Notas de Aula: Curso de Verão UFPR 2020 - Álgebra Linear

1 Aula 1 - Paleari

Definição 1.1 *Um **grupo** é um par $(G, *)$ consistindo de um conjunto G com uma função $* : G \times G \to G$ que possui as seguintes propriedades*

- *Associatividade: $*(*(a,b),c) = *(a,*(b,c))$ para todos $a,b,c \in G$.*
- *Elemento neutro: Existe um elemento $e \in G$ tal que $*(e,a) = *(a,e) = a$ para todo $a \in G$.*
- *Elemento inverso: Para todo $a \in G$ existe um elemento $b \in G$ tal que $*(a,b) = *(b,a) = e$.*

Se além das 3 propriedades acima, o par $(G, *)$ também satisfaz:

- *Comutatividade: Para todos $a, b \in G$ vale $*(a,b) = *(b,a)$,*

*dizemos que $(G, *)$ é **abeliano**.*

A notação $*(a,b)$ não é realmente usada, é comum usarmos a notação $a * b := *(a,b)$. Reescreva as condições de grupo acima usando essa notação. Informalmente, dizemos que o conjunto G é um grupo se ele possui uma operação binária que é associativa, possui elemento neutro e todo elemento possui um inverso.

Exercício 1.1 (Coerência) *Prove se $(G, *)$ é um par consistindo de um conjunto G com uma operação binária $* : G \times G \to G$ que é associativa, então se existe um elemento neutro para $*$, necessariamente ele é único. Prove em seguida que se $(G, *)$ é conjunto com uma operação binária associativa e que possui elemento neutro, então quando um elemento possui inverso, esse inverso é necessariamente único. Escreva em símbolos essas afirmações.*

O exercício acima nos diz que não há redundância na frase "o elemento neutro de G"ou na frase em "o elemento inverso de $a \in G$". Isso permite dar uma notação padrão para o neutro e para inversos de elementos de G como alguns livros textos fazem, embora não haja uma notação padrão pois ela depende do contexto. Em vários textos o inverso de um elemento $a \in G$ é denotado por a^{-1}. Mas se por exemplo o grupo em questão é o conjunto dos números reais com a operação sendo a adição $(\mathbb{R}, +)$, então o inverso de $a \in \mathbb{R}$ é seu oposto aditivo, que é denotado por $-a$. Por isso, não vamos padronizar a notação para inversos para não causar confusão. Porém, é comum usarmos a notação $+$ para representar a operação binária (genérica) em G, mas isso não deve ser confundido com a operação de adição usual, mesmo porque ela pode nem ter sentido no conjunto G (vide os exemplos abaixo). Isso nós faremos e o leitor deve se atentar para não causar confusão. Quando isso é feito, o elemento neutro de G é denotado por 0 (que não é necessariamente o número zero, pelo menos motivo já discutido).

Exemplo 1.1 *Alguns exemplos simples de grupos (verifique):*

- *O conjunto dos números racionais,reais ou complexos com operação de adição: $(\mathbb{K}, +)$, com $\mathbb{K} = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$;*
- *O conjunto dos números racionais,reais ou complexos não nulos com operação de produto: $(\mathbb{K} \setminus \{0\}, \cdot)$, com $\mathbb{K} = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$.*
- *O conjunto dos números inteiros módulo n com operação de adição: $(\mathbb{Z}_n, +)$.*

Definição 1.2 *Um **corpo** é uma terna $(\mathbb{K}, +, \cdot)$ consistindo de um conjunto $\mathbb{K}$ e duas operações binárias $+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ e $\cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ tais que:*

- $(\mathbb{K}, +)$ *é um grupo abeliano;*
- $(\mathbb{K} \setminus \{0\}, \cdot)$ *é um grupo*
- *Para todos* $x, y, z \in \mathbb{K}$ *vale* $x \cdot (y + z) = x \cdot y + x \cdot z$.

Exemplo 1.2 *Alguns exemplos de conjuntos que são e que não são corpos (verifique)*

- $(\mathbb{K}, +, \cdot)$ *com* $\mathbb{K} = \mathbb{Q}, \mathbb{R}$ *ou* $\mathbb{C}$ *são corpos com as operações usuais de adição e produto.*
- $(\mathbb{Z}, +, \cdot)$ *não é um corpo.*
- $(\mathbb{Z}_n, +, \cdot)$ *é um corpo com as operações usuais de soma e produto de classes módulo* n *se, e somente se,* n *é primo.*

Por causa dos corpos que serão usados na prática nesse curso, as notações para as operações binárias envolvidas serão as que imitam os exemplos básicos acima: "$+$" e "$\cdot$". O elemento neutro do grupo $(\mathbb{K} \setminus \{0\}, \cdot)$ é geralmente denotado por 1, o que também é coerente com os principais corpos que serão usados nesse curso.

Definição 1.3 *Seja* $(\mathbb{K}, +, \cdot)$ *um corpo. Um* $\mathbb{K}$*-espaço vetorial é um terno* $(V, +, \cdot)$ *consistindo de um conjunto* V *com funções* $+ : V \times V \to V$ *e* $\cdot : \mathbb{K} \times V \to V$ *satisfazendo:*

- $(V, +)$ *é um grupo abeliano.*
- *Para todos* $x, y \in \mathbb{K}$ *e* $v \in V$ *vale* $(x + y) \cdot v = x \cdot v + y \cdot v$.
- *Para todos* $x \in \mathbb{K}$ *e* $v, w \in V$ *vale* $x \cdot (v + w) = x \cdot v + x \cdot w$.
- *Para todos* $x, y \in \mathbb{K}$ *e* $v \in V$ *vale* $(x \cdot w) \cdot v = x \cdot (y \cdot w)$.
- *Para todo* $v \in V$ *vale* $1 \cdot v = v$.

Informalmente, um $\mathbb{K}$-espaço vetorial consiste de um grupo V com uma operação de multiplicação por escalares em $\mathbb{K}$ que é compatível com a operação de grupo de V. Observe que o símbolo $+$ está sendo usado para dois propósitos diferentes. Ora ele é a operação binária do grupo $(\mathbb{K}, +)$ ora ele é a operação binária do grupo $(V, +)$. O leitor deve possui maturidade para distinguir os dois pelo contexto, percebendo que não há perigo algum de confusão. Isso acontecerá diversas vezes ao longo do texto com outras operações também.

Exemplo 1.3 *Seja* $(\mathbb{K}, +, \cdot)$ *um corpo e* $V = \{x\}$ *um conjunto unitário. Então* V *é um* $\mathbb{K}$*-espaço vetorial com as únicas operações possíveis:* $x + x := x$ *e* $k \cdot x = x$ *para todo* $k \in \mathbb{K}$. *Todas as propriedades são facilmente verificadas e é claro que* x *deve ser o elemento neutro da operação* $+$. *Assim, podemos escrever simplesmente* $V = \{0\}$. *Chamamos este espaço vetorial de* ***espaço vetorial trivial****.*

Exemplo 1.4 *Seja* $(\mathbb{K}, +, \cdot)$ *um corpo e* $n \in \mathbb{N}$. *Denotamos por* $\mathbb{K}^n := \underbrace{\mathbb{K} \times \mathbb{K} \times \cdots \times \mathbb{K}}_{n \text{ vezes}}$ *o conjunto das* n*-uplas de elementos em* $\mathbb{K}$. *Definimos as operações* $+ : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}^n$ *e* $\cdot : \mathbb{K} \times \mathbb{K}^n \to \mathbb{K}^n$ *por:*

$$(x_1, ..., x_n) + (y_1, ..., y_n) := (x_1 + y_1, ..., x_n + y_n)$$

e

$$k \cdot (x_1, ..., x_n) := (k \cdot x_1, ..., k \cdot x_n)$$

para $(x_1, ..., x_n), (y_1, ..., y_n) \in \mathbb{K}^n$ *e* $k \in \mathbb{K}$. *Verifica-se facilmente que* $(\mathbb{K}^n, +, \cdot)$ *é um* $\mathbb{K}$*-espaço vetorial.*

Exemplo 1.5 *Sejam $\mathbb{K}$ um corpo e $m, n \in \mathbb{N}$. Uma matriz de ordem $m \times n$ com entradas em $\mathbb{K}$ é uma função $A : \{1, 2, ..., m\} \times \{1, 2, ..., n\} \to \mathbb{K}$. Para cada $i \in \{1, 2, ..., m\}$ e $j \in \{1, 2, ..., n\}$ definimos $A_{ij} := A(i, j)$. A notação usual de matriz mostra o conjunto imagem de A organizado em uma tabela como abaixo:*

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}.$$

O conjunto de todas as matrizes com ordem $m \times n$ (m linhas e n colunas) e entradas em $\mathbb{K}$ é denotado por $M_{m\times n}(\mathbb{K})$ ou as vezes por $\mathbb{K}^{m\times n}$. Geralmente escrevemos $A = (A_{ij})$. Para definir uma matriz, basta então definir cada uma de suas entradas. Dadas $A, B \in M_{m\times n}(\mathbb{K})$ e $k \in \mathbb{K}$ definimos as matrizes $A + B$ e $k \cdot A$ pelas suas entradas por

$$(A + B)_{ij} := A_{ij} + B_{ij}$$

e

$$(k \cdot A)_{ij} := k \cdot A_{ij}$$

para cada $i \in \{1, 2, ..., m\}$ e $j \in \{1, 2, ..., n\}$. É claro que $A + B \in M_{m\times n}(\mathbb{K})$ e $k \cdot A \in M_{m\times n}(\mathbb{K})$. Verifica-se facilmente que $(M_{m\times n}(\mathbb{K}), +, \cdot)$ é um $\mathbb{K}$-espaço vetorial. No caso particular $m = n$ escrevemos $M_{n\times n}(\mathbb{K})$ simplesmente como $M_n(\mathbb{K})$ e dizemos que seus elementos são matrizes quadradas de ordem n. Além dessas operações, existe uma operação de produto entre matrizes com ordens específicas. Se $A \in M_{m\times n}(\mathbb{K})$ e $B \in M_{n\times p}(\mathbb{K})$, podemos definir $A \cdot B \in M_{m\times p}(\mathbb{K})$ por

$$(A \cdot B)_{ij} := A_{i1} \cdot B_{1j} + A_{i2} \cdot B_{2j} + ... + A_{in} \cdot B_{nj} = \sum_{k=1}^{n} A_{ik} \cdot B_{kj}.$$

para cada $i \in \{1, 2, ..., m\}$ e $j \in \{1, 2, ..., p\}$. Verifica-se que sempre que fizer sentido, o produto de matrizes é associativo e distributivo com relação a soma de matrizes e associativo com o produto por escalares em $\mathbb{K}$. Em particular, para cada $n \in \mathbb{N}$, o produto de matrizes está bem definido em $M_n(\mathbb{K})$, porém, se $A, B \in M_n(\mathbb{K})$, então em geral $A \cdot B \neq B \cdot A$. Se $n \in \mathbb{N}$, escreveremos $Id_n \in M_n(\mathbb{K})$ a matriz $(Id_n)_{ij} = 1$ se $i = j$ e $(Id_n)_{ij} = 0$ se $i \neq j$. A matriz Id_n é chamada matriz ***identidade*** *de ordem n. Dizemos que duas matrizes $A, B \in M_n(\mathbb{K})$ são* ***inversas*** *se $A \cdot B = B \cdot A = Id_n$. Nesse caso dizemos que A (ou B) é* ***invertível*** *e denotamos $B =: A^{-1}$ (note que a inversa, se existir, é única). O conjunto das matrizes invertíveis de ordem n é denotado por $GL(n, \mathbb{K})$.*

Exemplo 1.6 *Sejam $\mathbb{K}$ um corpo e X um conjunto não vazio. Denotamos por $\mathcal{F}(X, \mathbb{K})$ o conjunto de todas as funções de X em $\mathbb{K}$. Dadas $f, g \in \mathcal{F}(X, \mathbb{K})$ e $k \in \mathbb{K}$, definimos para cada $x \in X$*

$$(f + g)(x) := f(x) + g(x),$$

e

$$(k \cdot f)(x) := k \cdot f(x).$$

Então $f + g, k \cdot f \in \mathcal{F}(X, \mathbb{K})$. Verifica-se facilmente que $(\mathcal{F}(X, \mathbb{K}), +, \cdot)$ é um $\mathbb{K}$-espaço vetorial.

2 Aula 2 - Paleari

Daqui em diante $\mathbb{K}$ sempre denotará um corpo. Se V é um $\mathbb{K}$-espaço vetorial, o elemento neutro do grupo $(V, +)$ sempre será denotado por 0.

Definição 2.1 *Seja $(V, +, \cdot)$ um $\mathbb{K}$-espaço vetorial. Dizemos que um sub-conjunto $S \subset V$ é um sub-espaço vetorial de V se as seguintes condições são cumpridas:*

- *$0 \in S$;*
- *S é fechado para adição: Para todos $v, w \in S$, vale $v + w \in S$;*
- *S é fechado para produtos por escalares: Para todos $v \in S$ e $k \in S$ vale $k \cdot v \in S$.*

Informalmente, $S \subset V$ é um sub-espaço vetorial é não-vazio e as operações do espaço vetorial ambiente se restringem a S.

Proposição 2.1 *Se $S \subset V$ é um sub-espaço vetorial, então $(S, +, \cdot)$ é um $\mathbb{K}$-espaço vetorial, em que as operações usadas em S são as herdadas do ambiente V.*

Demonstração: Basta perceber que as propriedades do ambiente V valem em qualquer subconjunto de V, logo basta que as operações se restrinjam a S. ∎

Definição 2.2 *Um **sistema linear** com m equações e n incógnitas em $\mathbb{K}$ é uma equação da forma $A \cdot X = b$, em que $A \in M_{m\times n}(\mathbb{K})$, $X \in M_{n\times 1}(\mathbb{K})$ e $b \in M_{m\times 1}(\mathbb{K})$ com A e b matrizes conhecidas e X uma matriz a ser determinada. Se $A = (A_{ij})$,*

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix},$$

então a equação $A \cdot X = b$ pode ser descrita como o conjunto de m equações abaixo:

$$\begin{cases} A_{11}x_1 & + & \cdots & + & A_{1n}x_n & = & b_1 \\ \vdots & & & & & & \vdots \\ A_{m1}x_1 & + & \cdots & + & A_{mn}x_n & = & b_m \end{cases}$$

Por hora identificaremos elementos de $\mathbb{K}^n$ com elementos de $M_{1\times n}(\mathbb{K})$, pois basicamente são n elementos de $\mathbb{K}$ dispostos ou em uma linha ou em uma coluna dependendo da interpretação. Mais tarde uma maneira mais formal de explicar essa identificação (e outras futuras) será apresentada.

Exemplo 2.1 (Soluções de Sistemas Lineares) *Definimos o conjunto solução do sistema linear $A \cdot X = b$ como o conjunto $S = \{X \in \mathbb{K}^n; A \cdot X = b\}$. No caso $b = 0$, dizemos que o sistema linear $A \cdot X = 0$ é **homogêneo**. Verifique que S é um sub-espaço vetorial de $\mathbb{K}^n$ se, e somente se, $b = 0$.*

Exemplo 2.2 (Matrizes Especiais) *Seja $A \in M_{m\times n}(\mathbb{K})$. Definimos $A^t \in M_{n\times m}(\mathbb{K})$ como a matriz*

$$(A^t)_{ij} := A_{ji}$$

*para cada $i \in \{1, 2, ..., m\}$ e $j \in \{1, 2, ..., n\}$. Chamamos A^t de matriz **transposta** de A. Se $\mathbb{K} = \mathbb{C}$ e $A \in M_{m\times n}(\mathbb{C})$, definimos $A^* \in M_{n\times m}(\mathbb{C})$, chamada a matriz **adjunta** de A, como $(A^*)_{ij} := \overline{A_{ji}}$, em que a barra indica o complexo conjugado do número A_{ji}. Dizemos que:*

- *$A \in M_n(\mathbb{C})$ é **normal** se $A \cdot A^* = A^* \cdot A$;*

- $A \in M_n(\mathbb{C})$ *é* ***auto-adjunta*** *se* $A^* = A$;
- $A \in M_n(\mathbb{C})$ *é* ***anti auto-adjunta*** *se* $A^* = -A$.
- $A \in M_n(\mathbb{C})$ *é* ***unitária*** *se* A *é invertível e* $A^{-1} = A^*$;
- $A \in M_n(\mathbb{R})$ *é* ***simétrica*** *se* $A = A^t$;
- $A \in M_n(\mathbb{R})$ *é* ***anti-simétrica*** *se* $A^t = -A$.
- $A \in M_n(\mathbb{R})$ *é* ***ortogonal*** *se* A *é invertível e* $A^{-1} = A^t$.

Em cada item acima, tome o conjunto das matrizes em $M_n(\mathbb{K})$ que possuem a propriedade indicada naquele item, verifique se esse conjunto é um sub-espaço vetorial de $M_n(\mathbb{K})$.

Exemplo 2.3 (Mais matrizes especiais) *Dizemos que uma matriz* $A \in M_{m\times n}(\mathbb{K})$ *é* ***triangular-superior*** *se* $A_{ij} = 0$ *sempre que* $i > j$. *Em outras palavras, todos os elementos da matriz abaixo da diagonal principal (elementos da forma* A_{ii}*) são todos nulos. Dizemos que* A *é* ***estritamente triangular-superior*** *se* $A_{ij} = 0$ *sempre que* $i \geq j$. *Definições análogas para matrizes triangulares-inferiores e estritamente triangulares-inferiores. Dizemos que* $A \in M_n(\mathbb{K})$ *é* ***diagonal*** *se* $A_{ij} = 0$ *sempre que* $i \neq j$, *isto é,* A *é triangular-superior e inferior ao mesmo tempo. Verifique que* $T_n = \{A \in M_n(\mathbb{K}); A \text{ é triangular-inferior}\}$ *e* $D_n = \{A \in M_n(\mathbb{K}); A \text{ é diagonal}\}$ *são sub-espaços vetoriais.*

Exemplo 2.4 *Seja* $\mathbb{K}$ *um corpo infinito. Denotamos por* $\mathbb{K}[x] \subset \mathcal{F}(\mathbb{K},\mathbb{K})$ *o conjunto das funções polinomiais de uma variável em* $\mathbb{K}$, *isto é,* $p \in \mathbb{K}[x]$ *se existem escalares* $a_0, a_1, ..., a_n \in \mathbb{K}$, *com* $n \in \mathbb{N}$ *sendo o maior natural para o qual* $a_n \neq 0$, *tais que* $p(x) = a_0 + a_1 \cdot x + ... + a_n x^n$ *para todo* $x \in \mathbb{K}$. *Pode-se mostrar nesse caso que a função* p *está completamente determinada pelos coeficientes* $a_0, a_1, ..., a_n$. *O número* n *está então bem definido e é chamado o grau do polinômio* p. *Verifique que* $\mathbb{K}[x]$ *é um sub-espaço vetorial de* $\mathcal{F}(\mathbb{K},\mathbb{K})$. *Para cada* $n \in \mathbb{K}$ *definimos* $\mathcal{P}_n(\mathbb{K}) \subset \mathbb{K}[x]$ *como o conjunto de todos os polinômios de grau menor do que ou igual a* n. *Verifique que* $\mathcal{P}_n(\mathbb{K})$ *também é um sub-espaço vetorial de* $\mathbb{K}[x]$.

Definição 2.3 *Sejam* X *e* Λ *conjuntos não vazios. Uma* ***família de elementos*** *de* X *indexada por* Λ *é uma função* $\Lambda \to V$. *A família é dita finita se o conjunto* Λ *for finito.*

Isto é, uma tal família consiste de uma escolha de um elemento de X para cada $\lambda \in \Lambda$, pensando como se o conjunto Λ fosse um conjunto de índices. Se para cada $\lambda \in \Lambda$ o elemento escolhido é denotado por $x_\lambda \in X$, podemos escrever a família usando a notação $(x_\lambda)_{\lambda\in\Lambda}$. No caso $\Lambda = \mathbb{N}$ a família é chamada uma **sequência** em V.

Definição 2.4 *Sejam* V *um* $\mathbb{K}$*-espaço vetorial e* $S \subset V$ *um sub-conjunto. Uma* ***combinação linear*** *de elementos de* S *é qualquer elemento da forma*

$$\sum_{\lambda\in\Lambda} a_\lambda x_\lambda$$

com $(x_\lambda)_{\lambda\in\Lambda}$ *uma família* ***finita*** *de elementos em* S *e* $a_\lambda \in \mathbb{K}$ *para todo* $\lambda \in \Lambda$.

Por exemplo, se $\Lambda = \{1, 2\}$, uma família de elementos em V indexada em Λ consiste simplesmente de um conjunto de até dois elementos $S = \{v_1, v_2\}$ de V. Uma combinação linear de elementos de S é então um elemento de V da forma $a_1v_1 + a_2v_2$ com $a_1, a_2 \in \mathbb{K}$. Generalize para $\Lambda = \{1, 2, ..., n\}$. E se $\Lambda = \mathbb{N}$?

Se V é um $\mathbb{K}$-espaço vetorial e $S \subset V$ é um sub-espaço vetorial, então toda combinação linear de elementos de S ainda resulta em um elemento de S por definição de sub-espaço vetorial (use indução na quantidade de elementos na combinação linear em conjunto com associatividade da adição).

Definição 2.5 *Seja* $S \subset V$ *um sub-conjunto qualquer. O* ***espaço gerado*** *por* S *é o conjunto de todas as possíveis combinações lineares de elementos de* S, *e é denotado por* **span** S.

Exemplo 2.5 *Se* $S = \{v_1, ..., v_n\} \subset V$*, isto é,* S *é uma família indexada no conjunto* $I_n = \{1, 2, ..., n\}$*, então*

$$\textbf{\textit{span}}\ S = \left\{ \sum_{k=1}^{n} a_k \cdot v_k, a_1, ..., a_n \in \mathbb{K} \right\}.$$

Exercício 2.1 *Verifique que para qualquer* $S \subset V$*, o conjunto* ***span*** S *é um subespaço vetorial de* V*.*

Exercício 2.2 *Verifique se* $S \subset V$ *já é um sub-espaço vetorial de* V*, então* $S =$ ***span*** S*.*

Definição 2.6 *Dizemos que um sub-conjunto* $S \subset V$ *é* ***linearmente independente*** *se a única maneira de escrever o vetor* 0 *como combinação linear de elementos de* S *é aquela em que todos os escalares são nulos. Em símbolos, toda vez que*

$$\sum_{\lambda \in \Lambda} a_\lambda x_\lambda = 0$$

com $(x_\lambda)_{\lambda \in \Lambda}$ *família finita de elementos de* S *e* $a_\lambda \in \mathbb{K}$ *para todo* $\lambda \in \Lambda$*, implicar* $a_\lambda = 0$ *para todo* $\lambda \in \Lambda$*. Se* S *não for linearmente independente, dizemos que* S *é* ***linearmente dependente****.*

Exemplo 2.6 *Sejam* $n \in \mathbb{N}$ *e* $V = \mathbb{K}^n$*. Para cada* $j = 1, 2, ..., n$ *definimos* $e_j = (0, ..., 0, 1, 0, ..., 0)$*, em que o elemento neutro* 1 *aparece exatamente na posição* j *e todas as outras entradas são nulas. O conjunto* $S = \{e_1, ..., e_n\}$ *é linearmente independente (verifique).*

Exemplo 2.7 *Sejam* $m, n \in \mathbb{N}$ *e* $V = M_{m \times n}(\mathbb{K})$*. Para cada* $i = 1, 2, ..., m$ *e* $j = 1, 2, ..., n$ *definimos* $E_{ij} \in M_{m \times n}(\mathbb{K})$ *como a matriz tal que* $(E_{ij})_{kl} = 1$ *se* $i = k$ *e* $j = l$ *e* $(E_{ij})_{kl} = 0$ *caso contrário. Isto é,* E_{ij} *é a matriz de ordem* $m \times n$ *que possui* 1 *na posição* ij *e* 0 *nas restantes. O conjunto* $S = \{E_{ij}, i = 1, 2, ..., m, j = 1, 2, ..., n\}$ *é linearmente independente (verifique).*

Exemplo 2.8 *Sejam* $\mathbb{K}$ *um corpo infinito e* $V = \mathbb{K}[x]$*. Para cada* $n \in \mathbb{N}_0$ *considere o polinômio* p_n *dado por* $p_n(x) = x^n$*. Então o conjunto* $S = \{p_n\}_{n \in \mathbb{N}_0}$ *é linearmente independente (verifique).*

Exercício 2.3 *Sejam* V *um* $\mathbb{K}$*-espaço vetorial e* $S_1, S_2 \subset V$ *sub-espaços vetoriais de* V*. Verifique que o sub-conjunto* $S_1 \cap S_2$ *ainda é um sub-espaço vetorial de* V*. Mostre entretanto que, em geral,* $S_1 \cup S_2$ *não é um sub-espaço vetorial de* V*.*

Definição 2.7 *Se* S_1 *e* S_2 *são sub-espaços vetoriais de* V*, definimos o* ***sub-espaço soma*** *de* S_1 *com* S_2 *como*

$$S_1 + S_2 := \textbf{\textit{span}}\ (S_1 \cup S_2).$$

Proposição 2.2 *Para cada par de sub-espaços vetoriais* S_1 *e* S_2 *de* V *vale*

$$S_1 + S_2 = \{u + v; u \in S_1 \textbf{ e } v \in S_2\}.$$

Demonstração: Dado $w \in S_1 + S_2 =$ **span** $(S_1 \cup S_2)$ segue da definição que $w = \sum_{\lambda \in \Lambda} a_\lambda x_\lambda$ com $(x_\lambda)_{\lambda \in \Lambda}$ uma família finita de elementos em $S_1 \cup S_2$ e $a_\lambda \in \mathbb{K}$ para todo $\lambda \in \Lambda$. Dividimos o conjunto de índices como uma união disjunta $\Lambda = \Lambda_1 \cup \Lambda_2$ tal que $x_\lambda \in S_i$ sempre que $\lambda \in \Lambda_i$ (mesmo que S_1 e S_2 tenham interseção, isso é possível), para $i = 1, 2$. Podemos escrever então $w = \sum_{\lambda \in \Lambda_1} a_\lambda x_\lambda + \sum_{\lambda \in \Lambda_2} a_\lambda x_\lambda$. Defina $\tilde{u} := \sum_{\lambda \in \Lambda_i} a_\lambda x_\lambda$ e $\tilde{v} := \sum_{\lambda \in \Lambda_2} a_\lambda x_\lambda$. Como S_1 é um sub-espaço vetorial de V, temos que $\tilde{u} \in S_1$. Analogamente, como S_2 é um sub-espaço vetorial de V, temos que $\tilde{v} \in S_2$. Logo $w = \tilde{u} + \tilde{v} \in \{u + v; u \in S_1 \text{ e } v \in S_2\}$. Reciprocamente, dado $w \in \{u + v; u \in S_1 \text{ e } v \in S_2\}$, segue que existem $\tilde{u} \in S_1$ e $\tilde{v} \in S_2$ tais que $w = \tilde{u} + \tilde{v}$. Como $\{\tilde{u}, \tilde{v}\} \subset S_1 \cup S_2$ e $w = 1 \cdot \tilde{u} + 1 \cdot \tilde{v}$ segue que w é uma combinação linear de elementos em $S_1 \cup S_2$, logo $w \in S_1 + S_2$. Concluímos então a igualdade do enunciado. ∎

Definição 2.8 *Quando* $V = S_1 + S_2$ *e* $S_1 \cap S_2 = \{0\}$*, dizemos que* V *é a* ***soma direta*** *de* S_1 *e* S_2*. Nesse caso escrevemos* $V = S_1 \oplus S_2$*.*

3 Aula 3 - Paleari

Definição 3.1 *Seja V um $\mathbb{K}$-espaço vetorial. Dizemos que um sub-conjunto $S \subset V$ é* ***gerador*** *de V se* ***span*** *$S = V$.*

Revisite os exemplos 2.6, 2.7 e 2.8 e verifique que os conjuntos S definidos em cada um deles são geradores do espaço vetorial em questão.

Observação 3.1 *Seja V um $\mathbb{K}$-espaço vetorial não trivial e $v \in V$, $v \neq 0$. Então o conjunto unitário $S = \{v\}$ é linearmente independente. De fato, escrever o vetor nulo como combinação linear de elementos de S significa a existência de um escalar $k \in \mathbb{K}$ tal que $k \cdot v = 0$. Porém, como $v \neq 0$, devemos ter $k = 0$, logo o único escalar da combinação deve ser nulo e portanto S é linearmente independente.*

Proposição 3.1 *Sejam V um $\mathbb{K}$-espaço vetorial e $S \subset V$ um sub-conjunto linearmente independente. Seja $v \in V$ tal que $v \notin$* ***span*** *S. Então $\tilde{S} := S \cup \{v\}$ é linearmente independente.*

Demonstração: Suponha que o vetor 0 possa ser escrito como uma combinação linear de elementos em $\tilde{S}$, isto é, que existam escalares $a_1, ..., a_n, b \in \mathbb{K}$ e vetores $v_1, ..., v_n \in S$ tais que

$$a_1v_1 + ... + a_nv_n + bv = 0.$$

Se $b \neq 0$, então podemos escrever

$$v = \left(-\frac{a_1}{b}\right)v_1 + ... + \left(-\frac{a_n}{b}\right)v_n,$$

o que implicaria $v \in$ **span** S, o que é uma contradição pela hipótese. Logo $b = 0$ e portanto $a_1v_1 + ... + a_nv_n = 0$. Como por hipótese S é linearmente independente, isso implica $a_1 = ... = a_n = 0$. Logo $a_1 = ... = a_n = b = 0$ e assim $\tilde{S}$ é linearmente independente. ∎

A observação e a proposição acima combinadas nos dizem que em qualquer espaço vetorial não trivial existem conjuntos linearmente independentes e que toda vez que um conjunto linearmente independente S é tal que **span** $S \subsetneq V$, então S está contido em outro conjunto linearmente independente $\tilde{S}$ estritamente maior. Assim, podemos incrementar conjuntos linearmente independentes enquanto eles não forem geradores de V. A questão é: esse processo eventualmente vai produzir um conjunto linearmente independente que também seja gerador de V? A resposta a essa pergunta consiste no chamado **Lema de Zorn**, que é uma proposição equivalente ao chamado **Axioma da Escolha**. Supor a validade do Axioma da Escolha é uma opção e nos direciona para a teoria (ZFC) (Zermelo-Fraenkel com Axioma da Escolha). Grande parte dos textos em matemática contemporânea optam por esse caminho, mas ele está longe de ser obrigatório. Cursos padrões de Álgebra Linear (como o nosso) seguem esse caminho, pois, como veremos a seguir, supor a validade do Axioma da Escolha nos permite provar que todo espaço-vetorial possui base, uma ferramenta essencial no estudo da Álgebra linear.

Sejam V um $\mathbb{K}$-espaço vetorial não trivial e $\mathcal{P}(V)$ o conjunto das partes de V. Este conjunto é parcialmente ordenado pela inclusão $\subset$. O mesmo vale então para $L(V) \subset \mathcal{P}(V)$ o conjunto de todos os sub-conjuntos linearmente independentes de V. Como já vimos, $L(V) \neq \varnothing$. Seja $\mathcal{S} = (S_\lambda)_{\lambda \in \Lambda}$ uma família totalmente ordenada de elementos em $L(V)$, isto é, todo par de elementos $S_\lambda, S_{\tilde{\lambda}}$ em $\mathcal{S}$ é comparável: $S_\lambda \subset S_{\tilde{\lambda}}$ ou $S_{\tilde{\lambda}} \subset S_\lambda$. Defina $M_{\mathcal{S}} := \cup_{\lambda \in \Lambda} S_\lambda$. Isto é, $M_{\mathcal{S}}$ é a união de todos os conjuntos da família $\mathcal{S}$. O fato da família $\mathcal{S}$ ser totalmente ordenada implica que $M_{\mathcal{S}}$ é um sub-espaço vetorial de V. De fato, sejam $u, v \in M_{\mathcal{S}}$ e $k \in \mathbb{K}$. Então existem $\lambda, \tilde{\lambda} \in \Lambda$ tais que $u \in S_\lambda$ e $v \in \tilde{\lambda}$. Como $\mathcal{S}$ é totalmente ordenado, podemos supor sem perda de generalidade que $S_\lambda \subset S_{\tilde{\lambda}}$. Assim $u, v \in S_{\tilde{\lambda}}$. Como $S_{\tilde{\lambda}}$ é um sub-espaço vetorial de V, segue que $u + kv \in S_{\tilde{\lambda}} \subset M_{\mathcal{S}}$. Além disso, é claro que $0 \in M_{\mathcal{S}}$, de modo que $M_{\mathcal{S}}$ é então um sub-espaço vetorial de V, ou seja, $M_{\mathcal{S}} \in L(V)$. Por definição, $S_\lambda \subset M_{\mathcal{S}}$ para todo $\lambda \in \Lambda$. Isso nos diz que $M_{\mathcal{S}}$ é um supremo para a família totalmente ordenada $\mathcal{S}$. Assim, mostramos que toda conjunto totalmente ordenado em $L(V)$ possui um supremo em $L(V)$. Pelo Lema de Zorn, $L(V)$ possui um elemento maximal $\mathcal{B} \in L(V)$, isto é, se $S \in L(V)$ e $\mathcal{B} \subset S$, então $\mathcal{B} = S$. Afirmamos que **span** $\mathcal{B} = V$. De fato, suponha por absurdo que **span** $\mathcal{B} \subsetneq V$, poderíamos tomar $v \in V$ com $v \notin$ **span** $\mathcal{B}$. Pela proposição 3.1, o conjunto $\mathcal{B} \cup \{v\}$ é linearmente independente, ou seja, $\mathcal{B} \cup \{v\} \in L(V)$ e $\mathcal{B} \subsetneq \mathcal{B} \cup \{v\}$, o que é uma contradição com a maximalidade de $\mathcal{B}$. Logo **span** $\mathcal{B} = V$.

Definição 3.2 *Uma **base** (de Hamel) para um $\mathbb{K}$-espaço vetorial V é um conjunto linearmente independente $\mathcal{B} \subset V$ que é gerador de V.*

Teorema 3.1 *[com ZFC] Todo $\mathbb{K}$-espaço vetorial V possui uma base.*

Revisite mais uma vez os exemplos 2.6, 2.7 e 2.8 e verifique que o conjunto S em cada um daqueles exemplos é base do espaço vetorial correspondente.

Proposição 3.2 *Seja V um $\mathbb{K}$-espaço vetorial e $A \subset V$ um conjunto linearmente independente. Então existe uma base $\mathcal{B}$ de V tal que $A \subset \mathcal{B}$.*

Demonstração: Exercício! Modifique a demonstração do Teorema 3.1 para famílias de conjuntos linearmente independentes que contenham A.

Definição 3.3 *Sejam V um $\mathbb{K}$-espaço vetorial e $S \subset V$ um sub-espaço vetorial. Um **complemento** para S é qualquer sub-espaço vetorial $\tilde{S} \subset V$ tal que $V = S \oplus \tilde{S}$.*

Proposição 3.3 *Todo sub-espaço vetorial possui um complemento.*

Demonstração: Basta tomar uma base $\mathcal{B} = \{v_\lambda; \lambda \in \Lambda\}$ de S e completar para uma base $\tilde{\mathcal{B}} = \{v_\lambda; \lambda \in \Lambda\} \cup \{\tilde{v}_\mu; \mu \in M\}$ de V e definir $\tilde{S} :=$ **span** $\{\tilde{v}_\mu; \mu \in M\}$. ∎

Proposição 3.4 *Seja V um $\mathbb{K}$-espaço vetorial, $A \subset V$ um conjunto linearmente independente com n elementos e $B \subset V$ um conjunto gerador de V com m elementos. Então $n \leq m$.*

Demonstração: Escreva $A = \{a_1, ..., a_n\}$ e $B = \{b_1, ..., b_m\}$. Como A é linearmente independente, em particular temos $a_i \neq 0$ para todo $i = 1, 2, ..., n$. Suponha por absurdo que $n > m$. Como **span** $B = V$, o vetor a_1 pode ser escrito como combinação linear dos elementos de B. Nem todos os coeficientes nesta combinação linear podem ser nulos pois isso implicaria $a_1 = 0$, logo algum deles deve ser não nulo. Reordenando os $b's$ se necessário, podemos supor que o coeficiente não nulo é o que acompanha b_1 na combinação linear. Isso implica que b_1 pode ser escrito como combinação linear de a_1 e de $b_2, ..., b_m$. Como $V =$ **span** B, segue que $V =$ **span** $\{a_1, b_2, ..., b_m\}$. Em particular, podemos escrever agora o vetor a_2 como combinação linear de $a_1, b_2, ..., b_m$. Se nessa combinação todos os coeficientes dos $b's$ fossem nulos, a_2 seria obtido a partir de a_1, o que contradiz A ser linearmente independente. Assim, algum coeficientes dos $b's$ nessa combinação deve ser nulo, e novamente reordenando se necessário, podemos supor que o coeficiente de b_2 seja não nulo. Isso implica que b_2 pode ser escrito como combinação linear de $a_1, a_2, b_3, ..., b_m$. Assim $V =$ **span** $\{a_1, a_2, b_3, ..., b_m\}$. Continuamos procedendo desta forma até que consigamos eliminar todos os $b's$, o que é possível pois estamos supondo $n > m$, de modo que $V =$ **span** $\{a_1, ..., a_m\}$. Mas isso é uma contradição pois $m < n$ e A é linearmente independente (por exemplo, a_n seria combinação linear dos demais $a's$). Logo $n \leq m$.

Proposição 3.5 *Seja V um $\mathbb{K}$-espaço vetorial, $A \subset V$ um conjunto linearmente independente e $B \subset V$ um conjunto gerador de V. Então $\#A \leq \#B$.*

Demonstração: A proposição 3.4 garantiu que o resultado vale no caso em que ambos A e B são finitos. Se A é finito e B é infinito, não há nada o que fazer. Se A é infinito, então B não pode ser finito. De fato, caso B fosse finito com, digamos, n elementos, podemos tomar um conjunto linearmente independente com $n + 1$ elementos $\{a_1, ..., a_{n+1}\} \subset A$. Isto é, B é gerador de V com n elementos e $\{a_1, ..., a_{n+1}\} \subset V$ é um conjunto linearmente independente com mais elementos do que B, uma contradição com a proposição 3.4. Logo podemos supor que ambos A e B são infinitos. Suponha por absurdo que $\#A > \#B$. Pela proposição 3.2 podemos completar A para uma base de V, o que resulta em um conjunto com cardinalidade maior do que ou igual a de A. Assim, podemos supor sem perda de generalidade que A já é uma base de V. Escreva $A = \{a_i; i \in I\}$, com $\#A = \#I$ e $B = \{b_j; j \in J\}$ com $\#B = \#J$. Para cada $j \in J$, como A é uma base de V, podemos escrever $b_j = \sum_{i \in E_j} \lambda_{ij} a_i$ para certos escalares $\lambda_{ij} \in \mathbb{K}$ e $E_j \subset I$ um conjunto finito. Como J é infinito, a cardinalidade de $\bigcup_{j \in J} E_j$ é igual a de J, logo $\# \bigcup_{j \in J} E_j < \#I$, o que implica $\bigcup_{j \in J} E_j \subsetneq I$. Assim, existe $i_i \in I$ que não pertence a E_j para todo $j \in J$. Como B gera V, o elemento da base correspondente a_{i_0} pode ser escrito como combinação linear dos elementos de B, e cada um deles, por sua vez, pode ser escrito como combinação linear dos a_i com $i \in \bigcup_{j \in J} E_j$, de modo que $\{a_{i_0}\} \cup \{a_i; i \in \bigcup_{j \in J} E_j\}$ é linearmente dependente, o que contradiz o fato de A ser linearmente independente. ∎

Corolário 3.1 (Dimensão) *Sejam V um $\mathbb{K}$-espaço vetorial e $\mathcal{B}_1$ e $\mathcal{B}_2$ duas bases de V. Então $\#\mathcal{B}_1 = \#\mathcal{B}_2$.*

Definição 3.4 *Seja V um $\mathbb{K}$-espaço vetorial. A* ***dimensão*** *de V é a cardinalidade de qualquer base de V. A dimensão é denotada por $\dim V$. Dizemos que V possui dimensão finita se V possui uma base com uma quantidade finita de elementos, caso contrário dizemos que V possui dimensão infinita e escrevemos $\dim V = \infty$.*

Exemplo 3.1 *Se $m, n \in \mathbb{N}$, então $\dim \mathbb{K}^n = n$, $\dim M_{m\times n} = m \cdot n$ e $\dim \mathcal{P}_n(\mathbb{K}) = n + 1$. Por outro lado, $\dim \mathbb{K}[x] = \infty$.*

Suponha agora que V seja um $\mathbb{K}$-espaço vetorial de dimensão finita, digamos $n \in \mathbb{N}$. Seja $\mathcal{B} = \{v_1, ..., v_n\}$ uma base de V e $v \in V$. Como $\mathcal{B}$ é uma base de V, em particular, é um conjunto gerador, existem escalares $a_1, ..., a_n \in \mathbb{K}$ tais que $v = a_1v_1 + ... + a_nv_n$. Suponha por um momento que o mesmo vetor v possa ser escrito como combinação linear de $\mathcal{B}$ com outro conjunto de escalares. Assim, suponha que existam $b_1, ..., b_n \in \mathbb{K}$ tais que $v = b_1v_1 + ... + b_nv_n$. Assim

$$v = a_1v_1 + ... + a_nv_n = b_1v_1 + ... + b_nv_n,$$

o que implica

$$(a_1 - b_1)v_1 + ... + (a_n - b_n)v_n = 0.$$

Assim, o vetor nulo 0 pode ser escrito como uma combinação linear de $\mathcal{B}$. Como por hipótese $\mathcal{B}$ é também linearmente independente, segue que $(a_1 - b_1) = ... = (a_n - b_n) = 0$, isto é, $a_i = b_i$ para todo $i = 1, 2, ..., n$. Concluímos assim que quando escrevemos um vetor como combinação linear dos elementos de uma base de V, os escalares que acompanham cada vetor de $\mathcal{B}$ estão completamente determinados (não é possível escrever um mesmo vetor como combinação linear dos elementos de uma base fixada de duas formas diferentes). O único detalhe que temos que tomar cuidado é com a ordem com que os elementos da base estão dispostos. Como conjuntos, não se importa se escrevemos $\{v_1, v_2, v_3, ..., v_n\}$ ou $\{v_2, v_1, v_3, ..., v_n\}$, mas os escalares que acompanham cada vetor trocam de posição. Por isso, daqui em diante usaremos o termo $\mathcal{B}$ é uma **base ordenada** e escreveremos $\mathcal{B} =< v_1, ..., v_n >$ para indicar que a ordem com que os vetores da base aparecem já está fixada. Desta forma, os escalares acompanhando cada vetor seguirão esta ordem. Assim, se $v = b_1v_1 + ... + b_nv_n$, escreveremos

$$[v]_{\mathcal{B}} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

chamado o **vetor de coordenadas** de v na base ordenada $\mathcal{B}$.

Como se relacionam as coordenadas de um mesmo vetor mas em duas bases diferentes? Suponha agora que $\mathcal{C} = \{u_1, ..., u_n\}$ seja outra base de V. Note que podemos escrever cada um dos vetores $v_1, ..., v_n$ como combinação linear de $\mathcal{C}$. Assim, para cada $j = 1, 2, ..., n$ escreva $v_j = a_{1j}u_1 + a_{2j}u_2 + ... + a_{nj}u_n$. Com isso montamos a chamada

matriz mudança de base de $\mathcal{B}$ para $\mathcal{C}$ definida por $M_{\mathcal{B}}^{\mathcal{C}} := [a_{ij}]_{1 \leq i,j \leq n}$. Em outras palavras, as colunas de $M_{\mathcal{B}}^{\mathcal{C}}$ são os vetores de coordenadas de cada um dos vetores $v's$ na base $\mathcal{C}$:

$$M_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} | & | & & | \\ [v_1]_{\mathcal{C}} & [v_2]_{\mathcal{C}} & \cdots & [v_n]_{\mathcal{C}} \\ | & | & & | \end{bmatrix}$$

Se $v = b_1 v_1 + ... + b_n v_n$, substituindo a expressão de cada um dos $v's$ escritos como combinação linear dos $u's$ nessa igualdade, obtemos

$$\begin{aligned} v = b_1 v_1 + ... + b_n v_n &= b_1 \left(\sum_{k=1}^{n} a_{k1} u_k \right) + ... + b_n \left(\sum_{k=1}^{n} a_{kn} u_k \right) \\ &= (a_{11} b_1 + ... + a_{1n} b_n)\, u_1 + ... + (a_{n1} b_1 + ... + a_{nn} b_n)\, u_n \\ &= (M_{\mathcal{B}}^{\mathcal{C}} \cdot [v]_{\mathcal{B}})_{11} u_1 + ... + (M_{\mathcal{B}}^{\mathcal{C}} \cdot [v]_{\mathcal{B}})_{n1} u_n. \end{aligned}$$

Isto é

$$[v]_{\mathcal{C}} = \begin{bmatrix} (M_{\mathcal{B}}^{\mathcal{C}} \cdot [v]_{\mathcal{B}})_{11} \\ (M_{\mathcal{B}}^{\mathcal{C}} \cdot [v]_{\mathcal{B}})_{12} \\ \vdots \\ (M_{\mathcal{B}}^{\mathcal{C}} \cdot [v]_{\mathcal{B}})_{1n} \end{bmatrix} = M_{\mathcal{B}}^{\mathcal{C}} \cdot [v]_{\mathcal{B}}.$$

Assim, para obter as coordenadas de v na base $\mathcal{C}$, sabendo as coordenadas de v na base $\mathcal{B}$, multiplicamos a matriz que muda de base de $\mathcal{B}$ para $\mathcal{C}$ pelo vetor de coordenadas de v na base $\mathcal{B}$. Vamos apenas resumir o discutido acima como uma proposição.

Proposição 3.6 *Vale a relação*

$$[v]_{\mathcal{C}} = M_{\mathcal{B}}^{\mathcal{C}} \cdot [v]_{\mathcal{B}}.$$

4 Aula 4 - Paleari

Definição 4.1 *Sejam V e W dois $\mathbb{K}$-espaços vetoriais. Uma função $T : V \to W$ é uma **transformação linear** se T satisfaz as condições:*

- *$T(v_1 + v_2) = T(v_1) + T(v_2)$ para todos $v_1, v_2 \in V$;*
- *$T(k \cdot v) = k \cdot T(v)$ para todos $k \in \mathbb{K}$ e $v \in V$.*

Em suma, uma função entre dois espaços vetoriais é uma transformação linear se ela preserva as operações algébricas de espaços vetoriais. Antes de começarmos os exemplos, vamos discutir um pouco mais de teoria básica.

Exercício 4.1 *Sejam $T : V \to W$ uma transformação linear, $v_1, ..., v_n \in V$ e $a_1, ..., a_n \in \mathbb{K}$. Verifique que*

$$T(a_1 v_1 + ... + a_n v_n) = a_1 T(v_1) + ... + a_n T(v_n).$$

São perguntas básicas em qualquer área da matemática procurar saber quando certas funções são injetivas ou sobrejetivas. No contexto de álgebra linear não é diferente. Particularmente, a pergunta sobre quando uma transformação linear é injetiva é relativamente simples. Sejam $T : V \to W$ uma transformação linear e suponha que existam $v_1, v_2 \in V$ tais que $T(v_1) = T(v_2)$. Mas daí $T(v_1) - T(v_2) = 0$ e pelo exercício acima, teremos $T(v_1 - v_2) = 0$.

Definição 4.2 *Seja $T : V \to W$ uma transformação linear. Definimos o **Núcleo** de T (ou Kernel de T), como o conjunto*

$$Ker\ T = \{v \in V; T(v) = 0\}.$$

Exercício 4.2 *Verifique que $Ker\ T \subset V$ é um sub-espaço vetorial.*

Proposição 4.1 *Uma transformação linear $T : V \to W$ é injetiva se, e somente se, $Ker\ T = \{0\}$.*

Demonstração: Primeiro observe o seguinte fato geral. Seja qual for a transformação linear, $T(0) = 0$. De fato, por linearidade, $T(0) = T(0 \cdot 0) = 0 \cdot T(0) = 0$. Logo, se T é injetiva, então o único vetor cuja imagem é o vetor nulo de W deve ser o vetor nulo de V, ou seja, Ker $T = \{0\}$. Reciprocamente, suponha Ker $T = \{0\}$ e que existam $v_1, v_2 \in V$ tais que $T(v_1) = T(v_2)$. Então, como vimos, $v_1 - v_2 \in$ Ker T. Mas se Ker $T = \{0\}$, temos então $v_1 - v_2 = 0$, o que implica $v_1 = v_2$. Assim, T é injetiva. ∎

Exemplo 4.1 *Seja $T : \mathbb{R}^4 \to \mathbb{R}^3$ a função definida por $T(x_1, x_2, x_3, x_4) = (2x_1 + x_2, -x_1 + x_3 + x_4, 2x_4)$. É fácil verificar que T é uma transformação linear. Para encontrar $Ker\ T$ precisamos encontrar o conjunto dos $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ para os quais*

$$\begin{cases} 2x_1 + x_2 &= 0 \\ -x_1 + x_3 + x_4 &= 0 \\ 2x_4 &= 0 \end{cases}$$

Dá ultima equação obtemos $x_4 = 0$, e portanto $x_3 = x_1$ e $x_2 = -2x_1$, com $x_1 \in \mathbb{R}$. Isto é,

$$Ker\ T = \{(x_1, -2x_1, x_1, 0); x_1 \in \mathbb{R}\} = \textbf{\textit{span}}\ \{(1, -2, 1, 0)\}.$$

Logo T não é injetiva e $\dim Ker\ T = 1$.

Exemplo 4.2 *Sejam $n \in \mathbb{N}$ e $\mathrm{Tr} : M_n(\mathbb{K}) \to \mathbb{K}$ a função definida por $Tr(A) = \sum_{i=1}^{n} A_{ii}$ para cada $A = (A_{ij}) \in M_n(\mathbb{K})$, chamada a função **traço**. É fácil ver que Tr é uma transformação linear. O núcleo de Tr consiste então das matrizes $A = (A_{ij})$ para as quais $\sum_i A_{ii} = 0$. Se $n \geq 2$, temos então que*

$$Ker\ Tr = \left\{ A = (A_{ij}) \in M_n(\mathbb{K}); A_{nn} = -\sum_{i=1}^{n-1} A_{ii} \right\}.$$

Assim, $\dim Ker\ Tr = n^2 - 1$. Se $n = 1$, então $M_n(\mathbb{K}) = \mathbb{K}$ e $Tr(A) = A$ para todo $A \in \mathbb{K}$, logo $Ker\ Tr = \{0\}$ nesse caso.

Exemplo 4.3 *Seja $\mathbb{K}$ um corpo infinito. Para cada $p \in \mathbb{K}[x]$, $p(x) = a_0 + a_1x^1 + a_2x^2 + ... + a_nx^n$, definimos $p'(x) := a_1 + 2a_2x^1 + 3a_3x^2 + ... + na_nx^{n-1} \in \mathbb{K}[x]$ a **derivada** de p. Definimos daí a função $D : \mathbb{K}[x] \to \mathbb{K}[x]$, $D(p) = p'$. É fácil verificar que D é linear. Note que se $p \in \mathbb{K}[x]$ é tal que $D(p) = 0$, então $p(x) = a_0$, isto é, p precisa ter grau 0, deste modo Ker $D = \{p \in \mathbb{K}[x]; p(x) = a_0\} =$ **span** $\{1\}$.*

Mas e quanto a sobrejetividade? Se $T : V \to W$ é uma transformação linear, então T é sobrejetiva se Im $T = W$.

Exercício 4.3 *Seja $T : V \to W$ uma transformação linear. Mostre que Im $T \subset W$ é um sub-espaço vetorial de W.*

Exercício 4.4 *Discuta a sobrejetividade das transformações lineares dos exemplos dados acima.*

Definição 4.3 *Uma transformação linear $T : V \to W$ é dita um **isomorfismo** (se $\mathbb{K}$-espaços vetoriais) se T é bijetiva.*

Exercício 4.5 *Mostre que se $T : V \to$ é um isomorfismo, então $T^{-1} : W \to V$ também é uma transformação linear.*

Exemplo 4.4 *Seja $n \in \mathbb{N}$. A função $T : \mathbb{K}^n \to M_{n\times 1}(\mathbb{K})$ definida por*

$$T(x_1, ..., x_n) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

é um isomorfismo.

Exemplo 4.5 *Sejam V um $\mathbb{K}$-espaço vetorial de dimensão n e $\mathcal{B}$ uma base ordenada de V. Defina $C : V \to M_{1\times n}(\mathbb{K})$ por $C(v) = [v]_{\mathcal{B}}$. Então C é um isomorfismo.*

Exemplo 4.6 *Seja $n \in \mathbb{N}$. Defina $T : \mathbb{K}^{n+1} \to P_n(\mathbb{K})$ por $T(a_0, a_1, ..., a_n) := a_0 + a_1x + ... + a_nx^n$. Então T é um isomorfismo.*

Quando existe um isomorfismo entre os espaços vetoriais V e W, dizemos do ponto de vista de álgebra linear que esses espaços são o mesmo e denotamos este fato com o símbolo $V \cong W$. Como conjuntos eles podem ser completamente diferentes, mas se existir uma correspondência 1-1 entre seus elementos que respeite as operações de espaços vetoriais (adição de vetores e produto de vetores por escalares), podemos interpretá-los como o mesmo espaço vetorial. Em particular, espaços vetoriais isomorfos devem possuir a mesma dimensão.

Exercício 4.6 *Sejam $T : V \to W$ um isomorfismo de $\mathbb{K}$-espaços vetoriais e $\mathcal{B} = \{v_\lambda\}_{\lambda\in\Lambda}$ uma base de V. Mostre que $T(\mathcal{B}) := \{T(v_\lambda)\}_{\lambda\in\Lambda}$ é uma base de W. Em particular, como bijeções preservam cardinalidade, $\dim V = \dim W$.*

Há uma relação forte entre o núcleo e a imagem de uma transformação linear qualquer. O chamado Primeiro Teorema do Isomorfismo nos diz em essência que Ker T determina Im T e vice-versa. A relação é mais forte ainda no caso de dimensão finita. Informalmente, uma função identifica seu domínio com sua imagem, a menos da sua parte não injetiva, que resulta em redundâncias nessa identificação. Como vimos, no caso de transformações lineares toda informação sobre a falta de injetividade está no núcleo. Portanto, o domínio de T a menos de seu núcleo (como se todo o núcleo fosse apenas um ponto) será identificado com sua imagem. Mas para formalizar o que acabamos de discutir, precisamos definir quocientes de espaços vetoriais.

Sejam V um $\mathbb{K}$-espaço vetorial e $S \subset V$ um sub-espaço vetorial. Definimos em V uma relação de equivalência da seguinte forma: dizemos os vetores $v_1, v_2 \in V$ são equivalentes se $v_1 - v_2 \in S$. Esta relação em V é:

- Reflexiva: para todo $v \in V$, $v - v = 0 \in S$ pois S é um sub-espaço vetorial.
- Simétrica: Se $v_1, v_2 \in V$ são tais que $v_1 - v_2 \in S$, então $v_2 - v_1 = -(v_1 - v_2) \in S$ pois S é fechado para multiplicação por escalares.
- Transitiva: Se $v_1, v_2, v_3 \in V$ são tais que $v_1 - v_2 \in S$ e $v_2 - v_3 \in S$, então $v_1 - v_3 = (v_1 - v_2) + (v_2 - v_3) \in S$ pois é uma soma de vetores em S e este último é um sub-espaço vetorial.

Portanto essa relação é de fato de equivalência. Dado $v \in V$, denotamos sua classe de equivalência por $[v]$. O conjunto das classes de equivalência dessa relação será denotado por V/S, chamado o **espaço quociente** de V por S. Note, em particular, que $v \in S$ vale $[v] = [0]$ pois $v - 0 = v \in S$. Assim, todos os vetores de S são identificados com apenas um ponto em V/S. Acontece que V/S não é apenas um conjunto, mas também tem uma estrutura natural de $\mathbb{K}$-espaço vetorial, isto é, existe uma maneira natural de definir uma operação de adição e de produto por escalares em V/S.

Dados $[v_1], [v_2] \in V/S$ e $k \in \mathbb{K}$, gostaríamos de definir

$$[v_1] + [v_2] := [v_1 + v_2], \quad k \cdot [v_1] := [k \cdot v_1].$$

A leitura das definições acima é a seguinte: na adição, tomamos um representante de cada uma das classes de equivalência, somamos os representantes em V e tomamos a classe de equivalência da soma obtida. Similar para o produto por escalar. Entretanto, o processo envolve escolhas, o resultado final não pode depender da escolha dos representantes pois caso contrário não estaremos definindo uma função. Por isso, para garantirmos que está tudo bem, precisamos provar que a classe $[v_1 + v_2]$ não depende da escolha dos representantes v_1 de $[v_1]$ e v_2 de $[v_2]$. Similarmente para $[k \cdot v_1]$. Assim, suponha que $\tilde{v_1}$ e $\tilde{v_2}$ sejam outros representantes de $[v_1]$ e de $[v_2]$ respectivamente. Por definição da relação, isso nos diz que $v_1 - \tilde{v_1} = s_1 \in S$ e $v_2 - \tilde{v_2} = s_2 \in S$. Daí $(v_1 + v_2) - (\tilde{v_1} + \tilde{v_2}) = (v_1 - \tilde{v_1}) + (v_2 - \tilde{v_2}) = s_1 + s_2 \in S$ pois S é sub-espaço vetorial. Logo $[v_1 + v_2] = [\tilde{v_1} + \tilde{v_2}]$ e a operação de adição está bem definida. A prova é inteiramente análoga para o produto por escalares e será deixada como exercício. A partir daí, prova-se facilmente que essas operações assim definidas satisfazem todos os axiomas de espaço vetorial. Além disso, por definição da operação, a função

$$\begin{array}{rccc} \pi: & V & \to & V/S \\ & v & \mapsto & [v]. \end{array}$$

é uma transformação linear (sobrejetiva por definição). Esta transformação linear é chamada **projeção canônica**. Além disso, também por definição, $\operatorname{Ker} \pi = S$.

Proposição 4.2 *Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita e $S \subset V$ um sub-espaço vetorial de V. Então V/S tem dimensão finita e* $\dim V/S = \dim V - \dim S$.

Demonstração: Digamos que $\dim S = k$ e $\dim V = n$. Seja $\{v_1, ..., v_k\}$ uma base de S e seja $\{u_{k+1}, ..., u_n\} \subset V$ tal que $\{v_1, ..., v_k, u_{k+1}, ..., u_n\}$ seja base de V (3.2). Defina $\tilde{\mathcal{B}} := \{[u_{k+1}], ..., [u_n]\} \subset V/S$. Afirmamos que $\tilde{\mathcal{B}}$ é uma base de V/S. Para provarmos que é linearmente independente, suponha que existam $a_{k+1}, ..., a_n \in \mathbb{K}$ tais que $a_{k+1}[u_{k+1}] + ... + a_n[u_n] = [0]$, o que significa $[a_{k+1}u_{k+1} + ... + a_n u_n] = [0]$, logo $s := a_{k+1}u_{k+1} + ... + a_n u_n \in S$. Por outro lado, $\{v_1, ..., v_k\}$ é base de S, portanto existem escalares $a_1, ..., a_k \in \mathbb{K}$ tais que $s = a_1 v_1 + ... + a_k v_k$. Combinando as duas maneiras de escrever s obtemos

$$a_1 v_1 + ... + a_k v_k - a_{k+1} u_{k+1} - ... - a_n u_n = 0,$$

mas como $\{v_1, ..., v_k, u_{k+1}, ..., u_n\}$ é linearmente independente, segue em particular que $a_{k+1} = ... = a_n = 0$. Para ver que $\tilde{\mathcal{B}}$ é gerador de V/S seja $v \in V$ e considere sua classe $[v] \in V/S$. Como $\{v_1, ..., v_k, u_{k+1}, ..., u_n\}$ é gerador de V, existem escalares $a_1, ..., a_n \in \mathbb{K}$ tais que $v = a_1 v_1 + ... + a_k v_k + a_{k+1} u_{k+1} + ... + a_n v_n$. Como $\{v_1, ..., v_k\} \subset S$ e $\operatorname{Ker} \pi = S$, temos que

$$[v] = [a_1 v_1 + ... + a_k v_k + a_{k+1} u_{k+1} + ... + a_n v_n] = a_{k+1}[u_{k+1}] + ... + a_n[u_n],$$

e assim $\tilde{\mathcal{B}}$ é gerador de V/S. Em particular, $\dim V/S = n - k = \dim V - \dim S$. ∎

Teorema 4.1 (Teorema do Isomorfismo) *Seja $T : V \to W$ uma transformação linear. Então a função $\tilde{T}$: $V/Ker\, T \to Im\, T$ definida por $T([v]) := T(v)$ está bem definida e é um isomorfismo de $\mathbb{K}$-espaços vetoriais.*

Demonstração: Se $[v_1] = [v_2]$, então $v_1 - v_2 = s \in \text{Ker } T$, logo $T(v_1) - T(v_2) = T(v_1 - v_2) = T(s) = 0$, assim $T(v_1) = T(v_2)$. Isto é, $\tilde{T}$ está bem definida pois sua definição não depende da escolha de um representante de cada classe. A função $\tilde{T}$ é claramente linear, segue facilmente da definição das operações em $V/\text{Ker } T$ e de que T é linear. Além disso, se $T([v]) = T(v) = 0$, então $v \in \text{Ker } T$ e portanto $[v] = [0]$, logo $\tilde{T}$ é injetiva. Finalmente, $\tilde{T}$ é obviamente sobrejetiva. ∎

Corolário 4.1 (Teorema Nucleo-Imagem) *Seja $T : V \to W$ uma transformação linear entre espaços vetoriais de dimensão finita. Então*

$$\dim V = \dim \text{Ker } T + \dim \text{Im } T.$$

Corolário 4.2 *Seja $T : V \to W$ uma transformação linear entre espaços vetoriais de dimensão finita.*

- *Se T é injetiva, então $\dim V \leq \dim W$.*
- *Se T é sobrejetiva, então $\dim V \geq \dim W$.*
- *Se T é um isomorfismo, então $\dim V = \dim W$.*
- *Se $\dim V = \dim W$, então T é injetiva se, e somente se, é sobrejetiva.*

5 Aula 5 - Paleari

Seja $T : V \to W$ uma transformação linear entre espaços vetoriais de dimensão finita. Suponha que $\mathcal{B} =< v_1, ..., v_n >$ seja uma base ordenada de V e que $\mathcal{C} =< w_1, ..., w_m >$ base uma base ordenada de W. Dado qualquer vetor $v \in V$ podemos encontrar escalares $a_1, ..., a_n \in \mathbb{K}$ tais que $v = a_1v_1 + ... + a_nv_n$. Da linearidade de T, temos que

$$T(v) = T(c_1v_1 + ... + c_nv_n) = c_1T(v_1) + ... + c_nT(v_n).$$

Em outras palavras, para conhecer o valor de T em um vetor arbitrário, é suficiente conhecer os valores de T nos elementos de alguma base do domínio de T. Por sua vez, cada $T(v_j)$ é completamente determinado pelas suas coordenadas na base ordenada $\mathcal{C}$. Isso nos permite caracterizar qualquer transformação linear, dadas uma base do domínio e uma do contradomínio, por um conjunto de números, o qual será organizado em uma matriz, como veremos abaixo. Para cada $j = 1, 2..., n$ escreva

$$T(v_j) = a_{1j}w_1 + ... + a_{mj}w_m,$$

e defina

$$[T]_{\mathcal{B}}^{\mathcal{C}} := [a_{ij}] = \begin{bmatrix} | & | & & | \\ [T(v_1)]_{\mathcal{C}} & [T(v_2)]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} \\ | & | & & | \end{bmatrix},$$

chamada a **matriz de T no par de bases $\mathcal{B}$ e $\mathcal{C}$**. Quando $V = W$ e $\mathcal{B} = \mathcal{C}$, escreveremos simplesmente $[T]_{\mathcal{B}} := [T]_{\mathcal{B}}^{\mathcal{B}}$ e chamamos esta de matriz de T na base $\mathcal{B}$.

Voltando ao caso geral, note que se $v = c_1v_1 + ... + c_nv_n$, então

$$\begin{aligned} T(v) &= c_1(a_{11}w_1 + ... + a_{m1}w_m) + ... + c_n(a_{1n}w_1 + ... + a_{mn}w_m) \\ &= (a_{11}c_1 + ... + a_{1n}c_n)w_1 + ... + (a_{m1}c_1 + ... + a_{mn}c_n)w_n \end{aligned}$$

Assim,

$$[T(v)]_{\mathcal{C}} = [T]_{\mathcal{B}}^{\mathcal{C}} \cdot [v]_{\mathcal{B}}. \tag{1}$$

Proposição 5.1 *Sejam V e W espaços vetoriais, $\{v_\lambda\}_{\lambda\in\Lambda}$ uma base de V e $\{w_\lambda\}_{\lambda\in\Lambda}$ uma família de vetores em W indexada no mesmo conjunto Λ. Então existe uma única transformação linear $T : V \to W$ tal que $T(v_\lambda) = w_\lambda$ para todo $\lambda \in \Lambda$.*

Demonstração: Como cada elemento $v \in V$ pode ser escrito de maneira única como $v = \sum_{\lambda\in\Lambda'} a_\lambda v_\lambda$ com $a_\lambda \in \mathbb{K}$ e $\Lambda' \subset \Lambda$ finito, e queremos que $T : V \to W$, devemos ter

$$T(v) = T\left(\sum_{\lambda\in\Lambda'} a_\lambda v_\lambda\right) = \sum_{\lambda\in\Lambda'} a_\lambda T(v_\lambda) = \sum_{\lambda\in\Lambda'} a_\lambda w_\lambda.$$

A fórmula na última expressão acima define T completamente e fica claro que é o único jeito de satisfazer as condições requeridas da proposição. ∎

Exemplo 5.1 *Sejam $V = W = \mathbb{R}^2$, $a \in \mathbb{R}$ e $r \subset \mathbb{R}^2$ a reta de equação $y = ax$. Afirmamos que existe uma transformação linear $R_r : \mathbb{R}^2 \to \mathbb{R}^2$ tal que para cada $v \in \mathbb{R}^2$, o vetor $R_r(v)$ é a reflexão de v com respeito a reta r. Como discutido acima, basta descobrir o que R_r faz com os elementos de algum base de $\mathbb{R}^2$. A base mais simples de todas certamente é a base canônica $\mathcal{C} =< e_1, e_2 >$, porém, como não sabemos quem é $a \in \mathbb{R}$, não conhecemos $R_r(e_1)$ nem $R_r(e_2)$. Precisamos de uma base mais adequada e que "visualize" o que essa transformação faz. Defina $u := (1, a)$. Note que $u \in \mathbb{R}^2$ é um vetor diretor da reta r. Além disso, como esse vetor mora em cima da reta, temos que $R_r(u) = u$. Mas precisamos de um segundo vetor, linearmente independente com u, tal que também saibamos o que R_r faz. Uma escolha natural é tomar um vetor perpendicular a u. Podemos por exemplo definir*

$v = (-a, 1)$. Note que $\mathcal{B} :=< u, v >$ é uma base de $\mathbb{R}^2$ e que $R_r(v) = -v$. Desta forma, $R_r(u) = u = 1 \cdot u + 0 \cdot v$ e $R_r(v) = -v = 0 \cdot u + (-1) \cdot v$ e portanto

$$[R_r]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Assim, se soubermos escrever um dado $(x, y) \in \mathbb{R}^2$ como combinação linear de $\mathcal{B}$, fica muito fácil descobrir quem é $R_r(x, y)$. Mas podemos já aproveitar o que foi descoberto e escrever explicitamente uma fórmula para $R_r(x, y)$. Basta encontrarmos as coordenadas de (x, y) na base $\mathcal{B}$ e aplicarmos a transformação:

$$(x, y) = \lambda u + \mu v = \lambda(1, a) + \mu(-a, 1) = (\lambda - a\mu, a\lambda + \mu),$$

o que nos leva ao sistema linear

$$\begin{cases} \lambda - a\mu & = & x \\ a\lambda + \mu & = & y \end{cases}$$

cuja solução é $\lambda = \frac{x+ay}{a^2+1}$ e $\mu = \frac{y-ax}{a^2+1}$. Logo

$$\begin{aligned} R_r(x, y) &= R_r(\lambda u + \mu v) \\ &= \lambda R_r(u) + \mu R_r(v) \\ &= \lambda u - \mu v \\ &= \left(\frac{x + ay}{a^2 + 1}\right)(1, a) + \left(\frac{y - ax}{a^2 + 1}\right)(a, -1) \\ &= \frac{1}{a^2 + 1}\big((1 - a^2)x + 2ay, 2ax + (a^2 - 1)y\big). \end{aligned}$$

Em particular,

$$[R_r]_{\mathcal{C}} = \frac{1}{a^2 + 1}\begin{bmatrix} (1 - a^2) & 2a \\ 2a & (a^2 - 1) \end{bmatrix}.$$

Suponha agora que U, V, W sejam $\mathbb{K}$-espaços vetoriais de dimensão finita, $T : U \to V$, $S : V \to W$ sejam transformações lineares e $\mathcal{B}, \mathcal{C}$ e $\mathcal{D}$ sejam bases de U, V e W respectivamente. Podemos formar a composição $S \circ T : U \to W$. É fácil ver que $S \circ T$ é também uma transformação linear.

Proposição 5.2 *Vale a relação*

$$[S \circ T]^{\mathcal{D}}_{\mathcal{B}} = [S]^{\mathcal{D}}_{\mathcal{C}} \cdot [T]^{\mathcal{C}}_{\mathcal{B}}.$$

Demonstração: Escreva $\mathcal{B} =< u_1, ..., u_n >$, $\mathcal{C} =< v_1, ..., v_m >$ e $\mathcal{D} =< w_1, ..., w_p >$. Por definição de matriz de transformação, a coluna j de $[S \circ T]^{\mathcal{D}}_{\mathcal{B}}$ é $[(S \circ T)(u_j)]_{\mathcal{D}}$. Por outro lado, pela relação (1), temos que

$$[(S \circ T)(u_j)]_{\mathcal{D}} = [S(T(u_j))]_{\mathcal{D}} = [S]^{\mathcal{D}}_{\mathcal{C}} \cdot [T(u_j)]_{\mathcal{C}} = [S]^{\mathcal{D}}_{\mathcal{C}} [T]^{\mathcal{C}}_{\mathcal{B}} \cdot [u_j]_{\mathcal{B}} = [S]^{\mathcal{D}}_{\mathcal{C}} [T]^{\mathcal{C}}_{\mathcal{B}} \cdot e^t_j.$$

Mas dada uma matriz A, a matriz $A \cdot e^t_j$ é justamente a coluna j de A e portanto o resultado está provado. ∎

Proposição 5.3 *Seja V um $\mathbb{K}$-espaço vetorial de dimensão finita, $T : V \to V$ uma transformação linear e $\mathcal{B}, \mathcal{C}$ duas bases de V. Então vale a relação*

$$[T]_{\mathcal{B}} = M^{\mathcal{B}}_{\mathcal{C}} \cdot [T]_{\mathcal{C}} \cdot M^{\mathcal{C}}_{\mathcal{B}}.$$

Demonstração: Escreva $\mathcal{B} =< u_1, ..., u_n >$ e seja $j \in \{1, 2, ..., n\}$. A coluna j de $[T]_{\mathcal{B}}$ é por definição $[T(u_j)]_{\mathcal{B}}$. Por outro lado,

$$\begin{aligned} M^{\mathcal{B}}_{\mathcal{C}} \cdot [T]_{\mathcal{C}} \cdot M^{\mathcal{C}}_{\mathcal{B}} \cdot e^t_j &= M^{\mathcal{B}}_{\mathcal{C}} \cdot [T]_{\mathcal{C}} \cdot [u_j]_{\mathcal{C}} \\ &= M^{\mathcal{B}}_{\mathcal{C}} \cdot [T(u_j)]_{\mathcal{C}} \\ &= [T(u_j)]_{\mathcal{B}}. \end{aligned}$$

Portanto as colunas de ambas as matrizes são iguais e o resultado está provado. ∎

Repare que todo espaço vetorial V possui ao menos duas transformações lineares canônicas associadas a ele. Definimos $\mathrm{Id}_V : V \to V$ por $\mathrm{Id}_V(v) = v$ para todo $v \in V$, chamada a **transformação identidade** de V. É claro que se trata de uma transformação linear. Se V possui dimensão finita e $\mathcal{B}$ e $\mathcal{C}$ são bases de V, note que

$$[\mathrm{Id}_V]^{\mathcal{C}}_{\mathcal{B}} = M^{\mathcal{C}}_{\mathcal{B}}.$$

Em particular, $[\mathrm{Id}_V]^{\mathcal{B}}_{\mathcal{B}} = \mathrm{Id}_{\dim V}$.

Proposição 5.4 *Sejam V e W espaços vetoriais de dimensão finita, $\mathcal{B}$ e $\mathcal{C}$ bases ordenadas de V e W respectivamente. Se $T : V \to W$ é um isomorfismo, então*

$$[T^{-1}]^{\mathcal{B}}_{\mathcal{C}} = \left([T]^{\mathcal{C}}_{\mathcal{B}}\right)^{-1}.$$

Demonstração: Como $T^{-1} \circ T = \mathrm{Id}_V$, obtemos $\mathrm{Id}_{\dim V} = [\mathrm{Id}_V]_{\mathcal{B}} = [T^{-1}]^{\mathcal{B}}_{\mathcal{C}} \cdot [T]^{\mathcal{C}}_{\mathcal{B}}$. ∎

Corolário 5.1 *Seja V um $\mathbb{K}$-espaço vetorial de dimensão finita e $\mathcal{B}$ e $\mathcal{C}$ bases de V. Então*

$$M^{\mathcal{B}}_{\mathcal{C}} = \left(M^{\mathcal{C}}_{\mathcal{B}}\right)^{-1}.$$

A segunda transformação linear canônica é $0 : V \to W$ definida por $0(v) = 0$ para todo $v \in V$, chamada **transformação nula**. É claro que se trata de uma transformação linear. Se V e W possuem dimensão finita, então $[0]^{\mathcal{C}}_{\mathcal{B}} = 0$ sejam quais forem as bases $\mathcal{B}$ e $\mathcal{C}$ de V e W respectivamente.

Se V e W são $\mathbb{K}$-espaços vetoriais, $T, S : V \to W$ são transformações lineares e $k \in \mathbb{K}$, podemos definir

$$\begin{array}{rccc} T + S : & V & \to & W \\ & v & \mapsto & T(v) + S(v), \end{array}$$

e

$$\begin{array}{rccc} k \cdot T : & V & \to & W \\ & v & \mapsto & k \cdot T(v) \end{array}.$$

É fácil ver que $S + T$ e $k \cdot T$ também são transformações lineares. Denotamos por $L(V, W)$ o conjunto de todas as transformações linear de V em W. Com as operações definidas acima, verifica-se que $L(V, W)$ é também um $\mathbb{K}$-espaço vetorial.

Proposição 5.5 *Se V e W tem dimensão finita, então $L(V, W)$ também tem dimensão finita e* $\dim L(V, W) = \dim V \cdot \dim W$.

Demonstração: Fixe uma base $\mathcal{B} =< v_1, ..., v_n >$ de V e $\mathcal{C} =< w_1, ..., w_m >$ uma base de W. Para cada $i \in \{1, ..., n\}$ e $j \in \{1, ..., m\}$ defina $T_{ij} : V \to W$ a única transformação linear cuja ação na base $\mathcal{B}$ é dada por $T_{ij}(v_k) = \delta_{ik} w_j$ para todo $k = 1, ..., n$. Seja agora $T : V \to W$ uma transformação linear qualquer. Para cada $k \in \{1, ..., n\}$ escreva

$$T(v_k) = a_{1k} w_1 + ... + a_{mk} w_m.$$

Note que

$$T(v_k) = a_{1k} w_1 + ... + a_{mk} w_m = a_{1k} T_{k1}(v_k) + ... + a_{mk} T_{km}(v_k) = (a_{1k} T_{k1} + ... + a_{mk} T_{km})(v_k).$$

Como $T_{ij}(v_k) = 0$ se $i \neq k$, para todo j, temos

$$T(v_k) = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} T_{ij} \right) (v_k)$$

para todo $k \in \{1, ..., n\}$, logo $\{T_{ij}\}_{i,j}$ gera $L(V, W)$. É fácil verificar que $\{T_{ij}\}_{i,j}$ é também linearmente independente. ■

Proposição 5.6 *Se $T, S : V \to W$ são transformações lineares entre espaços vetoriais de dimensão finita, $\mathcal{B}$ é base de V, $\mathcal{C}$ é base de W e $k \in \mathbb{K}$, então*

$$[T + k \cdot S]_{\mathcal{B}}^{\mathcal{C}} = [T]_{\mathcal{B}}^{\mathcal{C}} + k \cdot [S]_{\mathcal{B}}^{\mathcal{C}}.$$

Exercício 5.1 *Prove a proposição acima.*

6 Aula 6 - Paleari

Definição 6.1 *Seja V um $\mathbb{K}$-espaço vetorial. Definimos $V^* := L(V, \mathbb{K})$, chamado o **espaço dual** de V. Cada elemento de V^* é chamado um **funcional linear** em V.*

Se V possui dimensão finita e $\mathcal{B} =< v_1, ..., v_n >$ é uma base de V, então a demonstração da proposição 8.1 fornece uma base para V^*. Tomamos $\mathcal{C} =< 1 >$ como base de $\mathbb{K}$ (note que $m = 1$) e renomeamos as transformações lá definidas por $v^i := T_{i1} \in V^*$, $i = 1, ..., n$. Isto é, $v^i : V \to \mathbb{K}$ é o único funcional linear tal que $v^i(v_j) = \delta_{ij}$ para todo $j = 1, ..., n$. A base $\mathcal{B}^* :=< v^1, ..., v^n > \subset V^*$ é chamada **base dual** de $\mathcal{B}$. Em particular, $\dim V = \dim V^*$.

Exemplo 6.1 *Seja $V = \mathbb{K}^n$ e $\mathcal{C} =< e_1, ..., e_n >$ a base canônica de $\mathbb{K}^n$. Dado $f \in (\mathbb{K}^n)^*$ e $u = (x_1, ..., x_n) \in \mathbb{K}^n$, temos*

$$f(u) = f(x_1e_1 + ... + x_ne_n) = x_1f(e_1) + ... + x_nf(e_n).$$

Se escrevermos $a_j := f(e_j)$ para cada $j = 1, ..., n$, então os elementos $a_1, ..., a_n \in \mathbb{K}$ são constantes que determinam f. Assim, todo funcional linear em $(\mathbb{K}^n)^*$ é da forma

$$f(x_1, ..., x_n) = a_1x_1 + ... + a_nx_n,$$

*com $a_1, ..., a_n \in \mathbb{K}$ constantes. Além disso, pelo Teorema Núcleo-Imagem, ou $f = 0$ ou $\dim \mathrm{Ker}\, f = n - 1$. Um sub-espaço vetorial de $\mathbb{K}^n$ com dimensão $n - 1$ é chamado um **hiper-plano** em $\mathbb{K}^n$.*

Seja $T : V \to W$ uma transformação linear. A transformação T induz uma função entre os duais desses espaços, mas no sentido contrário. Defina $T^t : W^* \to V^*$ colocando $T^t(f) := f \circ T$ para cada $f \in W^*$. Como composição de transformações lineares ainda é linear e f toma valores em $\mathbb{K}$, segue que de fato $T^t(f) \in V^*$. Além disso, T^t é linear pois se $f, g \in W^*$, $k \in \mathbb{K}$ e $v \in V$, então

$$\begin{aligned} T^t(f + k \cdot g)(v) &= (f + k \cdot g)(T(v)) \\ &= f(T(v)) + k \cdot g(T(v)) \\ &= T^t(f)(v) + k \cdot T^t(g)(v) \\ &= (T^t(f) + k \cdot T^t(g))(v). \end{aligned}$$

Logo $T^t(f + k \cdot g) = T^t(f) + k \cdot T^t(g)$. A transformação linear T^t é chamada **aplicação transposta** de T. A próxima proposição justifica o nome dado a essa aplicação.

Proposição 6.1 *Seja $T : V \to W$ uma transformação linear entre espaços vetoriais de dimensão finita e sejam $\mathcal{B}$ e $\mathcal{C}$ bases de V e W respectivamente. Então*

$$[T^t]^{\mathcal{B}^*}_{\mathcal{C}^*} = \left([T]^{\mathcal{C}}_{\mathcal{B}}\right)^t.$$

Demonstração: Escreva $\mathcal{B} =< v_1, ..., v_n >$ e $\mathcal{C} =< w_1, ..., w_m >$ e seja $j \in \{1, ..., m\}$. Temos que escrever $T^t(w^j) = w_j \circ T \in V^*$ como combinação linear de $\mathcal{B}^*$. Escreva $T^t(w^j) = b_{1j}v^1 + ... + b_{nj}v^n$. Para cada $i \in \{1, ..., n\}$, temos que

$$T^t(w^j)(v_i) = w^j(T(v_i)) = w^j(a_{1i}w_1 + ... + a_{mi}w_m) = a_{ji}.$$

Por outro lado, $(b_{1j}v^1 + ... + b_{nj}v^n)(v_i) = b_{ij}$. Logo $b_{ij} = a_{ji}$ para todos i, j e a relação está provada. ∎

Dada $T : V \to W$, será que é possível explicitar $\mathrm{Ker}\, T^t \subset W^*$ e $\mathrm{Im}\, T^t \subset V^*$? Note que $f \in \mathrm{Ker}\, T^t$ significa $f \circ T = 0$, isto, é, $f(T(v)) = 0$ para todo $v \in V$. Em outras palavras, f **anula** a imagem de T. Reciprocamente, se $f|_{\mathrm{Im}\, T} = 0$, então para todo $v \in V$ temos $f(T(v)) = 0$, ou seja, $T^t(f) = f \circ T = 0$. Isso sugere a seguinte definição.

Definição 6.2 *Seja V um $\mathbb{K}$-espaço vetorial e $S \subset V$ um subconjunto qualquer. Definimos o **anulador** de S como*

$$S^0 := \{f \in V^*; f(v) = 0\ \forall v \in V\}.$$

Exercício 6.1 *Mostre que mesmo que S não seja um sub-espaço vetorial de V, o anulador de S é sempre um sub-espaço vetorial de V^*.*

Proposição 6.2 *Se $T : V \to W$ é uma transformação linear, então:*

a) $\text{Ker } T^t = (\text{Im } T)^0$*;*

b) $\text{Im } T^t = (\text{Ker } T)^0$.

Demonstração: O item a) já foi mostrado acima, resta mostrarmos o item b). Dados $f \in W^*$ e $v \in \text{Ker } T$, temos $(T^t f)(v) = f(T(v)) = f(0) = 0$ pois $v \in \text{Ker } T$, logo $T^t(f) \in (\text{Ker } T)^0$, assim $\text{Im } T^t \subset (\text{Ker } T)^0$. Seja agora $g \in (\text{Ker } T)^0$, queremos definir $f \in W^*$ tal que $g = f \circ T$. Seja $(v_\lambda)_{\lambda \in \Lambda}$ uma base de Ker T e $(\tilde{v}_\mu)_{\mu \in M}$ tal que $\{v_\lambda, \tilde{v}_\mu\}_{\lambda,\mu}$ seja base de V. Segue então que $(T(\tilde{v}_\mu))_{\mu \in M}$ é uma base de $\text{Im } T \subset W$. Seja agora $(w_\alpha)_{\alpha \in A}$ tal que $\mathcal{C} = \{T(\tilde{v})_\mu, w_\alpha\}_{\mu,\alpha}$ seja base de W. Defina $f \in W^*$ nos elementos da base $\mathcal{B}$ colocando $f(T(\tilde{v}_\mu)) := g(\tilde{v}_\mu)$ para cada $\mu \in M$ e $f(w_\alpha) := 0$ para cada $\alpha \in A$ (aqui usamos a proposição 5.1). Pela própria definição de f fica claro que $g = f \circ T$, isto é, $g \in \text{Im } T^t$ e assim $(\text{Ker } T)^0 \subset \text{Im } T^t$. ∎

Podemos seguir um passo adiante e tomar o espaço de todos os funcionais lineares definidos em V^*.

Definição 6.3 *Seja V um $\mathbb{K}$-espaço vetorial. Definimos o* ***bidual*** *de V como o espaço vetorial $(V^*)^* = \{F : V^* \to \mathbb{K}; F$ é linear$\}$.*

Apesar de parecer abstrato, o bidual $(V^*)^*$ está relacionado de uma maneira bem próxima com o próprio V. Dado $v \in V$, definimos $F_v : V^* \to \mathbb{K}$ a função dada por $F_v(f) := f(v)$ para cada $f \in V^*$. Afirmamos que F_v é um funcional linear. De fato, dados $f, g \in V^*$ e $k \in \mathbb{K}$, então

$$F_v(f + k \cdot g) = (f + k \cdot g)(v) = f(v) + k \cdot g(v) = F_v(f) + k \cdot F_v(g).$$

Assim, $F_v \in (V^*)^*$ para cada $v \in V$. A aplicação F_v é também chamada de **morfismo de avaliação em** v.

Teorema 6.1 *Seja V um $\mathbb{K}$-espaço vetorial. A aplicação*

$$\begin{array}{rccc} F: & V & \to & (V^*)^* \\ & v & \mapsto & F_v \end{array}$$

é linear e injetiva.

Demonstração: Para verificar a linearidade, sejam $v_1, v_2 \in V$ e $k \in \mathbb{K}$. Então

$$\begin{aligned} F(v_1 + kv_2)(f) &= F_{v_1+kv_2}(f) \\ &= f(v_1 + kv_2) \\ &= f(v_1) + kf(v_2) \\ &= F_{v_1}(f) + kF_{v_2}(f) \\ &= (F_{v_1} + fF_{v_2})(f) \end{aligned}$$

para todo $f \in V^*$, logo $F(v_1 + kv_2) = F(v_1) + kF(v_2)$. Agora suponha que exista $v \in V$, $v \neq 0$, tal que $F(v) = F_v = 0$. Isto é, para todo $f \in V^*$, vale $F_v(f) = f(v) = 0$. Como $v \neq 0$, podemos completar para uma base $\mathcal{B} = \{v, v_\alpha\}_{\alpha \in \Lambda}$ de V. Se $\mathcal{B}^* = \{g, v^\alpha\}_{\alpha \in \Lambda}$ é a base dual de $\mathcal{B}$, então $g(v) = 1$, uma contradição com $f(v) = 0$ para todo $f \in V^*$, logo F é injetiva. ∎

Corolário 6.1 *Se V é um $\mathbb{K}$-espaço vetorial de dimensão finita, então a aplicação $F : V \to (V^*)^*$ definida acima é um isomorfismo.*

Demonstração: Como vimos anteriormente, se V tem possui dimensão finita, então $\dim V = \dim V^*$. Pelo mesmo motivo, temos que $\dim V^* = \dim(V^*)^*$, logo $F : V \to (V^*)^*$ é uma transformação linear injetiva entre espaços vetoriais de mesma dimensão, e portanto um isomorfismo. ∎

Assim, para um espaço vetorial de dimensão finita, o seu bidual é **naturalmente** isomorfo ao próprio espaço. O termo "naturalmente"é usado para indicar que o isomorfismo definido não depende de nenhuma escolha de bases.

Para encerrar, vejamos com qual objeto um sub-espaço vetorial $S \subset V$ se identifica em $(V^*)^*$.

Proposição 6.3 *Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita e $S \subset V$ um sub-espaço vetorial. Então* $\dim V = \dim S + \dim S^0$.

Demonstração: Seja $< v_1, ..., v_k >$ uma base de S e complete para uma base $\mathcal{B} =< v_1, ..., v_k, v_{k+1}, ..., v_n >$ de V. Seja $\mathcal{B}^* =< v^1, ..., v^n >\subset V^*$ a base dual de $\mathcal{B}$. Afirmamos que $\mathcal{C} =< v^{k+1}, ..., v^n >$ é uma base de S^0. Primeiramente, segue diretamente da definição de base dual que $\mathcal{C} \subset S^0$. Agora, dado $f \in S^0$, como $\mathcal{B}^*$ é uma base de V^*, existem escalares $a_1, ..., a_n \in \mathbb{K}$ tais que

$$f = a_1 v^1 + ... + a_n v^n.$$

Como $f \in S^0$, temos que $f(v_j) = 0$ para todo $j \in \{1, ..., k\}$. Por outro lado, temos que

$$f(v_j) = a_1 v^1(v_j) + ... + a_k v^k(v_j) + a_{k+1} v^{k+1}(v_j) + ... + a_n v^n(v_k) = a_j.$$

Logo $a_1 = ... = a_k = 0$ e assim $f = a_{k+1} v^{k+1} + ... + a_n v^n \in$ **span** $\mathcal{C}$. É claro que $\mathcal{C}$ é linearmente independente, portanto $\mathcal{C}$ é base de S^0 e $\dim V = n = k + (n - k) = \dim S + \dim S^0$. ∎

Proposição 6.4 *Seja V um $\mathbb{K}$-espaço vetorial, $S \subset V$ um sub-espaço vetorial e $F : V \to (V^*)^*$ a transformação linear definida em 6.1. Então, $F(S) \subset (S^0)^0$.*

Demonstração: Sejam $v \in S$ e $f \in S^0$. Por definição de F, temos $F(v)(f) = F_v(f) = f(v) = 0$ pois $f \in S^0$ e $v \in S$. Logo F_v anula todos os elementos em S^0, ou seja, $F(v) \in (S^0)^0$. ∎

Corolário 6.2 *Se V possui dimensão finita e $S \subset V$ é um sub-espaço vetorial, então $F(S) = (S^0)^0$. Em particular,* $\dim S = \dim(S^0)^0$.

Demonstração: Já vimos que $F(S) \subset (S^0)^0$, então basta mostrar que $\dim F(S) = \dim(S^0)^0$. Como F é um isomorfismo, temos que $\dim F(S) = \dim(S)$. Pela proposição 6.3, temos que $\dim V = \dim S + \dim S^0$ e, pelo mesmo motivo, $\dim V^* = \dim S^0 + \dim(S^0)^0$. Como $\dim V = \dim V^*$, combinando as igualdades obtemos $\dim S = \dim(S^0)^0$ como queríamos. ∎

7 Aula 7 - Paleari

Nessa aula vamos discutir formalmente o conceito de determinante de uma matriz quadrada de ordem qualquer.

Leitura Complementar

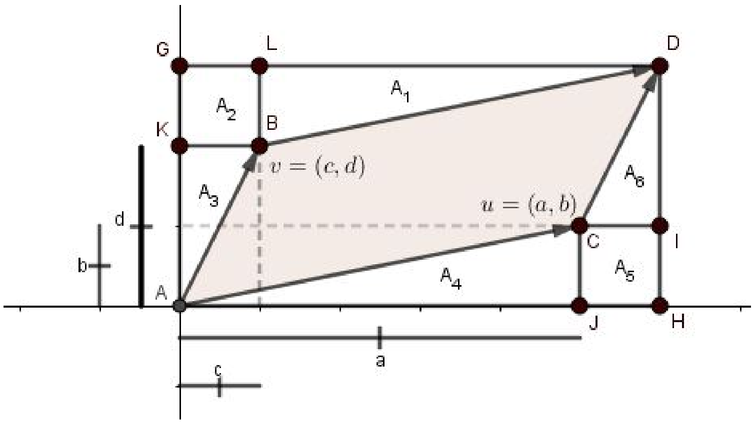
O caso trivial é a definição de determinante de matrizes 1×1. Neste caso, $\mathbb{K}^{1\times 1} = \mathbb{K}$ e definimos $\det : \mathbb{K}^{1\times 1} \to \mathbb{K}$ por $\det(A) = A$ para toda $A \in \mathbb{K}^{1\times 1}$. Por ora vamos usar a definição de determinante que o leitor já conhece para os casos de ordem 2 e 3, e vamos usá-los como motivação para discutir a definição geral em seguida. No caso $n = 2$, se

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix},$$

definimos

$$\det A := ad - bc.$$

O determinante 2×2 possui uma importante interpretação geométrica. Suponha que $u = (a, b)$ e $v = (c, d)$ sejam vetores linearmente independentes em $\mathbb{R}^2$ (ou seja, não são múltiplos). Deste modo, $\{u, v\}$ formam um paralelogramo no plano. Qual é a área desse paralelogramo em função das coordenadas de u e v?



A figura acima mostra o paralelogramo gerado pelos vetores u e v. Podemos calcular a área do paralelogramo $ABCD$ completando a figura com o retângulo $AHDG$, cujos lados medem $a + c$ e $b + d$ e retirar as 6 áreas menores indicadas na figura. Assim

$$\begin{aligned}
\textbf{Área } ABCD &= (a+c)(b+d) - 2 \cdot \frac{ab}{2} - 2 \cdot bc - 2 \cdot \frac{cd}{2} \\
&= ab - ab + bc - 2 \cdot bc + ad + cd - cd \\
&= ad - bc \\
&= \det \begin{bmatrix} a & c \\ b & d \end{bmatrix}.
\end{aligned}$$

Assim, $\det A$ calcula a área do paralelogramo gerado pelos vetores u e v colocados nas colunas da matriz A (nessa ordem). Se trocarmos a posição dos vetores nas colunas da matriz A note que

$$\det \begin{bmatrix} c & a \\ d & b \end{bmatrix} = bc - ad = -\det \begin{bmatrix} a & c \\ b & d \end{bmatrix},$$

isto é, o sinal do determinante muda. Isto está relacionado com o fato de que $\{u, v\}$ como dispostos na figura formam uma base com a mesma orientação da base canônica de $\mathbb{R}^2$ e determinantes são sensíveis a orientação de bases. Além disso, se $\{u, v\}$ fosse linearmente dependente, então nesse caso existe $k \in \mathbb{R}$ tal que $v = k \cdot u = (ka, kb)$. Daí,

$$\det \begin{bmatrix} a & ka \\ b & kb \end{bmatrix} = abk - abk = 0.$$

De fato, se u e v são múltiplos não há paralelogramo formado, portanto sua área deve ser 0. Assim, o fato do determinante 2×2 ser 0 ou não identifica se os vetores nas colunas de A são linearmente dependentes ou não. Desta forma, podemos pensar $\det : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ como uma função de 2 variáveis onde cada variável é um vetor que será colocado na coluna correspondente da matriz cujo determinante será calculado. Se $u, v \in \mathbb{R}^2$, vimos que $\det(u, v) = -\det(v, u)$ e que se $\{u, v\}$ é linearmente dependente, então $\det(u, v) = 0$. Mas temos mais um ponto importante a ser discutido, que a fórmula do det nos permite observar e será essencial. Se fixarmos por exemplo a segunda variável (deixarmos fixo $v = (c, d)$), e pensarmos a função resultante apenas como função de $u = (a, b)$, temos um funcional linear: $(a, b) \mapsto d \cdot a + (-c) \cdot b$. A observação análoga vale se fixarmos u e pensarmos a função resultante como função da variável v, teremos outro funcional linear. Assim, det pode ser pensado como uma função de duas variáveis, que é linear em cada variável e que troca de sinal quando as variáveis são trocadas de posição. Além disso, note que $\det(e_1, e_2) = 1$.

Caminhando agora para o caso $n = 3$. Se

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

é conhecido do leitor que $\det A$ pode ser calculado pela chamada Regra de Sarrus ou desenvolvimento de Laplace, obtendo

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Qual é a analogia geométrica nesse caso? Sejam $u, v, w \in \mathbb{R}^3$ vetores linearmente independentes. Então, u, v, w formam um prisma com bases sendo os paralelogramos gerados pelos vetores u, v e suas translações. Qual é o volume desse prima em termos das coordenadas de u, v, w? Podemos fazer um raciocínio idêntico ao caso $n = 2$, colocando o prisma dentro de um paralelepípedo reto-retângulo, calculando seu volume e removendo as partes que não interessam (paralelepípedos e tetraedros). Se as colunas de $A \in M_3(\mathbb{R})$ são as coordenadas dos vetores u, v, w respectivamente e $< u, v, w >$ tem a mesma orientação da base canônica, então verifica-se (com uma conta muita mais longa do que no caso $n = 2$) que $\det A$ é o volume do prisma procurado. Vamos nos convencer disso de uma forma diferente mas instrutiva. Movemos o prisma ao longo do espaço $\mathbb{R}^3$ e colocamos o paralelogramo base gerado pelos vetores u, v em cima do plano do chão x, y. Esse processo exige translações e rotações no espaço, que são operações que não alteram volumes de sólidos, portanto o volume do prisma colocado nessa posição será o mesmo que o original. Dessa forma, podemos supor que $u = (a, b, 0)$, $v = (c, d, 0)$ e $w = (e, f, g)$. Mas nesse caso é evidente que

$$\det \begin{bmatrix} a & c & e \\ b & d & f \\ 0 & 0 & g \end{bmatrix} = \det \begin{bmatrix} a & c \\ b & d \end{bmatrix} \cdot g = (ad - bc) \cdot g.$$

Note que $ad - bd$ é a área do paralelogramo gerado por u, v e g é a altura do prisma, de modo que o seu volume é o produto da área da base pela altura. Verifique que trocar colunas de posição no determinante em $n = 3$ também altera o sinal a cada troca. Além disso, se $\{u, v, w\}$ é linearmente dependente, então não há prisma gerado e portanto seu volume será 0. Como vemos isso usando o determinante? Suponha por exemplo que w é combinação linear de u e v. Como a última coordenada de u e de v são nulas e w é combinação linear deles, então $g = 0$ e portanto $\det A = 0$. E em analogia, podemos pensar $\det : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ como uma função de 3 variáveis que é linear em cada variável (observe pela fórmula geral) e que muda de sinal sempre que trocamos dois vetores de posição. Além disso, $\det(e_1, e_2, e_3) = 1$.

Definição Formal de Determinante

Ao longo de toda essa seção, $\mathbb{K}$ não precisa ser um corpo, basta ser um anel comutativo com unidade.

Relembrando: Seja $m \in \mathbb{N}$ e X um conjunto com m elementos. Uma **permutação** de X é qualquer bijeção $\sigma : X \to X$. Se $X = I_m = \{1, 2, ..., m\}$, usamos a seguinte notação para descrever diretamente o conjunto imagem de σ:

$$\begin{pmatrix} 1 & 2 & ... & m \\ \sigma(1) & \sigma(2) & ... & \sigma(m) \end{pmatrix}.$$

O conjunto de todas as permutações do conjunto I_m é denotado por S_m. É fácil verificar que S_m é um grupo com a operação de composição de funções. Uma permutação $\sigma \in S_m$ é uma **transposição** se existem $j, k \in \{1, ..., m\}$ tais que $\sigma(j) = k, \sigma(k) = j$ e $\sigma(i) = i$ para todo $i \neq j, k$. Ou seja, uma transposição apenas troca de posições dois elementos de I_m e mantém os outros fixos. A transposição é **adjacente** se $j \in \{1, ..., m-1\}$ e $k = j+1$. É possível mostrar que toda permutação $\sigma \in S_m$ pode ser escrita como composição de transposições adjacentes e que a quantidade destas em qualquer decomposição sempre tem a mesma paridade, isto é, todas as decomposições envolvem uma quantidade par de transposições ou uma quantidade ímpar delas. Isso permite definir o **sinal** da permutação σ, denotado por $(-1)^\sigma$, como 1 se σ é par ou -1 de σ é ímpar. Em particular, qualquer transposição adjacente é ímpar.

Sejam V é um $\mathbb{K}$-espaço vetorial, $m \in \mathbb{N}$ e $\sigma \in S_m$. Definimos $\sigma : V^m \to V^m$ por $\sigma(v_1, ..., v_m) := (v_{\sigma(1)}, v_{\sigma(2)}, ..., v_{\sigma(m)})$ para cada $(v_1, ..., v_m) \in V^m$. Repare no abuso de notação aos usarmos o mesmo símbolo σ para outra função.

Definição 7.1 *Sejam V e W dois $\mathbb{K}$-espaços vetoriais e $m \in \mathbb{N}$. Uma função $F : \underbrace{V \times ... \times V}_{m \text{ vezes}} \to W$ é chamada m-linear (ou simplesmente **multi-linear**) se para todo $j \in \{1, ..., m\}$ e para todos $v_1, ..., v_{j-1}, v_{j+1}, ..., v_m \in V$, a função $V \to W$, $v \mapsto F(v_1, ..., v_{j-1}, v, v_{j+1}, ..., v_m)$ é linear. Além disso, dizemos que:*

- *F é **simétrica** se para todos $(v_1, ..., v_m) \in V^m$ e $\sigma \in S_m$ vale $F(v_{\sigma(1)}, ..., v_{\sigma(m)}) = F(v_1, ..., v_m)$.*
- *F é **anti-simétrica** se para todos $(v_1, ..., v_m) \in V^m$ e $\sigma \in S_m$ vale $F(v_{\sigma(1)}, ..., v_{\sigma(m)}) = (-1)^\sigma F(v_1, ..., v_m)$.*

Suponha agora que $F : V^m \to W$ seja multi-linear e anti-simétrica, $j \in \{1, ...m-1\}$, $v_1, ..., v_{j-1}, v_{j+2}, ..., v_m, v \in V$. Então

$$F(v_1, ..., v_{j-1}, v, v, v_{j+2}, ..., v_m) = 0,$$

De fato, a transposição adjacente σ tal que $\sigma(j) = j+1$ é ímpar e como os vetores nas posições j e $j+1$ são os mesmos, a permutação preserva a lista de vetores, logo, F é anti-simétrica, temos

$$F(v_1, ..., v_{j-1}, v, v, v_{j+2}, ..., v_m) = -F(v_1, ..., v_{j-1}, v, v, v_{j+2}, ..., v_m) \Rightarrow F(v_1, ..., v_{j-1}, v, v, v_{j+2}, ..., v_m) = 0.$$

Isso implica que se $i < j \in \{1, ..., m\}$ e $v_1, ..., v_{i-1}, v_{i+1}, ..., v_{j-1}, v_{j+1}, ..., v_m, v \in V$, então

$$F(v_1, ..., v_{i-1}, v, v_{i+1}, ..., v_{j-1}, v, v_{j+1}, ..., v_m) = 0,$$

isto é, F se anula sempre que houver pelo menos dois vetores iguais na lista, não importando suas posições. Isso é claro pois basta usar uma permutação σ tal que $\sigma(i) = j-1$ e usar a anti-simetria de F e o resultado já provado para posições adjacentes acima.

Exercício 7.1 *Prove a recíproca da afirmação acima. Isto é, se $F : V^m \to W$ é multilinear e se anula sempre que dois vetores ou mais se repetem nas variáveis, então F é anti-simétrica.*

Suponha agora que $W = \mathbb{K}$ e que $\dim V = m$ e seja $\mathcal{B} =< v_1, ..., v_m >$ uma base ordenada de V. Sejam $u_1, ..., u_m \in V$ vetores quaisquer. Escreva $u_1 = a_{11}v_1 + ... + a_{1m}v_m$. Então, como F é linear na primeira variável:

$$\begin{aligned} F(u_1, u_2, ..., u_m) &= F(a_{11}v_1 + ... + a_{1m}v_m, u_2, ..., u_m) \\ &= a_{11}F(v_1, u_2, ..., u_m) + ... + a_{1m}F(v_m, u_2, ..., u_m) \\ &= \sum_{i_1=1}^{m} a_{1i_1}F(v_{i_1}, u_2, ..., u_m) \end{aligned}$$

Analogamente, escreva $u_2 = a_{21}v_1 + ... + a_{2m}v_m$, daí usando a linearidade de F na segunda variável obtemos

$$\begin{aligned} F(u_1, ..., u_m) &= \sum_{i_1=1}^{m} a_{1i_1}F(v_{i_1}, u_2, ..., u_m) \\ &= \sum_{i_1=1}^{m}\sum_{i_2=1}^{m} a_{i_1 1}a_{2i_2}a_{1i_1}a_{2i_2 2}F(v_{i_1}, v_{i_2}, u_3, ..., u_m). \end{aligned}$$

Continuando com esse processo vamos obter

$$F(u_1, ..., u_m) = \sum_{i_1,...,i_m} a_{1i_1}...a_{mi_m}F(v_{i_1}, ..., v_{i_m}).$$

Como $\{v_1, ..., v_m\}$ é uma base, esse conjunto possui exatamente m elementos distintos. A última soma acima é feita sobre todas as combinações possíveis de índices $\{i_1, ..., i_m\} \subset \{1, ..., m\}$, incluindo repetições. Porém, como F é anti-simétrica, na soma acima todas as parcelas com vetores repetidos se anulam e assim só restarão termos cujos índices $i_1, ..., i_m$ correspondem a uma permutação de $\{1, ..., m\}$. Podemos escrever então

$$F(u_1, ..., u_m) = \sum_{\sigma\in S_m} a_{1\sigma(1)}...a_{m\sigma(m)}F(v_{\sigma(1)}, ..., v_{\sigma(m)}).$$

Porém, novamente como F é anti-simétrica

$$\begin{aligned} F(u_1, ..., u_m) &= \sum_{\sigma\in S_m} (-1)^{\sigma}a_{1\sigma(1)}...a_{m\sigma(m)}F(v_1, ..., v_m) \\ &= F(v_1, ..., v_m)\cdot \sum_{\sigma\in S_m} (-1)^{\sigma}a_{1\sigma(1)}...a_{m\sigma(m)}. \end{aligned}$$

Note que o número $F(v_1, ..., v_m)$ é apenas uma constante e correspondente ao valor da aplicação multi-linear na base ordenada escolhida. Em particular, se $V = \mathbb{K}^m$ e $\mathcal{C} =< e_1, ..., e_m >$ é a base canônica de $\mathbb{K}^m$, então qualquer aplicação multi-linear e anti-simétrica $F : \mathbb{K}^m \to \mathbb{K}$ é da forma

$$F(u_1, ..., u_n) = k \cdot \sum_{\sigma\in S_m} (-1)^{\sigma}a_{1\sigma(1)}...a_{m\sigma(m)}, \tag{2}$$

em que $k = F(e_1, ..., e_m)$ e $u_j = (a_{j1}, ..., a_{jm})$ para cada $j \in \{1, ..., m\}$. É fácil ver qualquer F definida com uma fórmula como acima é de fato multi-linear e anti-simétrica.

Definição 7.2 *O **determinante** em ordem m é a única aplicação multi-linear e anti-simétrica* $\det : \mathbb{K}^m \to \mathbb{K}$ *tal que* $\det(e_1, ..., e_m) = 1$. *Podemos escrever*

$$\det(u_1, ..., u_m) = \sum_{\sigma\in S_m} (-1)^{\sigma}a_{1\sigma(1)}...a_{m\sigma(m)},$$

sendo $u_j = (a_{j1}, ..., a_{jm})$ para cada $j \in \{1, ..., m\}$.

Definição 7.3 *Seja $A = (a_{ij}) \in M_m(\mathbb{K})$. Definimos*

$$\det A := \det(u_1, ..., u_m),$$

onde $u_j := (a_{j1}, ..., a_{jm})$ para cada $j = 1, ..., m$.

Exercício 7.2 *Escreva uma fórmula para o determinante de matrizes de ordens 2 e 3.*

Proposição 7.1 *Sejam $A \in M_m(\mathbb{K})$ e $k \in \mathbb{K}$.*

- $\det A = \det A^t$;
- $\det(k \cdot A) = k^m \det A$ *(interprete geometricamente)*;
- *Se A possui duas linhas ou duas colunas múltiplas, ou ainda uma linha ou uma coluna de zeros, então* $\det A = 0$.

Demonstração: Segue diretamente da fórmula do determinante e das propriedades de multi-linearidade e anti-simetria. ∎

Sejam $m \in \mathbb{N}$ e $A \in M_m(\mathbb{K})$. Para cada $i, j \in \{1, ..., m\}$ denotamos por $A(i|j)$ a matriz de ordem $m-1$ obtida por remover a linha i e a coluna j de A. Para cada $j \in \{1, ..., m\}$ defina $D_j : \mathbb{K}^m \to \mathbb{K}$ por

$$D_j(u_1, ..., u_m) := \sum_{i=1}^{m} (-1)^{i+j} a_{ij} \det A(i|j).$$

onde $u_k = (a_{k1}, ..., a_{km})$ para cada $k \in \{1, ..., m\}$. Note que D_j é multi-linear. De fato, sem perda de generalidade vamos mostrar que D_j é linear na variável u_1. Fixe as variáveis $u_2, ..., u_m$. Então os números $a_{2j}, ..., a_{mj}$ são constantes e $\det A(2|j), ..., \det A(m|j)$ são determinantes com linhas $2, ..., m$ fixadas e apenas a primeira variando, e como det é multilinear, segue que $(-1)^{2+j} a_{2j} \det A(2|j) + ... + (-1)^{m+j} a_{mj} \det A(m|j)$ é uma combinação linear de aplicações lineares em u_1, e portanto linear. Com respeito ao primeiro termo da soma, $\det A(1|j)$ é uma constante (não envolve a linha u_1), e a_{1j} é variável, logo $((-1)^{1+j} \det A(1|j)) \cdot a_{1j}$ é linear na variável u_1. O raciocínio é análogo com respeito as outras variáveis, de modo que D_j é multi-linear. Vejamos agora que D_j é anti-simétrica. Basta mostrar que D_j se anula sempre que duas linhas adjacentes sejam iguais. Suponha que $u_k = u_{k+1}$. Se $i \neq k$ e $i \neq k+1$, então a matriz $A(i|j)$ possui linhas iguais, portanto $\det A(i|j) = 0$. Assim

$$D_j(u_1, ..., u_m) = (-1)^{k+j} a_{kj} \det A(k|j) + (-1)^{k+1+j} a_{(k+1)j} \det A(k+1|j).$$

Como $u_k = u_{k+1}$, temos que $a_{kj} = a_{(k+1)j}$ e $\det A(k|j) = \det A(k+1|j)$. Logo

$$D_j(u_1, ..., u_m) = a_{kj} \det A(k|j)(-1)^{k+j} \cdot (+1-1) = 0.$$

Portanto D_j é anti-simétrica. Finalmente, se $(u_1, ..., u_m) = (e_1, ..., e_m)$, então $a_{ij} = 0$ se $i \neq j$ e $a_{jj} = 1$ e $A(j|j) = \mathrm{Id}_{m-1}$. Logo

$$D_j(e_1, ..., e_m) = (-1)^{j+j} a_{jj} \det A(j|j) = 1.$$

Assim, para cada $j \in \{1, ..., m\}$, a aplicação multi-linear D_j é anti-simétrica e vale 1 na base canônica. Da unicidade do determinante, temos então que para todo $j \in \{1, ..., m\}$ vale

$$\det A = \sum_{i=1}^{m} (-1)^{i+j} a_{ij} \det A(i|j).$$

Isso nos dá uma fórmula recursiva para calcular determinantes. Calculamos um determinante de ordem m como uma combinação de determinantes de ordem $m-1$, e o mesmo pode ser feito para cada um desses. Essa fórmula é chamada **Desenvolvimento de Laplace** para o cálculo do determinante. Note que como $\det A = \det A^t$, o desenvolvimento pode ser feito ao longo de qualquer linha ou coluna de A. Para cada $i, j \in \{1, ..., m\}$, o escalar $(-1)^{i+j} \det A(i|j)$ é chamado (i, j)-ésimo **cofator** de A.

Proposição 7.2 *Seja $A = (a_{ij}) \in M_m(\mathbb{K})$ uma matriz triangular superior (ou inferior). Então*

$$\det A = a_{11} \cdot a_{22} \cdot ... \cdot a_{mm}.$$

Demonstração: Basta fazer o desenvolvimento de Laplace repetidas vezes ao longo da primeira linha de A, depois ao longo da segunda linha e assim por diante. ∎

8 Aula 8 - Paleari

Aqui $\mathbb{K}$ ainda é um anel comutativo com unidade.

Proposição 8.1 *Sejam* $A, B \in M_m(\mathbb{K})$, *então* $\det(AB) = \det A \cdot \det B$.

Demonstração: Fixe uma matriz $B \in M_m$ e defina $F : \mathbb{K}^m \to \mathbb{K}$ por $F(A) := \det(AB)$. Se as linhas de A são denotadas por $\alpha_1, ..., \alpha_m$, então as linhas de AB são $\alpha_1 B, \alpha_2 B, ..., \alpha_m B$. Logo

$$F(\alpha_1, ..., \alpha_m) = \det(\alpha_1 B, ..., \alpha_m B).$$

Como o produto de matrizes é distributivo com soma e escalares, B está fixada e det é multi-linear, segue que F é multi-linear. Além disso, se $\alpha_k = \alpha_{k+1}$, então $\alpha_k B = \alpha_{k+1} B$, logo se duas linhas de A são iguais, temos que F se anula pois det é anti-simétrica. Logo $F : \mathbb{K}^m \to \mathbb{K}$ é multi-linear e anti-simétrica e pela equação (refdet) vale

$$F(A) = F(e_1, ..., e_m) \det(A).$$

Por outro lado, $F(e_1, ..., e_m) = \det(\mathrm{Id} \cdot B) = \det B$. Logo $F(A) = \det(A) \cdot \det(B)$ e assim $\det(AB) = \det(A) \cdot \det(B)$. ∎

Corolário 8.1 *Se* $\mathbb{K}$ *é um corpo e* $A \in M_m(\mathbb{K})$ *é invertível, então* $\det A \neq 0$ *e* $\det(A^{-1}) = \det(A)^{-1}$.

Demonstração: Basta usar a proposição 8.1 na relação $A \cdot A^{-1} = \mathrm{Id}_m$. ∎

Observação 8.1 *Observe a seguinte consequência da multilinearidade e anti-simetria do determinante. Seja* $A \in M_m(\mathbb{K})$ *com linhas* $\alpha_1, ..., \alpha_m$, $k \in \mathbb{K}$, $i, j \in \{1, ..., m\}$ *com* $i < j$ *e* B *uma matriz com as mesmas linhas de* A *exceto a linha* i, *que vale* $\alpha_i + k\alpha_j$. *Então* $\det B = \det A$. *De fato,*

$$\det B = \det(\alpha_1, ..., \alpha_{i-1}, \alpha_i + k\alpha_j, a_{i+1}, ..., \alpha_m) = \det(\alpha_1, ..., \alpha_m) + k \det(\alpha_1, ..., \alpha_j, ..., \alpha_j, ..., \alpha_m) = \det A + k \cdot 0 = \det A.$$

Proposição 8.2 *Sejam* $m, r \in \mathbb{N}$, $A \in M_r(\mathbb{K})$, $B \in M_{r\times s}(\mathbb{K})$, $C \in M_s(\mathbb{K})$ *e* $0 \in M_{s\times r}(\mathbb{K})$, *então*

$$\det \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = \det A \cdot \det C.$$

Demonstração: Defina

$$D(A, B, C) := \det \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}.$$

Fixando A e B e pensando D como uma função de C, como o bloco a esquerda de C é nulo, temos que D é multi-linear e anti-simétrica como função das linha de C. Pela relação (2) temos que

$$D(A, B, C) = (\det C) D(A, B, \mathrm{Id}_s).$$

Por outro lado, observe que a matriz

$$\begin{bmatrix} A & B \\ 0 & \mathrm{Id} \end{bmatrix}$$

pode ser obtida da matriz

$$\begin{bmatrix} A & 0 \\ 0 & \mathrm{Id} \end{bmatrix}$$

como na observação 8.1 (acrescentando os $b's$ aproveitando os elementos da identidade no bloco abaixo). Logo $D(A, B, \mathrm{Id}) = D(A, 0, \mathrm{Id})$. Por outro lado, como o bloco a direita de A é nulo, a aplicação $A \mapsto D(A, 0, \mathrm{Id})$ é multi-linear e anti-simétrica como função das linhas de A. Logo, novamente pela relação (2) teremos $D(A, 0, \mathrm{Id}) = \det(A) \cdot D(\mathrm{Id}, 0, \mathrm{Id})$. Como claramente $D(\mathrm{Id}, 0, \mathrm{Id}) = 1$, teremos

$$D(A, B, C) = (\det C)D(A, B, \mathrm{Id}) = (\det C)D(A, 0, \mathrm{Id}) = (\det C)(\det A)D(\mathrm{Id}, 0, \mathrm{Id}) = \det A \det C.$$

■

Lembramos que dada $A = (a_{ij}) \in M_m(\mathbb{K})$ e dados $i, j \in \{1, ..., m\}$ o (i, j)-ésimo cofator de A é o escalar $c_{ij} := (-1)^{i+j} \det A(i|j)$. Pelo desenvolvimento de Laplace temos que

$$\det A = \sum_{i=1}^{m} a_{ij} c_{ij}.$$

Afirmamos que se $j \neq k$, então

$$\sum_{i=1}^{m} a_{ik} c_{ij} = 0.$$

Para ver isso, seja $B = (b_{ij})$ a matriz que possui as mesmas colunas de A exceto a coluna j, onde colocamos uma cópia da coluna k de A. Assim B possui duas colunas iguais e portanto $\det B = 0$. Por outro lado, note que por definição de B, temos $B(i|j) = A(i|j)$ e pelo desenvolvimento de Laplace pela coluna j de B teremos:

$$\begin{aligned} 0 &= \det B \\ &= \sum_{i=1}^{m} (-1)^{i+j} b_{ij} \det B(i|j) \\ &= \sum_{i=1}^{m} (-1)^{i+j} a_{ik} \det A(i|j) \\ &= \sum_{i=1}^{m} a_{ik} c_{ij}. \end{aligned}$$

Concluímos que

$$\sum_{i=1}^{m} a_{ik} c_{ij} = \sum_{i=1} c_{ij} a_{ik} = \delta_{jk} \det A.$$

Definimos

$$\mathrm{adj}(A)_{ij} := c_{ji} = (-1)^{i+j} \det A(j|i),$$

a chamada **adjunta clássica** de A. Por definição, $\mathrm{adj}(A)$ satisfaz

$$(\mathrm{adj} A)A = (\det A)\mathrm{Id}.$$

Vejamos que também vale a relação $A(\mathrm{adj}(A)) = (\det A)\mathrm{Id}$. Note que para cada i, j temos $(\mathrm{adj}A)_{ji} = (-1)^{j+i} \det A(i|j) = (-1)^{i+j} \det A^t(j|i) = (\mathrm{adj}A^t)_{ij}$, isto é, $(\mathrm{adj}A)^t = \mathrm{adj}A^t$. Aplicando a relação já conhecida para A^t, obtemos

$$(\mathrm{adj}A^t)A^t = (\det A^t)\mathrm{Id},$$

o que implica $(\mathrm{adj}A)^t A^t = (A \cdot \mathrm{adj}(A))^t = (\det A)\mathrm{Id}$. Como $(\det A)\mathrm{Id}$ é simétrica, obtemos $A \cdot \mathrm{adj}(A) = (\det A)\mathrm{Id}$. Em particular, se $\det A \neq 0$, então A é invertível e

$$A^{-1} = \frac{1}{\det A} \mathrm{adj}(A).$$

Resumimos o que acabamos de provar no seguinte.

Teorema 8.1 *Seja $A \in M_m(\mathbb{K})$, então*

$$A \cdot (adjA) = (adjA) \cdot A = (\det A)Id,$$

em particular, A é invertível se, e somente se, $\det A \neq 0$.

Definição 8.1 *Dizemos que duas matrizes $A, B \in M_m(\mathbb{K})$ são* ***semelhantes*** *se existe $C \in M_n(\mathbb{K})$ invertível tal que $B = C^{-1}AC$.*

Exercício 8.1 *Verifique que a relação de semelhança em $M_m(\mathbb{K})$ é uma relação de equivalência.*

Pela propriedade 8.1, note que se A e B são semelhantes, então

$$\det A = \det(C^{-1}BC) = \det(C^{-1}) \det B \det C = \frac{1}{\det C} \det B \det C = \det B,$$

isto é, matrizes semelhantes tem mesmo determinante.

Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita, $T : V \to V$ uma transformação linear e $\mathcal{B}$ e $\mathcal{C}$ são bases de V. Lembramos que

$$[T]_{\mathcal{B}} = (M^{\mathcal{C}}_{\mathcal{B}})^{-1}[T]_{\mathcal{C}}M^{\mathcal{C}}_{\mathcal{B}}.$$

Ou seja, $[T]_{\mathcal{B}}$ e $[T]_{\mathcal{C}}$ são semelhantes. Assim, faz sentido definir o **determinante da transformação linear** T como $\det T := \det[T]_{\mathcal{B}}$, em que $\mathcal{B}$ é qualquer base de V. O conceito está bem definido e independe da escolha da base pelo que acabamos de ver.

Para finalizar, vamos apresentar a **Regra de Cramer** para a solução de sistemas lineares da forma $AX = Y$ com $A \in M_n(\mathbb{K})$ e $Y^t = (y_1, ..., y_m)$ dados. Note que $AX = Y$ implica $(\text{adj}A)AX = (\text{adj}A)Y$, e portanto $\det A \cdot X = (\text{adj}A)Y$. Logo, se $\det A \neq 0$, então a única solução do sistema $AX = Y$ é dada matricialmente por

$$X = \frac{1}{\det A}(\text{adj}A)Y.$$

Em coordenadas, para cada $j \in \{1, ..., m\}$ temos

$$x_j = \frac{1}{\det A}\sum_{i=1}^{m}(\text{adj}A)_{ji}y_i = \frac{1}{\det A}\sum_{i=1}^{m}(-1)^{i+1}y_i \det A(i|j).$$

Note que pelo desenvolvimento de Laplace do determinante (ao longo da coluna j), a expressão $\sum_{i=1}^{m}(-1)^{i+j}y_i \det A(i|j)$ é exatamente o determinante da matriz B_j que possui as mesmas colunas de A, exceto a coluna j onde é colocado o vetor Y no lugar da coluna j de A. Assim

$$x_j = \frac{\det B_j}{\det A}.$$