Álgebra Linear

Rodrigo Ramos

5 de maio de 2021

Sumário

Sobre estas anotações

Este documento consiste exclusivamente de anotações minhas para acompanhar um curso introdutório de Álgebra Linear a nível de graduação. Cada indivíduo tem preferência por um estilo de abordagem deste assunto: alguns preferem uma ênfase maior no ponto de vista de sistemas lineares, matrizes e uso precoce do determinante; outros preferem o ponto de vista de operadores lineares, postergando a representação matricial e a introdução do determinante para um momento mais apropriado. Nestas notas, segue-se a última abordagem, utilizando principalmente [1] como fonte de estudos.

1 Espaços Vetoriais

1.1 Espaços e Subespaços Vetoriais

Definição 1.1. *Um **espaço vetorial**[1] é uma quádrupla* $(E, \mathbb{F}, +, \cdot)$*, onde* E *é um conjunto;* $\mathbb{F}$ *é um corpo;* $+\colon E \times E \to E$ *e* $\cdot\colon E \times \mathbb{F} \to E$ *são operações binárias, chamadas respectivamente de soma e multiplicação por escalar, tais que para todos* $u, v, w \in E$ *e para todos* $\alpha, \beta \in \mathbb{F}$ *os seguintes axiomas são respeitados:*

1. $u + v = v + u$ *(**comutatividade**).*
2. $(u + v) + w = u + (v + w)$ *e* $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$ *(**associatividade**).*
3. $\exists! 0 \in E$ *tal que* $\forall v \in E$, $v + 0 = 0 + v = v$ *(**vetor nulo**).*
4. $\forall v \in E\ \exists u \in E$ *tal que* $v + u = 0$*. Por convenção, denota-se* u *por* $-v$*. (**inverso aditivo**).*
5. $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$ *e* $\alpha \cdot (v + u) = \alpha \cdot v + \alpha \cdot u$ *(**distributividade**).*
6. *Se* $1 \in \mathbb{F}$ *é o elemento neutro da multiplicação em* $\mathbb{F}$*, então* $\forall v \in E$ *tem-se* $e \cdot v = v$ *(**multiplicação por 1**).*

*Os elementos de um espaço vetorial são chamados de **vetores**.*

Observação. *0 sempre representará tanto o vetor nulo como o elemento neutro da soma em* $\mathbb{F}$*. Os corpos mais usuais a serem escolhidos são* $\mathbb{R}$ *e* $\mathbb{C}$*. Nestas notas, todos os espaços vetoriais são reais, ou seja,* $\mathbb{F} = \mathbb{R}$*. Por último, para não carregar a notação, simplesmente denota-se o espaço vetorial pelo conjunto* E *onde está definida a soma e omite-se o símbolo da operação* $\cdot$*.*

Exemplo 1.1. *Claramente,* $\mathbb{F}$ *é um espaço vetorial sobre si mesmo quando consideramos a soma e a multiplicação por escalar como sendo respectivamente a soma e a multiplicação em* $\mathbb{F}$*.*

Exemplo 1.2. $E = \{0\}$ *também é um espaço vetorial se restringirmos a soma em* $\mathbb{F}$ *ao conjunto* $\{0\} \times \{0\}$ *e se restringirmos a multiplicação a* $\{0\} \times \mathbb{F}$*.*

Exemplo 1.3. *Seja* $n \in \mathbb{N}$*. Então,* $\mathbb{R}^n$ *é um espaço vetorial (real) quando definimos a soma e a multiplicação por escalar como:*

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n);$$

$$\alpha(x_1, \ldots, x_n) := (\alpha x_1, \ldots, \alpha x_n).$$

para todos $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$ *e para todo* $\alpha \in \mathbb{R}$*.*

Exemplo 1.4. *O conjunto* $\mathbb{R}^{\mathbb{N}}$*, cujos elementos são sequências de números reais, torna-se um espaço vetorial quando definimos a soma e a multiplicação por escalar componente a componente, similarmente ao exemplo anterior.*

[1] Na literatura, um nome alternativo é espaço linear.

Exemplo 1.5. *Seja $M(m \times n)$ o conjunto de todas as matrizes m por n de entradas reais. Se definirmos para todos $\mathbf{a} = [a_{ij}]$ e $\mathbf{b} = [b_{ij}]$ em $M(m \times n)$ e para todo $\alpha \in \mathbb{R}$ que*

$$\mathbf{a} + \mathbf{b} := [a_{ij} + b_{ij}],$$

$$\alpha\mathbf{a} := [\alpha a_{ij}],$$

$1 \leq i \leq m$, $1 \leq j \leq n$, então $M(m \times n)$ torna-se um espaço vetorial.

Exemplo 1.6. *Seja X um conjunto não-vazio. Denotando o conjunto de todas as funções de X em $\mathbb{R}$ por $\mathcal{F}(X;\mathbb{R})$, então se para todas $f, g\colon X \to \mathbb{R}$ e todo $\alpha \in \mathbb{R}$ definirmos*

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

$$(\alpha f)(x) := \alpha f(x),$$

então $\mathcal{F}(X;\mathbb{R})$ torna-se um espaço vetorial.

Exemplo 1.7. *Sejam E uma variedade diferenciável, $p \in E$ e $C^\infty(E)$ o conjunto de todas as funções suaves definidas em E. Uma derivação em p é um mapa $v\colon C^\infty(E) \to \mathbb{R}$ que satisfaz as seguintes propriedades:*

$$\begin{aligned} v(f + \alpha g) &= vf + \alpha vg,\ \forall f, g \in C^\infty(E), \forall \alpha \in \mathbb{R}; \\ v(fg) &= f(p)vg + g(p)vf,\ \forall f, g \in C^\infty(E). \end{aligned}$$

Assim, o conjunto de todas as derivações em p é um espaço vetorial quando definimos a soma e a multiplicação por escalar da maneira usual. Este conjunto é chamado de ***espaço tangente a E no ponto p*** *e T_pE é uma notação comum para ele.*

Exemplo 1.8. *Sejam $V_1, \ldots, V_k$ e W espaços vetoriais sobre $\mathbb{R}$. Um mapa $F\colon V_1 \times \ldots \times V_k \to W$ é dito* ***multilinear*** *quando for linear em cada uma de suas entradas quando as demais estão fixadas. Em outras palavras, para cada $i \in \{1, \ldots, k\}$, para todos $v_i, v_i' \in V_i$ e para todos $\alpha, \beta \in \mathbb{R}$ vale que*

$$F(v_1, \ldots, \alpha v_i + \beta v_i', \ldots, v_k) = \alpha F(v_1, \ldots, v_i, \ldots, v_k) + \beta F(v_1, \ldots, v_i', \ldots, v_k).$$

O conjunto de todas os mapas multilineares de $V_1 \times \ldots \times V_k$ em W é denotado por $L(V_1, \ldots, V_k; W)$. Tal conjunto torna-se um espaço vetorial sobre $\mathbb{R}$ quando definidas a soma pontual usual e a multiplicação por escalar usual.

A partir dos axiomas estabelecidos na Definição 1.1, temos o seguinte lema.

Lema 1.1. *Seja E um espaço vetorial (real). Então, vale que*

a) Sejam $u, v, w \in E$ tais que $w + u = w + v$. Então, $u = v$.

b) Para todo $v \in E$, o inverso aditivo $-v$ é único.

c) *Para todo* $v \in E$, *temos que* $0v = 0 \in E$. *Para todo* $\alpha \in \mathbb{R}$, *temos que* $\alpha 0 = 0$.

d) *Sejam* $v \in E$ *e* $\alpha \in \mathbb{R}$ *tais que* $v \neq 0$ *e* $\alpha \neq 0$, *então* $\alpha v \neq 0$.

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

Demonstração.
Para provar o item a, veja que

$$\begin{aligned} w + u = w + v &\implies -w + (w + u) = -w + (w + v) \\ &\implies (-w + w) + u = (-w + w) + v \\ &\implies 0 + u = 0 + v \\ &\implies u = v. \end{aligned}$$

Para o item b, por absurdo, suponha que existe $w \in E$ tal que $w \neq -v$ e $v + w = 0$. Então, segue

$$v + w = 0 \implies v + w = v + (-v).$$

Daí, pelo item a, conclui-se que $w = -v$. Uma contradição. Com relação ao item c, basta notar que

$$v + 0v = (1 + 0)v = 1v = v.$$

Pelo item a segue que $0v = 0$. Para mostrar o outro resultado, perceba que

$$\alpha 0 = \alpha(0 + 0) = \alpha 0 + \alpha 0.$$

Pelo item a, segue que $\alpha 0 = 0$.
Pelo item c, segue que

$$\alpha v = 0 \implies \alpha = 0 \text{ ou } v = 0.$$

Daí, pela contrapositiva obtém-se o resultado do item d.
Por último, temos que

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

Pela unicidade do inverso aditivo, conclui-se que $(-1)v = -v$. ■

Exercício 1.1. *Sejam* $u = (a, b)$ *e* $v = (c, d)$ *elementos de* $\mathbb{R}^2$ *e* $u \neq 0$. *Mostre que existe* $\alpha \in \mathbb{R}$ *tal que* $v = \alpha u$ *se e somente se* $ad - bc = 0$.

Exercício 1.2. *Sejam $E_1, \ldots, E_n$ espaços vetoriais. Mostre que o produto cartesiano $E_1 \times \ldots \times E_n$ é um espaço vetorial se definirmos que*

$$(v_1, \ldots, v_n) + (u_1, \ldots, u_n) := (v_1 + u_1, \ldots, v_n + u_n)$$

$$\alpha(v_1, \ldots, v_n) := (\alpha v_1, \ldots, \alpha v_n),$$

onde $v_i + u_i$ e αv_i são a soma e a multiplicação por escalar definidas no i-ésimo espaço vetorial E_i.

Em determinadas situações, um espaço vetorial E pode ser um ambiente muito geral, de forma que pouca informação possa ser extraída dele. Nesses casos, é mais vantajoso estudarmos certos subconjuntos de E que herdem a sua estrutura de espaço vetorial. Em vista disso, fazemos uso da

Definição 1.2. *Sejam E um espaço vetorial e F um subconjunto de E. Diz-se que F é um* ***subespaço vetorial de E*** *quando:*

1. $0 \in F$.
2. *Se $u, v \in F$, então $u + v \in F$.*
3. *Se $\alpha \in \mathbb{R}$ e $v \in F$, então $\alpha v \in F$.*

Em outras palavras, um subespaço vetorial de E é um espaço vetorial que contém o vetor nulo de E e cujas operações são definidas a partir da restrição das operações em E ao subconjunto F.

Exemplo 1.9. *Seja E um espaço vetorial. O subconjunto $\{0\} \subset E$ é um subespaço de E quando definimos as operações nele pela restrição das operações de soma e multiplicação por escalar aos conjuntos $\{0\} \times \{0\}$ e $\{0\} \times \mathbb{R}$, respectivamente. Diz-se que este é um* ***subespaço trivial de E****, sendo E o outro subespaço trivial.*

Exemplo 1.10. *Se $v \in E$, então $F = \{\alpha v \colon \alpha \in \mathbb{R}\}$ é um subespaço vetorial de E.*

Exemplo 1.11. *Seja $E = \mathcal{F}(\mathbb{R}; \mathbb{R})$. O conjunto $C^k(\mathbb{R}) \subset \mathcal{F}(\mathbb{R}; \mathbb{R})$ de todas funções $f \colon R \to \mathbb{R}$ k vezes continuamente deriváveis é um subespaço vetorial de E. Outros subespaços de E são $C^\infty(\mathbb{R})$; $\mathcal{P}$, o conjunto dos polinômios com coeficientes reais; e $\mathcal{P}_n$, o conjunto de polinômios de grau menor igual a n com coeficientes reais.*

Exemplo 1.12. *Se X é um conjunto não-vazio qualquer, então o conjunto das funções $f \colon X \to \mathbb{R}$ que são limitadas é um subespaço vetorial de $\mathcal{F}(X; \mathbb{R})$.*

Exemplo 1.13. *Se $U \subset \mathbb{R}^2$ for um conjunto aberto, claramente $F = C^2(U)$ é um subespaço de $\mathcal{F}(\mathbb{R}^2; \mathbb{R})$. Considerando F como um espaço vetorial, fixado $c \in \mathbb{R}$, seja $\mathcal{S} \subset F$ o conjunto das soluções da E.D.P abaixo*

$$\frac{\partial^2 f}{\partial t^2} - c^2 \frac{\partial^2 f}{\partial x^2} = 0.$$

Então, $\mathcal{S}$ é subespaço vetorial de $C^2(U)$ quando herda as operações definidas nele.

Exemplo 1.14. *Fixados $a_1, \ldots, a_n \in \mathbb{R}$, o conjunto H de todos vetores $(x_1, \ldots, x_n) \in \mathbb{R}^n$ tais que*

$$a_1x_1 + \ldots + a_nx_n = 0$$

é um subespaço de $\mathbb{R}^n$ e define o que chama-se de hiperplano que passa pela origem de $\mathbb{R}^n$.

Exercício 1.3. *Seja X um conjunto qualquer. Uma relação de equivalência em X, usualmente denotada por $\sim$, é um subconjunto de $X \times X$ tal que*

1. *$a \sim$ [2]a para todo $a \in X$.*
2. *Se $a \sim b$, então $b \sim a$ para todos $a, b \in X$.*
3. *Se $a \sim b$ e $b \sim c$, então $a \sim c$ para todos $a, b, c \in X$.*

Dada uma relação de equivalência $\sim$ em um conjunto X, para cada $a \in X$ define-se a classe de equivalência de a como sendo $[a] := \{b \in X \colon a \sim b\}$. O conjunto de todas as classes de equivalência é denotado por $X/\sim$ e chama-se de conjunto quociente de X por $\sim$.

Se $X = E$ for um espaço vetorial e F for subespaço de E, mostre que o conjunto $E_F = \{(u,v) \in V \times V \colon u - v \in F\}$ é uma relação de equivalência em E, denotada também por $\sim_F$.

O quociente de E por $\sim_F$ é denotado por E/F e podemos torná-lo um espaço vetorial se definirmos que $\forall u, v \in E$ e $\forall \alpha \in \mathbb{R}$:

$$[u] + [v] := [u + v]$$

$$\alpha[u] := [\alpha u].$$

*Tal espaço vetorial chama-se **espaço quociente de E por F**.*

Mostre que as operações acima estão bem definidas (isto é, independem dos representantes da classe de equivalência) e que de fato obedecem aos axiomas de um espaço vetorial.

A noção de "quocientizar" um conjunto é bastante importante em teoria de grupos por conta de um resultado denominado primeiro teorema fundamental de isomorfismos, que permite identificações legais entre certos grupos[3]. No mais, se V e W são espaços vetoriais, tal resultado também é utilizado para definir a propriedade universal do produto tensorial $V \otimes W$ destes espaços, que é linearizar mapas multilineares.

Lema 1.2. *Sejam E um espaço vetorial e Λ um conjunto de índices. Se para cada $\lambda \in \Lambda$ tem-se que F_λ é um subespaço vetorial de E, então o conjunto*

$$F = \bigcap_{\lambda \in \Lambda} F_\lambda$$

[2]Isto é mera notação para simbolizar que $(a, a) \in R$, onde R é a relação de equivalência em X. Em acréscimo, $x \sim y$ lê-se "x equivalente a y".

[3]A título de exemplo, $SU(2)/\{-\mathbb{1}, \mathbb{1}\} \simeq SO(3)$

é um subespaço vetorial de E.

Demonstração.

Basta provar que F satisfaz as propriedades da definição 1.2. A primeira propriedade é imediata, pois $0 \in F_\lambda$ para todo $\lambda \in \Lambda$. Com relação à segunda propriedade, se $u, v \in F$, por definição temos que $u, v \in F_\lambda$ para todo $\lambda \in \Lambda$. Daí, $u + v \in F_\lambda$ para todo $\lambda \in \Lambda$, donde conclui-se que $u + v \in F$. O caso para αv é análogo. ■

Pelo lema 1.2, segue que a solução de um sistema linear

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

é um subespaço vetorial de $\mathbb{R}^n$, pois trata-se da interseção de m hiperplanos, cada um definido por uma das equações acima.

A definição a seguir nos permite construir um subespaço vetorial a partir de qualquer subconjunto de um espaço vetorial E.

Definição 1.3. *Seja $X \subset E$ um subconjunto qualquer do espaço vetorial E. O **subespaço vetorial de E gerado por X**, também chamado de span de X, é por definição o conjunto de todas as combinações lineares*

$$\alpha_1 x_1 + \ldots + \alpha_n x_n$$

de elementos de X. É comum denotar esse conjunto por $S(X)$ ou por span(X).

No caso especial em que $S(X) = E$, diz-se que X é um conjunto de geradores de E.

Exercício 1.4. *Mostre que $S(X)$ é um subespaço vetorial.*

Exercício 1.5. *Mostre que $S(X)$ é a interseção de todos os subespaços vetoriais de E que contém X. Logo, $S(X)$ é o menor subespaço vetorial que contém X.*

Um exemplo de geradores de $\mathbb{R}^n$ é o conjunto $\{e_1, \ldots, e_n\}$, onde e_i é o vetor cuja i-ésima entrada vale 1 e as demais valem zero.

Definição 1.4. *Se F_1 e F_2 são dois subespaços vetoriais de E tais que $F_1 \cap F_2 = \{0\}$, o subespaço vetorial $S(F_1 \cup F_2) \equiv F_1 \oplus F_2$ é chamado de soma direta de F_1 e F_2.*

Teorema 1.1. *Sejam F, F_1, F_2 subespaços vetoriais de E com $F_1 \subset F$ e $F_2 \subset F$. São equivalentes:*

a) $F = F_1 \oplus F_2$.

b) *Todo elemento $w \in F$ existem $v_1 \in F_1$ e $v_2 \in F_2$, únicos, tais que $w = v_1 + v_2$.*

Demonstração.

Assumindo que vale o item a), segue que $F_1 \cap F_2 = \{0\}$. Pela definição de soma direta, se $w \in F$, então $\exists v_1 \in F_1$ e $\exists v_2 \in F_2$ tais que $w = v_1 + v_2$. Supondo que $w = u_1 + u_2$ para outros $u_1 \in F_1$ e $u_2 \in F_2$, conclui-se que

$$u_1 - v_1 = v_2 - u_2.$$

Como F_1 e F_2 são subespaços, cada lado da igualdade acima pertence a um dos subespaços. Visto que a interseção deles é $\{0\}$, é imediato que $u_1 = v_1$ e que $u_2 = v_2$.

Para mostrar a volta, seja $w \in F_1 \cap F_2$. Por um lado, temos que $w + 0 = w \in F_1$ e $0 \in F_2$. Por outro, temos que $0 + w = w \in F_2$, $0 \in F_1$. Pela hipótese b), a unicidade da representação implica em $w = 0$. Consequentemente $F_1 \cap F_2 = \{0\}$ e portanto $F = F_1 \oplus F_2$. ■

Exemplo 1.15. *Uma matriz* $\mathbf{a} \in M(n \times n)$ *é dita simétrica quando* $a_{ij} = a_{ji}$. *De forma análoga,* $\mathbf{a}$ *é anti-simétrica se* $a_{ij} = -a_{ji}$. *Os conjuntos das matrizes simétricas e das matrizes anti-simétricas,* S *e* A *respectivamente, são subespaços vetoriais de* $M(n \times n)$ *e temos que* $M(n \times n) = \mathrm{S} \oplus \mathrm{A}$.

Exemplo 1.16. *Sejam* E *e* F *espaços vetoriais. Sejam* $A = \{f \in \mathcal{F}(E;F)\colon \forall v \in E,\ f(v) = f(-v)\}$ *e* $B = \{f \in \mathcal{F}(E;F)\colon \forall v \in E,\ -f(v) = f(-v)\}$. *Temos que* $\mathcal{F}(E;F) = A \oplus B$.

Definição 1.5. *Seja* E *um espaço vetorial. Fixados* $x, y \in E$, *a reta que une estes pontos é por definição o conjunto*

$$r = \{(1-t)x + ty\colon t \in \mathbb{R}\}.$$

Definição 1.6. *Seja* V *um subconjunto de um espaço vetorial* E. *Diz-se que* V *é uma* ***variedade afim*** *se para quaisquer dois pontos de* V, *a reta que une estes pontos está contida em* V.

Em vista da definição acima, o conjunto de soluções para um sistema linear não-homogêneo pode ser interpretado como uma variedade afim.

Teorema 1.2. *Seja* V *uma variedade afim não-vazia no espaço vetorial* E. *Existe um único subespaço vetorial* $F \subset E$ *tal que para todo* $x \in V$ *tem-se*

$$V = x + F \coloneqq \{x + v\colon v \in F\}.$$

Demonstração.

Escolhido $x \in V$ qualquer, basta definir F como sendo o conjunto cujos elementos são da forma $y - x$ com $y \in V$. ■

Como consequência direta do teorema 1.2, se as soluções de um sistema linear são uma variedade afim $V \neq \emptyset$, então se x_0 é uma solução particular dessa sistema, temos que a solução mais geral possível é a soma de x_0 com um elemento de F, que é o conjunto de soluções do sistema homogêneo associado.

1.2 Bases

Definição 1.7. *Seja E um espaço vetorial. Um conjunto $X \subset E$ é dito **linearmente independente** (L.I.) quando nenhum de seus elementos é combinação linear dos demais. Um conjunto que não é L.I. é dito **linearmente dependente** (L.D.).*

No caso em que X possui apenas um elemento v, por convenção, determina-se que X é $L.I.$ se e somente se $v \neq 0$.

Para identificar se um conjunto é $L.I.$ ou $L.D.$, é conveniente utilizar o seguinte critério

Teorema 1.3. *Sejam E um espaço vetorial e $X \subset E$ um conjunto. X é $L.I.$ se e somente se a única combinação linear nula[4] de elementos de X é aquela cujos coeficientes são todos nulos.*

Demonstração.

Para demonstrar a ida, suponha que exista uma combinação linear nula de elementos de X com coeficientes não todos nulos:

$$\alpha_1 v_1 + \ldots + \alpha_m v_m = 0.$$

Sem perda de generalidade, se $\alpha_1 \neq 0$, então podemos reescrever v_1 como sendo

$$v_1 = -\frac{\alpha_2}{\alpha_1} v_2 - \ldots - \frac{\alpha_m}{\alpha_1} v_m.$$

Donde conclui-se que X é $L.D.$. Pela contrapositiva, obtém-se o resultado desejado.

Para mostrar provar a volta, faz-se novamente uso da contrapositiva. Se X for $L.D.$, segue que existe $v \in X$ tal que

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

para certos $\alpha_1, \ldots \alpha_m \in \mathbb{R}$ – nem todos nulos – e certos $v_1, \ldots, v_m \in X$ distintos de v. Daí, segue que

$$v - \alpha_1 v_1 - \ldots - \alpha_m v_m = 0$$

é uma combinação linear nula de elementos de X onde nem todos os coeficientes são nulos. ∎

Corolário 1.1. *Se $v = \alpha_1 v_1 + \ldots + \alpha_m v_m = \beta_1 v_1 + \ldots + \beta_m v_m$ e $\{v_1, \ldots, v_m\}$ é L.I., então $\alpha_i = \beta_i$ para $i = 1, \ldots, m$.*

Exemplo 1.17. *Em $\mathbb{R}^n$, o conjunto dos vetores canônicos, $\{e_1, \ldots, e_n\}$ é linearmente independente. Já em $\mathcal{P}_n$, os monômios $1, x, \ldots, x^n$ também formam um conjunto L.I..*

[4] Uma combinação linear nula é definida por $\sum_{\lambda} \alpha_\lambda v_\lambda = 0$.

Exercício 1.6. *Mostre que os conjuntos do exemplo acima de fato são linearmente independentes em seus devidos espaços vetoriais.*

A independência linear dos conjuntos citados no exemplo anterior é uma aplicação direta do

Teorema 1.4. *Sejam $v_1, \ldots, v_m$ vetores não-nulos do espaço vetorial E. Se nenhum deles for combinação linear dos anteriores, então $X = \{v_1, \ldots, v_m\}$ é L.I..*

Demonstração.
Por absurdo, suponha que X seja L.D.. Então, existem $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ não todos nulos tais que

$$\alpha_1 v_1 + \ldots + \alpha_m v_m = 0.$$

Se $\alpha_r v_r$ for o último vetor não nulo desta soma, então segue que

$$v_r = -\frac{\alpha_1}{\alpha_r} v_1 - \ldots - \frac{\alpha_{r-1}}{\alpha_r} v_{r-1}.$$

Daí, v_r é combinação linear dos vetores anteriores, uma contradição. ■

Definição 1.8. *Uma **base** em um espaço vetorial E é um conjunto $\mathrm{B} \subset E$ L.I. e tal que $S(\mathrm{B}) = E$.*

Definição 1.9. *Seja E um espaço vetorial. Diz-se que E tem **dimensão finita n** se ele admite uma base B tal que $\#B = n \in \mathbb{N}_0$. Nesta situação, denota-se a dimensão de E por $\dim E \equiv \#B$.*

Pelo corolário do teorema 1.3, segue que se B é uma base num espaço vetorial E, então os coeficientes α_i que determinam um vetor $v \in E$ são unicamente determinados nesta base. Devido a isto, estes coeficientes são chamados de *coordenadas de v na base* B.

Para o próximo teorema, é necessário o seguinte resultado:

Lema 1.3. *Todo sistema linear homogêneo cujo número de incógnitas é maior do que o número de equações admite uma solução diferente da trivial, i.e., $(x_1, \ldots, x_n) \neq (0, \ldots, 0)$.*

Demonstração.
Considere o sistema a seguir

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

onde $m < n$.

Por indução em m, no caso em que $m = 1$, tem-se

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0,$$

onde $a_{1i} \neq 0$ para algum $1 \leq i \leq n$. Sem perda de generalidade, se $\alpha_{1n} \neq 0$, então

$$x_n = -\frac{\alpha_{11}}{\alpha_{1n}}x_1 - \ldots - \frac{\alpha_{1n-1}}{\alpha_{1n}}x_{n-1}.$$

Atribuindo valores não-nulos para $x_1, \ldots, x_{n-1}$, obtém-se uma solução não trivial para o caso em que $m = 1$.

Pelo passo indutivo, suponha que o lema é verdadeiro quando há $m-1$ equações. No caso em que temos m equações, sem perda de generalidade, segue da última equação do sistema que

$$x_n = -\frac{\alpha_{11}}{\alpha_{mn}}x_1 - \ldots - \frac{\alpha_{1n-1}}{\alpha_{mn}}x_{n-1}.$$

Substituindo isto nas demais equações, obtém-se um sistema com $m-1$ equações e $n-1$ incógnitas. Como $m-1 < n-1$, a partir da hipótese de indução segue que este sistema tem solução não trivial $(y_1, \ldots, y_{n-1})$. Definindo

$$y_n = -\frac{\alpha_{11}}{\alpha_{mn}}y_1 - \ldots - \frac{\alpha_{1n-1}}{\alpha_{mn}}y_{n-1},$$

obtém-se uma solução não-trivial para o sistema inicial. ■

Teorema 1.5. *Sejam E um espaço vetorial e $m \in \mathbb{N}$ tal que $\{v_1, \ldots, v_m\}$ é um conjunto de geradores de E. Qualquer subconjunto de E que tenha mais de m elementos é linearmente dependente.*

Demonstração.

Sejam n um natural maior que m e $\{w_1, \ldots, w_n\}$ um subconjunto de E. Como $\{v_1, \ldots, v_m\}$ é um conjunto de geradores de E, então para cada $1 \leq j \leq n$ temos $w_j = \sum_i^n \alpha_{ij}v_i$. Para garantir a existência de coeficientes $y_1, \ldots, y_n \in \mathbb{R}$ não todos nulos e tais que

$$y_1w_1 + \ldots + y_nw_n = 0,$$

uma condição suficiente é que cada um dos somatórios abaixo seja igual a zero

$$\left(\sum_{j=1}^{n} \alpha_{1j}y_j\right) v_1 + \ldots \left(\sum_{j=1}^{n} \alpha_{mj}y_j\right) v_m = 0.$$

Daí, temos um sistema com m equações e $n > m$ incógnitas. Pelo lema 1.3, segue que este sistema possui solução não-trivial $(y_1, \ldots, y_n)$. ■

Corolário 1.2. *Suponha que $\{v_1, \ldots, v_m\}$ gera E. Se $\{u_1, \ldots, u_n\}$ é L.I., então $n \leq m$.*

Demonstração.

Contrapositiva de "*se $m < n$, então $\{u_1, \ldots, u_n\}$ é L.D.*". ■

Corolário 1.3. *A dimensão de um espaço vetorial de dimensão finita é bem definida.*

Demonstração.
Suponha que E é um espaço vetorial de dimensão finita n e que B e C são bases de E com respectivamente n e m. Pela definição de base, temos que B gera E e que C é L.I., portanto $m \leq n$ devido ao corolário 1.3. Um pensamento análogo mostra que $n \leq m$. A partir das duas desigualdades, conclui-se que $m = n$. ■

Corolário 1.4. *Seja E um espaço vetorial de dimensão finita,* $\dim E = n$*. Um conjunto com n vetores gera E se e somente se for L.I..*

Demonstração.
Para mostrar a suficiência, suponhamos que $X = \{v_1, \ldots, v_n\}$ é um conjunto $L.D.$ de geradores de E com n elementos. Sem perda de generalidade, se v_n for combinação linear dos demais, segue que $X \setminus \{v_n\}$ é uma base de E com $n - 1$ elementos, uma contradição, pois $\dim E = n$.

Para mostrar a necessidade, suponha que exista $v \in E$ tal que $v \notin S(X)$. Disto segue que $X \cup \{v\}$ é um conjunto L.I. (teorema 1.4) com mais de n elementos. Como $\dim E = n$, dada qualquer base de E, segue do teorema 1.5 que qualquer conjunto com mais de n elementos é L.D.. Obtém-se assim uma contradição e portanto $S(X) = E$. ■

Exemplo 1.18. *Os vetores canônicos em $\mathbb{R}^n$ constituem uma base deste espaço vetorial. Em $M(m \times n)$, o conjunto das matrizes $\mathbf{e_{ij}}$, i.e., aquelas cuja ij-ésima entrada vale 1 e as demais valem zero, é uma base deste espaço. Em vista disso, claramente temos que* $\dim \mathbb{R}^n = n$ *e* $\dim M(m \times n) = mn$.

Teorema 1.6. *Seja E um espaço vetorial de dimensão finita n. Então:*

a) Todo conjunto X de geradores de E contém uma base.

b) Todo conjunto L.I. $\{v_1, \ldots, v_m\} \subset E$ está contido numa base.

c) Todo subespaço vetorial $F \subset E$ tem dimensão finita $\dim F \leq n$.

d) Seja $F \subset E$ um subespaço vetorial. Se $\dim F = n$, *então $F = E$.*

Demonstração.

a) Seja $Y = \{v_1, \ldots, v_m\} \subset X$ um conjunto L.I. com o número máximo de elementos. Se existe $v \in X$ tal $v \in S(Y)$, então $Y \cup \{x\}$ é um subconjunto L.I. (teorema 1.4) de X com cardinalidade maior que Y, o que contraria a maximalidade de Y. Disto segue que $X \subset S(Y)$ e sendo $S(X) = E$, temos que $E \subset S(Y)$. Facilmente conclui-se que Y é uma base.

b) Se $m = n$, basta utilizar o corolário 1.4. Se $m < n$, então $S(\{v_1, \ldots, v_m\}) \neq E$, pois $\{v_1, \ldots, v_m\}$ não é uma base. Daí, se $v \notin S(\{v_1, \ldots, v_m\})$, então $\{v_1, \ldots, v_m, v\}$ é L.I.. Se $\#\{v_1, \ldots, v_m, v\} < n$, repete-se o processo até obter-se um conjunto L.I. com n vetores e aplica-se o corolário 1.4.

c) Seja $Y = \{v_1, \ldots, v_m\} \subset F$ um conjunto L.I. com o número máximo de elementos. Pelo item anterior, segue que Y está contido numa base B de E. Daí, pelo corolário 1.2 temos que $m \leq n$. Além disso, como F é subespaço, podemos concluir que $S(Y) = F$. Sendo assim, $\dim F = m \leq n$.

d) Seja Y o mesmo conjunto do item anterior. Sendo $\dim F = n$, pelo corolário 1.4 segue que Y é uma base de E, ou seja, vale que $S(Y) = E$. Como $S(Y) = F$, por transitividade concluímos que $F = E$.

■

Exercício 1.7. *Mostre que o conjunto* $\{e^x, e^{2x}, x^3, x^2, x\} \subset C^\infty(\mathbb{R})$ *é linearmente independente.*

2 Transformações Lineares

2.1 Transformações Lineares

Definição 2.1. *Sejam E e F espaços vetoriais. Uma* ***transformação linear*** *é um mapa $A\colon E \to F$, $v \mapsto A(v) \equiv Av \equiv A \cdot v$, que satisfaz as seguintes propriedades:*

1. $A(u + v) = Av + Au$.

2. $A(\alpha v) = \alpha \cdot Av$

para todos $u, v \in E$ e para todo $\alpha \in \mathbb{R}$. Por vezes, diz-se que Av é a imagem de v pela transformação A.

É imediato que se $A\colon E \to F$ é uma transformação linear, então $A(0) = 0$, $A(\sum_i^n \alpha_i v_i) = \sum_i^n \alpha_i A v_i$ e que $A(-v) = -Av$.

Exercício 2.1. *Seja $L(E; F)$ o conjunto das transformações lineares de E em F. Mostre que esse conjunto é um espaço vetorial*[5] *quando dotado das operações usuais de soma e multiplicação por escalar.*

Exercício 2.2. *Sejam E e F espaços vetoriais e $A \in L(E; F)$. Mostre que a imagem de E por A, $\mathcal{I}m(A) := \{u \in F\colon \exists v \in E \text{ tal que } A \cdot v = u\}$, é um subespaço vetorial de F.*

Definição 2.2. *Uma transformação linear $A \in L(E; E)$ é chamada de* ***operador linear****. Para facilitar a notação, deste ponto em diante, escreveremos $L(E)$ em vez de $L(E; E)$, sendo $End(E)$ outra notação possível* [6].

[5] No caso, ele é um subespaço vetorial de $\mathcal{F}(E; F)$.

[6] Neste caso, $End(E)$ representa o conjunto de endomorfismos de E em E, ou seja, o conjunto de aplicações de E em si mesmo que preservam a estrutura algébrica de E. Pela generalidade da notação, é possível encontrá-la em contextos diversos, que fogem do escopo da álgebra linear.

Definição 2.3. *O* ***dual algébrico*** *de um espaço vetorial* E *é o conjunto* $L(E;\mathbb{R})$*, que deste ponto em diante será denotado por* E^**. Os elementos de* E^* *são denominados* ***funcionais lineares****.*

A presença do termo *algébrico* na definição acima pode ser um tanto pedante, mas ocorre que nas situações em que E é também um espaço normado, pode-se definir o *dual topológico de* E como sendo o conjunto dos funcionais lineares contínuos (na topologia induzida pela norma). No caso de dimensão finita, não há distinção entre os duais algébrico e topológico de um espaço vetorial normado, pois qualquer funcional linear é contínuo (exercício 2.3).

Exercício 2.3. *Seja* X *um conjunto. Uma topologia em* X *é uma coleção* τ *de subconjuntos (nomeados abertos) de* X *que satisfaz aos seguintes critérios:*

1. X *e* $\emptyset$ *pertencem a* τ*;*
2. *Dada uma família indexada* $\{A_\lambda\}_{\lambda\in\Lambda}$ *tal que* $A_\lambda \in \tau$ *para cada índice, tem-se* $\bigcup_{\lambda\in\Lambda} A_\lambda \in \tau$*;*
3. *Dados* $A_1, ..., A_n \in \tau$*, tem-se que* $A_1 \cap ... \cap A_n \in \tau$*.*

Daí, um espaço topológico é uma dupla (X,τ)*, onde* τ *é uma topologia no conjunto* X*.*

Uma função $f\colon X \to Y$ *entre espaços topológicos é dita contínua se para todo aberto* $A \in \tau_Y$*, sua pré-imagem* $f^{-1}(A) \subset X$ *é um elemento de* τ_X*, i.e., é um aberto de* X*.*

Um espaço vetorial topológico E *é, portanto, um espaço vetorial onde a soma e a multiplicação por escalar são operações contínuas com relação às topologias produto*[7] *em* $E \times E$ *e* $\mathbb{R} \times E$*.*

Dado um espaço vetorial E*, uma norma neste espaço é uma função* $\|\cdot\|\colon E \to \mathbb{R}$ *que respeita as seguintes propriedades:*

1. $\|\alpha v\| = |\alpha|\|v\|$ *para todos* $v \in E$ *e todos* $\alpha \in \mathbb{R}$*;*
2. $\|v\| = 0 \iff v = 0$*;*
3. $\|v + u\| \leq \|v\| + \|u\|$ *para todos* $u, v \in E$*.*

Dados $r > 0$ *e* $u \in E$*, podemos definir em* E *a bola aberta de raio* r *centrada em* u *como sendo o conjunto*

$$B_r(u) := \{v \in E\colon \|v - u\| < r\}.$$

Com isto, podemos induzir uma topologia em E *ao impormos que os abertos desta topologia são, por definição, uniões de bolas abertas.*

Mostre que num espaço vetorial normado E *de dimensão finita* n*, todo funcional linear* $f \in E^*$ *é contínuo na topologia induzida pela norma.*

[7] A topologia produto é aquela cujos abertos são uniões de elementos da forma $U \times V \in \tau_E \times \tau_E$ ou $(a,b) \times V \in \tau_{\mathbb{R}} \times \tau_E$.

Dica: Fixe uma base em E e defina a seguinte norma em E:

$$\|v\| = \sum_{i=1}^{n} |\alpha_i|,$$

onde α_i são os coeficientes que determinam v na base fixada. Mostre que dado $f \in E^*$, existe $M > 0$ finito tal que para todo $v \in E$:

$$|f(v)| \leq M\|v\|.$$

Isto significa que todo funcional linear é limitado. Daí, mostre que um funcional limitado é contínuo (no sentido visto em Cálculo mesmo).

Teorema 2.1. *Sejam E, F espaços vetoriais e B uma base de E. Se para cada $u \in \mathrm{B}$ escolhemos um vetor $u' \in F$, então existe uma única transformação linear $A\colon E \to F$ tal que $A \cdot u = u'$ para cada $u \in \mathrm{B}$.*

Demonstração.
Fixada uma base B em E, por conta do corolário 1.1 segue que todo $v \in E$ é unicamente determinado por coeficientes $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ – assumindo, claro, que E tem dimensão finita –. Daí, sob as hipóteses do teorema, basta definir $A\colon E \to F$ como sendo o seguinte mapa:

$$v = \sum_{i=1} \alpha_i u_i \in E \mapsto \sum_{i=1}^{n} \alpha_i u_i' \in F,$$

onde $u_i' = A \cdot u_i$. A linearidade e a unicidade são imediatas. ∎

Em outras palavras, o teorema 2.1 informa que uma transformação linear fica inteiramente definida a partir da forma como ela atua em uma base de E. Além do mais, em virtude deste teorema, podemos escolher certos funcionais lineares em E^* tal que vale o resultado a seguir:

Teorema 2.2. *Seja $\mathrm{B} = \{v_1, \ldots, v_n\}$ uma base de um espaço vetorial E de dimensão finita. Para cada $i = 1, \ldots, n$, seja $f_i\colon E \to \mathbb{R}$ o funcional linear determinado por $f_i(v_j) = \delta_{ij}$, onde δ_{ij} representa o delta de Kronecker. Nestas condições, o conjunto $\{f_1, \ldots, f_n\}$ é uma base de E^*.*

Demonstração.
Primeiramente, mostraremos que $S(\{f_1, \ldots, f_n\}) = E^*$. Com efeito, seja $\omega \in E^*$ qualquer. Fixada a base B em E, seja $v = \sum_i^n \alpha_i v_i \in E$ arbitrário. Pela linearidade de ω segue que

$$\omega(v) = \sum_{i=1}^{n} \alpha_i \omega(v_i).$$

Veja que pela definição dos f_i, para cada $1 \leq i \leq n$ vale que

$$f_i(v) = \sum_{j=1}^{n} \alpha_j f_i(v_j) = \alpha_i.$$

Desta forma, temos que

$$\omega(v) = \sum_{i=1}^{n} \omega(v_i) f_i(v).$$

Como os valores $\omega(v_i)$ dependem apenas da base fixada, temos que

$$\omega = \sum_{i=1}^{n} \beta_i f_i,$$

onde $\beta_i \equiv \omega(v_i)$. Sendo ω qualquer, podemos concluir que $\{f_1, \ldots, f_n\}$ é um conjunto de geradores de E^*.

Por último, mostraremos a independência linear. Sejam $\beta_1, \ldots, \beta_n \in \mathbb{R}$ tais que

$$\sum_{i=1}^{n} \beta_i f_i = 0.$$

Temos que para todo $v_j \in$ B vale:

$$0 = \sum_{i=1}^{n} \beta_i f_i(v_j) = \beta_j.$$

Logo, $\{f_1, \ldots, f_n\}$ é L.I..

Das conclusões acima, segue que $\{f_1, \ldots, f_n\}$ é uma base para E^*. ■

A base definida no teorema 2.2 é denominada de *base dual de* B. Igualmente, é possível mostrar que dada uma base em E^*, existe uma base V em E tal que B seja a base dual de V (fica como exercício). Facilmente, também pode ser mostrado que a base dual é única para espaços vetoriais de dimensão finita (também fica como exercício).

Exemplo 2.1. *A integração de uma função contínua sobre um intervalo fechado pode ser vista como um funcional linear* $\varphi\colon C^0([a,b]) \to \mathbb{R}$ *que mapeia uma função* $f \in C^0([a,b])$ *no número real*

$$\varphi(f) = \int_a^b f(x)\mathrm{d}x.$$

Já em $C^\infty(\mathbb{R})$*, a derivada pode ser tratada como um operador linear* $D\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ *que mapeia* f *em sua função derivada* $Df = f'$*.*

Exemplo 2.2. *Sejam* (X, Σ, μ) *um espaço de medida e* $\mathcal{L}^p(X,\mu)$ *o conjunto das funções* $f\colon X \to \mathbb{R}$ *tais que* $\int_X |f|^p \mathrm{d}\mu < \infty$ *. Se impormos que* $f \sim g \iff f = g$ *q.t.p*$[\mu]$*, ou seja, a medida do conjunto onde* $f \neq g$ *é zero, então o conjunto*

$$L^p(X,\mu) := \mathcal{L}^p(X,\mu)/\sim$$

é o espaço de Lebesgue L^p, onde consideramos $1 \leq p < \infty$. No caso em que $X = \mathbb{R}$, μ é a medida de Lebesgue e $p = 1$, a transformada de Fourier é o mapa $F\colon L^1(\mathbb{R}) \to C^0(\mathbb{R})$ definido por

$$Ff(k) \equiv \hat{f}(k) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)e^{-ikx}\mathrm{d}x.$$

Como $L^1(\mathbb{R})$ torna-se um espaço vetorial quando dotado das operações usuais, segue que F é uma transformação linear.

Exemplo 2.3. *Seja $A \in L(\mathbb{R}^n; \mathbb{R}^m)$. Pelo teorema 2.1, se fixarmos a base canônica em ambos os espaços e definirmos que, para cada $j = 1, \ldots, n$, $A \cdot e_j = v_j \in \mathbb{R}^m$, onde $v_j = (a_{1j}, \ldots, a_{mj})$, então a transformação linear A está inteiramente definida e para $v = (x_1, \ldots, x_n)$ temos que*

$$A \cdot v = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}x_j e_j = \sum_{j=1}^{m} y_i e_i.$$

Pela unicidade dos coeficientes, $A \cdot v$ definas seguintes igualdades:

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

Tratando os vetores em $\mathbb{R}^n$ e em $\mathbb{R}^m$ como vetores-coluna, pelas igualdades acima podemos inferir que $A\cdot$ equivale à multiplicação de uma matriz **a** *pelo vetor-coluna v:*

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

Perceba que a j-ésima coluna da matriz **a** *tem como entrada justamente as coordenadas da imagem de e_j pela transformação A. Por costume, chama-se* **a** *de **matriz da transformação A relativa às bases de** $\mathbb{R}^n$ **e** $\mathbb{R}^m$.*

Exemplo 2.4. *Seja $\theta \in [0, 2\pi)$ fixado. A transformação linear $\mathrm{R}(\theta)\colon \mathbb{R}^2 \to \mathbb{R}^2$ definida por*

$$\mathrm{R}(\theta) \cdot (x, y) := (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$$

representa geometricamente a rotação no sentido anti-horário pelo ângulo θ do vetor $v = (x, y)$ em torno da origem. Na base canônica, temos que a matriz de $\mathrm{R}(\theta)$ é

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

Exemplo 2.5. *Se* $\{(x', \alpha x') \in \mathbb{R}^2 \colon x \in \mathbb{R}\}$ *é uma reta que passa pela origem, então a projeção de um vetor* $v = (x, y) \in \mathbb{R}^2$ *sobre esta reta é obtida por meio do operador* $P \colon \mathbb{R}^2 \to \mathbb{R}^2$*, dado por*

$$P \cdot v = \left(\frac{x + \alpha y}{1 + \alpha^2}, \frac{\alpha x + \alpha^2 y}{1 + \alpha^2} \right).$$

Na base canônica, sua matriz é

$$\begin{pmatrix} \dfrac{1}{1+\alpha^2} & \dfrac{\alpha}{1+\alpha^2} \\ \dfrac{\alpha}{1+\alpha^2} & \dfrac{\alpha^2}{1+\alpha^2} \end{pmatrix}.$$

Exemplo 2.6. *No caso em que deseja-se refletir um vetor* $v \in \mathbb{R}^2$ *em torno de uma reta, podemos fazer uso de um operador linear* $S \colon \mathbb{R}^2 \to \mathbb{R}^2$ *que pode ser expresso como*

$$S = 2P - I,$$

onde P *é o operador do exemplo anterior e* $I \colon \mathbb{R}^2 \to \mathbb{R}^2$ *é o mapa identidade.*

Exemplo 2.7. *Seja* V *um espaço vetorial sobre* $\mathbb{R}$ *e considere* $\omega, \eta \in V^*$*, onde* V^* *é o dual algébrico de* V*. Defina uma função* $\omega \otimes \eta \colon V \times V \to \mathbb{R}$ *por*

$$\omega \otimes \eta(v_1, v_2) := \omega(v_1)\eta(v_2).$$

Como ω *e* η *são funcionais lineares, facilmente mostra-se que* $\omega \otimes \eta$ *é uma forma bilinear e portanto* $\omega \otimes \eta \in L(V, V; \mathbb{R})$*.*

Generalizando, suponha que $V_1, \ldots, V_k$ *e* $W_1, \ldots, W_l$ *são espaços vetoriais sobre* $\mathbb{R}$*. Se* $F \in L(V_1, \ldots, V_k; \mathbb{R})$ *e* $G \in L(W_1, \ldots, W_l; \mathbb{R})$*, então a função* $F \otimes G \colon V_1 \times \ldots \times V_k \times W_1 \times \ldots \times W_l \to \mathbb{R}$ *definida por*

$$F \otimes G(v_1, \ldots, v_k, w_1, \ldots, w_l) := F(v_1, \ldots, v_k)G(w_1, \ldots, w_l)$$

é um elemento de $L(V_1, \ldots, V_k, W_1, \ldots, W_l; \mathbb{R})$ *e chama-se* $F \otimes G$ *de* ***produto tensorial de*** $\boldsymbol{F}$ ***e*** $\boldsymbol{G}$*.*

Exercício 2.4. *Sejam* $V_1, \ldots, V_k$ *espaços vetoriais sobre* $\mathbb{R}$*. Para cada* $j \in \{1, \ldots, k\}$*, suponha que* $\{E_1^{(j)}, \ldots, E_{n_j}^{(j)}\}$ *é uma base para* V_j *e seja* $\{\varepsilon_{(j)}^1, \ldots, \varepsilon_{(j)}^{n_j}\}$ *a base dual para* V_j^**. Então o conjunto*

$$\mathcal{B} := \{\varepsilon_{(1)}^{i_1} \otimes \ldots \otimes \varepsilon_{(k)}^{i_k} \colon 1 \leq i_1 \leq n_1, \ldots, 1 \leq i_k \leq n_k\}$$

é uma base para $L(V_1, \ldots, V_k; \mathbb{R})$*.*

2.1.1 Produto de Transformações Lineares

Sejam E, F, G espaços vetoriais. Se $A \in L(E; F)$ e $B \in L(F, G)$, podemos definir a transformação linear $BA \in L(E; G)$ como sendo aquela que para todo $v \in E$:

$$BA(v) := B(Av).$$

Como corolário da definição, valem as propriedades a seguir

A1) $\alpha(B \circ A) = (\alpha B) \circ A = B \circ (\alpha A)$.

A2) $B \circ (A + C) = B \circ A + B \circ C$, onde $C \in L(E; F)$.

A3) $(B + C) \circ A = B \circ A + C \circ A$, onde $C \in L(F; G)$.

A4) $C \circ (B \circ A) = (C \circ B) \circ A$, onde $C \in L(G; H)$, sendo H outro espaço vetorial.

No caso particular em que $E = F = G = H$, a operação binária $\circ\colon L(E) \times L(E) \to L(E)$ transforma $L(E)$ no que chamamos de **álgebra associativa sobre** $\mathbb{R}$. Mais geralmente, uma álgebra sobre $\mathbb{R}$ é um espaço vetorial E dotado com um mapa $(u, v) \in E \times E \mapsto u \circ v \in E$ que satisfaz às três primeiras propriedades acima. Alguns exemplos banais de álgebras são $\mathbb{R}^3$ com o produto vetorial e $C^\infty(\mathbb{R})$ com o produto pontual $(fg)(x) = f(x)g(x)$.

Exercício 2.5. *Para mostrar que a composição de transformações lineares não é necessariamente comutativa, mostre que $R(\pi/2)P \neq PR(\pi/2)$. Além disso, usando o operador P, mostre que não vale a lei do corte.*

Definição 2.4. *Um operador $A \in L(E)$ é dito **nilpotente** se existe $n \in \mathbb{N}$ tal que $A^n = 0$.*

Exemplo 2.8. *O operador derivação $D\colon \mathcal{P}_n \to \mathcal{P}_n$ é nilpotente, visto que para qualquer $p \in \mathcal{P}_n$, $D^{n+1}p = 0$.*

2.1.2 Núcleo e Imagem

Se A é uma transformação linear do espaço E no espaço F, já sabemos que sua imagem, $\mathcal{I}m(A)$, é um subespaço vetorial de F(exercício 2.2). Esse subespaço ao lado do subespaço $\ker A$, denominado núcleo da transformação A, nos permite caracterizar quando A é injetiva, sobrejetiva ou inversível. Torna-se, portanto, imprescindível apresentar os resultados relacionados a tais subespaços.

Exemplo 2.9. *Seja $\varphi\colon E \to \mathbb{R}$ um funcional linear relativo a um espaço vetorial E. Os únicos subespaços da reta são os triviais, ou seja, $\{0\}$ e o próprio $\mathbb{R}$. Deste modo, φ ou é nulo ou é sobrejetivo, visto que $\mathcal{I}m(\varphi)$ é um subespaço de $\mathbb{R}$.*

Definição 2.5. *Seja $A\colon E \to F$ uma transformação linear. Uma **inversa à direita de A** é uma transformação linear $B\colon F \to E$ tal que $AB = I_F$, onde I_F é o mapa identidade em F.*

Teorema 2.3. *Uma transformação linear $A\colon E \to F$ entre espaços de dimensão finita admite inversa à direita se e somente se for sobrejetora.*

Demonstração.
Suponha que $A \in L(E; F)$ admite inversa à direita $B\colon F \to E$. Se $w \in F$ é qualquer, então para $v = Bw \in E$, temos que $Av = A(Bw) = w$. Donde conclui-se que w é imagem de v pela transformação A e portanto A é sobrejetora.

Agora, suponha que $A \in L(E;F)$ é sobrejetora. Fixada uma base $\mathrm{B} = \{w_1, \ldots, w_m\}$ em F. Pela sobrejetividade, existem $v_1, \ldots, v_m \in E$ distintos e tais que $Av_j = w_j$ para $1 \leq j \leq m$. Se para cada w_j associarmos o respectivo vetor v_j, está garantida a existência de uma única transformação linear $B\colon F \to E$ tal que $Bw_j = v_j$ (teorema 2.1). Daí, se $w = \sum_{j=1}^{m} \beta_j w_j \in F$ é qualquer, segue que

$$\begin{aligned} A(Bw) &= A\left(\sum_{j=1}^{m} \beta_j v_j\right) \\ &= \sum_{j=1}^{m} \beta_j A v_j \\ &= \sum_{j=1}^{m} \beta_j w_j \\ &= w. \end{aligned}$$

Desta forma, B é uma inversa à direita de A. ■

Definição 2.6. *Seja $A\colon E \to F$ uma transformação linear. O **núcleo de A** por definição é o conjunto*

$$\ker A := \{v \in E\colon Av = 0\}.$$

Exercício 2.6. *Seja $A \in L(E;F)$. Mostre que $\ker A$ é um subespaço de E quando herda sua estrutura de espaço vetorial.*

Exemplo 2.10. *Fixado $a \in \mathbb{R}$ positivo, seja $A\colon \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear definida por $A(x,y) := (ay, 0)$. Tanto seu núcleo como sua imagem são o eixo x.*

Exemplo 2.11. *Seja $A\colon C^0(\mathbb{R}) \to C^0(\mathbb{R})$ a transformação linear $f \mapsto Af$, onde Af é a função contínua definida por*

$$Af(x) := \int_0^x f(t)\mathrm{d}t.$$

Se $f \in \ker A$, temos que $\int_0^x f(t)\mathrm{d}t = 0$ para todo $x \in \mathbb{R}$. Daí, pelo teorema fundamental do cálculo, segue que $(Af)'(x) = f(x) = 0$ para todo $x \in \mathbb{R}$ e portanto $\ker A = \{0\}$. A partir da definição de $\mathcal{I}m(A)$ e do teorema fundamental do cálculo, podemos concluir facilmente que a imagem de A é o conjunto $X = \{f \in C^1(\mathbb{R})\colon f(0) = 0\} \subset C^0(\mathbb{R})$.

Lema 2.1. *Para que uma transformação linear $A \in L(E;F)$ seja injetora é necessário e suficiente que $\ker A = \{0\}$.*

Demonstração.

Para mostrar a necessidade, suponha que $A \in L(E;F)$ é injetora. Se $v, u \in \ker A$, pela definição de $\ker A$, temos que $Av = Au$. Daí pela injetividade de A

segue que $Av = Au \implies u = v$, logo o núcleo de A possui apenas um elemento. Como é trivial que $0 \in \ker A$, segue que $\ker A = \{0\}$.

Para mostrar a suficiência, suponha que $\ker A = 0$. Se $u, v \in E$ são tais que $Av = Au$, então $A(u - v) = 0$ e portanto $u - vin \ker A$. Pela hipótese sobre o núcleo de A, só podemos ter $u = v$. ∎

Lema 2.2. *Para que uma transformação linear $A \in L(E; F)$ seja injetora é necessário e suficiente que a imagem de todo conjunto L.I. em E seja também um conjunto L.I. em F*

Demonstração.

Necessidade: suponha que $A\colon E \to F$ é linear e injetora. Se $\{v_1, \ldots, v_n\} \subset E$ é L.I., então dada uma combinação linear nula

$$\sum_{i=1}^{n} \alpha_i A v_i = 0,$$

pela linearidade de A e pelo teorema 1.3, conclui-se que $\alpha_i =$ para cada $i = 1, \ldots, n$. Daí, $\{Av_1, \ldots, Av_n\}$ é L.I. em F.

Suficiência: se $u, v \in E$ são distintos, então $\{u, v\}$ é L.I. e consequentemente $Au \neq Av$ por conta da hipótese. Logo, A é injetora. ∎

Definição 2.7. *Seja $A\colon E \to F$ uma transformação linear. Uma* ***inversa à esquerda de A*** *é uma transformação linear $B\colon F \to E$ tal que $BA = I_E$, onde I_E é o mapa identidade em E.*

Teorema 2.4. *Uma transformação linear $A\colon E \to F$ entre espaços de dimensão finita admite inversa à esquerda se e somente se A for injetora.*

Demonstração.

Suponha que $A \in L(E; F)$ admite uma inversa à esquerda $B \in L(F; E)$. Se $v \in \ker A$, então $Av = 0 \in F$. Daí, $B(Av) = B0 = 0 \in E$ devido à linearidade de B. Por outro lado, visto que B é uma inversa à esquerda, $B(Av) = v$. Daí, conclui-se que $v = 0$ e pelo lema 2.1 segue que A é injetora.

Suponha que A é injetora. Fixada uma base $\{v_1, \ldots, v_n\} \subset E$, segue do lema 2.2 que $\{Av_1, \ldots, Av_n\} \subset F$ é linearmente independente e pelo teorema 1.6 este conjunto está contido numa base $\{u_1, \ldots, u_n, u_{n+1}, \ldots, u_m\} \subset F$, onde $u_i = Av_i$ para $1 \leq i \leq n$. A partir da associação $u_i \mapsto v_i$ para $1 \leq i \leq n$ e $u_j \mapsto 0$ para $n + 1 \leq j \leq m$, pelo teorema 2.1 está definida uma única transformação linear $B\colon F \to E$. Se $v = \sum_{k=1}^{n} \alpha_k v_k$, segue que

$$\begin{aligned} B(Av) &= \sum_{k=1}^{n} \alpha_k BA(v_k) \\ &= \sum_{k=1}^{n} \alpha_k v_k \\ &= v. \end{aligned}$$

Logo, B é uma inversa à esquerda de A. ∎

Definição 2.8. *Uma transformação linear* $A\colon \in L(E;F)$ *é dita* ***invertível*** *se existe* $B \in L(F;E)$ *tal que* $BA = I_F$ *e* $AB = I_E$. *Diz-se que* B *é a inversa de* A *e por convenção denota-se* B *por* A^{-1}.

Definição 2.9. *Sejam* E *e* F *espaços vetoriais de dimensão finita. Diz-se que* E *é isomorfo a* F *se existe uma transformação linear* $A\colon E \to F$ *invertível. Neste caso, diz-se que* A *é um isomorfismo entre os espaços* E *e* F.

A noção de isomorfismo não restringe-se apenas a espaços vetoriais. No contexto de *teoria de grupos*[8], um isomorfismo é um mapa $\varphi\colon G \to H$ bijetor (isto é, injetivo e sobrejetivo) entre grupos G e H quaisquer e que preserva a estrutura de grupo, isto é, se g e p pertencem a G, então $\varphi(g \cdot_G p) = \varphi(g) \cdot_H \varphi(p)$, onde $\cdot_G$ e $\cdot_H$ são as operações de grupo em G e H respectivamente. Num sentido ainda mais geral, a ideia por trás dos isomorfismos é obter um mapa entre objetos de mesma estrutura tal que este mapa seja invertível e preservador de estruturas; neste contexto homeomorfismos entre espaços topológicos, difeomorfismos entre variedades diferenciáveis, homomorfismos invertíveis entre fibrados vetoriais e isometrias entre espaços métricos podem ser interpretados como isomorfismos.

Exercício 2.7. *Sejam* E, F *e* G *espaços vetoriais. Se* $A \in L(E;F)$ *e* $B \in L(F;G)$ *são isomorfismos, mostre que* $A^{-1} \in L(F;E)$ *e* $BA \in L(G;E)$ *são isomorfismos. Além disso,* $(AB)^{-1} = A^{-1}B^{-1}$.

Lema 2.3. *Seja* $A\colon E \to F$ *uma transformação linear que possui uma inversa à esquerda* $B \in L(F;E)$ *e uma inversa à direita* $C \in L(E;F)$. *Então* $B = C$ *e* E *é isomorfo a* F.

Demonstração.
Se B e C são respectivamente inversas à esquerda e à direita de A, então:

$$B = BI_F = B(AC) = (BA)C = I_E C = C.$$

Logo, vale que $AB = I_F$ e já que hipótese $BA = I_E$, conclui-se que $B = A^{-1}$. Por definição de espaços isomorfos, conclui-se que $E \simeq F$. ∎

Lema 2.4. *Sejam* E *e* F *espaços vetoriais de dimensão finita.* E *é isomorfo a* F *se e somente se* $\dim E = \dim F$.

Demonstração.
Suponha $E \simeq F$, onde ambos os espaços vetoriais possuem dimensão finita. Sendo A um isomorfismo de E em F e $\mathcal{B} \subset E$ uma base, segue do lema 2.2 que a imagem de $\mathcal{B}$ por A é um conjunto L.I. de F. Como A é sobrejetivo, facilmente conclui-se que qualquer vetor w pertencente a F é combinação linear de elementos imagem de B por A. Pelas definições de base e de dimensão de um espaço vetorial, é trivial ver que $\dim E = \dim F$.

[8]Grosso modo, um *grupo* é uma dupla $(G, \cdot_G)$ onde G é um conjunto e $\cdot_G$ é uma operação binária em G que possui as seguintes propriedades: $\cdot_G$ é associativa; existe um único elemento neutro $e \in G$ tal que $e \cdot_G g = g \cdot_G e = g$; e para cada g existe um único $g^{-1} \in G$ tal que $g \cdot_G g^{-1} = g^{-1} \cdot_G g = e$. Claramente, espaços vetoriais são grupos com relação à soma.

Suponha que $\dim E = n = \dim F$, sendo que $n \in \mathbb{N}$. Primeiramente, mostraremos que qualquer espaço V de dimensão finita n é isomorfo a $\mathbb{R}^n$. Para tanto, fixada uma base $\{v_1, \ldots, v_n\} \subset V$, seja $\varphi\colon V \to \mathbb{R}^n$ a aplicação linear dada por $v_i \mapsto e_i$. Claramente, φ define um isomorfismo entre V e $\mathbb{R}^n$. Em vista disso, se $A\colon E \to \mathbb{R}^n$ e $B\colon F \to \mathbb{R}^n$ são isomorfismos, então $B^{-1}A\colon E \to F$ também é um isomorfismo. Logo, $E \simeq F$. ■

Exemplo 2.12. *Uma matriz* **a** *é dita triangular inferior se todos os elementos acima de sua diagonal principal são nulos, ou seja, $a_{ij} = 0$ quando $i < j$.*

Para a demonstração de um dos teoremas mais importantes desta seção, faremos uso do lema abaixo.

Lema 2.5. *Sejam E e F espaços vetoriais de dimensão finita e seja $A\colon E \to F$ uma transformação linear. Se $\{v_1, \ldots, v_n\}$ e $\{Av_{n+1}, \ldots, Av_{n+m}\}$ são respectivamente bases de $\ker A$ e de $\mathcal{I}m(A)$, então $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+m}\}$ é uma base de E.*

Demonstração.
Primeiramente, mostraremos que $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+m}\}$ é L.I.. Considere que existem $m + n$ coeficientes α_i tais que

$$\sum_{i=1}^{n+m} \alpha_i v_i = 0.$$

Pela linearidade de A e por $A \cdot 0 = 0$, segue que

$$\sum_{i=1}^{n+m} \alpha_i A v_i = 0.$$

Como $Av_i = 0$ para os primeiros n termos e sendo $\{Av_{n+1}, \ldots, Av_{n+m}\}$ L.I., conclui-se que $\alpha_i = 0$ para $n+1 \leq i \leq n+m$. Como $\{v_1, \ldots, v_n\}$ também é L.I., então os coeficientes presentes no somatório restante $\sum_{i=1}^{n} \alpha_i v_i = 0$ também são nulos, donde concluímos que $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+m}\}$ é linearmente independente.

Por último, $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+m}\}$ gera E. Com efeito, para $v \in E$ qualquer, temos que $Av = \sum_{i=1}^{m} \lambda_{n+i} A v_{n+i}$ para determinados coeficientes λ_{n+i}, pois $\{Av_{n+1}, \ldots, Av_{n+m}\}$ é base de $\mathcal{I}m(A)$. Disto, segue que $v - \sum_{i=1}^{m} \lambda_{n+i} v_{n+i} \in \ker A$. Visto que $\{v_1, \ldots, v_n\}$ é base do núcleo, temos podemos reescrever v como

$$v = \sum_{i=1}^{n+m} \lambda_i v_i.$$

Assim, facilmente concluímos que $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+m}\}$ de fato é um conjunto de geradores de E. ■

Teorema 2.5 (Núcleo e imagem). *Sejam E e F espaços vetoriais de dimensão finita. Para toda transformação linear $A\colon E \to F$ vale que $\dim E = \dim(\ker A) + \dim(\mathcal{I}m(A))$.*

Demonstração.
O teorema segue imediatamente do lema 2.5. ■

Corolário 2.1. *Sejam E e F espaços vetoriais de dimensão finita n. Para que uma transformação linear $A\colon E \to F$ seja injetiva é necessário e suficiente que ela seja sobrejetiva.*

Demonstração.
Suponha que $A\colon E \to F$ é uma transformação linear entre espaços vetoriais de dimensão finita $\dim E = \dim F$. Veja que:

$$\begin{aligned} A \textit{ é injetiva} &\iff \dim(\ker A) = 0 \\ &\iff \dim E = \dim(\mathcal{I}m(A)) \\ &\iff \dim F = \dim(\mathcal{I}m(A)) \\ &\iff A \textit{ é sobrejetiva}, \end{aligned}$$

conforme desejado. ■

Corolário 2.2. *Sejam E e F são espaços vetoriais de dimensão finita n. Se as transformações lineares $A\colon E \to F$, $B\colon F \to E$ são tais que $BA = I_E$ então $AB = I_F$ e $B = A^{-1}$.*

Demonstração.
Pelas hipóteses, B é uma inversa à esquerda de A, portanto A é injetora (teorema 2.4). Como $\dim E = n = \dim F$, segue do corolário 2.1 que A é sobrejetora. Daí, pelo lema 2.3, temos que $AB = I_F$ e portanto $B = A^{-1}$. ■

Definição 2.10. *Seja E um espaço vetorial de dimensão finita. O espaço vetorial $L(E^*;\mathbb{R})$ é chamado de **bi-dual de** E e denota-se este espaço por E^{**}.*

Lema 2.6. *Seja E um espaço vetorial de dimensão finita. Se $v \in E$ é tal que $f(v) = 0$ para todo $f \in E^*$, então $v = 0$.*

Demonstração.
Pela contrapositiva, suponha que $v \in E$ é não-nulo. Então $\{v\}$ é um conjunto L.I. e pode ser estendido a uma base $\{v, u_1, \ldots, u_n\} \subset E$ (teorema 1.6). Daí, o mapa dado por $v \mapsto 1$ e $u_i \mapsto 0$ para $1 \le i \le n$ define um único funcional linear $f \in E^*$ (teorema 2.1) tal que $f(v) = 1$ e $f(u_i) = 0$. ■

Teorema 2.6. *Todo espaço vetorial de dimensão finita E é isomorfo ao seu bi-dual E^{**}.*

Demonstração.
Seja $\varphi\colon E \to E^{**}$ o mapa $v \mapsto \varphi(v) \equiv v^{**}$, onde $v^{**}\colon E^* \to \mathbb{R}$ é definido por

$$v^{**}f := f(v).$$

Claramente, φ é linear, pois para qualquer $f \in E^*$ vale o seguinte:

$$(\lambda v^{**} + u^{**})f = f(\lambda v + u) = \lambda f(v) + f(u) = \lambda v^{**}f + u^{**}f.$$

Além disso, se $v \in \ker \varphi$, então para todo $f \in E^*$ vale que

$$f(v) = 0.$$

Disto podemos concluir que $v = 0$ e portanto φ é injetora. Segue do teorema 2.2 que qualquer espaço vetorial de dimensão finita é isomorfo ao seu dual algébrico, portanto E^{**} tem a mesma dimensão de E. Satisfeitas as hipóteses do corolário 2.1, conclui-se que φ é sobrejetora e consequentemente $E \simeq E^{**}$. ■

O teorema acima mostra que, ao menos em dimensão finita, existe uma identificação natural entre um espaço vetorial E e o seu bi-dual E^{**}. Isso é útil para trabalharmos com versões menos abstratas de produtos tensoriais da forma $V_1^{**} \otimes \ldots \otimes V_k^{**}$ (proposição 6.5).

2.1.3 Representação Matricial

No exemplo 2.3, vê-se que uma transformação linear de $\mathbb{R}^n$ em $\mathbb{R}^m$ pode ser representada por uma matriz $\mathbf{a} \in M(m \times n)$ uma vez fixadas as bases canônicas em cada espaço. Tal exemplo trata-se de uma aplicação do

Teorema 2.7. *Sejam E e F espaços vetoriais de dimensão finita. Fixadas as bases ordenadas $\mathcal{V} = \{v_1, \ldots, v_n\} \subset E$ e $\mathcal{W} = \{w_1, \ldots, w_m\} \subset F$, existe um isomorfismo entre os espaços vetoriais $L(E; F)$ e $M(m \times n)$.*

Demonstração.

Fixe as bases ordenadas $\mathcal{V} = \{v_1, \ldots, v_n\} \subset E$ e $\mathcal{W} = \{w_1, \ldots, w_m\} \subset F$. Para cada $1 \leq j \leq n$, existem m coeficientes a_{ij} tais que

$$Av_j = \sum_{i=1}^{m} a_{ij} w_i.$$

O mapa $\varphi\colon L(E; F) \to M(m \times n)$ definido por $A \mapsto \varphi(A) = \mathbf{a}$, onde $\mathbf{a}$ é a matriz m por n cuja ij-ésima entrada é o coeficiente a_{ij}, é um isomorfismo.

Visto que $(\alpha A + B)v_j = \alpha A v_j + B v_j$ para quaisquer $A, B \in L(E; F)$ e para qualquer $\alpha \in \mathbb{R}$, a linearidade de φ segue de imediato.

Se $\mathbf{a} = [a_{ij}] \in M(m \times n)$, então para cada $1 \leq j \leq n$ pode-se definir o vetor $u_j = \sum_{i=1}^{m} a_{ij} w_i$. Pelo teorema 2.1, o mapa $v_j \mapsto u_j$ pode ser estendido linearmente para uma única transformação linear $A \in L(E; F)$ tal que $Av_j = \sum_{i=1}^{m} a_{ij} w_i$. Logo, $\varphi(A) = \mathbf{a}$. Supondo que $\mathbf{a}$ é qualquer, facilmente conclui-se que φ é sobrejetivo.

Se $A \in \ker \varphi$, então $Av_j = 0$ para $j = 1, \ldots, n$. Como $\{v_1, \ldots, v_n\}$ é base, consequentemente tem-se que $\forall v \in E$ $Av = 0$ e portanto $A = 0$. Sendo $\ker \varphi = \{0\}$, segue do lema 2.1 que φ é injetora.

Sendo uma transformação linear sobrejetora e injetora, conclui-se que φ é um isomorfismo. ■

Por meio do isomorfismo φ definido na demonstração acima, a matriz da transformação linear A nas bases ordenadas $\mathcal{V}$ e $\mathcal{W}$ é aquela cuja j-ésima coluna tem como entradas os coeficientes que determinam o vetor Av_j.

Exemplo 2.13. *A matriz do operador derivação* $D\colon \mathcal{P}_n \to \mathcal{P}_n$ *com relação à base ordenada* $\{1, t, t^2, \dots, t^n\}$ *é:*

$$\mathbf{d} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & n \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Exemplo 2.14. *No espaço* F *gerado por pela base ordenada* $\{\cos x, \sin x\}$*, a matriz do operador derivação com relação a esta base é*

$$\mathbf{d} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Já no espaço G *gerado pela base ordenada*

$$\{e^x \cos x, e^x \sin x, e^{2x} \cos x, e^{2x} \sin x, e^{3x} \cos x, e^{3x} \sin x\},$$

a matriz de $D\colon G \to G$ *com relação a esta base é*

$$\mathbf{d} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & -1 & 3 \end{pmatrix}.$$

Para quaisquer naturais m, n e p podemos definir a operação $\cdot\colon M(m \times n) \times M(n \times p) \to M(m \times p)$, denominada *produto matricial*, definida por $(\mathbf{b}, \mathbf{a}) \mapsto \mathbf{b} \cdot \mathbf{a} \equiv \mathbf{c} \in M(m \times p)$, onde $\mathbf{c}$ é a matriz cuja ij-ésima entrada é dada por

$$c_{ij} = \sum_{k=1}^{n} b_{ik} a_{kj}.$$

Em outras palavras, c_{ij} é obtido a partir do produto entrada a entrada da i-ésima linha de $\mathbf{b}$ pela j-ésima coluna de $\mathbf{a}$.

Teorema 2.8. *Sejam* E, F *e* G *espaços vetoriais de dimensão finita. Fixadas as bases ordenadas* $\mathcal{U} = \{u_1, \dots, u_p\} \subset E, \mathcal{V} = \{v_1, \dots, v_n\} \subset F$ *e* $\mathcal{W} = \{w_1, \dots, w_m\} \subset G$*, se* $A\colon E \to F$ *e* $B\colon F \to G$ *são transformações lineares cujas matrizes nas respectivas bases são* $\mathbf{a}$ *e* $\mathbf{b}$*, então a matriz da transformação linear* $BA\colon E \to G$ *nas bases* $\mathcal{U}$ *e* $\mathcal{W}$ *é o produto* $\mathbf{ba}$*.*

Demonstração.

Por um lado, sabe-se que para cada $1 \leq j \leq p$ $(BA)u_j = \sum_{i=1}^{m} c_{ij} w_i$. Por outro lado, tem-se $(BA)u_j = \sum_{i=1}^{m} \sum_{k=1}^{n} b_{ik} a_{kj} w_i$. Segue da unicidade dos coeficientes (corolário 1.1) que a matriz de BA é dada pela igualdade desejada. ■

Por conta do isomorfismo do teorema 2.7, podemos traduzir alguns resultados da seção 2.1.2 para a linguagem matricial. No que segue, um vetor-coluna da matriz $\mathbf{a} = [a_{ij}] \in M(m \times n)$ é a m-úpla $(a_{1j}, a_{2j}, \ldots, a_{mj})$ para algum $1 \leq j \leq n$.

Lema 2.7. *Para que uma matriz $\mathbf{a} \in M(m \times n)$ possua uma inversa à esquerda é necessário e suficiente que seus vetores-coluna sejam linearmente independentes.*

Demonstração.
Suponha que $\mathbf{b} \in M(n \times m)$ é uma matriz inversa à esquerda de $\mathbf{a}$, i.e., $\mathbf{ba} = \mathbf{I}_n$. Pelo teorema 2.7 e pelo teorema 2.8 é fácil concluir que $\mathbf{b}$ e $\mathbf{a}$ são matrizes na base canônica de transformações lineares $B \in L(\mathbb{R}^m; \mathbb{R}^n)$ e $A \in L(\mathbb{R}^n; \mathbb{R}^m)$ e também que $BA = I_{\mathbb{R}^n}$. Visto que B é uma inversa à esquerda de A se e somente se A for injetora (teorema 2.4), segue que $\{Ae_1, \ldots, Ae_n\} \subset \mathbb{R}^m$ é linearmente independente.

Claramente, se $\{Ae_1, \ldots, Ae_n\} \subset \mathbb{R}^m$ é linearmente independente, então A é injetora e portanto possui uma inversa à esquerda B tal que matricialmente $\mathbf{ba} = \mathbf{I}_n$ (aplicação do teorema 2.8). ■

Lema 2.8. *Para que uma matriz $\mathbf{a} \in M(m \times n)$ possua uma inversa à direita é necessário e suficiente que seus vetores-coluna sejam linearmente gerem $\mathbb{R}^m$.*

Demonstração.
Suponha que $\mathbf{b}$ é uma matriz inversa à direita de $\mathbf{a}$. Seguindo as ideias da demonstração anterior, essas matrizes são representações de transformações lineares $B \in L(\mathbb{R}^m; \mathbb{R}^n)$ e $A \in L(\mathbb{R}^n; \mathbb{R}^m)$ tais que B é uma inversa à direita de A. Conclui-se que A é sobrejetora (teorema 2.3) e portanto todo $w \in \mathbb{R}^m$ é combinação linear de vetores Ae_j, $j = 1, \ldots, n$. Aplicando o raciocínio inverso, conclui-se que $\mathbf{a}$ possui uma matriz inversa à direita. ■

Matricialmente, uma transformação linear $A \colon \mathbb{R}^n \to \mathbb{R}^m$ ser invertível significa que sua matriz na base canônica $\mathbf{a}$ é quadrada($m = n$) e existe $\mathbf{a}^{-1}$ tal que $\mathbf{aa}^{-1} = \mathbf{a}^{-1}\mathbf{a} = \mathbf{I}_n$.

Lema 2.9. *Toda matriz $\mathbf{a} \in M(m \times n)$ que possui inversa à direita $\mathbf{c}$ e inversa à esquerda $\mathbf{b}$ é quadrada, invertível e $\mathbf{b} = \mathbf{c} = \mathbf{a}^{-1}$.*

Demonstração.
Sejam A, B e C as transformações lineares cujas representações matriciais são respectivamente $\mathbf{a}$, $\mathbf{b}$ e $\mathbf{c}$. Como $\mathbf{b}$ e $\mathbf{c}$ são matrizes inversas à esquerda e à direita respectivamente, segue que B e C são respectivamente inversas à esquerda e à direita de A. Daí, sendo A uma transformação de $\mathbb{R}^n$ em $\mathbb{R}^m$, segue do lema 2.3 que $n = m$ e $B = C = A^{-1}$. ■

O último resultado fica como exercício:

Exercício 2.8. *Toda matriz $a \in M(n \times n)$ que admite inversa à esquerda também admite inversa à direita. Daí, $\mathbf{a}$ é invertível e as inversas laterais são iguais a $\mathbf{a}^{-1}$.*

No teorema 2.7, vimos que para cada aplicação linear $A\colon E \to F$ entre espaços vetoriais de dimensão finita podemos associar uma determinada matriz se fixarmos bases $\mathcal{V}$ e $\mathcal{W}$ em E e F. Vejamos agora como a matriz de A se porta quando realizamos mudanças de bases nos espaços E e F.

Sejam E e F espaços vetoriais de dimensão finita. Se $\mathcal{V} = \{v_1, \ldots, v_n\}$ e $\mathcal{V}' = \{v'_1, \ldots, v'_n\}$ são bases de E e se $\mathcal{W} = \{w_1, \ldots, w_m\}$ e $\mathcal{W}' = \{w'_1, \ldots, w'_m\}$ são bases de F, então existem matrizes $\mathbf{p} \in M(n \times n)$ e $\mathbf{q} \in M(m \times m)$ tais que

$$v'_j = \sum_{k=1}^{n} p_{kj} v_k \qquad e \qquad w'_r = \sum_{i=1}^{m} q_{ir} w_i. \tag{2.1}$$

Se $A\colon E \to F$ é uma transformação linear, então nas bases $\mathcal{V}'$ e $\mathcal{W}'$ segue que para cada $v'_j \in \mathcal{V}'$:

$$Av'_j = \sum_{r=1}^{m} a'_{rj} w'_r.$$

Por um lado, 2.1 implica em

$$\begin{aligned} Av'_j &= \sum_{k=1}^{n} p_{kj} A(v_k) \\ &= \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik} p_{kj} w_i, \end{aligned}$$

onde fez-se uso de $Av_k = \sum_{i=1}^{m} a_{ik} w_i$.

Por outro lado, 2.1 também implica em

$$\sum_{r=1}^{m} a'_{rj} w'_r = \sum_{i=1}^{m} \sum_{r=1}^{m} q_{ir} a'_{rj} w_i.$$

Por unicidade dos coeficientes, segue que

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

Pela definição de produto matricial, a igualdade acima equivale a

$$\mathbf{ap} = \mathbf{qa}',$$

onde $\mathbf{a}$ é a matriz de A nas bases $\mathcal{V}$ e $\mathcal{W}$ enquanto que $\mathbf{a}'$ é a matriz de A nas bases $\mathcal{V}'$ e $\mathcal{W}'$. Observe que as matrizes $\mathbf{p}$ e $\mathbf{q}$ são invertíveis, pois são representações nas bases $\mathcal{V}$ e $\mathcal{W}$, respectivamente, de operadores $P\colon E \to E$ e $Q\colon F \to F$ que levam bases em bases: $Pv_j = v'_j$ e $Qw_r = w'_r$. Daí, vale que

$$\mathbf{a}' = \mathbf{q}^{-1}\mathbf{ap} \quad e \quad \mathbf{a} = \mathbf{qa}'\mathbf{p}^{-1}.$$

A matriz $\mathbf{p}$ é denominada *matriz de passagem da base* $\mathcal{V}$ *para a base* $\mathcal{V}'$ e $\mathbf{q}$ é denominada *matriz de passagem da base* $\mathcal{W}$ *para a base* $\mathcal{W}'$.

Exemplo 2.15. *Na base canônica de* $\mathbb{R}^2$ *a matriz do operador linear* A *que reflete um vetor* v *em torno da reta* $y = \alpha x$ *é*

$$\mathbf{a} = \begin{pmatrix} \frac{1-\alpha^2}{1+\alpha^2} & \frac{2\alpha}{1+\alpha^2} \\ \frac{2\alpha}{1+\alpha^2} & \frac{\alpha^2-1}{1+\alpha^2} \end{pmatrix}.$$

Na base ordenada $\{(1, \alpha), (-\alpha, 1)\}$*, temos que*

$$\mathbf{a}' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Definição 2.11. *Seja* $A\colon E \to F$ *uma transformação linear entre espaços vetoriais de dimensão finita. O* ***posto de A*** *é a dimensão de* $\mathcal{I}m(A)$*.*

Pelo teorema do Núcleo e da Imagem, é imediato que o posto de A é igual a $\dim E$ se e somente se A for injetora. Também é imediato que o posto de A é igual a $\dim F$ se e somente se A for sobrejetora.

Se $\mathbf{a} \in M(m \times n)$ é a representação de um operador linear A em bases $\mathcal{U}$ e $\mathcal{V}$, então o número máximo de colunas L.I. desta matriz é igual ao posto de A. Tal número é nomeado *posto segundo colunas de* ***a***. De maneira similar, o *posto segundo linhas de* ***a*** é o número máximo de linhas L.I. da matriz $\mathbf{a}$. No mais, com relação a esses números, vale o

Teorema 2.9. *Para toda matriz* $\mathbf{a} \in M(m \times n)$*, o posto segundo colunas é igual ao posto segundo linhas.*

Para a demonstração do teorema, faremos uso da

Definição 2.12. *Seja* $a \in M(m \times n)$*. A* ***matriz transposta de a*** *é a matriz* $\mathbf{a}^\intercal = [a^\intercal_{ij}] \in M(n \times n)$*, onde* $a^\intercal_{ij} := a_{ji}$*.*

Demonstração do Teorema 2.9.

Seja $\mathbf{a} \in M(m \times n)$ e l o posto segundo linhas de $\mathbf{a}$. Se o posto segundo colunas de $\mathbf{a}$ é p, então existem p vetores $w_k = (b_{1k}, \ldots, b_{mk}) \in \mathbb{R}^m$ tais que para cada vetor-coluna $v_j = (a_{1j}, \ldots, a_{mj}) \in \mathbb{R}^m$, vale que

$$v_j = \sum_{k=1}^{p} c_{kj} w_k.$$

Desta forma, para a i-ésima coordenada de v_j tem-se que

$$a_{ij} = \sum_{k=1}^{p} b_{ik} c_{kj}$$

para $1 \leq i \leq m,\ 1 \leq j \leq n.$.

Por outro lado, a igualdade acima mostra que cada vetor-linha $u_i = (a_{i1}, \ldots, a_{in})$ pode ser escrito como combinação linear dos p vetores $z_k = (c_{k1}, \ldots, c_{kn})$. Pela maximalidade do posto segundo linhas de $\mathbf{a}$, segue que $l \leq p$.

Para mostrar a outra desigualdade, veja que para $\mathbf{a}^\intercal$ vale que $u_i^\intercal = \tilde{v}_i = (a_{i1}, \ldots, a_{in})$, onde $\tilde{v}_i$ é o i-ésimo vetor-coluna de $\mathbf{a}^\intercal \in M(n \times m)$. Por um raciocínio análogo ao do caso anterior, existem l vetores $\tilde{w}_k = (w_{1k}, \ldots, \tilde{w}_{nk}) \in \mathbb{R}^n$ tais que

$$u_i^\intercal = \sum_{k=1}^{l} \tilde{c}_{ki} \tilde{w}_k.$$

Desta forma, para a j-ésima componente de $u_i^\intercal$, vale que

$$a_{ji} = \sum_{k=1}^{l} \tilde{w}_{jk} \tilde{c}_{ki}.$$

Por outro lado, o j-ésimo vetor-linha de $\mathbf{a}^\intercal$ é dado por $\tilde{u}_j = v_j^\intercal = (a_{j1}, \ldots, a_{jm}) \in \mathbb{R}^m$. Em vista disso, a igualdade acima mostra que cada vetor-linha vale que

$$v_j^\intercal = \sum_{k=1}^{l} \tilde{w}_{jk} \tilde{z}_k,$$

onde $\tilde{z}_k = (\tilde{c}_{k1}, \ldots, \tilde{c}_{km})$. Logo, pela maximalidade de s, segue que $p \leq l$.

De $l \leq p$ e $p \leq l$, conclui-se que $p = l$. ■

Implicitamente na demonstração acima, foi utilizado que o posto segundo colunas de $\mathbf{a}$ é igual ao posto segundo linhas de $\mathbf{a}^\intercal$ e o fato análogo para o posto segundo linhas de $\mathbf{a}$. Acredito que isso seja evidente.

Exercício 2.9. *Sejam $\mathbf{a}, \mathbf{b} \in M(m \times n)$ e $\alpha \in \mathbb{R}$ quaisquer. Mostre que:*

1. $(\alpha\mathbf{a} + \mathbf{b})^\intercal = \alpha\mathbf{a}^\intercal + \mathbf{b}^\intercal$.
2. $(\mathbf{a}^\intercal)^\intercal = \mathbf{a}$.
3. $(\mathbf{ab})^\intercal = \mathbf{b}^\intercal\mathbf{a}^\intercal$.
4. *Se $\mathbf{a}$ for invertível, então $(\mathbf{a}^{-1})^\intercal = (\mathbf{a}^\intercal)^{-1}$.*

Definição 2.13. *Seja $a \in M(n \in n)$. O **traço de a** é o número real determinado pelo funcional linear $\operatorname{tr}\colon M(n \times n) \to \mathbb{R}$, $\mathbf{a} \mapsto \operatorname{tr} a \coloneqq a_{11} + \ldots + a_{nn}$.*

Da definição acima, é trivial mostrar que para todas matrizes quadradas $\mathbf{a}, \mathbf{b} \in M(n \times n)$ temos que $\operatorname{tr}(\mathbf{ab}) = \operatorname{tr}(\mathbf{ba})$. Em vista deste resultado, se $A\colon E \to E$ é um operador linear definido num espaço vetorial de dimensão finita, podemos definir o valor real $\operatorname{Tr} A$ como sendo o traço de qualquer representação matricial de A. Claramente, este valor é único para cada $A \in L(E)$, pois se $\mathbf{a}$ e $\mathbf{a}'$ são representações de A em quaisquer bases de E, como $\mathbf{a}' = \mathbf{p}^{-1}\mathbf{ap}$, segue que

$$\operatorname{tr} \mathbf{a}' = \operatorname{tr}\left(\mathbf{p}^{-1}(\mathbf{ap})\right) = \left(\mathbf{a}(\mathbf{pp}^{-1})\right) = \operatorname{tr} \mathbf{a},$$

onde $\mathbf{p}$ é a matriz de passagem de base.

2.2 Eliminação Gaussiana

3 Produto Interno

Um produto interno em um espaço vetorial de dimensão finita é uma aplicação que permite a introdução de algumas noções geométricas em álgebra linear, como o ângulo entre dois vetores, o comprimento de um vetor e a distância entre dois vetores.

Provavelmente, os espaços com produto interno de maior interesse para quem estuda mecânica quântica são os *espaços de Hilbert*. Infelizmente, tais espaços não serão introduzidos nestas notas, mas existe uma literatura infindável acerca destes. Espaços vetoriais com produto interno também são ubíquos geometria Riemanniana, pois o espaço tangente à uma variedade Riemanniana num ponto p é um espaço vetorial munido com produto interno.

Definição 3.1. *Seja E um espaço vetorial de dimensão finita. Um **produto interno em E** é um mapa $\langle \cdot, \cdot \rangle \colon E \times E \to \mathbb{R}$ que satisfaz às seguintes propriedades*

1. *Para todo $u, v, w \in E$ e para todo $\alpha \in \mathbb{R}$:*

$$\begin{aligned} \langle u, v + w \rangle &= \langle u, v \rangle + \langle u, w \rangle, \\ \langle v + w, u \rangle &= \langle v, u \rangle + \langle w, u \rangle, \\ \langle \alpha u, v \rangle &= \langle u, \alpha v \rangle = \alpha \langle u, v \rangle. \end{aligned}$$

2. *Para todos $u, v \in E$:*

$$\langle u, v \rangle = \langle v, u \rangle.$$

3. *Seja $u \in E$. Se $u \neq 0$, então $\langle u, u \rangle > 0$.*

Em suma, um produto interno em E é um mapa bilinear, simétrico e positivo-definido.

Como corolário da definição, vê-se que se $u \in E$ é tal que $\langle u, v \rangle = 0$ para todo $v \in E$, então $u = 0$. Suponha por absurdo que $\langle u, v \rangle = 0$ para todo $v \in E$ e $u \neq 0$; da hipótese segue que $\langle u, u \rangle = 0$ para $u \neq 0$, uma contradição.

Definição 3.2. *Um espaço vetorial munido com produto interno é um par ordenada $(E, \langle \cdot, \cdot \rangle)$, onde E é um espaço vetorial e $\langle \cdot, \cdot \rangle$ é um produto interno em E.*

Para não carregar a notação, diz-se apenas que E é um espaço munido com produto interno, deixando subentendida a presença de $\langle \cdot, \cdot \rangle$.

4 Decomposição Espectral

Nesta seção estão reunidos algumas definições e resultados relativos à adjunta de uma transformação linear, subespaços invariantes e operadores auto-adjuntos.

Esses fatos são apresentados como forma de preparar o leitor para o resultado culminante desta seção: o teorema espectral. Em suma, seu enunciado informa que todo operador auto-adjunto num espaço vetorial (real) de dimensão finita munido de produto interno pode ser representado numa base formada por seus auto-vetores. A compreensão deste teorema pode ser visto como um primeiro passo para um entendimento mais formal de resultados utilizados em cursos de mecânica quântica a nível de graduação.

Vale atentar-se que para traduzir o conteúdo desta seção para espaços vetoriais sobre o corpo $\mathbb{C}$, pequenas modificações em algumas definições e resultados precisam ser feitas.

4.1 Adjunta

No teorema 2.2, vê-se que é possível encontrar um isomorfismo entre um espaço vetorial de dimensão finita E e seu dual E^*. Tal isomorfismo, no entanto, só existe mediante a presença de uma base $\mathcal{U}$ em E. Por conta desta dependência, é comum dizer que este isomorfismo não é intrínseco a E: o funcional associado a um elemento $v \in E$ depende da base estabelecida. Para contornar esta situação e estabelecer um isomorfismo $\varphi\colon E \to E^*$ independente de bases, é preciso que E seja um espaço com produto interno $\langle \cdot, \cdot \rangle$.

Teorema 4.1. *Seja E um espaço vetorial de dimensão finita munido de produto interno $\langle \cdot, \cdot \rangle$. O mapa $\varphi\colon E \to E^*$ que a cada $u \in E$ associa o funcional $\varphi(u) \in E^*$, cuja ação sobre cada $v \in E$ é dada por*

$$\varphi(u)(v) := \langle u, v \rangle,$$

define um isomorfismo entre E e E^*.

Demonstração.

Como o produto interno $\langle \cdot, \cdot \rangle$ é bilinear, a linearidade de φ é imediata.

Se $v \in \ker \varphi$, então para todo $u \in E$ vale que

$$\langle v, u \rangle = 0.$$

Escolhendo $u = v$, conclui-se que $v = 0$ graças a $\langle \cdot, \cdot \rangle$ ser positivo-definido. Daí, segue que $\ker \varphi = \{0\}$ e portanto φ é injetora (lema 2.1).

Sendo uma aplicação linear injetora entre espaços vetoriais de dimensão finita, φ é um isomorfismo (corolário 2.1). ■

Tal resultado é conhecido como *teorema da representação de Riesz*. No caso, esta é sua forma quando nos restringimos a espaços vetoriais de dimensão finita. Sua demonstração para o caso mais abrangente pode ser encontrada em qualquer bom livro de análise funcional.

Veja que por conta do teorema 4.1, dado um funcional $f \in E^*$, existe um único $u \in E$ tal que $f(v) = \langle u, v \rangle$ para todo $v \in E$. Esta observação é essencial para estabelecermos o que é a adjunta de uma transformação linear.

Definição 4.1. *Sejam* $(E; \langle \cdot, \cdot \rangle_E)$ *e* $(F; \langle \cdot, \cdot \rangle_F)$ *espaços vetoriais de dimensão finita munidos de produto interno. Se* $A\colon E \to F$ *é uma transformação linear, a* ***adjunta de A*** *é a transformação linear* $A^*\colon F \to E$ *que satisfaz à seguinte propriedade:*

$$\langle Av, w \rangle_F = \langle v, A^* w \rangle_E$$

para todo $v \in E$ *e para todo* $w \in F$.

Proposição 4.1. *Seja* $A\colon E \to F$ *uma transformação linear entre espaços vetoriais de dimensão finita munidos de produto interno. A adjunta de* A *existe e é única.*

Demonstração.
Existência: seja $\xi\colon F \to E^*$ o mapa definido por $\xi(w) \coloneqq \langle A\cdot, w \rangle$. Pelo teorema da representação de Riesz (teorema 4.1), existe um único $u \in E$ tal que $\xi(w)(v) = \langle u, v \rangle$ para todo $v \in E$. Pela definição de $\xi(w)$, a igualdade anterior significa que existe um único $u \in E$ tal que

$$\langle Av, w \rangle_F = \langle v, u \rangle_E$$

para todo $v \in E$. Desta forma, basta definir $A^*\colon F \to E$, como sendo $A^* w \coloneqq u$. Veja que para quaisquer $w, w' \in F$, $v \in E$ e qualquer $\alpha \in \mathbb{R}$:

$$\begin{aligned}
\langle v, A^*(\alpha w + w') \rangle_E &= \langle Av, \alpha w + w' \rangle_F \\
&= \alpha \langle Av, w \rangle_F + \langle Av, w' \rangle_F \\
&= \langle v, \alpha A^* w + A^* w' \rangle_E.
\end{aligned}$$

Consequentemente

$$\langle v, A^*(\alpha w + w') - \alpha A^* w - A^* w' \rangle_E = 0$$

para todo $v \in E$. Como o produto interno é positivo-definido, a segunda entrada na igualdade acima é o vetor nulo e portanto vale que $A^*(\alpha w + w') = \alpha A^* w + A^* w'$. Logo, A^* é linear.

Unicidade: suponha que exista outra transformação $B^*\colon F \to E$ que satisfaz

$$\langle Av, B^* w \rangle_E = \langle Av, w \rangle_F$$

para quaisquer $v \in E$, $w \in F$. É trivial mostrar que

$$\langle v, (A^* - B^*) w \rangle_E = 0$$

para todo $v \in E$. Daí, como $\langle \cdot, \cdot \rangle_E$ é positivo-definido, conclui-se facilmente que $A^* w = B^* w$. Sendo w qualquer em F, segue que $A^* = B^*$.

∎

Em espaços vetoriais reais de dimensão finita, vale o seguinte resultado com relação à representação matricial da adjunta:

Proposição 4.2. *Seja $A\colon E \to F$ uma transformação linear entre espaços vetoriais de dimensão finita munidos de produto interno. Fixadas as bases $\mathcal{U} = \{u_1, \ldots, u_n\} \subset E$ e $\mathcal{V} = \{v_1, \ldots, v_m\} \subset F$, a representação matricial de $A^*\colon F \to E$ nas bases $\mathcal{V}$ e $\mathcal{U}$ é $\mathbf{a}^\intercal \in M(n \times m)$, onde $\mathbf{a}^\intercal \in M(m \times n)$ é a transposta da representação matricial de A nas bases $\mathcal{U}$ e $\mathcal{V}$.*

Demonstração.
Seja $\mathbf{a}^* \in M(n \times m)$ a matriz de A^* nas bases $\mathcal{V}$ e $\mathcal{U}$. Usando a propriedade que define a adjunta, veja que para quaisquer $1 \leq i \leq n$ e $1 \leq j \leq m$

$$a^*_{ij} = \langle u_i, A^* v_j \rangle = \langle A u_i, v_j \rangle = a_{ji}.$$

Logo, $\mathbf{a}^* = \mathbf{a}^\intercal$. ∎

Corolário 4.1. *O posto de A^* é igual ao posto de A^**

Demonstração.
Evidente. ∎

Por conta da proposição 4.2, todas as propriedades do exercício 2.9 são traduzidas para a adjunta da seguinte forma:

1. $(\alpha A + B)^* = \alpha A^* + B^*$.
2. $(A^*)^* = A^*$.
3. $(AB)^* = B^* A^*$.
4. Se A for invertível, então $(A^{-1})^* = (A^*)^{-1}$.

Definição 4.2. *Seja E um espaço vetorial munido de produto interno. Se $X \subset E$ é um subconjunto qualquer, o* ***complemento ortogonal de*** $\mathbf{X}$ *é o subconjunto de E definido por*

$$X^\perp := \{v \in E\colon \langle v, x \rangle = 0 \text{ para todo } x \in X\}.$$

É trivial mostrar que $X^\perp$ é um subespaço vetorial quando X for não-vazio.

Lema 4.1. *Seja E um espaço vetorial munido de produto interno. Se X e Y são subconjuntos de E tais que $X \subset Y$, então $Y^\perp \subset X^\perp$. No mais, vale que $X \cap X^\perp = \{0\}$.*

Demonstração.
Suponha que $X \subset Y$. Se $v \in Y^\perp$, então $\langle v, y \rangle = 0$ para todo $y \in Y$. Em particular, isto vale para todo $x \in X$, já que $x \in X \implies x \in Y$. Logo, $Y^\perp \subset X^\perp$.
Por absurdo, suponha que $v \in X \cap X^\perp$ e $v \neq 0$. Segue das hipóteses que $\langle v, v \rangle = 0$, uma contradição. ∎

Um outro resultado trivial é mostrar que se X for não-vazio, então $X^\perp = S(X)^\perp$, onde $S(X)^\perp$ é o complemento ortogonal do subespaço gerado por X.

Proposição 4.3. *Seja E um espaço vetorial de dimensão finita munido de produto interno. Se F é um subespaço vetorial de E, então vale a decomposição $E = F \oplus F^{\perp}$.*

Demonstração.
Seja $\mathcal{U} = \{u_1, \ldots, u_r\}$ uma base (ordenada) de F. Estendendo-a a uma base (ordenada) de E $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_n\}$, por meio do processo de Gram-Schmidt obtém-se uma base (ordenada) ortogonal de E $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ cujos primeiros r vetores são uma base de F e os demais são ortogonais ao subespaço F. Daí, para todo $v = \alpha_1 v_1 + \ldots + \alpha_n v_n \in E$, temos que

$$v = u + w$$

onde $u = \alpha_1 v_1 + \ldots + \alpha_r v_r \in F$ e $w = \alpha_{r+1} v_{r+1} + \ldots + \alpha_n v_n \in F^{\perp}$. Sendo v qualquer, segue que $E = F \oplus F^{\perp}$(teorema 1.1). ■

Corolário 4.2. $\dim E = \dim F + \dim F^{\perp}$.

Demonstração.
Pela proposição 4.3, para cada $v \in E$ que existem $u \in F$ e $w \in F^{\perp}$ tais que $v = u+w$. Seja $P_F \colon E \to E$ a projeção sobre F, isto é, para todo $u+w = u \in E$, $P_F(v) = u$. Como $\ker P_F = F^{\perp}$ e $\mathcal{I}m(P_F) = F$, segue do teorema do núcleo e da imagem que $\dim E = \dim F + \dim F^{\perp}$. ■

Corolário 4.3. *Para todo subespaço vetorial $F \subset E$, $(F^{\perp})^{\perp} = F$.*

Demonstração.
Claramente, para todo subconjunto $X \subset E$, $X \subset (X^{\perp})^{\perp}$. Em particular, temos que $F \subset (F^{\perp})^{\perp}$. Pelo corolário anterior é fácil ver que $\dim F = \dim (F^{\perp})^{\perp}$. Pelo item d do teorema 1.6, segue que $F = (F^{\perp})^{\perp}$. ■

Exemplo 4.1. *Considere o seguinte produto interno em $M(n \times n)$:*

$$\langle \mathbf{a}, \mathbf{b} \rangle \coloneqq \sum_{i,j} a_{ij} b_{ij}.$$

Lembrando que $M(n \times n) = S \oplus A$, onde S é o subespaço das matrizes quadradas simétricas e A é o subespaço das matrizes quadradas antissimétricas, pelo proposição 4.3 fica claro que A é o complemento ortogonal de S e vice-versa.

Exemplo 4.2. *Em $C^0([-1,1])$, sejam F e G os subespaços vetoriais das funções pares e ímpares, respectivamente. Com relação ao produto interno*

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) \mathrm{d}x,$$

é fácil verificar que $G = F^{\perp}$ e $F = G^{\perp}$.

A partir de agora, quando estiver claro p subespaço $F \subset E$, a projeção ortogonal sobre F será denotada por $P \colon E \to E$ no lugar de P_F.

Teorema 4.2. *Seja $A\colon E \to F$ uma transformação linear entre espaços de dimensão finita munidos de produto interno. Nestas condições, vale que*

1. $\ker A^* = \mathcal{I}m(A)^\perp$.
2. $\mathcal{I}m(A^*) = (\ker A)^\perp$.
3. $\ker A = \mathcal{I}m(A^*)^\perp$.
4. $\mathcal{I}m(A) = (\ker A^*)^\perp$.

Demonstração.
Basta demonstrar a primeira igualdade. As demais são consequências da primeira.
Veja que

$$\begin{aligned} w \in \ker A^* &\iff \langle v, A^* w\rangle = 0 \ \forall v \in E \\ &\iff \langle Av, w\rangle = 0 \ \forall v \in E \\ &\iff \langle u, w\rangle \ \forall u \in \mathcal{I}m A \\ &\iff w \in \mathcal{I}m(A)^\perp. \end{aligned}$$

■

Corolário 4.4. *A fim de que o sistema de m equações lineares com n incógnitas*

$$\sum_{j=1}^{m} a_{ij} x_j = b_i \ (i = 1, \ldots, m)$$

possua solução, é necessário e suficiente que o vetor $b = (b_1, \ldots, b_m)$ seja perpendicular às soluções $y = (y_1, \ldots, y_n)$ do sistema transposto homogêneo

$$\sum_{j=1}^{m} a_{ji} y_j = 0 \ (i = 1, \ldots, n).$$

Demonstração.
Usando o item 4 do teorema 4.2, a demonstração é imediata. ■

4.2 Subespaços Invariantes

Tal qual a adjunta, a noção de subespaços invariantes por operadores lineares é essencial para a compreensão do teorema espectral.

Definição 4.3. *Seja $A\colon E \to E$ um operador linear definido num espaço vetorial de dimensão finita. Um subespaço vetorial F de E é dito* ***invariante por*** **A** *se $A(F) \subset F$.*

Exemplos triviais de subespaços invariantes por quaisquer transformações lineares são $\{0\}$, E, $\ker A$ e $\mathcal{I}m(A)$.

Exemplo 4.3. *Se $F \subset E$ for um subespaço de dimensão 1, então ele é invariante por uma transformação linear $A\colon E \to E$ se para todo vetor não-nulo $u \in F$ existe $\lambda \in \mathbb{R}$ tal que $Au = \lambda u$.*

No caso em que F é um subespaço de dimensão 2 e invariante por A, se $\{u, v\} \subset F$ for linearmente independente, então existem constantes $\lambda, \gamma, \mu, \nu \in \mathbb{R}$ tais que $Au = \lambda u + \gamma v$ e $Av = \mu u + \nu v$.

Exemplo 4.4. *Os subespaços F e G gerados por $\{\cos x, \sin x\}$ e $\{e^x, xe^x, x^2e^x\}$ são invariantes pelo operador derivação $D\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$.*

Definição 4.4. *Seja $A\colon E \to E$ um operador linear definido num espaço vetorial E. Um **autovetor do operador A** é um vetor não-nulo $v \in E$ para o qual existe $\lambda \in \mathbb{R}$ tal que*

$$Av = \lambda v.$$

*O valor λ é chamado de **autovalor do operador A** e diz-se que v é um autovetor correspondente ao autovalor λ.*

Definição 4.5. *Sejam $A\colon E \to E$ um operador linear definido num espaço vetorial E e λ um autovalor de A. O **auto-subespaço** correspondente a λ é o conjunto*

$$E_\lambda := \{v \in E\colon Av = \lambda v\}.$$

É trivial mostrar que E_λ é um subespaço invariante.

Mostraremos a seguir que todo operador linear definido num espaço vetorial de dimensão finita possui um subespaço invariante.

Definição 4.6. *Seja $p(x) \in \mathcal{P}_n$. Diz-se que $p(x)$ é mônico se o coeficiente do termo de mais alto grau for igual a 1. No mais, diz-se que $p(x)$ é irredutível (sobre $\mathbb{R}$) se não pode ser fatorado como o produto de dois polinômios não-constantes com coeficientes reais e graus menores que $Gr(p(x))$.*

Teorema 4.3 (Fundamental da álgebra). *Todo polinômio real $p(x)$ de grau n admite uma fatoração $p(x) = p_1(x)p_2(x)\ldots p_k(x)$, onde cada $p_i(x)$ é um polinômio mônico irredutível de grau 1 ou 2.*

Dados um polinômio $p(x) = a_0 + a_1x + \ldots a_nx^n$ e um operador $A\colon E \to E$, $p(A)$ é a notação utilizada para o operador

$$p(A) := a_0I_n + a_1A + \ldots + a_nA^n.$$

Lema 4.2. *Para todo operador linear $A\colon E \to E$ num espaço vetorial de dimensão finita, existem um polinômio mônico irredutível p, de grau 1 ou 2, e um vetor não nulo $v \in E$, tais que $p(A) \cdot v = 0$.*

Demonstração.

Se E é n-dimensional, então $\dim L(E) = n^2$. Daí, $\{I_n, A, \ldots, A^{n^2}\}$ é linearmente dependente, ou seja, existem coeficientes reais nem todos nulos, $\alpha_0, \alpha_1, \ldots, \alpha_{n^2}$, tais que

$$\alpha_0I_n + \alpha_1A + \ldots + \alpha_{n^2}A^{n^2} = 0;$$

Se α_m é o coeficiente não-nulo do maior índice, então ao dividir a expressão acima por este coeficiente, podemos obter um polinômio $p(x) = \beta_0 + \beta_1 x + \ldots + x^m$ tal que $p(A) = 0$. Devido ao teorema 4.3, este polinômio admite uma fatoração em polinômios reais mônicos irredutíveis de grau 1 ou 2:

$$p(x) = p_1(x)p_2(x)\ldots p_k(x).$$

Daí, $p(A) = p_1(A)p_2(A)\ldots p_k(A) = 0$. Como 0 não é um operador invertível, segue que existe $1 \leq i \leq k$ tal que $p_i(A)$ não é invertível. Sendo $p_i(A)$ não invertível, seu núcleo é não-trivial, ou seja, existe $v \in E$ não-nulo tal que $p_i(A) \cdot v = 0$. ■

Teorema 4.4. *Seja $A \colon E \to E$ um operador linear definido num espaço vetorial de dimensão finita. Existe um subespaço invariante por A de dimensão 1 ou 2.*

Demonstração.
Seja $A \colon E \to E$ linear num espaço vetorial de dimensão finita. Pelo lema 4.2, existe um polinômio (real) mônico irredutível de grau 1 ou 2 $p(x)$ tal e um vetor não-nulo $v \in E$ tais que $p(A) \cdot v = 0$.

Se $p(x)$ tem grau 1, então $p(A) = a_0 I_n + a_1 A$. Daí:

$$p(A) \cdot v = a_0 v + Av.$$

Como $p(A)$ é zero, então $Av = -a_0 v$. Daí, v é um autovetor de A correspondente ao autovalor $\lambda = -a_0$ e portanto $E_\lambda \neq \emptyset$.

Se $p(x) = x^2 + a_1 x + a_0$, então $A^2 v + a_1 Av + a_0 v = 0$. Daí, $A(Av) = -a_1 Av - a_0 v$ e, portanto, o subespaço gerado por $\{v, Av\}$ é invariante por A. No mais, $\{v, Av\}$ é L.I, pois do contrário λ seria uma raiz real de $p(x)$, o que é uma contradição com o fato de $p(x)$ ser irredutível de grau 2. Disto, conclui-se que $S(\{v, Av\})$ é um subespaço invariante de dimensão 2. ■

Teorema 4.5. *Seja $A \colon E \to E$ um operador linear definido num espaço vetorial de dimensão finita. Os autovetores correspondentes a autovalores distintos de A são linearmente independentes.*

Demonstração.
Dado um operador $A \colon E \to E$, sejam $v_1, \ldots, v_m \in E$ autovetores correspondentes aos autovalores distintos dois a dois $\lambda_1, \ldots, \lambda_m$, respectivamente. A demonstração será feita por indução em m. Para $m = 1$, $\{v_1\}$ é L.I. por definição. Suponhamos agora que a afirmação é verdadeira para os primeiros $m-1$ autovalores. Sejam $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ coeficientes tais que

$$\sum_{i=1}^{m} \alpha_i v_i = 0. \tag{4.1}$$

Aplicando A em ambos os lados da igualdade acima e usando que $Av_i = \lambda v_i$, obtém-se que

$$\sum_{i=1}^{m} \lambda_i \alpha_i v_i = 0. \tag{4.2}$$

Multiplicando 4.1 por λ_m e subtraindo isto de 4.2, segue que

$$\sum_{i=1}^{m-1}(\lambda_i - \lambda_m)\alpha_i v_i = 0.$$

Segue da hipótese de indução que $(\lambda_i - \lambda_m)\alpha_i = 0$ para $1 \leq i \leq m-1$. No mais, como $i \neq m \implies \lambda_i \neq \lambda_m$, conclui-se que $\alpha_i = 0$ para $1 \leq i \leq m-1$. Consequentemente, eq. (4.1) resume-se somente a $\alpha_m v_m = 0$. Visto que v_m é não-nulo, segue que $\alpha_m = 0$. Como $\alpha_1, \dots, \alpha_m$ são todos nulos, conclui-se que $\{v_1, \dots, v_m\}$ é L.I.. ■

Corolário 4.5. *Seja* $\dim E = n$. *Se um operador linear* $A\colon E \to E$ *admite* n *autovalores distintos, então existe uma base em* E *formada por autovetores correspondentes a cada autovalor.*

Demonstração.
Evidente. ■

Uma observação importante: suponha que v é um autovetor de um operador $A\colon E \to E$. Segue que $Av = \lambda v$ pode ser reinterpretado como a igualdade $(A - \lambda I_n)v = 0$. Daí, v pertence ao núcleo do operador $A - \lambda I_n$. Como $v \neq 0$, segue que este operador não pode ser invertível (já que seu núcleo não é trivial). Também vale a recíproca, se $A - \lambda I_n$ não é invertível, então $v \in \ker(A - \lambda I_n)$ é um autovetor de A.

Exemplo 4.5. *Em dimensão 2, dados* $A\colon E \to E$ *um operador linear e uma base* $\{u, v\} \subset E$, *sendo* $Au = au + cv$ *e* $Av = bu + dv$, *a matriz de* A *nesta base é*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Daí, $(A - \lambda I_2)u = (a - \lambda)u + cv$ *e* $(A - \lambda I_2)v = (b - \lambda)u + dv$. *Para que* $A - \lambda I_2$ *não seja invertível, é necessário e suficiente que* $\{(A - \lambda I_2)u, (A - \lambda I_2)v\}$ *seja L.D.. Isto é equivalente a* $(a - \lambda)(d - \lambda) - bc = 0$, *ou seja,* λ *é raiz do polinômio*

$$p(\lambda) = \lambda^2 - (a + d)\lambda + ad - bc,$$

denominado polinômio característico do operador A. *Apesar de parecer o caso, o polinômio característico independe das bases escolhidas; a presença do traço de* A, *que é* $a + d$, *num dos coeficientes é um indício disto.*

Exemplo 4.6. *O polinômio característico do operador* $R(\theta)\colon \mathbb{R}^2 \to \mathbb{R}^2$ *é*

$$p(\lambda) = \lambda^2 - (2\cos\theta)\lambda + 1.$$

É fácil perceber que R *só possui auto-valores reais se e somente se* $\theta = 0$ *ou* $\theta = \pi$.

4.3 Teorema Espectral

Definição 4.7. *Seja E um espaço vetorial com produto interno. Um operador linear $A \colon E \to E$ é dito* ***auto-adjunto*** *se $A = A^*$. Em outras palavras, A é auto-adjunto se vale que $\langle Av, w\rangle = \langle v, Aw\rangle$ para todos $v, w \in E$.*

Claramente, se $A, B \colon E \to E$ são auto-adjuntos e $\alpha \in \mathbb{R}$ é qualquer, então αA e $A + B$ são auto-adjuntos. Com relação ao produto AB vale o seguinte lema:

Lema 4.3. *Sejam $A, B \colon E \to E$ operadores lineares auto-adjuntos definidos num espaço vetorial com produto interno. O produto AB é auto-adjunto se e somente se $AB = BA$.*

Demonstração.
Como A e B são auto-adjuntos, então para todos $v, w \in E$ temos que $\langle ABv, w\rangle = \langle v, BAw\rangle$. Daí, AB é auto-adjunto se e somente se $AB = BA$. ∎

Proposição 4.4. *Seja $A \colon E \to E$ um operador auto-adjunto definido num espaço vetorial de dimensão finita munido com produto interno. Se $\{e_1, \dots, e_n\} \subset E$ é uma base ortonormal qualquer, então a matriz de A nesta base é simétrica.*

Demonstração.
Como A é auto-adjunto, em particular vale que $\langle Ae_i, e_j\rangle = \langle e_i, Ae_j\rangle$ para $1 \leq i, j \leq n$. Por ortonormalidade, segue que $\langle Ae_i, e_j\rangle = e_{ji}$ e $\langle e_i, Ae_j\rangle = a_{ij}$. Logo, $a_{ji} = a_{ij}$ e portanto a matriz de A na base $\{e_1, \dots, e_n\}$ é simétrica. ∎

Exemplo 4.7. *Se $F \subset E$ é um subespaço vetorial, então a projeção ortogonal sobre F é um operador auto-adjunto. Também vale a recíproca, se $E = F_1 \oplus F_2$ e $A \colon E \to E$ é tanto auto-adjunto como uma projeção sobre F_1, então $F_2 = F_1^\perp$.*

Teorema 4.6. *Seja $A \colon E \to E$ um operador linear. Se $F \subset E$ é um subespaço invariante por A, então $F^\perp$ é um subespaço invariante por $A^* \colon E \to E$.*

Demonstração.
Seja $F \subset E$ um subespaço vetorial invariante pelo operador $A \colon E \to E$. Pela decomposição $E = F \oplus F^\perp$, para qualquer $w \in F^\perp$ temos que $A^*w = u + v$, onde $u \in F$ e $v \in F^\perp$. Daí, dado $z \in F$ qualquer, por um lado temos que

$$\langle A^*w, z\rangle = \langle u, z\rangle.$$

Por outro lado, como F é invariante por A:

$$\langle A^*w, z\rangle = \langle w, Az\rangle = 0.$$

Logo, $\langle u, z\rangle = 0$. Como $z \in F$ é qualquer, conclui-se facilmente que $u = 0$ e portanto $A^*w = v \in F^\perp$. Sendo w qualquer, segue que $F^\perp$ é invariante por A^*. ∎

Corolário 4.6. *Se $F \subset E$ é um subespaço invariante por um operador auto-adjunto $A \colon E \to E$, então $F^\perp$ também é invariante por A.*

Demonstração.
Aplicação direta do teorema 4.6. ■

No contexto de operadores auto-adjuntos, o teorema 4.5 assume a seguinte forma:

Teorema 4.7. *Se $\lambda_1, \dots \lambda_m$ são autovalores dois a dois distintos de um operador auto-adjunto $A\colon E \to E$, então seus respectivos autovetores são dois a dois ortogonais.*

Demonstração.
Se $i \neq j$, então:

$$\begin{aligned}(\lambda_i - \lambda_j)\langle v_i, v_j\rangle &= \langle \lambda_i v_i, v_j\rangle - \langle v_i, \lambda_j v_j\rangle \\ &= \langle A v_i, v_j\rangle - \langle v_i, A v_j\rangle \\ &= \langle A(v_i - v_i), v_j\rangle \\ &= 0.\end{aligned}$$

Como $\lambda_i - \lambda_j \neq 0$, então $\langle v_i, v_j\rangle = 0$, donde conclui-se que v_i é ortogonal a v_j. ■

Lema 4.4. *Seja E um espaço vetorial de dimensão finita munido de produto interno. Todo operador auto-adjunto $A\colon E \to E$ possui um autovetor.*

O lema é corolário do seguinte resultado:

Proposição 4.5. *Seja $A\colon E \to E$ um operador auto-adjunto definido um espaço vetorial de dimensão 2, munido de produto interno. Existe uma base ortonormal $\{u_1, u_2\} \subset E$ formada por autovetores de A.*

Demonstração da proposição.
Seja $\{v, w\} \subset E$ uma base ortonormal qualquer. Pela proposição 4.2, necessariamente temos que $Av = av + bw$ e $Aw = bv + cw$. Os autovalores de A são as raízes reais do polinômio característico $p(\lambda) = \lambda^2 - (a+c)\lambda + ac - b^2$. Veja que o discriminante $\delta = (a-c)^2 + 4b^2$ é maior ou igual a zero. No primeiro caso, $\delta = 0$, temos que $a = c$, $b = 0$ e $A = aI_n$. Desta forma, todo vetor não-nulo é um autovetor de A. Se $\delta > 0$, então A possui dois autovalores distintos λ_1, λ_2, portanto existem vetores unitários $u_1, u_2 \in E$ que são ortogonais. Logo, $\{u_1, u_2\}$ é uma base de E. ■

Demonstração do lema.
Pelo teorema 4.4, existe um subespaço $F \subset E$ de dimensão 1 ou 2 que é invariante por $A\colon E \to E$. Se $\dim F = 1$, todo vetor $v \in F$ é autovetor de A. Se $\dim F = 2$, então basta aplicar a proposição 4.5 na restrição $A|_F$. ■

Teorema 4.8 (Espectral). *Seja E um espaço vetorial de dimensão finita munido de produto interno. Para todo operador auto-adjunto $A\colon E \to E$ existe uma base ortonormal $\{u_1, \dots, u_n\} \subset E$ formada por autovetores de A.*

Demonstração.

A demonstração será feita por indução em n. Se $\dim E = 1$, então todo vetor não-nulo $v \in E$ é autovetor de A. Supondo agora que o teorema é válido para o caso $n-1$, considere que $\dim E = n$. Pelo lema 4.4, A possui um autovetor $u_n \in E$ e portanto o subespaço F gerado por u_n é invariante por A. Sendo A auto-adjunto, da invariância de F por A segue que $F^\perp$ também é invariante por A. Visto que $\dim F^\perp = n-1$, pela hipótese de indução existe uma base ortonormal $\{u_1, \ldots, u_{n-1}\} \subset F^\perp$ formada por autovetores de A. Daí, $\{u_1, \ldots, u_n\}$ é uma base ortonormal de E formada por autovetores de A. ■

5 Determinante

6 Produto Tensorial de Espaços Vetoriais

6.1 Produto tensorial

Nesta seção, recomenda-se que o(a) leitor(a) esteja acostumado(a) com a notação de Einstein. Além disso, caso queira entender a demonstração da propriedade universal do produto tensorial, é necessário saber o *primeiro teorema de isomorfismos*[9]. Por último, esta seção apresenta conteúdos que não fazem parte da ementa padrão de um curso de álgebra linear e, em vista disso, esta leitura pode ser postergada para um momento mais apropriado.

A noção de produto tensorial apresentada anteriormente pode ser generalizada para espaços vetoriais. Essa generalização pode ser um tanto confusa ou até mesmo abstrata demais, então vale a pena esclarecer de antemão o que será feito a seguir. Grosso modo, a partir de um conjunto de espaços vetoriais $V_1, \dots, V_k$, queremos construir um novo espaço vetorial, simbolizado por $V_1 \otimes \dots \otimes V_k$, cuja dimensão é o produto das dimensões dos espaços vetoriais utilizados. Um elemento desse novo espaço é um determinado tipo de combinação linear de objetos $v_1 \otimes \dots \otimes v_k$, onde $v_j \in V_j$, que por sua vez dependem linearmente dos v_j por construção. Sugiro a leitura de [3] para um melhor entendimento das motivações que levam à definição abstrata do produto tensorial.

O primeiro passo a ser feito é esclarecer qual é o "determinado tipo de combinação linear" mencionado acima. Trata-se de uma **combinação linear formal**. Dado um conjunto S, diz-se que uma expressão da forma $\sum_{i=1}^{m} a_i x_i$, onde $a_i \in \mathbb{R}$ e $x_i \in S$, é uma combinação linear formal de elementos de S. A priori, S não possui nenhuma estrutura algébrica; o ponto aqui é ter uma expressão caracterizada inteiramente por certos elementos de S juntamente dos coeficientes ligados a estes elementos. Se quisermos, podemos ser um pouco mais rigorosos e definir o seguinte

Definição 6.1. *Seja S um conjunto. Uma **combinação linear formal de elementos de S** é uma função $f\colon S \to \mathbb{R}$ tal que $\{x \in S\colon f(x) \neq 0\}$ é um conjunto finito.*

Definição 6.2. *Seja S um conjunto. O **espaço vetorial livre sobre S**, $\mathcal{F}(S)$, é por definição o conjunto de todas as combinações lineares formais de elementos de S.*

Introduzindo a soma pontual usual e a multiplicação por escalar usual neste conjunto, facilmente verifica-se que $\mathcal{F}(S)$ torna-se um espaço vetorial.

De certa maneira, podemos entender S como um subconjunto de $\mathcal{F}(S)$, a partir do mapa $S \ni x \mapsto \delta_x \in \mathcal{F}(S)$, onde δ_x é a função característica do conjunto $\{x\}$. Como esse mapa é injetivo, abusando um pouco a notação, podemos entender δ_x como o próprio x e portanto pensar em S como um subconjunto do espaço (vetorial) livre. Neste contexto, podemos ver que S é uma base para

[9] Recomendo a leitura do capítulo 2 das notas de aula do professor João Barata, disponíveis em [2].

$\mathcal{F}(S)$, pois como cada elemento deste espaço é uma função que não se anula apenas para um conjunto finito de pontos, podemos escrever f como

$$f = \sum_{i=1}^{m} a_i x_i,$$

onde $a_i = f(x_i)$ e x_i é o ponto onde f não se anula[10].

Proposição 6.1. *Sejam S um conjunto e W um espaço vetorial. Qualquer mapa $A\colon S \to W$ admite uma única extensão linear $A'\colon \mathcal{F}(S) \to W$.*

Demonstração.

Primeiramente, vamos mostrar que $A\colon S \to W$ admite uma extensão linear. Para tanto, considere $f \in \mathcal{F}(S)$ e sejam $x_1, \ldots, x_m \in S$ os pontos onde f não se anula. Afirmo que o mapa $A'\colon \mathcal{F}(S) \to W$ definido por

$$A'(f) := \sum_{i=1}^{m} a_i A(x_i)$$

é uma extensão linear de A. Com efeito, sejam $g \in \mathcal{F}(S)$ e $\alpha \in \mathbb{R}$ quaisquer. Sendo $\mathcal{F}(S)$ um espaço vetorial, então a função $\alpha f + g$ está bem definida e

$$A'(\alpha f + g) := \sum_{k=1}^{p} c_k A(z_k)$$

onde $c_k = \alpha f(z_k) + g(z_k)$. Logo, segue que

$$A'(\alpha f + g) = \alpha \sum_{k=1}^{p} f(z_k) A(z_k) + \sum_{k=1}^{p} g(z_k) A(z_k).$$

Veja que se $z_k \in S$ é tal que $f(z_k) = 0$ ou $g(z_k) = 0$, então os somatórios acima podem ser reescritos como $\alpha A'(f)$ e $A'(g)$. Portanto, A' de fato é uma extensão linear.

Para mostrar unicidade, assuma que $B'\colon \mathcal{F}(S) \to W$ seja uma outra extensão linear do mapa $A\colon S \to W$. Dada $f \in \mathcal{F}(S)$ qualquer, por linearidade de B', segue que

$$B'(f) = \sum_{i=1}^{m} a_i B'(x_i) = \sum_{i=1}^{m} a_i A(x_i) = A'(f).$$

Sendo a segunda igualdade obtida observando que $B'|_S = A$. Sendo f qualquer, conclui-se que $B' = A'$. ■

[10] Está subentendido que x_i está representando a função δ_{x_i}. Ademais, a soma na verdade é indexada em S, mas como f se anula nos demais elementos, podemos escrever a soma da forma como foi feita.

Sejam $V_1, \ldots, V_k$ espaços vetoriais. Considere o espaço vetorial $\mathcal{F}(V_1 \times \ldots \times V_k)$ e seja $\mathcal{R}$ o subespaço vetorial formado pelo span de elementos da forma

$$(v_1, \ldots, av_i, \ldots, v_k) - a(v_1, \ldots, v_i, \ldots, v_k),$$

$$(v_1, \ldots, v_i + v_i', \ldots, v_k) - (v_1, \ldots, v_i, \ldots, v_k) - (v_1, \ldots, v_i', \ldots, v_k).$$

Sendo $v_j, v_j') \in V_j$, $i \in \{1, \ldots, k\}$ e $a \in \mathcal{R}$.

O **produto tensorial dos espaços** $V_1, \ldots, V_k$, simbolizado por $V_1, \otimes \ldots \otimes V_k$ é o conjunto

$$V_1 \otimes \ldots \otimes V_k := \mathcal{F}(V_1 \times \ldots \times V_k)/\mathcal{R}.$$

Seja $\Pi\colon \mathcal{F}(V_1 \times \ldots \times V_k) \to V_1 \otimes \ldots \otimes V_k$ a projeção natural, então

$$\Pi(v_1, \ldots, v_k) = v_1 \otimes \ldots \otimes v_k \tag{6.1}$$

para todo $(v_1, \ldots, v_k) \in \mathcal{F}(V_1 \times \ldots \times V_k)$.

Segue da definição que

$$v_1 \otimes \ldots \otimes av_i \otimes \ldots \otimes v_k = a(v_1, \otimes \ldots \otimes v_k)$$

$$\begin{aligned} v_1 \otimes \ldots \otimes (v_i + v_i') \otimes \ldots \otimes v_k &= v_1 \otimes \ldots \otimes v_i \otimes \ldots \otimes v_k \\ &\quad + v_1 \otimes \ldots \otimes v_i' \otimes \ldots \otimes v_k. \end{aligned}$$

Proposição 6.2 (**Propriedade universal do produto tensorial**)**.** *Sejam $V_1, \ldots, V_k$ espaços vetoriais (de dimensão finita) sobre $\mathbb{R}$. Se $A\colon V_1 \times \ldots \times V_k \to X$ é um mapa multilinear em um espaço vetorial X, então existe um único mapa linear $\tilde{A}\colon V_1 \otimes \ldots \otimes V_k \to X$ tal que o seguinte diagrama comuta:*

$$\begin{tikzcd} V_1 \times \ldots \times V_k \arrow[r, "A"] \arrow[d, "\pi"] & X \\ V_1 \otimes \ldots \otimes V_k \arrow[ur, dashed, "\tilde{A}"'] & \end{tikzcd}$$

sendo π o mapa definido por $\pi(v_1, \ldots, v_k) = v_1 \otimes \ldots \otimes v_k$.

Demonstração.

Pela Proposição 6.1, A possui uma única extensão linear $A'\colon \mathcal{F}(V_1 \times \ldots \times V_k) \to X$. Facilmente verifica-se que $\mathcal{R} \subseteq \ker \tilde{A}$ e portanto $V_1 \otimes \ldots \otimes V_k \subseteq \mathcal{F}(V_1 \times \ldots \times V_k)/\ker A'$. Temos que A' é um homomorfismo por conta da linearidade. Ademais, $\ker A'$ e $Im(A')$ são subespaços vetoriais, o que significa que em particular são subgrupos abelianos do espaço vetorial livre e de X, respectivamente. Pelo primeiro teorema de isomorfismos, existe um único mapa $\psi\colon \mathcal{F}(V_1 \times \ldots \times V_k)/\ker A' \to Im(A')$ tal que $\mathcal{F}(V_1 \times \ldots \times V_k)/\ker A' \simeq Im(A')$, definido por $\psi([x]) := A'(x)$. Logo, basta definir $\tilde{A} := \psi|_{V_1 \otimes \ldots \otimes V_k}$ para obter o mapa desejado, pois $\tilde{A} \circ \pi = A$. ∎

Proposição 6.3 (**Base para** $V_1 \otimes \ldots \otimes V_k$)**.** *Suponha que* $V_1, \ldots, V_k$ *são espaços vetoriais sobre* $\mathbb{R}$ *tal que* $\dim V_j = n_j$*, para* $j = 1, \ldots, k$*. Além disso, para cada* j*, seja* $\{E_1^{(j)}, \ldots, E_{n_j}^{(j)}\}$ *uma base para* V_j*. Então o conjunto*

$$\mathcal{C} = \left\{ E_{i_1}^{(1)} \otimes \ldots \otimes E_{i_k}^{(k)} \colon 1 \leq i_1 \leq n_1, \ldots 1 \leq i_k \leq n_k \right\}$$

é uma base para $V_1 \otimes \ldots \otimes V_k$*.*

Demonstração.

Segue da definição de $V_1 \otimes \ldots \otimes V_k$ que este espaço é o span de elementos da forma $v_1 \otimes \ldots \otimes v_k$, onde $v_j \in V_j$. Escrevendo cada v_j em termos dos elementos da base de seu respectivo espaço, segue que $span(\mathcal{C}) = V_1 \otimes \ldots \otimes V_k$. Resta que o conjunto é L.I.. Para tanto, suponha que

$$\alpha^{i_1 \ldots i_k} E_{i_1}^{(1)} \otimes \ldots \otimes E_{i_k}^{(k)} = 0.$$

Para cada k-úpla $m_1, \ldots, m_k$ de índices, seja $\tau^{m_1 \ldots m_k} \colon V_1 \times \ldots \times V_K \to \mathbb{R}$ o mapa definido por

$$\tau^{m_1 \ldots m_k}(v_1, \ldots, v_k) := \varepsilon_{(1)}^{m_1}(v_1) \ldots \varepsilon_{(k)}^{m_k}(v_k).$$

Claramente este é um mapa multilinear, portanto, pela Proposição 6.2(propriedade universal), existe o mapa linear $\tilde{\tau}^{m_1 \ldots m_k} \colon V_1 \otimes \ldots \otimes V_K \to \mathbb{R}$ tal que

$$\begin{aligned} \tilde{\tau}^{m_1 \ldots m_k} \left(\alpha^{i_1 \ldots i_k} E_{i_1}^{(1)} \otimes \ldots \otimes E_{i_k}^{(k)} \right) &= \alpha^{i_1 \ldots i_k} \tau^{m_1 \ldots m_k} \left(E_{i_1}^{(1)}, \ldots, E_{i_k}^{(k)} \right) \\ &= \alpha^{m_1 \ldots m_k} = 0. \end{aligned}$$

■

Corolário 6.1. $\dim(V_1 \otimes \ldots \otimes V_k) = \prod_{j=1}^{k} n_j$.

Pelo corolário acima, podemos ver que o produto tensorial de espaços vetoriais $V_1, \ldots V_k$ é bem maior que o produto cartesiano destes espaços. Uma evidência disto é que nem todo elemento do produto tensorial dos espaços acima é da forma $v_1 \otimes \ldots \otimes v_k$; existem elementos que são combinações lineares de objetos da forma anterior. Quando um tensor pode ser descrito apenas por $v_1 \otimes \ldots \otimes v_k$, diz-se que ele é puro/decomponível. Parece que noção pode ser estendida à mecânica quântica no contexto de entrelaçamento quântico. Talvez valha a pena dar uma pesquisada.

Proposição 6.4. *Sejam* V_1*,* V_2 *e* V_3 *espaços vetoriais de dimensão finita sobre* $\mathbb{R}$*. Então existem isomorfismos únicos*

$$V_1 \otimes (V_2 \otimes V_3) \simeq V_1 \otimes V_2 \otimes V_3 \simeq (V_1 \otimes V_2) \otimes V_3,$$

com $v_1 \otimes (v_2 \otimes v_3$*,* $v_1 \otimes v_2 \otimes v_3$ *e* $(v_1 \otimes v_2) \otimes v_3$ *sendo correspondentes dois a dois.*

Demonstração.
Seja $A\colon V_1 \times V_2 \times V_3 \to (V_1 \otimes V_2) \otimes V_3$ definido por

$$A(v_1, v_2, v_3) := (v_1 \otimes v_2) \otimes v_3.$$

Sendo A multilinear, então pela propriedade universal, segue que existe o mapa linear $\tilde{A}\colon V_1 \otimes V_2 \otimes V_3 \to (V_1 \otimes V_2) \otimes V_3$ definido por

$$\tilde{A}(v_1 \otimes v_2 \otimes v_3) = (v_1 \otimes v_2) \otimes v_3.$$

Claramente $\tilde{A}$ é sobrejetivo. Logo, pelo Teorema do Núcleo e da Imagem, segue que $\dim \ker(\tilde{A}) = 0$. Donde conclui-se, da linearidade de $\tilde{A}$, que $\ker(\tilde{A} = \{0\}$ e portanto esta aplicação é um isomorfismo. Como qualquer outro isomorfismo teria de concordar com $\tilde{A}$ nos elementos $v_1 \otimes v_2 \otimes v_3$, está garantida a unicidade do isomorfismo. ■

Proposição 6.5 (Produto tensorial abstrato versus produto tensorial de covetores). *Sejam $V_1, \ldots, V_k$ espaços vetoriais de dimensão finita. Então existe um isomorfismo canônico*

$$V_1^* \otimes \ldots \otimes V_k^* \simeq L(V_1, \ldots, V_K; \mathbb{R}), \tag{6.2}$$

tal que cada $(v_1 \otimes \ldots \otimes v_k$ corresponde ao produto tensorial de covetores[11] $\omega^1 \otimes \ldots \otimes \omega^k$.

Demonstração.
Seja $\psi\colon V_1^* \times \ldots \times V_k^* \to L(V_1, \ldots, V_k; \mathbb{R})$ definido pontualmente por

$$\psi(\omega^1, \ldots, \omega^k)(v_1, \ldots, v_k) := \omega^1(v_1) \ldots \omega^k(v_k).$$

Como ψ é um mapa multilinear e pela propriedade universal do produto tensorial, conclui-se que existe um único mapa linear $\tilde{psi}\colon V_1^* \otimes \ldots \otimes V_k^* \to L(V_1, \ldots, V_k; \mathbb{R})$. Definindo bases em cada um dos V_j, visto que $\tilde{\psi}$ leva a base de $V_1^* \otimes \ldots \otimes V_k^*$ na base (induzida) de $L(V_1, \ldots, V_k; \mathbb{R})$, conclui-se que $\tilde{\psi}$ é um isomorfismo. ■

Como os espaços na Proposição 6.5 são de dimensão finita, existe o isomorfismo canônico entre V_j^{**} e V_j para cada j e portanto, pela mesma proposição, temos também que

$$V_1 \otimes \ldots \otimes V_k \simeq L(V_1^*, \ldots, V_k^*; \mathbb{R}). \tag{6.3}$$

Doravante, podemos entender um tensor tanto como uma classe de equivalência em quocientes à la $\mathcal{F}(V_1 \times \ldots \times V_k)/\mathcal{R}$ como uma função multilinear em conjuntos à la $L(V_1^*, \ldots, V_k^*; \mathbb{R})$.

[11] Funcionais lineares, elementos do dual algébrico de cada V_j.

6.2 Tensores covariantes, contravariantes, simétricos e alternantes

Definição 6.3. *Seja V um espaço vetorial sobre $\mathbb{R}$ de dimensão finita. Se k é um inteiro positivo, então um **tensor covariante do tipo k sobre V** é um elemento de $V_1^* \otimes \ldots \otimes V_k^*$, onde $V_j = V$ para todo j. Pela proposição 6.5, este elemento pode ser visto como uma função k-linear*

$$\alpha\colon V \times \ldots \times V \to \mathbb{R}.$$

*O número k é chamado de **rank de $\boldsymbol{\alpha}$**.*

Por convenção, um tensor covariante de tipo 0 é um número real. O espaço vetorial dos tensores covariantes do tipo k sobre V também é denotado por

$$T^k(V^*) \equiv V_1^* \otimes \ldots \otimes V_k^*.$$

Exemplo 6.1. *Seja V um espaço vetorial de dimensão finita. Então*

a) Qualquer funcional linear $\omega\colon V \to \mathbb{R}$ é um elemento de $T^1(V^*)$.

*b) Todo produto interno $\langle\cdot|\cdot\rangle : V \times V \to \mathbb{R}$ é um tensor covariante do tipo 2, comumente chamado de **forma bilinear**.*

c) O determinante de um operador linear $A\colon V \to V$ é um elemento de $T^n(V^*)$ quando $\dim V = n$.

Definição 6.4. *Seja V um espaço vetorial sobre $\mathbb{R}$ de dimensão finita. O espaço dos **tensores contravariantes de rank k sobre V** é o espaço vetorial*

$$T^k(V) \equiv V_1 \otimes \ldots \otimes V_k,$$

onde $V_j = V$ para todo j.

Por convenção, assume-se que $T^0(V) = \mathbb{R}$ e pela definição acima, temos que $T^1(V) = V$.

Pela proposição 6.5, temos que cada vetor contravariante de rank k é uma função multilinear definida no produto cartesiano de k cópias de V^*.

Definição 6.5. *Sejam k e l inteiros não negativos. O espaço dos **tensores mistos do tipo (k,l) sobre V** é definido como*

$$T^{(k,l)}(V) \equiv V_1 \otimes \ldots \otimes V_k \otimes V_1^* \otimes \ldots \otimes V_l^*,$$

onde $V_j = V$, $j = 1, \ldots, k$ e $V_i^* = V^*$, $i = 1, \ldots, l$.

Observe que

$$T^{(0,0)} = T^0(V) = T^0(V^*) = \mathbb{R}$$

$$T^{(0,1)} = T^1(V^*) = V^*$$

$$T^{(1,0)} = T^1(V) = V$$

$$T^{(0,k)} = T^k(V^*)$$

$$T^{(k,0)} = T^k(V)$$

Observação: as igualdades acima só valem devido à forma como definimos $T^{(k,l)}(V)$. Existem textos que invertem os papéis de k e l. Outra notação possível para os tensores mistos conforme definidos aqui é $T^l_k(V)$.

Corolário 6.2. *Seja V um espaço vetorial sobre $\mathbb{R}$ tal que* $\dim V = n \in \mathbb{N}_0$. *Suponha que a coleção $\{E_i\}$ é uma base para V e que $\{\varepsilon^j\}$ seja a base dual. Então, vale que os conjuntos:*

$$\{\varepsilon^{i_1} \ldots \otimes \ldots \otimes \varepsilon^{i_k} \colon 1 \leq i_1, \ldots, i_k \leq n\},$$

$$\{E_{i_1} \ldots \otimes \ldots \otimes E_{i_k} \colon 1 \leq i_1, \ldots, i_k \leq n\},$$

$$\{E_{i_1} \ldots \otimes \ldots \otimes E_{i_k} \otimes \varepsilon^{j_1} \ldots \otimes \ldots \otimes \varepsilon^{j_l} \colon 1 \leq i_1, \ldots, i_k, j_1, \ldots, j_l \leq n\}$$

são bases para $T^k(V^*)$, $T^k(V)$ e $T^{(k,l)}(V)$, respectivamente.

Definição 6.6. *Seja V um espaço vetorial real de dimensão finita. Um tensor k-covariante α em V é dito* ***simétrico*** *quando seu valor é inalterado por trocas de posição dois a dois de seus argumentos, ou seja:*

$$\alpha(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = \alpha(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k)$$

para $1 \leq i < j \leq k$.

Proposição 6.6. *Seja α um tensor k-covariante sobre V. As afirmações a seguir são equivalentes:*

a) α é simétrico.

b) Para quaisquer $v_1, \ldots, v_k \in V$, o valor de $\alpha(v_1, \ldots, v_k)$ é inalterado por qualquer troca de posição de seus argumentos.

c) Para qualquer base fixada em V, as componentes de α permanecem inalteradas por qualquer permutação da ordem dos índices.

Demonstração.

a $\implies$ b) Seja $\sigma\colon \{1,\ldots,k\} \to \{1,\ldots,k\}$ uma permutação. Defina

$$(\sigma\alpha)(v_1,\ldots,v_k) := \alpha(v_{\sigma(1)},\ldots,v_{\sigma(k)}).$$

Podemos escrever σ como o produto de r transposições[12] τ_l, $l = 1,\ldots,r$. Pela simetria de α, segue que $\tau_l\alpha = \alpha$ para todo l. Desta forma, $\sigma\alpha = (\tau_1 \circ \ldots \circ \tau_r)\alpha = \alpha$. Apesar de a representação de σ como produto de transposições não ser única, este resultado independe deste fato, pois se escrevermos σ como produto de $s \neq r$ transposições, basta aplicar o mesmo raciocínio.

b $\implies$ c) Seja $\{e_1,\ldots,e_n\}$ uma base qualquer em V. Fixando-a, sabemos que

$$\alpha(e_{l_1},\ldots,e_{l_k}) = \alpha_{l_1\ldots l_k},\ 1 \leq l_1,\ldots,l_k \leq n.$$

Por hipótese, se σ é uma permutação de $\{1,\ldots,k\}$, temos que

$$\alpha(e_{l_1},\ldots,e_{l_k}) = \alpha(e_{\sigma(l_1)},\ldots,e_{\sigma(l_k)}).$$

Logo, $\alpha_{l_1\ldots l_k} = \alpha_{\sigma(l_1)\ldots\sigma(l_k)}$.

c $\implies$ a) Seja $\{e_1,\ldots,e_n\}$ uma base qualquer em V fixada.

Sejam $v_1,\ldots,v_i,\ldots,v_j,\ldots,v_k \in V$, com $i < j$. Usando a notação de Einstein, sabemos que

$$\alpha(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_k) = \alpha(v_1^{l_1}e_{l_1},\ldots,v_i^{l_i}e_{l_i},\ldots,v_j^{l_j}e_{l_j},\ldots,v_k^{l_k}e_{l_k}).$$

Pela multilinearidade de α, segue que

$$\alpha(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_k) = v_1^{l_1}\ldots v_i^{l_i}\ldots v_j^{l_j}\ldots v_k^{l_k}\alpha(e_{l_1},\ldots,e_{l_i},\ldots,e_{l_j},\ldots,e_{l_k}).$$

Portanto

$$\alpha(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_k) = v_1^{l_1}\ldots v_i^{l_i}\ldots v_j^{l_j}\ldots v_k^{l_k}\alpha_{l_1\ldots l_i\ldots l_j\ldots l_k}.$$

Por hipótese, as componentes de α não se alteram por quaisquer permutações dos índices, portanto $\alpha_{l_1\ldots l_i\ldots l_j\ldots,l_k} = \alpha_{l_1\ldots l_j\ldots l_i\ldots l_k}$. Além disso, como os $v_m^{l_m}$ são números reais, por comutatividade do produto, concluímos que

$$\alpha(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_k) = \alpha(v_1,\ldots,v_j,\ldots,v_i,\ldots,v_k).$$

Logo, α é simétrico.

■

O espaço dos tensores k-contravariantes simétricos é um subespaço linear de $T^k(V^*)$. Denotando-o por $\Sigma^k(V^*)$, define-se a projeção $Sim\colon T^k(V^*) \to \Sigma^k(V^*)$, chamada de **simetrização**, que a cada tensor $\alpha \in T^k(V^*)$ associa o tensor

$$Sim\ \alpha := \frac{1}{k!}\sum_{\sigma\in S_k}\sigma\alpha,$$

onde σ é um elemento do grupo de permutações S_k.

[12]Permutações de apenas dois elementos de $\{1,\ldots,k\}$.

Exercício 6.1 (Propriedades da simetrização.). *Seja α um tensor contravariante de rank k. Então*

a) $Sim\ \alpha$ é simétrico.

b) $Sim\ \alpha = \alpha \iff \alpha \in \Sigma^k(V^*)$.

Exercício 6.2. *Se α e β são respectivamente tensores simétricos contravariantes de ranks k e l, não necessariamente $\alpha \otimes \beta$ é um tensor simétrico. Apesar disso, usando a simetrização, podemos definir uma operação chamada* ***produto simétrico*** *que associa ao par ordenado (α, β) um tensor simétrico de rank $k+l$. Definimos esse novo tensor como sendo*

$$\alpha\beta := Sim(\alpha \otimes \beta).$$

Explicitamente, temos que

$$\alpha\beta(v_1, \ldots, v_{k+l}) = \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\beta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}).$$

Mostre que o produto simétrico satisfaz às seguintes propriedades:

a) O produto simétrico é bilinear.

b) Se α e β são covetores, i.e., $k = 1$, então

$$\alpha\beta = \frac{1}{2}(\alpha \otimes \beta + \beta \otimes \alpha).$$

Definição 6.7. *Seja V um espaço vetorial de dimensão finita. Um tensor k-covariante em V é dito* ***alternante***[13] *quando qualquer mudança de posição entre dois argumentos provoca uma alteração de sinal, ou seja:*

$$\alpha(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -\alpha(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k),$$

para $1 \leq i < j \leq k$.

Outros nomes possíveis para estes tensores são k-covetores, multi-covetores e formas exteriores. O conjunto dos tensores k-covariantes alternantes é um subespaço vetorial de $T^k(V^*)$ simbolizado por $\Lambda^k(V^*)$.

O exercício a seguir é análogo à proposição 6.6.

Exercício 6.3. *Seja α um tensor k-covariante sobre V. As afirmações a seguir são equivalentes:*

a) α é anti-simétrico.

[13] Às vezes chamado de anti-simétrico.

b) *Para quaisquer* $v_1, \ldots, v_k \in V$ *e qualquer permutação* $\sigma \in S_k$*:*

$$\alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = (sgn\ \sigma)\alpha(v_1, \ldots, v_k).$$

Lembrando que o sinal de uma permutação é $sgn\ \sigma = (-1)^m$*, onde* m *é o número de transposições em que* σ *foi decomposta.*

c) *Para qualquer base fixada em* V*, as componentes de* α *mudam de sinal para qualquer transposição de índices.*

Algumas observações simplórias sobre tensores covariantes: todo tensor de rank 0 é simétrico e alternante, pois são números reais; todo tensor de rank 1 também é simétrico e alternante, pois existem 0 transposições a serem feitas; todo tensor de rank 2 β é combinação linear de tensores simétricos e alternantes, pois

$$\beta(v, w) = \frac{1}{2}(\beta(v, w) - \beta(w, v)) + \frac{1}{2}(\beta(v, w) + \beta(w, v)) = \alpha(v, w) + \eta(v, w),$$

onde α é alternante e η é simétrico.

Referências

[1] Elon Lages Lima. *Álgebra Linear*. Coleção Matemática Universitária. IMPA, 2018.

[2] J. C. A. Barata. *Notas para um Curso de Física-Matemática*. http://denebola.if.usp.br/~jbarata/Notas_de_aula/capitulos.html. Acessado em 04/03/2021.

[3] William T. Gowers. *How to lose your fear of tensor products*. https://www.dpmms.cam.ac.uk/~wtg10/tensors3.html. Acessado em 06/12/2020.