matriz A). Por um lado, se $A = 0$ então o espaço coluna de A será formado apenas pelo vetor $b = 0$; no outro extremo, qualquer matriz inversível em $\mathbb{R}^{n\times n}$ terá todo o espaço $\mathbb{R}^n$ como espaço coluna (qualquer lado direito em $\mathbb{R}^n$ define uma solução para o sistema).

Definição 19. *Seja $T : E \to F$ uma transformação linear definida entre espaços vetoriais E, F sobre um mesmo corpo $\mathbb{K}$. Então a* imagem *de T é um subespaço vetorial definido como*

$$\mathcal{I}m(T) = \{y \in F \,|\, y = T(x), \text{ para algum } x \in E\}.$$

Verificamos se esse subconjunto é realmente um subespaço:

(i) Suponha que $y_1, y_2 \in \mathcal{I}m(T)$. Então, existem $x_1, x_2 \in E$ tais que $T(x_1) = y_1$ e $T(x_2) = y_2$. Logo, podemos escrever

$$y_1 + y_2 = T(x_1) + T(x_2) = T(x_1 + x_2),$$

onde a última igualdade é válida pois T é linear. Assim, $y_1 + y_2$ pode ser escrito como $T(w)$, com $w = x_1 + x_2 \in E$ (pois E é espaço vetorial), e assim também pertence à imagem de T.

(ii) Seja $\alpha \in \mathbb{K}$ e $y \in \mathcal{I}m(T)$. Então, $\alpha y = \alpha T(x)$ para algum $x \in E$. Isto implica que $\alpha y = T(\alpha x)$, (já que T é linear) e portanto, αy também pertence à imagem de T.

Portanto, a imagem é, de fato, um subespaço linear.

Definição 20. *Seja $T : E \to F$ uma transformação linear como acima. Então chamamos a dimensão da imagem de T de* posto *de T (ou* rank *de T).*

Se a transformação linear T tem representação como uma matriz A com relação a uma base de E e uma base de F, então podemos falar da imagem de A como sendo

$$\mathcal{I}m(A) = \{b \in F : b = Ax, \text{ para algum } x \in E\}.$$

Isso significa que b está na imagem de A se e somente se o sistema $Ax = b$ tem solução. Neste caso, podemos escrever

$$\begin{aligned} b &= Ax \\ &= A(:,1)x_1 + A(:,2)x_2 + \ldots + A(:,n)x_n, \end{aligned}$$

em que $A(:, j)$ representa a coluna j de A. Desta forma, $b \in \mathcal{I}m(A)$ se e somente se b pode ser escrito como combinação linear das colunas de A (em que os coeficientes da combinação linear são as entradas de x). Com isso, o posto da matriz A é igual ao número de colunas l.i. de A (que é igual ao número de linhas l.i. de A por um teorema anterior).

Como encontrar uma base para a imagem de uma matriz A? Se a matriz U é a matriz na forma escada obtida depois de aplicarmos a A as operações elementares da eliminação gaussiana, então a imagem de A é diferente da imagem de U, mas estes dois espaços têm a mesma dimensão. Para vermos isso, basta observarmos que o número de colunas l.i. de U sempre será igual ao número de colunas l.i. de A, visto que os pivôs aparecem em U justamente nas colunas que eram l.i. em A.

Para recuperarmos uma base para a imagem de A, basta realizarmos a eliminação para obtermos U; os índices das colunas de U que formam uma base para a sua imagem (as colunas que tem os pivôs da eliminação) correspondem aos índices das colunas de A que são base para sua imagem. Isto acontece pois o sistema homogêneo $Ax = 0$ equivale ao sistema homogêneo $Ux = 0$; portanto, $Ax = 0$ determina, assim como $Ux = 0$, a dependência entre os vetores coluna da matriz A (e de U), com coeficientes x, idênticos aos dois sistemas. Se um conjunto de colunas de A é l.i., então as colunas correspondentes de U também são l.i.

Exemplo 5.1. *Finalmente, chegamos ao caso mais fácil, em que o posto é o menor possível (exceto pela matriz nula que possui posto 0).*

Exemplo 5.2.

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{pmatrix}.$$

Cada linha é um múltiplo da primeira, e assim o espaço gerado pelas linhas de A é unidimensional. De fato, podemos escrever esta matriz como um produto entre *um vetor linha e um vetor coluna:*

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 4 \\ -1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \end{pmatrix}.$$

Ao mesmo tempo, as colunas também são múltiplos do mesmo vetor coluna, e assim o espaço gerado pelas colunas também tem dimensão 1.

Teorema 17. *Toda matriz de posto 1 pode ser escrita como $A = uv^T$.*

Núcleo

Agora, queremos analisar as possíveis soluções de um sistema $Ax = 0$. Obviamente, a solução $x = 0$ sempre é possível, mas podem haver outros $x \in \mathbb{R}^n$ que satisfaçam $Ax = 0$, inclusive infinitas soluções deste tipo (isto sempre acontece quando temos mais variáveis do que equações, ou seja, $n > m$ - Teorema anterior). O conjunto das soluções do sistema homogêneo $Ax = 0$ é também um subespaço vetorial.

Definição 21. *Seja $T : E \to F$ uma transformação linear. Então o* núcleo *de T (ou* espaço nulo*, ou* kernel *de T) é o subespaço vetorial definido como*

$$\mathcal{N}(T) = \{x \in E \,|\, T(x) = 0\}.$$

Analogamente ao caso da imagem, podemos calcular o núcleo da matriz que representa uma transformação linear encontrando todas as soluções do sistema homogêneo definido por essa matriz.

É fácil encontrarmos o núcleo para a matriz do exemplo anterior: basta tomarmos

$$\begin{pmatrix} 1 & 0 \\ 5 & 4 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

para percebermos que, da primeira equação, temos que $u = 0$, e pela segunda equação devemos ter igualmente $v = 0$. Portanto, neste caso, apenas o vetor nulo faz parte do núcleo da matriz.

Vamos supor agora que a matriz do sistema é

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 5 & 4 & 9 \\ 2 & 4 & 6 \end{pmatrix}.$$

A imagem desta matriz é igual à imagem de A, já que a terceira coluna é combinação linear das outras duas. No entanto, o núcleo desta matriz contém qualquer múltiplo do vetor $(1, 1, -1)$. Portanto, este subespaço é uma reta que passa pela origem.

3.6 Sistemas Retangulares

Considere a seguinte matriz.

$$A = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{pmatrix}$$

O primeiro pivô é $a_{11} \neq 0$, e assim teremos

$$A \to \begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 6 & 2 \end{pmatrix}$$

O candidato a segundo pivô é zero, e assim vamos procurar embaixo dele por uma entrada não-nula. No entanto, todas as entradas abaixo dele são também nulas, e assim poderíamos parar a eliminação por aqui. Mas como esta matriz é retangular, não precisamos declarar a matriz singular; podemos apenas continuar com a eliminação na próxima coluna. Assim, teremos

$$U = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

pois na quarta coluna o candidato a pivô é nulo, e não podemos fazer mais nenhuma operação.

Assim, podemos como nas matrizes quadradas decompor $A = LU$; neste caso $L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$. Podemos notar assim que L é uma matriz 3×3, ou seja, quadrada de dimensão igual ao número de linhas das matrizes A e U. Caso seja necessário trocar uma linha pela outra no processo e eliminação para evitarmos um pivô nulo, podemos encontrar uma matriz de permutação P tal que $PA = LU$. Resumimos tudo isso no seguinte resultado.

A qualquer matriz $A \in \mathbb{R}^{m\times n}$ correspondem uma matriz de permutação P, uma matriz triangular inferior L com diagonal unitária, e uma matriz $m \times n$ na forma escada U tais que $PA = LU$.

Nosso objetivo no momento é olhar para esta forma final e verificar quais as possibilidades para o sistema $Ax = b$.

Se $b = 0$, então as operações da eliminação não tem efeito algum no vetor b; portanto as soluções de $Ax = b$ são as mesmas de $Ux = 0$:

$$Ux = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

As variáveis x_1, x_2, x_3 e x_4 podem ser concentradas em dois grupos: as variáveis básicas, que correspondem a colunas com pivôs não-nulos (neste caso, x_1 e x_3) e as variáveis livres, correspondentes a colunas sem pivôs (neste caso x_2 e x_4).

Para encontrarmos a solução deste sistema, basta então escolhermos valores arbitrários para as variáveis livres; assim, no nosso exemplo teremos

$$\begin{aligned} x_3 &= -\frac{1}{3}x_4 \\ x_1 &= -3x_2 - x_4. \end{aligned}$$

Portanto, este sistema tem infinitas soluções, e qualquer solução é uma combinação do tipo

$$x = \begin{pmatrix} -3x_2 - x_4 \\ x_2 \\ -\frac{1}{3}x_4 \\ x_4 \end{pmatrix} = x_2 \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ -\frac{1}{3} \\ 1 \end{pmatrix}.$$

Podemos ver então que o vetor $(-3, 1, 0, 0)$ é a solução quando $x_2 = 1, x_4 = 0$ e o outro vetor é a solução quando $x_2 = 0, x_4 = 1$. Portanto, qualquer outra solução será combinação linear destes vetores: eles formam a base para o espaço nulo de A (que é igual ao espaço nulo de U). O espaço nulo é um subespaço de mesma dimensão que o número de variáveis livres.

No caso não-homogêneo, a situação é diferente, pois depois da eliminação temos um sistema $Ux = c$ da forma

$$\begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 - 2b_2 + 5b_1 \end{pmatrix}.$$

É claro que a terceira equação implica que o sistema só pode ser consistente se $b_3 - 2b_2 + 5b_1 = 0$. Em outras palavras, o conjunto de vetores b gerados por Ax não pode ser todo o espaço, pois já temos uma restrição imposta. Mesmo com mais variáveis do que equações, pode ser que não tenhamos solução. Isto também pode ser visto do fato que o espaço das soluções é gerado pelas colunas de A, e assim é gerado pelos vetores

$$\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \\ -3 \end{pmatrix}, \begin{pmatrix} 3 \\ 9 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 0 \end{pmatrix}.$$

No entanto, o segundo vetor é 3 vezes o primeiro, e o quarto vetor equivale ao primeiro mais uma fração do terceiro! Portanto, estes vetores são l.d., e correspondem exatamente às colunas sem pivôs!

Escolhendo um vetor no plano $b_3 - 2b_2 + 5b_1 = 0$, por exemplo $b = (1, 5, 5)$, veremos que após a eliminação, temos o sistema $Ux = c$

$$\begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}.$$

Assim, a última equação pode ser eliminada ($0 = 0$) e as outras nos dão

$$\begin{aligned} x_3 &= 1 - \frac{1}{3}x_4 \\ x_1 &= -2 - 3x_2 - x_4. \end{aligned}$$

Mais uma vez, x_2, x_4 são variáveis livres, e qualquer solução geral pode ser escrita como

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ -\frac{1}{3} \\ 1 \end{pmatrix}.$$

Esta é uma soma do primeiro vetor com a solução de $Ax = 0$. *Toda solução de $Ax = b$ é a soma de uma solução particular (obtida quando tomamos as variáveis livres todas iguais a zero) e de uma solução do sistema homogêneo $Ax = 0$.* Estas soluções gerais estão em um conjunto que não é em um subespaço (pois zero não está contido nele) mas é paralelo ao espaço nulo, transladado pela solução particular. Assim, o cálculo da solução do sistema acima envolve os seguintes passos:

- Reduzir $Ax = b$ a $Ux = c$
- Tomar todas as variáveis livres (associadas a colunas sem pivôs) como zero e encontrar uma solução particular de $Ux = c$
- Tomar $c = 0$ e dar a cada variável livre, uma de cada vez, o valor 1 e zero para todas as outras. Encontrar uma solução homogênea através deste processo (ou seja, um vetor x no espaço nulo).

Vemos então que a eliminação nos revela o número de pivôs, e consequentemente o número de variáveis livres. Se r pivôs são não nulos, então existem r variáveis básicas e $n - r$ variáveis livres. Este número r é chamado *posto* (*rank*) da matriz A.

> **Resumo:** Suponha que através da eliminação transformamos $Ax = b$ em $Ux = c$. Seja o número de pivôs não nulos igual a r, as últimas $m - r$ linhas de U nulas. Então, só existe solução para o sistema se as últimas $m - r$ componentes de c também forem nulas. Se $r = m$, sempre existe solução.

A solução geral é a soma de uma solução particular (com todas as variáveis livres iguais a zero) e de uma solução homogênea (com as $n - r$ variáveis livres como parâmetros independentes). Se $r = n$, não existem variáveis livres e o espaço nulo contém somente o vetor $x = 0$.

Existem dois casos extremos:

- Se $r = n$, então não existem variáveis livres para x. ($\mathcal{N} = \{0\}$)
- Se $r = m$, então não existem linhas nulas em U. ($\mathcal{I}m(A) = \mathbb{R}^m$)

A dimensão do espaço coluna $\mathcal{I}m(A)$ é igual ao posto r, e uma base de $\mathcal{I}m(A)$ é formada pelas r colunas de A que correspondem, em U, às colunas contendo pivôs.

Para ver isso com mais clareza, considere um exemplo:

$$U = \begin{pmatrix} d_1 & * & * & * & * & * \\ 0 & 0 & 0 & d_2 & * & * \\ 0 & 0 & 0 & 0 & 0 & d_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

É claro que esta matriz tem três linhas independentes. Afirmamos que existem também apenas 3 colunas independentes, não mais, mostrando que as 3 colunas que contém pivôs são L.I. Suponha que existam c_1, c_2, c_3 tais que

$$c_1 \begin{pmatrix} d_1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} * \\ d_2 \\ 0 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} * \\ * \\ d_3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Então, como os pivôs são diferentes de zero, é claro que $c_1 = c_2 = c_3 = 0$. Portanto, estas colunas são l.i. e formam uma base para o espaço coluna de A.

O espaço nulo à esquerda de A, que é o espaço nulo de A^T, é um subespaço de $\mathbb{R}^m$. A dimensão deste espaço é fácil de ser encontrada, já que o número de variáveis básicas mais o número de variáveis livres deve ser igual ao número total de colunas: logo, $\mathcal{N}(A^T)$ tem dimensão $m - r$, já que

$$\dim(\mathcal{I}m(A)) = \dim(\mathcal{I}m(A^T)) = r.$$

Vamos ver a importância deste espaço mais à frente.

O espaço linha de A tem a mesma dimensão r do espaço linha de U, e eles tem a mesma base pois são o mesmo espaço. Isto ocorre pois as linhas de U são combinações lineares das linhas de A, e portanto as operações elementares não alteram o espaço, apenas revelam sua base, que é formada pelas linhas não nulas de U.

O espaço nulo de A é o espaço nulo de U, pois as soluções que satisfazem $Ax = 0$ também satisfazem $Ux = 0$. As restrições aos vetores do espaço nulo são dadas pelas linhas não nulas de U, e assim o número de linhas nulas de U indica a dimensão do espaço nulo ($n - r$, se tivermos r linhas não nulas). Esta dimensão é às vezes chamada de *nulidade* de A.

Resumindo todas estas considerações, chegamos ao seguinte resultado.

Teorema 18 (Fundamental da Álgebra Linear, Parte I)**.** *Seja $A \in \mathbb{R}^{m\times n}$ com r linhas l.i. (ou r colunas l.i.). Então,*

(i) $\mathcal{I}m(A)$ *tem dimensão* r

(ii) $\mathcal{N}(A)$ *tem dimensão* $n - r$

(iii) $\mathcal{I}m(A^T)$ *tem dimensão* r

(iv) $\mathcal{N}(A^T)$ *tem dimensão* $m - r$*.*

3.7 Existência de Inversas

Sabemos que se A tem uma inversa à esquerda ($BA = I$) e uma inversa à direita ($AC = I$) então as duas inversas são iguais. Agora, através do posto de uma matriz, podemos decidir se ela tem tais inversas ou não: uma inversa existe somente quando o posto da matriz é o maior possível. Sabemos que o posto satisfaz $r \leq m$ e $r \leq n$. Uma matriz $m \times n$ não pode ter mais do que m linhas l.i. ou n colunas l.i. Queremos mostrar então que, quando $r = n$ existe uma inversa à direita, e que quando $r = m$ existe uma inversa à esquerda. No primeiro caso, $Ax = b$ sempre tem solução; no segundo caso, se a solução existir ela é única. Somente uma matriz quadrada pode ter $r = m = n$, e assim somente uma matriz quadrada pode definir um sistema com solução única e existência garantida.

$$\boxed{\begin{array}{cccc} \star & \star & \star & \star \\ \star & \star & \star & \star \end{array}} \boxed{\begin{array}{cc} \star & \star \\ \star & \star \\ \star & \star \\ \star & \star \end{array}} = \boxed{\begin{array}{cc} \star & \star \\ \star & \star \end{array}}$$

- O sistema $Ax = b$ tem pelo menos uma solução x para cada $b$se e somente se as colunas de A geram $\mathbb{R}^m$, ou seja, $r = m$. Neste caso, existe uma inversa à direita $C \in \mathbb{R}^{n \times m}$ tal que $AC = I_m$, a matriz identidade de ordem m. Isto só é possível se $m \leq n$ (caso contrário, existiriam mais colunas l.i. do que n, absurdo).

- O sistema $Ax = b$ tem no máximo uma solução x para cada b se e somente se as colunas de A forem linearmente independentes, ou seja, $r = n$. Neste caso, o sistema homogêneo só tem solução trivial e a solução geral será apenas uma solução particular, que é única (não há variáveis livres). Assim, existe uma inversa à esquerda $B \in \mathbb{R}^{n \times m}$ tal que $BA = I_n$. Isto só é possível se $m \geq n$.

No primeiro caso, uma solução possível é $x = Cb$, já que assim teríamos $Ax = ACb = b$. Mas podem existir outras soluções se existirem outras inversas à direita.

No segundo caso, se a solução para $Ax = b$ existir, ela tem que ser $x = BAx = Bb$. Mas a solução pode não existir.

Existem fórmulas simples para as inversas à direita e à esquerda, se elas existirem:

$$B = (A^T A)^{-1} A^T \qquad \text{e} \qquad C = A^T (AA^T)^{-1}.$$

O que pode falhar nestas fórmulas é a existência das inversas de A^TA e de AA^T. Vamos mostrar (mais à frente) que estas inversas existem quando o posto de A for n ou m, respectivamente. Note também que estas inversas (à direita e à esquerda) não são únicas.

Exemplo 7.1 (Strang, p. 97)**.** *Considere a matriz*

$$A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix}.$$

Como o posto de A é 2, nossa análise acima sugere uma inversa à direita C. De fato, note que

$$AC = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix} \begin{pmatrix} 1/4 & 0 \\ 0 & 1/5 \\ c_{31} & c_{32} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

De fato, existem infinitas inversas à direita; a última linha de C é totalmente arbitrária. Por outro lado, não existe inversa à esquerda, já que a última coluna de BA será nula para qualquer B, o que nos impede de obter a matriz identidade com esse produto.

Para esse exemplo, se usarmos a fórmula $C = A^T(AA^T)^{-1}$, teremos

$$C = \begin{pmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/16 & 0 \\ 0 & 1/25 \end{pmatrix} = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/5 \\ 0 & 0 \end{pmatrix}.$$

Por outro lado, a transposta de A tem infinitas inversas à esquerda:

$$\begin{pmatrix} 1/4 & 0 & b_{13} \\ 0 & 1/5 & b_{23} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{pmatrix} = I_{2\times 2}.$$

Para uma matriz retangular, não é possível termos, ao mesmo tempo, existência e unicidade. Se $m \neq n$, não podemos ter $r = m$ e $r = n$. Para uma matriz quadrada, vale o oposto: não é possível termos existência sem unicidade. Pode-se dizer que, para que uma matriz quadrada $A \in \mathbb{R}^{n\times n}$ seja não-singular, cada uma das condições seguintes é necessária e suficiente:

1. As colunas de A geram $\mathbb{R}^n$, de modo que $Ax = b$ tenha ao menos uma solução para cada b.

2. As colunas são independentes, de modo que $Ax = 0 \Rightarrow x = 0$.

3.8 Transformações lineares inversíveis.

Primeiramente, vamos observar alguns exemplos.

Exemplo. *A transformação linear definida pela diferenciação de um polinômio de $\mathcal{P}_n$ e que resulta em um polinômio em $\mathcal{P}_{n-1}$ tem como núcleo o espaço unidimensional de polinômios constantes: $\frac{da_0}{dx} = 0$. Sua imagem é o espaço n dimensional $\mathcal{P}_{n-1}$. A soma da dimensão do núcleo (1) e do posto (n) nos dá a dimensão do espaço original.*

Exemplo. *A integração de 0 a x, vista como transformação linear de $\mathcal{P}_n$ a $\mathcal{P}_{n+1}$ tem núcleo $\{0\}$. Note que a integração não produz polinômios constantes, e portanto não gera todo o $\mathcal{P}_{n+1}$ (ou seja, sua imagem não é todo o espaço de chegada).*

Exemplo 8.1. *A multiplicação por um polinômio fixo como $2 + 3x$ é uma transformação linear:*

$$Ap = (2 + 3x)(a_0 + \ldots + a_n x^n) = 2a_0 + \ldots + 3a_n x^{n+1}.$$

Esta transformação leva $\mathcal{P}_n$ em $\mathcal{P}_{n+1}$, com núcleo contendo apenas $p = 0$.

Se quisermos investigar, agora de maneira teórica, como decidir se uma transformação linear possui inversa ou não, precisamos refinar alguns conceitos.

Definição 22. *Seja T uma transformação linear de E em F.*

- *Se $Im(T) = F$, dizemos que T é* sobrejetora.
- *Se o núcleo de T contiver apenas o vetor nulo, dizemos que esta transformação é* injetora. *Isto é equivalente a dizermos que se $T(u) = T(v)$ então $T(u - v) = 0 \Rightarrow u - v = 0$.*

Teorema 19. *Para que $T : E \to F$ seja inversível, é necessário e suficiente que T seja injetora e sobrejetora.*

Demonstração. Primeiramente, observe que T possui inversa à direita $Q : F \to E$ se e somente se pudermos escrever

$$T(Q(v)) = v, \quad \forall\, v \in F. \tag{3.2}$$

Além disso, T possui inversa à esquerda $S : F \to E$ se e somente se pudermos escrever

$$S(T(u)) = u, \quad \forall\, u \in E. \tag{3.3}$$

- **Afirmação: T é sobrejetora se e somente se possui inversa à direita.** Se T for sobrejetora, então $\mathcal{I}m(T) = F$. Desta forma, todos os vetores de F podem ser escritos como resultado da aplicação de T em algum vetor de E. Em particular, se tomamos uma base $\beta \subset F$, podemos escolher vetores u_i em E tais que $T(u_i) = v_i$, para cada $v_i \in \beta \subset F$. Como vimos que para definir uma transformação linear basta definirmos o resultado da sua aplicação em elementos de uma base do domínio, basta escolhermos a transformação Q que a cada v_i associa u_i como inversa à direita de T; desta forma, $T(Q(v_i)) = T(u_i) = v_i$ para todo $v_i \in F$. Por outro lado, se T admitir inversa à direita $Q : F \to E$, então para todo $v \in F$, $T(Q(v)) = v$, ou seja, $v = T(w)$ para todo $v \in F$ (com $w = Q(v)$) e assim T é sobrejetora (a imagem de T é todo o F).

- **Afirmação: T é injetora se e somente se possui inversa à esquerda.** Se T é injetora, então ela leva um conjunto de vetores l.i. em um conjunto de vetores l.i.: Sejam $\{u_1, \ldots, u_n\} \subset E$ l.i. Então,

$$\begin{aligned} &\alpha_1 T(u_1) + \ldots + \alpha_n T(u_n) = 0 \\ \Leftrightarrow\ &T(\alpha_1 u_1 + \ldots + \alpha_n u_n) = 0 \end{aligned}$$

  Como T é injetora, $T(u) = 0 \Rightarrow u = 0$ e assim

$$\alpha_1 u_1 + \ldots + \alpha_n u_n = 0$$

  Agora, como o conjunto $\{u_1, \ldots, u_n\}$ é l.i., isso implica que $\alpha_1 = \ldots = \alpha_n = 0$. Logo, $\{T(u_1), \ldots, T(u_n)\}$ formam um conjunto l.i. em F.

  Assim, se tomarmos uma base $\{u_i\} \subset E$, podemos definir uma base para F se tomarmos $\{T(u_i), v_i\}$, onde os v_i são acrescentados ao conjunto l.i. Av_i se necessário para completar a base; logo, para definirmos uma inversa à esquerda basta tomarmos $S : F \to E$ tal que $S(T(u_i)) = u_i$ e $S(v_i) = 0$ para todo i.

  Finalmente, se T possui inversa à esquerda $S : F \to E$, então $T(u) = 0 \Rightarrow u = S(T(u)) = S(0) = 0$, ou seja, T é injetora.

□

Se $T : E \to F$ é inversível, dizemos que ela é uma *bijeção* ou um *isomorfismo* entre E e F, e que E e F são isomorfos. Um isomorfismo $T : E \to F$ transforma uma base de E em uma base de F e se uma transformação linear leva uma base de E numa base de F, então ela é um isomorfismo. Desta forma, dois espaços vetoriais de dimensão finita isomorfos tem a mesma dimensão; por outro lado, suponha que E é um espaço vetorial de dimensão n. Se fixarmos uma base $\{v_i\} \subset E$, podemos definir $T : \mathbb{R}^n \to E$ tal que se $u = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$, $T(u) = T(\alpha_1, \ldots, \alpha_n) = \alpha_1 v_1 + \ldots + \alpha_n v_n$. Desta forma, $Ae_1 = v_1, \ldots, Ae_n = v_n$, em que os $e_i \in \mathbb{R}^n$ são os elementos da base canônica do $\mathbb{R}^n$. Portanto, T transforma a base canônica do $\mathbb{R}^n$ na base de E e portanto define um isomorfismo entre E e $\mathbb{R}^n$: em outras palavras, *todo espaço vetorial de dimensão n é isomorfo a* $\mathbb{R}^n$.

Exemplo 8.2. *O espaço $\mathcal{P}_n$ tem dimensão $n+1$ e portanto é isomorfo a $\mathbb{R}^{n+1}$. O espaço das matrizes $\mathcal{M}(m \times p)$ é isomorfo a $\mathbb{R}^{mp}$.*

3.8.1 Teorema do núcleo e da imagem de uma transformação linear.

Teorema 20 (do Núcleo e da Imagem)**.** *Sejam E, F espaços vetoriais de dimensão finita. Para toda transformação linear $T : E \to F$ temos que*

$$dim(E) = dim\mathcal{N}(T) + dim\mathcal{I}m(T).$$

Demonstração. Vamos mostrar que se $\{T(u_i)\}$ é uma base de $\mathcal{I}m(T)$ e $\{v_i\}$ é uma base de $\mathcal{N}(T)$, então $\{u_i, v_i\}$ é uma base de E.

Para isto, considere que se tivermos

$$\alpha_1 u_1 + \ldots + \alpha_p u_p + \beta_1 v_1 + \ldots + \beta_q v_q = 0 \tag{3.4}$$

então aplicando T dos dois lados da equação, teríamos que $\alpha_1 T(u_1) + \ldots + \alpha_p T(u_p) = 0$, já que $v_i \in \mathcal{N}(T)$. Mas, como $\{T(u_i)\}$ é uma base de $\mathcal{I}m(T)$, estes vetores são l.i., e assim $\alpha_i = 0$. Portanto, em (3.4) teríamos apenas $\beta_1 v_1 + \ldots + \beta_q v_q = 0$. No entanto, como $\{v_i\}$ é base de $\mathcal{N}(T)$, estes vetores são l.i. e logo $\beta_i = 0$. Portanto, $\{u_i, v_i\}$ são l.i.

Em seguida, considere um vetor arbitrário $w \in E$. Como $T(w) \in \mathcal{I}m(T)$, podemos escrever

$$\begin{aligned} T(w) &= \alpha_1 T(u_1) + \ldots + \alpha_p T(u_p) \\ &\Leftrightarrow T(w - (\alpha_1 u_1 + \ldots + \alpha_p u_p)) = 0. \end{aligned}$$

Logo, $w - (\alpha_1 u_1 + \ldots + \alpha_p u_p) \in \mathcal{N}(A)$, e portanto podemos escrever

$$w - (\alpha_1 u_1 + \ldots + \alpha_p u_p) = \beta_1 v_1 + \ldots + \beta_q v_q,$$

o que implica que $w = \alpha_1 u_1 + \ldots + \alpha_p u_p + \beta_1 v_1 + \ldots + \beta_q v_q$, e assim estes vetores geram E. $\square$

Corolário 8. *Sejam E, F espaços vetoriais de mesma dimensão finita n. Uma transformação linear $T : E \to F$ é injetiva se e somente se é sobrejetiva e portanto, é um isomorfismo.*

Demonstração. Com efeito, temos

$$n = dim\mathcal{N}(T) + dim\mathcal{I}m(T).$$

Logo, $dim\mathcal{N}(T) = 0 \Leftrightarrow dim\mathcal{I}m(T) = n$. $\square$

Exemplo 8.3. *Descrever a imagem e o núcleo de*

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

e

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 5 & 4 & 9 \\ 2 & 4 & 6 \end{pmatrix}$$

Note que $\mathcal{I}m(A) = span\{(1,0)\}$ e $\mathcal{N}(A) = span\{(x_1, x_1)\} = span\{(1,1)\}$. A dimensão de cada um dos espaços é 1 e sua soma é 2.

Para B, lembre que $\mathcal{N}(B) = \mathcal{N}(U)$ (em que U é a matriz obtida depois que escalonamos a matriz B) e que $\mathcal{I}m(B)$ não é igual a $\mathcal{I}m(U)$, mas as colunas que geram a imagem de B são as colunas correspondentes às colunas em que aparecem os pivôs de U. Logo, para resolver este problema primeiro escalonamos a matriz B. Assim, $\mathcal{N}(B) = span\{(1,1,-1)\} = span\{(-1,-1,1)\}$ e $\mathcal{I}m(B) = span\{(1,0,0),(0,1,1)\}$. Note ainda que $dim(\mathcal{N}(B)) + dim(\mathcal{I}m(B)) = 1 + 2 = 3$.

Exemplo 8.4.

$$A = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{pmatrix}$$

Observação Toda solução de um sistema $Ax = b$ é soma de uma solução particular (onde todas as variáveis livres valem 0) e de uma solução ao sistema homogêneo $Ax = 0$. Isto é equivalente a dizermos que o núcleo e a imagem de A se completam em dimensão para formar o espaço de saída.

Exemplo 8.5.

$$Ax = b : \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

solução geral = solução particular + solução homogênea:

$$x = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -2 \\ 1 \\ -1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

3.9 Mudança de base

Sejam $U = \{u_1, \ldots, u_n\} \subset E$ e $W = \{w_1, \ldots, w_m\} \subset F$ bases em relação às quais a matriz da transformação linear $T : E \to F$ é definida. Isto significa que para cada $u_j \in U$ definimos a transformação T como sendo $T(u_j) = v_j$, em que $v_j \in F$. Assim, v_j pode ser escrito como combinação linear dos elementos de W (pois W é base de F) e assim existem $a_{ij} \in \mathbb{R}$ tais que

$$T(u_j) = v_j = \sum_{i=1}^{m} a_{ij} w_i, \qquad (j = 1, \ldots, n).$$

Tomando novas bases $U' = \{u'_1, \ldots, u'_n\} \subset E$ e $W' = \{w'_1, \ldots, w'_m\} \subset F$, a transformação linear T tem uma nova matriz A' definida por

$$T(u'_j) = \sum_{r=1}^{m} a'_{rj} w'_r, \qquad (j = 1, \ldots, n). \tag{3.5}$$

Para obtermos a relação entre as matrizes A e A', consideramos as matrizes de mudança de base $P = I_U^{U'} \in \mathbb{K}^{n\times n}$ e $Q = I_W^{W'} \in \mathbb{K}^{m\times m}$ definidas pelas igualdades

$$u'_j = \sum_{k=1}^{n} p_{kj} u_k \qquad \text{e} \qquad w'_r = \sum_{i=1}^{m} q_{ir} w_i.$$

Agora, podemos escrever cada um dos membros de (3.5) em termos da base W, fazendo

$$\begin{aligned} T(u'_j) = \sum_{k=1}^{n} p_{kj} T(u_k) &= \sum_{k=1}^{n} p_{kj} \sum_{i=1}^{m} a_{ik} w_i \\ &= \sum_{k=1}^{n} \sum_{i=1}^{m} p_{kj} a_{ik} w_i \\ &= \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{ik} p_{kj} \right) w_i, \end{aligned}$$

e

$$\begin{aligned} \sum_{r=1}^{m} a'_{rj} w'_r &= \sum_{r=1}^{m} a'_{rj} \sum_{i=1}^{m} q_{ir} w_i \\ &= \sum_{r=1}^{m} \sum_{i=1}^{m} a'_{rj} q_{ir} w_i \\ &= \sum_{i=1}^{m} \left(\sum_{r=1}^{m} q_{ir} a'_{rj} \right) w_i. \end{aligned}$$

Igualando os coeficientes de w_i, temos que

$$\sum_{k=1}^{n} a_{ik} p_{kj} = \sum_{r=1}^{m} q_{ir} a'_{rj},$$

ou seja, $AP = QA'$.

Observe agora que toda matriz de mudança de base é inversível: leva uma base numa base. Assim, podemos concluir que

$$A' = Q^{-1} AP.$$

É útil também observar que se U é uma base de $\mathbb{R}^n$, então a matriz de mudança de base da base canônica para U é aquela cujas n colunas são os vetores $u_1, \dots, u_n$.

No caso particular de um operador $T : E \to E$ e de suas matrizes A e A' relativas às bases U e U', temos uma única matriz de mudança de base P, o que nos dá

$$A' = P^{-1} AP.$$

3.10 Soma direta

Como vimos que a imagem e o núcleo de uma transformação linear são complementares com relação à dimensão do espaço vetorial em que estão contidos, podemos imaginar uma decomposição do espaço nestes dois subespaços.

Vimos anteriormente que se F_1 e F_2 são subespaços do espaço vetorial E, o subespaço vetorial de E formado pela conjunção dos elementos de F_1 e de F_2 é o conjunto $F_1 + F_2$ de todas as somas $u+v$, onde $u \in F_1$ e $v \in F_2$. No caso particular em que $F_1 \cap F_2 = \{0\}$, escrevemos $F_1 \oplus F_2$ e chamamos esta operação de *soma direta*.

Existe uma noção análoga à de soma direta, que é o produto cartesiano de dois espaços vetoriais E_1 e E_2. Estes dois espaços não precisam fazer parte do mesmo espaço E. Os elementos do conjunto $E_1 \times E_2$ são pares ordenados $(u, v), u \in E_1, v \in E_2$. As operações que tornam este novo espaço um espaço vetorial são

$$(u, v) + (u', v') = (u + u', v + v'),$$
$$\alpha(u, v) = (\alpha u, \alpha v).$$

O vetor nulo é $(0, 0)$ e o inverso aditivo é $(-u, -v)$.

Se $\{u_1, \ldots, u_m\} \subset E_1$ e $\{v_1, \ldots, v_k\} \subset E_2$, é imediato observarmos que

$$\{(u_1, 0), \ldots, (u_m, 0), (0, v_1), \ldots, (0, v_k)\} \subset E_1 \times E_2$$

é uma base, de modo que $dim(E_1 \times E_2) = dimE_i + dimE_2$.

Se F_1, F_2 são subespaços de E com $F_1 \cap F_2 = \{0\}$, então a transformação linear

$$A : F_1 \times F_2 \to F_1 \oplus F_2$$

definida por $A(u, v) = u + v, u \in F_1, v \in F_2$ é um isomorfismo, pois se

$$\{u_1, \ldots, u_m\} \subset F_1$$

e

$$\{v_1, \ldots, v_n\} \subset F_2$$

são bases, então a base

$$\{(u_1, 0), \ldots, (u_m, 0), (0, v_1), \ldots, (0, v_n)\} \subset F_1 \times F_2$$

é transformada por A no conjunto

$$\{u_1, \ldots, u_m, v_1, \ldots, v_n\}$$

que é, por sua vez, uma base de $F_1 \oplus F_2$. Segue-se que $dim(F_1 \oplus F_2) = dimF_1 + dimF_2 = m + n$.

No caso mais geral, em que a interseção $F_1 \cap F_2$ não se reduz necessariamente ao vetor nulo, a soma $F_1 + F_2$ pode não ser mais uma soma direta, mas ainda podemos definir a transformação linear

$$A : F_1 \times F_2 \to F_1 + F_2$$

de forma que $A(u, v) = u + v$. Obviamente, A é sobrejetiva: todo elemento de $F_1 + F_2$ pode ser representado como imagem por A de um vetor em $F_1 \times F_2$. Seu núcleo é formado pelos pares (u, v) tais que $u + v = 0$, ou seja $v = -u$. Neste caso, ambos $u, v \in F_1 \cap F_2$. A correspondência $u \mapsto (u, -u)$ é um isomorfismo entre $F_1 \cap F_2$ e $\mathcal{N}(A)$. Pelo Teorema do Núcleo e da Imagem, temos então que

$$\begin{aligned} dimF_1 + dimF_2 &= dim(F_1 \times F_2) \\ &= dim\mathcal{N}(A) + dim(F_1 + F_2) \\ &= dim(F_1 \cap F_2) + dim(F_1 + F_2). \end{aligned}$$

Assim, enunciamos o seguinte teorema.

Teorema 21. *Sejam F_1, F_2 subespaços de dimensão finita de um espaço vetorial E. Temos que*

$$dimF_1 + dimF_2 = dim(F_1 \cap F_2) + dim(F_1 + F_2).$$

A noção de soma direta está intimamente ligada à noção de projeção. Se $E = F_1 \oplus F_2$ é a decomposição do espaço vetorial E como soma direta destes subespaços, definimos o operador $P : E \to E$, projeção de E sobre F_1 paralelamente a F_2 da seguinte maneira: todo vetor $w \in E$ se escreve, de modo único, como soma $w = u + v$ de $u \in F_1, v \in F_2$. Escolha então $Pw = u$.

O operador linear assim definido tem imagem F_1 e núcleo F_2. Além disso é fácil ver que P é idempotente (ou seja, $P^2 = P$). Vamos mostrar a seguir que todo operador linear idempotente é uma projeção. Note primeiramente que se $P^2 = P$, então para todo $w \in \mathcal{I}m(P)$, temos que $P(w) = w$ pois $w \in \mathcal{I}m(P) \Rightarrow w = P(v) \Rightarrow P(w) = P(P(v)) = P(v) = w$.

Teorema 22. *Seja $P : E \to E$ um operador linear. Se $P^2 = P$, então E é a soma direta do núcleo com a imagem de P. Além disso, P é a projeção sobre $\mathcal{I}m(P)$ paralelamente a $\mathcal{N}(P)$.*

Demonstração. Todo $v \in E$ se escreve como soma $v = (v - P(v)) + P(v)$, onde $P(v) \in \mathcal{I}m(P)$ e, como $P(v - P(v)) = P(v) - P(P(v)) = P(v) - P(v) = 0$, vemos que $v - P(v) \in \mathcal{N}(P)$. Portanto, $E = \mathcal{N}(P) + \mathcal{I}m(P)$. Se $w \in \mathcal{N}(P) \cap \mathcal{I}m(P)$, por um lado temos que $P(w) = 0$, e por outro $P(w) = w$; logo, $w = 0$. Assim, $\mathcal{N}(P) \cap \mathcal{I}m(P) = \{0\}$ e temos que $E = \mathcal{I}m(P) \oplus \mathcal{N}(P)$. O resto é óbvio. □

Exemplo 10.1. *Para todo operador linear $T : E \to E$ num espaço vetorial de dimensão finita vale a relação $dim(E) = dim\mathcal{N}(T) + dim\mathcal{I}m(T)$, mas não temos sempre que esta soma é direta: por exemplo, se $T : \mathbb{R}^2 \to \mathbb{R}^2$ é definido por $T(x, y) = (x - y, x - y)$ então tomando $w = (1, 1)$ temos que $w = T(v)$ com $v = (2, 1)$ e $T(w) = 0$, portanto $w \in \mathcal{I}m(T) \cap \mathcal{N}(T)$.*

Definição 23. *Uma* involução *é um operador linear $S : E \to E$ tal que $S^2 = I$, ou seja $S(S(v)) = v$ para todo $v \in E$.*

Em outras palavras, uma involução é um operador inversível, igual ao seu próprio inverso. Um exemplo é a reflexão (ortogonal) no plano em torno de uma reta que passa pela origem. Veremos agora que toda involução é a reflexão em torno de um subespaço, paralelamente a outro.

Teorema 23. *Seja $S : E \to E$ uma involução. Os conjuntos $F_1 = \{u \in E : S(u) = u\}$ e $F_2 = \{v \in E : S(v) = -v\}$ são subespaços vetoriais e $E = F_1 \oplus F_2$. Para todo $w = u + v$ com $u \in F_1, v \in F_2$ tem-se $S(w) = u - v$. Além disso, $P = \frac{1}{2}(S + I)$ é a projeção sobre F_1 paralelamente a F_2.*

Demonstração. Para todo $w \in E$, podemos escrever $w = u + v$, onde $u = \frac{(w+S(w))}{2}$ e $v = \frac{(w-S(w))}{2}$. Como $S^2 = I$, é claro que $S(u) = u$ e $S(v) = -v$, ou seja, $u \in F_1, v \in F_2$. É claro também que $F_1 \cap F_2 = \{0\}$, e que $w = u + v \Rightarrow S(w) = u - v$ se $u \in F_1$ e $v \in F_2$. Finalmente,

$$
\begin{aligned}
P = \frac{1}{2}(S + I) \Rightarrow P^2 &= \frac{1}{4}(S^2 + 2S + I) \\
&= \frac{1}{4}(2S + 2I) \\
&= \frac{1}{2}(S + I) = P.
\end{aligned}
$$

Pode-se ver facilmente que o núcleo de P é F_2 e a imagem de P é F_1. □

O caso mais comum de reflexão é aquele em que se tem $dim(E) = n$, $dim(F_1) = n - 1$ e $dim(F_2) = 1$, de modo que S é a reflexão em torno do hiperplano F_1 paralelamente à reta F_2.