Qualificação de Mestrado

Lucas Giraldi Almeida Coimbra

5 de março de 2024

Conteúdo

1 Álgebra Linear

1.1 Fundamentos e Dualidade

Fundamentos

Tome $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Um **espaço vetorial** é um conjunto V munido de duas operações

$$\begin{aligned} +\colon V \times V &\to V \\ (x,y) &\mapsto x+y \end{aligned} \quad \text{e} \quad \begin{aligned} \cdot\colon \mathbb{K} \times V &\to V \\ (\lambda,x) &\mapsto \lambda x \end{aligned} \tag{1}$$

tais que a operação $+$ (soma) é comutativa, associativa, possui identidade e todos os inversos, e a operação de $\cdot$ (produto por escalar) satisfaz as relações distributivas, $1x = x$ e $\lambda(\mu x) = (\lambda\mu)x$.

Um **subespaço vetorial** de um espaço vetorial V é um subconjunto $S \subset V$ tal que para todos $x, y \in S$ e $\lambda \in \mathbb{K}$, temos $x + y \in S$ e $\lambda x \in S$. Todo espaço vetorial é um subespaço vetorial de si mesmo, assim como $\{0\}$ é sempre um subespaço vetorial. Chamamos V e $\{0\}$ de **subespaços triviais**. Se $U, W \subset V$ são dois subespaços, então o conjunto $U + V = \{u + w \mid u \in U, w \in W\}$ é a **soma** desses subespaços, e é também um subespaço. Se $U \cap W = \{0\}$, então diremos que a soma desses espaços é uma **soma direta** e a denotamos por $U \oplus W$. A intersecção $U \cap W$ também é sempre um subespaço vetorial.

Uma **combinação linear** de vetores em V é uma soma finita

$$\sum_{i=1}^{n} \lambda^i x_i = 0 \tag{2}$$

onde cada $\lambda^i \in \mathbb{K}$ e cada $x_i \in V$. Dado um conjunto $S \subset V$, o conjunto $\langle S \rangle$ de todas as combinações lineares de elementos de S é um subespaço vetorial de V, chamado de **subespaço gerado por** S. O conjunto S é **gerador** de V se $\langle S \rangle = V$.

Dizemos que $x_1, \dots, x_n \in V$ são **linearmente independentes** se para quaisquer $\lambda^1, \dots, \lambda^n \in \mathbb{K}$ tais que

$$\sum_{i=1}^{n} \lambda^i x_i = 0, \tag{3}$$

então $\lambda_i = 0$ para todo i. Vetores que não são linearmente independentes são **linearmente dependentes**. Fica claro da definição que um conjunto de vetores é linearmente dependente se, e somente se, um dos vetores pode ser escrito como combinação linear dos outros. Além disso, é fácil ver que se uma subcoleção de vetores é linearmente dependente, então a coleção original também é. Mais ainda, qualquer coleção de vetores que contenha o 0 é linearmente dependente.

Lema 1. *Sejam $S = \{s_1, \dots, s_n\}$ um gerador de V e $v_1, \dots, v_m$ vetores linearmente independentes. Então, $m \leq n$.*

Demonstração. Suponha que $m > n$. Como S gera V, então existem $\lambda^1, \dots, \lambda^n$ tais que

$$y_1 = \sum_{i=1}^{n} \lambda^i s_i. \tag{4}$$

Como $y_1 \neq 0$ (pela independência linear), então algum λ_j é não nulo, ou seja, podemos substituir s_j por y_1 e o conjunto resultante ainda gera V. Pela independência linear dos y_i, podemos fazer essa operação mais $n-1$ vezes, garantindo que $y_1, \dots, y_n$ geram V. Porém, isso significa que $y_{n+1}, \dots, y_m$ são combinação linear de $y_1, \dots, y_n$, o que contradiz a independência linear. Segue então que $m \leq n$. $\square$

Um espaço V é **finitamente gerado** se existe um conjunto gerador finito. Uma **base** de V é um conjunto gerador linearmente independente.

Lema 2. *Todo espaço finitamente gerado possui uma base.*

Demonstração. Se $S = \{s_1, \ldots, s_n\}$ gera V, então se S é linearmente independente o trabalho acabou. Caso contrário, algum s_i é combinação linear dos outros, e então retiramos ele e o conjunto resultante ainda gera V. Fazemos isso até que os vetores que sobram em S sejam linearmente independentes, e assim temos uma base. □

A partir de agora vamos trabalhar apenas com espaços finitamente gerados e, caso queiramos falar em um contexto mais geral, iremos explicitar. A **dimensão** de um espaço V, denotada por $\dim V$, é o número de elementos de uma base. Pelo Teorema a seguir, esse número está bem definido.

Teorema 3. *Toda base possui mesmo número de elementos.*

Demonstração. Como bases são linearmente independentes e geradoras, o resultado segue facilmente do Lema 1. □

O Lema 2 assume implicitamente que o conjunto S que gera V é não vazio. Caso tenhamos $V = \langle \varnothing \rangle$, então $V = \{0\}$ e o chamamos de **espaço trivial**. Sua dimensão é, por definição, nula.

Teorema 4. *Todo conjunto linearmente independente pode ser estendido para uma base.*

Demonstração. Se S é um conjunto linearmente independente, então considere $\langle S \rangle$. Se $\langle S \rangle = V$, então o conjunto S já é uma base. Caso contrário, seja $v \in V \setminus \langle S \rangle$ e tome $S_1 = S \cup \{v\}$. Podemos agora testar se $\langle S_1 \rangle = V$ e, caso contrário, repetir o processo. Como o espaço V é finitamente gerado, esse processo obrigatoriamente acaba, que é quando adicionamos vetores o suficiente em S para que se torne uma base. □

Note que todo subespaço de um espaço com dimensão finita, possui dimensão finita (pelo Lema 1). Se W é um subespaço de V, um subespaço U de V é um **complemento** de W se $U \oplus W = V$.

Teorema 5. *Complementos sempre existem e são únicos.*

Demonstração. Se W é um subespaço, seja $v_1, \ldots, v_m$ uma base de W e a complete para uma base $v_1, \ldots, v_m$, $w_1, \ldots, w_n$ de V. Defina $U = \langle w_1, \ldots, w_n \rangle$. Se $x \in U \cap W$, então existem $\lambda^1, \ldots, \lambda^m$ e $\mu^1, \ldots, \mu^n$ em $\mathbb{K}$ tais que

$$x = \sum_{i=1}^{m} \lambda^i v_i = \sum_{j=1}^{n} \mu^j w_j, \tag{5}$$

ou seja,

$$\sum_{i=1}^{m} \lambda^i v_i + \sum_{j=1}^{n} \mu^j w_j = 0, \tag{6}$$

portanto cada λ^i e μ^j é nulo, da onde segue que $x = 0$ e assim $U \cap W = \{0\}$. Mais ainda, se $x \in V$, então podemos escrever

$$x = \sum_{i=1}^{m} \lambda^i v_i + \sum_{j=1}^{n} \mu^j w_j \tag{7}$$

e assim $x = v + w$ com $v \in U$ e $w \in W$, da onde segue que $V = U \oplus W$. □

Note que da demonstração acima tiramos um outro fato importante: se $V = U \oplus W$, então $\dim V = \dim U + \dim W$. A **codimensão** de um subespaço S em um espaço V é definida por $\operatorname{codim} S = \dim V - \dim S$. Um espaço de codimensão 1 é um **hiperplano** em V.

Uma **transformação linear** é um mapa $T \colon V \to W$ entre espaços vetoriais tal que $T(x+y) = T(x) + T(y)$ e $T(\lambda x) = \lambda T(x)$ para todos $x, y \in V$ e $\lambda \in \mathbb{K}$. Se T for bijetora, dizemos que é um **isomorfismo linear** e que V e W são **isomorfos**. A cada transformação linear estão associados dois subespaços vetoriais: o **núcleo** e a **imagem**:

$$\ker T = \{v \in V \mid T(v) = 0\} \quad \text{e} \quad \operatorname{im} T = \{w \in W \mid w = T(v) \text{ para algum } v \in V\}. \tag{8}$$

É importante notar que uma transformação linear pode ser unicamente determinada pelo seus valores em alguma base do domínio. De fato, se $e_1, \ldots, e_n$ é uma base de V, então dado $v \in V$ temos $v = \lambda^i e_i$ e portanto, por linearidade, $T(v) = \lambda^i T(e_i)$, assim basta sabermos as coordenadas de v e os valores de T na base para determinar $T(v)$. A partir de agora, será comum denotarmos Tv para $T(v)$ caso T seja linear.

Proposição 6. *Uma transformação linear $T: V \to W$ é injetora se, e somente se,* $\ker T = \{0\}$*. Além disso, transformações lineares preservam dependência linear, e transformações lineares injetoras preservam independência linear.*

Demonstração. Se T é injetora, então $\ker T = \{0\}$ pois só existe um vetor que é levado em $0 \in W$, que é $0 \in V$. Agora, se $\ker T = \{0\}$, então se $T(v) = T(w)$, temos $T(v) - T(w) = 0$ e assim $T(v - w) = 0$, portanto $v - w = 0$ e assim $v = w$.

Se $v_1, \ldots, v_n$ são linearmente dependentes, então existem $\lambda^1, \ldots, \lambda^n$ não toodos nulos e tais que

$$\sum_{i=1}^{n} \lambda^i v_i = 0. \tag{9}$$

Dessa forma, se T é linear, como $T(0) = 0$ temos

$$\sum_{i=1}^{n} \lambda^i T(v_i) = 0, \tag{10}$$

assim os vetores $T(v_i)$ são linearmente dependentes.

Se $v_1, \ldots, v_n$ são linearmente independentes e T é injetora, então considere uma combinação linear nula

$$\sum_{i=1}^{n} \lambda^i T(v_i) = 0. \tag{11}$$

Como T é linear, isso equivale a dizer que

$$T\left(\sum_{i=1}^{n} \lambda^i v_i\right) = 0 \tag{12}$$

e, como T é injetora, então

$$\sum_{i=1}^{n} \lambda^i v_i = 0, \tag{13}$$

assim cada $\lambda^i = 0$ e portanto os vetores $T(v_i)$ são linearmente independentes. $\square$

Corolário 7. *Se V e W tem dimensão finita, então são isomorfos se, e somente se, tem a mesma dimensão.*

Demonstração. Se $T: V \to W$ é isomorfismo, considere uma base $v_1, \ldots, v_n$ de V. Então $T(v_1), \ldots, T(v_n)$ são linearmente independentes e geram a imagem, afinal, se $w \in \operatorname{im} T$, então existe $v \in V$ com $T(v) = w$, assim

$$w = T\left(\sum_{i=1}^{n} \lambda^i v_i\right) = \sum_{i=1}^{n} \lambda^i T(v_i). \tag{14}$$

Como T é sobrejetor, $\operatorname{im} T = W$, portanto $T(v_1), \ldots, T(v_n)$ formam base de W, assim $\dim V = \dim W$.

Se $\dim V = \dim W = n$, então sejam $v_1, \ldots, v_n$ e $w_1, \ldots, w_n$ bases de V e W, respectivamente. Definimos $T: V \to W$ por $T(v_i) = w_i$ e o estendemos por linearidade, ou seja, se

$$v = \sum_{i=1}^{n} \lambda^i v_i, \tag{15}$$

definimos

$$T(v) = \sum_{i=1}^{n} \lambda^i w_i. \tag{16}$$

Esse mapa é injetor pois se $T(v) = 0$, $\lambda^i = 0$ e assim $v = 0$. O mapa é sobrejetor pois se

$$w = \sum_{i=1}^{n} \mu^i w_i \in W, \tag{17}$$

então

$$T\left(\sum_{i=1}^{n} \mu^i v_i\right) = w. \tag{18}$$

$\square$

Proposição 8. *Se $T\colon V \to W$ é um isomorfismo, então T^{-1} também é.*

Demonstração. T^{-1} é também bijetora, então basta mostrarmos sua linearidade. Se $v, w \in W$ e $\lambda \in \mathbb{K}$, então

$$T^{-1}(v+w) = T^{-1}(T(T^{-1}(v)) + T(T^{-1}(w))) = T^{-1}(T(T^{-1}(v) + T^{-1}(w))) = T^{-1}(v) + T^{-1}(w) \tag{19}$$

e, além disso,

$$T^{-1}(\lambda v) = T^{-1}(\lambda T(T^{-1}(v))) = T^{-1}(T(\lambda T^{-1}(v))) = \lambda T^{-1}(v). \tag{20}$$

□

Com a noção de isomorfismo, podemos generalizar a noção de soma direta. Note que $V = U \oplus W$ se, e somente se, o mapa

$$\begin{aligned} +\colon U \times W &\to V \\ (u, w) &\mapsto u + w \end{aligned} \tag{21}$$

é um isomorfismo. Generalizando, dizemos que V é a **soma direta** dos subespaços $V_1, \ldots, V_n$ se o mapa

$$\begin{aligned} \sum\colon V_1 \times \cdots \times V_n &\to V \\ (v_1, \ldots, v_n) &\mapsto v_1 + \cdots + v_n \end{aligned} \tag{22}$$

for um isomorfismo. Nesse caso, escrevemos

$$V = V_1 \oplus \cdots \oplus V_n \quad \text{ou} \quad V = \bigoplus_{i=1}^{n} V_i \tag{23}$$

e fica claro que $\dim V = \dim V_1 + \cdots + \dim V_n$. Além disso, segue da bijetividade do mapa em questão que se $v \in V$ então existem $v_1, \ldots, v_n$ com $v_i \in V_i$ únicos tais que $v = v_1 + \cdots + v_n$.

Fixado W um subespaço de um espaço V, podemos definir uma relação de equivalência, denotada por

$$u \equiv v \mod W \tag{24}$$

se, e somente se, $u - v \in W$. Denotamos a classe de equivalência de $v \in V$ por $[v]$ e o conjunto de todas as classes de equivalência por V/W, que será chamado de **quociente de V por W**. Esse conjunto possui estrutura de espaço vetorial utilizando as seguintes operações, que estão bem definidas:

$$[v] + [w] = [v + w] \quad \text{e} \quad \lambda[v] = [\lambda v]. \tag{25}$$

Proposição 9. *Se V é um espaço vetorial e W um subespaço, então $\dim V/W = \dim V - \dim W$. Mais precisamente, se $w_1, \ldots, w_n$ é uma base de V de maneira que $w_1, \ldots, w_n, v_1, \ldots, v_m$ é uma base de V, então $[v_1], \ldots, [v_m]$ é uma base de V/W.*

Demonstração. Primeiro, vamos verificar que $[v_1], \ldots, [v_m]$ são linearmente independentes. De fato, considere uma combinação linear nula

$$\sum_{i=1}^{n} \lambda^i [v_i] = [0]. \tag{26}$$

Pela definição das operaçções no quociente, temos que

$$\left[\sum_{i=1}^{n} \lambda^i v_i\right] = [0], \tag{27}$$

ou seja,

$$\sum_{i=1}^{m} \lambda^i v_i = w \in W. \tag{28}$$

Escrevendo w na base de W, temos

$$\sum_{i=1}^{m} \lambda^i v_i = \sum_{j=1}^{n} \mu^j w_j, \tag{29}$$

portanto $\lambda^i = \mu^j = 0$, da onde segue que $[v_1], \dots, [v_m]$ são linearmente independentes. O próximo passo é mostrar que esses vetores geram V/W. Se $[v] \in V/W$, então

$$v = \sum_{i=1}^{m} \lambda^i v_i + \sum_{j=1}^{n} \mu^j w_j \tag{30}$$

e assim

$$[v] = \left[\sum_{i=1}^{m} \lambda^i v_i + \sum_{j=1}^{n} \mu^j w_j \right] = \sum_{i=1}^{m} \lambda^i [v_i] + \sum_{j=1}^{n} \mu^j [w_j] = \sum_{i=1}^{m} \lambda^i [v_i] + \sum_{j=1}^{n} \mu^j [0] = \sum_{i=1}^{m} \lambda^i [v_i]. \tag{31}$$

□

Como corolário, se $\dim V = \dim W$, temos que $\dim V/W = 0$, portanto $V/W = \{0\}$ e assim $V = W$. O proxímo item é o que chamamos de teorema do isomorfismo, na sua versão linear.

Teorema 10. *Se $T\colon V \to W$ é linear, então o mapa*

$$\begin{aligned} \tilde{T}\colon V/\ker T &\to \operatorname{im} T \\ [v] &\mapsto T(v) \end{aligned} \tag{32}$$

está bem definido e é um isomorfismo linear. Como consequência, temos o teorema do núcleo e imagem:

$$\dim V = \dim \operatorname{im} T + \dim \ker T.$$

Demonstração. O mapa é sobrejetor, afinal, se $w \in \operatorname{im} T$, então existe $v \in V$ tal que $T(v) = w$, portanto $\tilde{T}([v]) = w$. O mapa é injetor, afinal, se $[v] \in \ker \tilde{T}$, então $\tilde{T}([v]) = 0$, assim $T(v) = 0$, portanto $v \in \ker T$ e assim $[v] = [0]$. Como $\dim V/\ker T = \dim V - \dim \ker T$ e $\tilde{T}$ é isomorfismo, então

$$\dim V - \dim \ker T = \dim \operatorname{im} T, \tag{33}$$

de onde segue o resultado. □

Se V e W são espaços vetoriais, o conjunto $V \times W$ é também um espaço vetorial com as operações

$$(x, y) + (v, w) = (x + v, y + w) \quad \text{e} \quad \lambda(x, y) = (\lambda x, \lambda y). \tag{34}$$

Se $V = U + W$, e $I = U \cap W$, então podemos tomar $T\colon U \times W \to V$ dada por $(u, w) \mapsto u + w$. Como $V = U + W$, o mapa é sobrejetor. Além disso, seu núcleo é $\{(x, -x) \mid x \in I\}$, que é isomorfo a I, assim segue do teorema do isomorfismo que

$$\dim U \times V = \dim V + \dim U \cap W \tag{35}$$

e, como $\dim U \times V = \dim U + \dim V$, segue que

$$\dim V = \dim U + \dim W - \dim U \cap W. \tag{36}$$

Proposição 11. *Se V e W possuem a mesma dimensão finita, então são equivalentes as seguintes afirmações sobre uma transformação linear $T\colon V \to W$:*

- *T é injetora;*
- *T é sobrejetora;*
- *T é um isomorfismo;*
- *T leva bases em bases;*

Demonstração. Se T é injetora, então $\ker T = \{0\}$ e assim, pelo teorema do núcleo e imagem, $\dim W = \dim V = \dim \ker T + \dim \operatorname{im} T = \dim \operatorname{im} T$, assim $\operatorname{im} T = W$ e T é sobrejetora.

Se T é sobrejetora, então $\dim \operatorname{im} T = \dim W = \dim V$, assim $\dim V = \dim \ker T + \dim V$, portanto $\dim \ker T = 0$ e portanto $\ker T = \{0\}$, da onde segue que T é injetora, e portanto um isomorfismo (pois já é sobrejetora).

Se T é isomorfismo, então se $v_1, \dots, v_n$ é base de V, então pela injetividade, $T(v_1), \dots, T(v_n)$ são linearmente independentes. Mais ainda, $T(v_1), \dots, T(v_n)$ geram a imagem de T, que é W, portanto esses vetores formam base de W.

Por fim, se T leva bases em bases, então se $v_1, \dots, v_n$ é base de V e $T(v) = 0$, então se

$$v = \sum_{i=1}^{n} \lambda^i v_i, \tag{37}$$

temos

$$0 = T(v) = \sum_{i=1}^{n} \lambda^i T(v_i), \tag{38}$$

da onde segue, pela independência linear de $T(v_1), \dots, T(v_n)$, que $\lambda^i = 0$, assim $v = 0$ e portanto T é injetora. □

Dualidade

A partir de agora, todo espaço será finitamente gerado, ou em outros termos, terá dimensão finita. Seja V um espaço vetorial sobre $\mathbb{K}$ (lembrando que $\mathbb{K} = \mathbb{R}, \mathbb{C}$). Um **funcional** ou **covetor** em V é um mapa linear $\varphi \colon V \to \mathbb{K}$. Denotamos o conjunto de todos os funcionais lineares em V por V^*. Esse conjunto se torna um espaço vetorial ao definirmos

$$(\varphi + \psi)(v) = \varphi(v) + \psi(v) \quad \text{e} \quad (\lambda\varphi)(v) = \lambda\varphi(v). \tag{39}$$

Fixada uma base $e = (e_1, \dots, e_n)$ de V (aqui, e representa uma lista de n vetores), se $v[e] = (v^1, \dots, v^n)$ são as coordenadas de v na base e, isso é,

$$v = \sum_{i=1}^{n} v^i e_i \tag{40}$$

, então podemos definir os mapas

$$\begin{aligned} \varepsilon^i \colon V &\to \mathbb{K} \\ v &\mapsto v^i \end{aligned}, \tag{41}$$

que vamos chamar de **diferenciais** com respeito a base e. A lista de vetores $\varepsilon = (\varepsilon^1, \dots, \varepsilon^n)$ é uma base de V^*, chamada de **base dual** de e. O fato dessa lista ser uma base implica diretamente que o mapa $e_i \mapsto \varepsilon^i$ é um isomorfismo entre V e V^*. Esse isomorfismo não é natural, isso é, ele depende da escolha de base e. Uma outra base gera outro isomorfismo, e não existe uma base canônica que podemos considerar.

Para um exemplo de isomorfismo natural, podemos considerar o **espaço bidual** de V, que é simplesmente V^{**}, isso é, o conjunto de todos os funcionais em V^*. Se $v \in V$ e $\varphi \in V^*$, podemos definir $v(\varphi) = \varphi(v)$ e portanto tratar cada $v \in V$ como um elemento de V^{**}. Essa identificação gera um isomorfismo entre V e V^{**} que não depende de nenhuma escolha arbitrária.

Seja W um subespaço de V. O **aniquilador** de W, denotado por $W^\perp$, é o subespaço de V^* consistido de todos os covetores que se anulam em W.

Proposição 12. *Se W é subespaço de V, então*

$$\dim W^\perp + \dim W = \dim V, \tag{42}$$

ou seja, $\dim W^\perp = \operatorname{codim} W$.

Demonstração. Considere uma base $w_1, \dots, w_n$ de W e um completamento $w_1, \dots, w_n, v_1, \dots, v_m$ para uma base de V. Defina $T \colon V \to V^*$ por $T(w_i) = 0$ e $T(v_i) = \nu_i$, onde ν_i são os elementos da base dual correspondentes a v_i. O núcleo de T é claramente W, basta mostrarmos que $\operatorname{im} T = W^\perp$.

Se $\varphi \in W^\perp$, então

$$\varphi = \sum_{i=1}^{n} \lambda^i \omega_i + \sum_{j=1}^{m} \mu^j \nu_j. \tag{43}$$

e assimm $\varphi(w_i) == \lambda^i$, mas $\varphi \in W^\perp$, portanto $\varphi(w_i) = 0$, assim

$$\varphi = \sum_{j=1}^{m} \mu^j \nu_j \tag{44}$$

e portanto

$$T\left(\sum_{j=1}^{m} \mu^j v_j\right) = \varphi, \tag{45}$$

assim $\varphi \in \operatorname{im} T$. Por outro lado, se $\varphi \in \operatorname{im} T$, então φ é combinação linear dos covetores $\nu_1, \dots, \nu_m$, que estão todos em $W^\perp$, portanto $\varphi \in W^\perp$, assim $\operatorname{im} T = W^\perp$. □

Proposição 13. *Se W é um subespaço de V, então o isomorfismo $v \mapsto (\varphi \mapsto \varphi(v))$ identifica W com $W^{\perp\perp}$.*

Demonstração. De fato, se $v \in W$, então precisamos mostrar que $v(\varphi) = 0$ para todo $\varphi \in W^\perp$, isso é, $w \in W^{\perp\perp}$. Isso, porém, é óbvio, já que $v(\varphi) = \varphi(v) = 0$, já que $v \in W$.

Agora, se $w \in W^{\perp\perp}$, então $w(\varphi) = 0$ para todo $\varphi \in W^\perp$, ou seja, $\varphi(w) = 0$ para todo $\varphi \in W^\perp$, assim $w \in W$. □

1.2 Mapas Lineares, Matrizes, Determinante e Traço

Mapas Lineares e Matrizes

Um fato utilizando anteriormente, e que não foi provado, é que a composição de mapas lineares é linear.

Proposição 14. *Composição de mapas lineares é linear e, em particular, composição de isomorfismos é isomorfismo. Mais ainda, se S e T são isomorfismos, $(S \circ T)^{-1} = T^{-1} \circ S^{-1}$.*

Demonstração. Se $T\colon V \to W$ e $S\colon W \to U$ são lineares, então dados $u, v \in V$ e $\lambda \in \mathbb{K}$, temos

$$(S \circ T)(u+v) = S(T(u+v)) = S(T(u) + T(v)) = S(T(u)) + S(T(v)) = (S \circ T)(u) + (S \circ T)(v) \tag{46}$$

e, além disso,

$$(S \circ T)(\lambda u) = S(T(\lambda u)) = S(\lambda T(u)) = \lambda S(T(u)) = \lambda (S \circ T)(u). \tag{47}$$

Se T e S forem isomorfismos, então

$$(S \circ T) \circ (T^{-1} \circ S^{-1}) = S \circ (T \circ T^{-1}) \circ S^{-1} = S \circ S^{-1} = \mathrm{Id}_U \tag{48}$$

e, além disso,

$$(T^{-1} \circ S^{-1}) \circ (S \circ T) = T^{-1} \circ (S^{-1} \circ S) \circ T = T^{-1} \circ T = \mathrm{Id}_V\,. \tag{49}$$

□

Além disso, a demonstração de que o núcleo de um mapa linear é um subespaço vetorial também nunca foi apresentada, mas isso é por que esse fato é um corolário de um resultado um pouco mais geral.

Proposição 15. *Se $T\colon V \to W$ é linear e $U \subset W$ é um subespaço, então $T^{-1}(U)$ é um subesaço*

Demonstração. De fato, se $u, v \in T^{-1}(U)$, então $T(u), T(v) \in W$, assim $T(u+v) = T(u) + T(v) \in W$, portanto $u + v \in T^{-1}(W)$. Mais ainda, se $\lambda \in \mathbb{K}$, então $T(\lambda u) = \lambda T(u) \in W$, assim $\lambda u \in T^{-1}(W)$. □

Podemos observar que, em coordenadas, todo mapa linear possui uma forma canônica.

Proposição 16. *Se $\dim V = n$, $\dim W = m$ e $T\colon V \to W$ é linear, então fixadas e e f bases de V e W temos*

$$Tv[f] = \left(\sum_{i_1=1}^{n} \lambda^1_{i_1} v[e]^{i_1}, \dots, \sum_{i_m=1}^{n} \lambda^m_{i_m} v[e]^{i_m}\right) \tag{50}$$

onde $Tv[f]$ são as coordenadas de Tv na base f e $v[e]^j$ é a j-ésima coordenada de v na base e.

Podemos organizar os números λ^i_j da Proposição 16 em um retângulo da forma

$$\begin{bmatrix} \lambda^1_1 & \lambda^1_2 & \cdots & \lambda^1_{n-1} & \lambda^1_n \\ \lambda^2_1 & \lambda^2_2 & \cdots & \lambda^2_{n-1} & \lambda^2_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda^{m-1}_1 & \lambda^{m-1}_2 & \cdots & \lambda^{m-1}_{n-1} & \lambda^{m-1}_n \\ \lambda^m_1 & \lambda^m_2 & \cdots & \lambda^m_{n-1} & \lambda^m_n \end{bmatrix} \tag{51}$$

Uma tabela nesse estilo será chamada de **matriz**. Cada coleção de números colocados na horizontal será uma **linha** da matriz, e cada coleção de números na vertical será uma **coluna** da matriz. Perceba que toda transformação da origem a uma matriz como a descrita acima, que vamos chamar de **matriz de T com respeito as bases e e f**. Para abreviar a notação, denotamos a matriz acima por $[T]_{e,f} = [\lambda^i_j]_{m\times n}$, onde m e n são o número de linhas e colunas, respectivamente.

Agora vamos introduzir as operações em matrizes. Se $A = [a^i_j]_{n\times m}$ e $B = [b^i_j]_{n\times m}$ são matrizes, então podemos somá-las e multiplicar uma delas por um escalar da seguinte forma:

$$A + B = [a^i_j + b^i_j]_{n\times m} \quad \text{e} \quad \lambda A = [\lambda a^i_j]_{n\times m}. \tag{52}$$

Além disso, se $C = [c^i_j]_{m\times k}$ então podemos definir um produto entre A e C (note que o número de linhas de C deve ser o mesmo número de colunas de A) fazendo

$$AC = \left[\sum_{s=1}^{m} a^i_s c^s_j\right]_{n\times k}. \tag{53}$$

O produto não é comutativo, afinal, se $k \neq n$ então CA pode nem estar definido. Se $k = n \neq m$, então necessariamente $AC \neq CA$, visto que AC é $n \times n$ e $CA = m \times m$. Por fim, mesmo que $n = m = k$, poderíamos ter $AC \neq CA$, por exemplo:

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 7 & -1 \\ 1 & 3 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 3 & -2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 1 \\ 2 & 4 \end{bmatrix}. \tag{54}$$

Proposição 17. *Toda matriz vem de um mapa linear.*

Demonstração. Se $A = [a^i_j]$ é uma matriz $n \times m$, podemos enxergá-la como um mapa $\mathbb{R}^m \to \mathbb{R}^n$ que pega vetores $x = (x^1, \ldots, x^n)$ e retorna Ax, onde x é visto como a matriz coluna

$$x = \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix} \tag{55}$$

e Ax, que também é uma matriz coluna $n \times 1$, é identificada com um vetor de $\mathbb{R}^n$ da mesma maneira. O mapa é linear, afinal,

$$A(x+y) = \left[\sum_{k=1}^{m} a^i_k (x^k + y^k)\right]_{n\times 1} = \left[\sum_{k=1}^{m} a^i_k x^k + \sum_{k=1}^{m} a^i_k y^k\right]_{n\times 1} = Ax + Ay \tag{56}$$

e, além disso,

$$A(\lambda x) = \left[\sum_{k=1}^{m} a^i_k (\lambda x^k)\right]_{n\times 1} = \left[\lambda \sum_{k=1}^{m} a^i_k x^k\right]_{n\times 1} = \lambda A x. \tag{57}$$

É fácil notar que a matriz de A com respeito às bases canônicas de $\mathbb{R}^m$ e $\mathbb{R}^n$ é a própria A. □

Todo sistema linear pode ser traduzido para uma equação de matrizes. De fato, considere o sistema linear

$$\begin{cases} a^1_1 x^1 + \cdots + a^1_m x^m = b^1 \\ \qquad\qquad \vdots \\ a^n_1 x^1 + \cdots + a^n_m x^m = b^n \end{cases} \tag{58}$$

e perceba que, denotando $x = (x^1, \dots, x^n)$ e $b = (b^1, \dots, b^n)$, podemos escrever esse sistema simplesmente como $Ax = b$, onde $A = [a^i_j]_{n\times m}$.

Se $n > m$, o sistema é dito **sobredeterminado**, e existem mais equações do que incógnitas. Nesse caso a matriz nunca pode ser sobrejetora, e sempre vão existir infinitos vetores b tais que $Ax = b$ não possui solução. Se $b \in \operatorname{im} A$, note que $A^{-1}(b) = y + \ker A$ para algum y tal que $Ay = b$, então ou toda solução é única (se A for injetora) ou nenhuma solução é única (caso contrário).

Se $m > n$, isso é, existe mais incógnitas do que equações, então A nunca é injetora, e assim sempre que existe solução para $Ax = b$, a solução nunca é única. A justificativa é simples, afinal, $\dim\ker A = m - \dim\operatorname{im} A$ e $\dim\operatorname{im} A \le n < m$, assim $\dim\ker A > 0$ e, se $b \in \operatorname{im} A$, então $A^{-1}(b) = y + \ker A$ para alguma solução y, assim $A^{-1}(b)$ é infinito.

Se $n = m$, então se $Ax = b$ sempre tem solução, ela é sempre única, pois sobrejetividade e injetividade de A são equivalentes. Nesse caso, A é um isomorfismo, denotamos sua inversa por A^{-1} e a solução do sistema é $x = A^{-1}b$. Note que checar a injetividade de A é determinar se $Ax = 0$ possui solução não nula. Assumimos que o leitor já sabe técnicas de escalonamento de sistemas e portanto achar soluções já seja uma ferramenta conhecida.

Proposição 18. *Se $S, T\colon V \to U$ e $R\colon U \to W$ são lineares e e, f e g são bases de V, U e W, respectivamente, então*

$$[T+S]_{e,f} = [T]_{e,f} + [S]_{e,f}, \quad [\lambda T]_{e,f} = \lambda [T]_{e,f} \quad e \quad [R\circ T]_{e,g} = [R]_{f,g}[T]_{e,f}. \tag{59}$$

Demonstração. Note que, se $[T]_{e,f} = [\lambda^i_j]$, $[S]_{e,f} = [\mu^i_j]$ e $[R]_{f,g} = [\nu^i_j]$, então

$$Te_j = \sum_{i=1}^{\dim U} \lambda^i_j f_i, \quad Se_j = \sum_{i=1}^{\dim U} \mu^i_j f_i \quad \text{e} \quad Rf_j = \sum_{i=1}^{\dim W} \nu^i_j g_i, \tag{60}$$

assim

$$(T+S)e_j = Te_j + Se_j = \sum_{i=1}^{\dim U} \lambda^i_j f_i + \sum_{i=1}^{\dim U} \mu^i_j f_i = \sum_{i=1}^{\dim U} (\lambda^i_j + \mu^i_j) f_i, \tag{61}$$

portanto o resultado da soma segue. Similarmente,

$$(\delta T)e_j = \delta T e_j = \delta \sum_{i=1}^{\dim U} \lambda^i_j f_i = \sum_{i=1}^{\dim U} (\delta\lambda^i_j) f_i, \tag{62}$$

portanto o resultado do produto por escalar segue. Por fim,

$$(R\circ T)e_j = RTe_j = R\left(\sum_{i=1}^{\dim U} \lambda^i_j f_i\right) = \sum_{i=1}^{\dim U} \lambda^i_j R f_i = \sum_{i=1}^{\dim U} \lambda^i_j \sum_{k=1}^{\dim W} \nu^k_i g_k = \sum_{k=1}^{\dim W}\left(\sum_{i=1}^{\dim U} \lambda^i_j \nu^k_i\right) g_k, \tag{63}$$

da onde segue o resultado da composição. □

Esse resultado é tudo que precisamos para tratar matrizes e mapas lineares como as mesmas entidades. Em particular, isso garante que $(AB)^{-1} = B^{-1}A^{-1}$ para matrizes também! Mesmo que o resultado tenha sido provado apenas para mapas.

Proposição 19. *Se $T, S\colon V \to U$ e $R\colon U \to W$ são lineares, então vale que $R\circ(T+S) = R\circ T + R\circ S$. Se $Q\colon W \to V$ é linear, então vale que $(T+S)\circ Q = T\circ Q + S\circ Q$.*

Demonstração. Na primeira distributiva, basta usar que R é linear e, na segunda, basta aplicar a definição da soma de mapas. □

Se $T\colon V \to U$ é linear, então induz um mapa $T^\top\colon U^* \to V^*$, chamado de **transposto** de T, dado por $T^\top(\varphi) = \varphi\circ T$. Ao mesmo tempo, dada uma matriz $A = [a^i_j]_{n\times m}$ definimos sua **matriz transposta** por $A^\top = [a^j_i]_{m\times n}$. Adivinhem?

Proposição 20. *Se e é uma base de V, f é uma base de U e ε e δ são suas bases duais, então*

$$[T']_{\delta,\varepsilon} = [T]^\top_{e,f} \tag{64}$$

Demonstração. Como $\dim V = \dim V^*$ e $\dim U = \dim U^*$, as dimensões das duas matrizes batem. Agora, note que se

$$Te_j = \sum_{i=1}^{\dim U} \lambda_j^i f_i, \tag{65}$$

então

$$(T^\top \delta^j)(e_k) = (\delta^j \circ T)(e_k) = \delta^j \left(\sum_{i=1}^{\dim U} \lambda_k^i f_i \right) = \lambda_k^j, \tag{66}$$

assim, se $v[e] = (v^1, \dots, v^n)$ então

$$(T^\top \delta^j)(v) = \sum_{i=1}^{n} \lambda_i^j v^i = \sum_{i=1}^{n} \lambda_i^j \varepsilon^i(v), \tag{67}$$

portanto

$$T^\top \delta^j = \sum_{i=1}^{n} \lambda_i^j \varepsilon^i, \tag{68}$$

provando assim o resultado. □

Vamos agora mostrar alguns fatos sobre a transposta de mapas lineares.

Proposição 21. *Se $T, S\colon V \to U$ e $R\colon U \to W$ são lineares, então $(R \circ T)^\top = T^\top \circ R^\top$, $(T+S)^\top = T^\top + S^\top$ e, se T for um isomorfismo, então $T^\top$ também é e $(T^\top)^{-1} = (T^{-1})^\top$.*

Demonstração. Note que

$$(R \circ T)^\top \varphi = \varphi \circ R \circ T = T^\top(\varphi \circ R) = T^\top(R^\top \varphi) \tag{69}$$

e, além disso,

$$(T + S)^\top \varphi = \varphi \circ (T + S) = \varphi \circ T + \varphi \circ S = T^\top \varphi + S^\top phi. \tag{70}$$

Por fim, se T for isomorfismo, então

$$(T^\top \circ (T^{-1})^\top)\varphi = \varphi \circ T^{-1} \circ T = \varphi \quad \mathrm{e}((T^{-1})^\top \circ T^\top)\varphi = \varphi \circ T \circ T^{-1} = \varphi, \tag{71}$$

portanto $(T^\top)^{-1} = (T^{-1})^\top$. □

Até agora, temos maquinário para provar um monte de coisas sobre matrizes, usando mapas lineares. Mas e o contrário, é possível? Podemos garantir, por exemplo, que $T^{\top\top} = T$ (usando a idenficação natural $V^{**} = V$) usando apenas que isso é óbvio para matrizes? Sim! Porém, devemos anter terminar a nossa correspondência entre mapas lineares e matrizes. Já sabemos que toda matriz é um mapa linear, e todo mapa linear pode ser representado por uma matriz. Resta mostrar que essa correspondência "vai e volta":

> Se $T\colon V \to U$ é um mapa linear, então o mapa induzido por uma matriz $[T]_{e,f}$ corresponde a T. Se A é uma matriz, então a matriz do mapa induzido por A, em alguma base, é A.

A segunda parte já concluímos anteriormente, então basta entendermos a primeira mais precisamente, e prová-la.

Proposição 22. *Se $T\colon V \to U$ é um mapa linear, $\dim V = n$, $\dim U = m$ e e e f são bases de V e U, então o mapa $\mathbb{R}^n \to \mathbb{R}^m$ induzido por $[T]_{e,f}$ leva $v[e]$ em $Tv[f]$.*

Demonstração. Sejam $[T]_{e,f} = [\lambda_j^i]_{m \times n}$ e $v[e] = (v^1, \dots, v^n)$. Então o mapa induzido por $[T]_{e,f}$ é dado por

$$[T]_{e,f}(v^1, \dots, v^n) = \left(\sum_{i=1}^{n} \lambda_i^1 v^i, \dots, \sum_{i=1}^{n} \lambda_i^m v^i \right). \tag{72}$$

Agora, note que

$$Tv = T\left(\sum_{i=1}^{n} v^i e_i \right) = \sum_{i=1}^{n} v^i T(e_i) = \sum_{i=1}^{n} v^i \sum_{j=1}^{m} \lambda_i^j f_j = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \lambda_i^j v^i \right) f_j, \tag{73}$$

o que conclui que $Tv[f] = [T]_{e,f} v[e]$. □

Corolário 23. *Se $T, S\colon V \to U$ são lineares e existem bases e e f de V e U tais que $[T]_{e,f} = [S]_{e,f}$, então $T = S$.*

Demonstração. Como $[T]_{e,f} = [S]_{e,f}$, então eles levam $v[e]$ em $Tv[f] = Sv[f]$, portanto $Tv = Sv$. □

Vamos usar isso ao nosso favor! Note que, em termos de matrizes, fica bem claro que, para qualquer matriz T, $T^{\top\top} = T$, já que estamos apenas trocando linhas por colunas, duas vezes. Isso indica, claro, que o mesmo vale para mapas lineares!

Proposição 24. *Fazendo as identificações naturais de $V^{**} = V$ e $U^{**} = U$, se $T\colon V \to U$ é linear, então $T^{\top\top} = T$.*

Demonstração. Como

$$[T]_{e,f}^{\top} = [T^{\top}]_{\delta,\varepsilon} \implies [T]_{e,f}^{\top\top} = [T^{\top}]_{\delta,\varepsilon}^{\top}, \tag{74}$$

e também

$$[T]_{e,f}^{\top\top} = [T]_{e,f} \quad \text{e} \quad [T^{\top}]_{\delta,\varepsilon}^{\top} = [T^{\top\top}]_{e,f}, \tag{75}$$

então

$$[T]_{e,f} = [T^{\top\top}]_{e,f}. \tag{76}$$

□

Prosseguindo com matrizes, vamos agora falar de posto. Dada uma matriz A de dimensões $n \times m$, o **posto** de A, denotado $\operatorname{rank} A$, e que é a dimensão da imagem de A. Também definimos o **posto de linhas** de A, que é o maior número de linhas de A linearmente independentes, quando consideradas como vetores de $\mathbb{R}^m$. Por fim, definimos o **posto de colunas** de A, que é o maior número de colunas de A linearmente independentes, quando consideradas como vetores de $\mathbb{R}^n$. A ideia é mostrarmos que todos esses são equivalentes.

Proposição 25. *As colunas de $A = [a^i_j]_{n\times m}$ geram $\operatorname{im} A$.*

Demonstração. Se $x \in \mathbb{R}^m$, então $Ax \in \mathbb{R}^n$ e, se $x = (x^1, \ldots, x^m)$, temos

$$Ax = \left(\sum_{j=1}^{m} a^1_j x^j, \ldots, \sum_{j=1}^{m} a^n_j x^j \right) = \sum_{j=1}^{m} x^j (a^1_j, \ldots, a^n_j). \tag{77}$$

□

Corolário 26. *O posto de colunas de uma matriz é igual ao posto dessa mesma matriz.*

A parte problemática é mostrar que o posto de linhas e colunas é o mesmo. Bom, o caminho para isso é simples: o posto de linhas de A é o posto de colunas de $A^{\top}$, que é o posto de $A^{\top}$. Se mostrarmos que $\operatorname{rank} A^{\top} = \operatorname{rank} A$, o trabalho terminou. Para isso, vamos utilizar a noção de aniquilador que vimos anteriormente.

Teorema 27. *Se $T\colon V \to U$ é linear, então $(\operatorname{im} T)^{\perp} = \ker T^{\top}$.*

Demonstração. Se $\varphi \in (\operatorname{im} T)^{\perp}$, então dado $v \in V$, $\varphi(Tv) = 0$, isso é, $T^{\top}\varphi = 0$, ou seja, $\varphi \in \ker T^{\top}$. Por outro lado, se $\varphi \in \ker T^{\top}$, então $T^{\top}\varphi = 0$, ou seja, se $u \in \operatorname{im} T$, então existe $v \in V$ com $u = Tv$, assim $\varphi(u) = \varphi(Tv) = 0$ e portanto $\varphi \in (\operatorname{im} T)^{\perp}$. □

Corolário 28. *Se $T\colon V \to U$ é linear, então $\dim \operatorname{im} T = \dim \operatorname{im} T^{\top}$.*

Demonstração. Pelo teorema do núcleo e imagem,

$$\dim \operatorname{im} T^{\top} + \dim \ker T^{\top} = \dim U^* \tag{78}$$

e, como

$$\dim (\operatorname{im} T)^{\perp} + \dim \operatorname{im} T = \dim U, \tag{79}$$

então o resultado segue, usando que $\dim U = \dim U^*$ e $\dim (\operatorname{im} T)^{\perp} = \dim \ker T^{\top}$. □

Vamos agora focar a discussão em transformações lineares da forma $T\colon V \to V$, que costumamos chamar de **operadores** em V. Denotamos o conjunto dos operadores por $\mathcal{L}(V,V)$. Por consequência, nossa discussão de matrizes irá também se limitar a **matrizes quadradas**, isso é, aquelas em que o número de linhas é igual ao número de colunas. Note que a composição de mapas em $\mathcal{L}(V,V)$ sempre está definida, assim como o produto de duas matrizes quadradas de mesmas dimensões, por mais que ele não seja comutativo. O mapa identidade $\mathrm{Id}\colon V \to V$ é claramente linear e, dada e uma base de V, como $\mathrm{Id}(e) = e$, fica claro que

$$[\mathrm{Id}]_{e,e} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \tag{80}$$

que a partir de agora será chamada de **matriz identidade** $n \times n$, em que $n = \dim V$.

A partir de agora, iremos utilizar que a composição de mapas corresponde ao produto de matrizes e omitir o sinal de composição em mapas lineares. Dados dois mapas lineares $T, S \in \mathcal{L}(V,V)$, definimos $T_S = STS^{-1}$. O mapa $T \mapsto T_S$ é uma **conjugação**. Dizemos que T_S é **similar** ou **conjugada** a T. Esse mapa é claramente um isomorfismo linear (basta ver que a conjugação por S^{-1} é o mapa inverso da conjugação por S).

Teorema 29. *Conjugação é uma relação de equivalência.*

Demonstração. Se $R = T_S$ para alguma T, então $T = R_{S^{-1}}$, assim conjugação é simétrica. Mais ainda, $T = T_{\mathrm{Id}}$, então conjugação é reflexiva. Por fim, se $R = T_S$ e $T = K_L$, então $R = T_S = (K_L)_S = K_{SL}$. □

Mais ainda, fica claro que se M é inversível (possui inversa) e conjugada a T, então T é inversível. De fato, se $M = T_S$, então $T = M_{S^{-1}} = M_S^{-1}$, portantoo $T^{-1} = M_S$.

Determinantes e Traço

Para definirmos a noção de determinante, precisamos de permutações. Se $X = \{x_1, \ldots, x_n\} \subset \mathbb{N}$, uma **permutação em** X é uma bijeção $\sigma\colon X \to X$. O **discriminante** de X é o número

$$D(x_1, \ldots, x_n) = \prod_{i<j} (x_i - x_j). \tag{81}$$

O **sinal** da permutação σ é o número $(-1)^\sigma$ que satisfaz

$$P(x_1, \ldots, x_n) = (-1)^\sigma P(X_{\sigma(1)}, \ldots, x_{\sigma(n)}). \tag{82}$$

Denotamos por S_n o conjunto de todas as permutações em $[n] = \{1, \ldots, n\}$. Se $A = [a^i_j]_{n\times n}$ é uma matriz, definimos o determinante

$$\det(A) = \sum_{\sigma \in S_n} (-1)^\sigma a^1_{\sigma(1)} \cdots a^n_{\sigma(n)}. \tag{83}$$

Proposição 30. *Se A é uma matriz $n \times n$, então $\det(A) = \det(A^\top)$.*

Demonstração. De fato, basta usar que $(-1)^\sigma = (-1)^{\sigma^{-1}}$ e temos

$$\det(A^\top) = \sum_{\sigma \in S_n} (-1)^\sigma a_1^{\sigma(1)} \cdots a_n^{\sigma(n)} \tag{84}$$

$$= \sum_{\sigma \in S_n} (-1)^{\sigma^{-1}} a^1_{\sigma^{-1}(1)} \cdots a^n_{\sigma^{-1}(n)} \tag{85}$$

$$= \det(A). \tag{86}$$

□

Teorema 31. *Valem as seguintes propriedades:*

1. *Se $a_j = (a^1_j, \ldots, a^n_j)$ e $a^i = (a^i_1, \ldots, a^i_n)$, então podemos enxergar* det *como uma função das colunas a_j ou das linhas a^i. Para toda $\sigma \in S_n$,*

$$\det(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) = (-1)^\sigma \det(a_1, \ldots, a_n) \quad e \quad \det(a^{\sigma(1)}, \ldots, a^{\sigma(n)}) = (-1)^\sigma \det(a^1, \ldots, a^n); \tag{87}$$

2. *Se duas colunas ou duas linhas de A forem iguais, então* $\det(A) = 0$;

3. *O* $\det$ *é multilinear, isso é, para todos* i, j,

$$\det(a_1, \ldots, a_{j-1}, \lambda a_j + b_j, a_{j+1}, \ldots, a_n) = \lambda \det(a_1, \ldots, a_n) + \det(a_1, \ldots, a_{j-1}, b_j, a_{j+1}, \ldots, a_n) \quad (88)$$

e

$$\det(a^1, \ldots, a^{i-1}, \lambda a^i + b^i, a^{i+1}, \ldots, a^n) = \lambda \det(a^1, \ldots, a^n) + \det(a^1, \ldots, a^{i-1}, b^i, a^{i+1}, \ldots, a^n); \quad (89)$$

4. $\det(\mathrm{Id}_{n\times n}) = 1$;

5. *Se* $a_1, \ldots, a_n$ *ou* $a^1, \ldots, a^n$ *forem linearmente dependentes, então* $\det(A) = 0$.

Demonstração. 1. Se $\sigma \in S^n$, então

$$\begin{aligned}
\det(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) &= \sum_{\rho\in S_n} (-1)^\rho a^1_{\rho(\sigma(1))} \cdots a^n_{\rho(\sigma(n))} &(90)\\
&= (-1)^\sigma \sum_{\rho\in S_n} (-1)^\rho (-1)^\sigma a^1_{\rho(\sigma(1))} \cdots a^n_{\rho(\sigma(n))} &(91)\\
&= (-1)^\sigma \sum_{\rho\in S_n} (-1)^{\rho\circ\sigma} a^1_{\rho(\sigma(1))} \cdots a^n_{\rho(\sigma(n))} &(92)\\
&= (-1)^\sigma \det(a_1, \ldots, a_n). &(93)
\end{aligned}$$

O resultado para a permutação de linhas sai do fato de que $\det(A) = \det(A^\top)$;

2. Se $a_i = a_j$, seja σ a permutação dada por $\sigma(i) = j$, $\sigma(j) = i$ e que fixa todo o resto. Temos

$$\det(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n) = -\det(a_1, \ldots, a_j, \ldots, a_i, \ldots, a_n) \quad (94)$$

e portanto, como $a_i = a_j$, temos $\det(A) = -\det(A)$, assim $\det(A) = 0$. O resultado para o caso em que duas linhas são iguais sai do fato de que $\det(A) = \det(A^\top)$;

3. De fato, temos

$$\begin{aligned}
\det(a^1, \ldots, \lambda a^i + b^i, \ldots, a^n) &= \sum_{\sigma\in S_n} (-1)^\sigma a^1_{\sigma(1)} \cdots (\lambda a^i_{\sigma(i)} + b^i_{\sigma(i)}) \cdots a^n_{\sigma(n)} &(95)\\
&= \lambda \sum_{\sigma\in S_n} (-1)^\sigma a^1_{\sigma(1)} \cdots a^i_{\sigma(i)} \cdots a^n_{\sigma(n)} + \sum_{\sigma\in S_n} (-1)^\sigma a^1_{\sigma(1)} \cdots b^i_{\sigma(i)} \cdots a^n_{\sigma(n)} &(96)\\
&= \lambda \det(a^1, \ldots, a^n) + \det(a^1, \ldots, b^i, \ldots, a^n). &(97)
\end{aligned}$$

O resultado análogo para o caso da multilinearidade nas colunas sai do fato de que $\det(A) = \det(A^\top)$;

4. Temos $a^i_j = 0$ se $i \neq j$, e a única prmutação que fixa todos os valores é a identidade, que tem sinal 1, portannto

$$\det(\mathrm{Id}_{n\times n}) = a^1_1 \cdots a^n_n = 1; \quad (98)$$

5. Se $a_i = \lambda^1 a_1 + \cdots + \lambda^{i-1} a_{i-1} + \lambda^{i+1} a_{i+1} + \cdots + \lambda^n a_n$, então

$$\begin{aligned}
\det(a_1, \ldots, a_i, \ldots, a_n) &= \det(a_1, \ldots, \lambda^1 a_1 + \cdots + \lambda^{i-1} a_{i-1} + \lambda^{i+1} a_{i+1} + \cdots + \lambda^n a_n, \ldots, a_n) &(99)\\
&= \sum_{j=1, j\neq i}^{n} \lambda^j \det(a_1, \ldots, a_j, \ldots, a_n) &(100)
\end{aligned}$$

e, como cada a_j é uma cópia de outra coluna, todos os determinantes dentro do somatório são nulos, assim $\det(A) = 0$. O resultado análogo para a dependência linear das linhas sai do fato de que $\det(A) = \det(A^\top)$.

□

Essas propriedades são importantes para fazer contas, porém, elas são mais importantes ainda pois podem ser utilizadas para definir o determinante!

Proposição 32. *As propriedades* 1, 3 *e* 4 *definem unicamente o determinante.*

Demonstração. Queremos mostrar que qualquer função $D(a_1, \ldots, a_n)$ que seja multilinear, alternada e que tenha valor 1 na base canônica de $\mathbb{R}^n$, é o determinante. Note que as propriedades 2 e 5 também valém para qualquer D desse tipo, já que são consequências das outras.

Como $a_j = (a_j^1, \ldots, a_j^n) = a^1 e_i + \cdots + a^n e_n$ onde e_i é a base canônica, então usando as propriedades 3 e 4 temos

$$D(a_1, \ldots, a_n) = \sum_{i_1=1}^{n} a_1^{i_1} D(e_{i_1}, a_2, \ldots, a_n) \tag{101}$$

$$= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_1^{i_1} a_2^{i_2} D(e_{i_1}, e_{i_2}, a_3, \ldots, a_n) \tag{102}$$

$$= \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} a_1^{i_1} \cdots a_n^{i_n} D(e_{i_1}, \ldots, e_{i_n}) \tag{103}$$

e, denotando por $\sigma_{i_1,\ldots,i_n}$ a permutação $\sigma(j) = i_j$ temos

$$D(e_{i_1}, \ldots, e_{i_n}) = (-1)^{\sigma_{i_1,\ldots,i_n}} D(e_1, \ldots, e_n) = (-1)^{\sigma_{i_1,\ldots,i_n}}. \tag{104}$$

Porém, como todas as combinações de $i_1, \ldots, i_n$ são atingidas nos somatórios, $\sigma_{i_1,\ldots,i_n}$ eventualmente se passa por todas as permutações, assim

$$D(a_1, \ldots, a_n) = \sum_{\sigma \in S_n} (-1)^{\sigma} a_1^{\sigma(1)} \cdots a_n^{\sigma(n)} = \det(A^{\top}) = \det(A). \tag{105}$$

$\square$

O próximo passo é mostrar que o determinante é multiplicativo. Para isso introduzimos uma notação. Se A é uma matriz $n \times n$ e $t \in \mathbb{R}$, então $A + t$ é definido como $A + t\,\mathrm{Id}_{n\times n}$.

Teorema 33. *Se A e B são matrizes $n \times n$, então* $\det(AB) = \det(A)\det(B)$.

Demonstração. Suponha que $\det(B) \neq 0$ e defina $D(A) = \det(AB)/\det(B)$. Seja $e_1, \ldots, e_n$ é a base canônica de $\mathbb{R}^n$. Se $f_i \in \mathbb{R}^n$, então existe uma matriz $n \times n$ tal que $f_i = Ce_i$. Assim,

$$D(Ae_1, \ldots, \lambda Ae_i + Ce_i, \ldots, Ae_n) = \frac{\det(ABe_1, \ldots, (\lambda A + C)Be_i, \ldots, ABe_n)}{\det(B)} \tag{106}$$

$$= \frac{\lambda \det(ABe_1, \ldots, ABe_n) + \det(ABe_1, \ldots, CBe_i, \ldots, ABe_n)}{\det(B)} \tag{107}$$

$$= \lambda D(Ae_1, \ldots, Ae_n) + D(Ae_1, \ldots, Ce_i, \ldots, Ae_n), \tag{108}$$

da onde segue a multilinearidade. Mais ainda,

$$D(\mathrm{Id}_{n\times n}) = \det(\mathrm{Id}_{n\times n} B)/\det(B) = \det(B)/\det(B) = 1 \tag{109}$$

e

$$D(Ae_{\sigma(1)}, \ldots, Ae_{\sigma(n)}) = \det(AB_{\sigma(1)}, \ldots, ABe_{\sigma(n)})/\det(B) \tag{110}$$

$$= (-1)^{\sigma} \det(ABe_1, \ldots, ABe_n)/\det(B) \tag{111}$$

$$= (-1)^{\sigma} D(Ae_1, \ldots, Ae_n), \tag{112}$$

portanto segue que $D(A) = \det(A)$ e assim $\det(AB) = \det(A)\det(B)$. Se $\det(B) = 0$, tome $B(t) = B + t$. Temos que $\det(B(t))$ é um polinômio não nulo (afinal, é mônico), portanto $\det(AB(t)) = \det(A)\det(B(t))$ e, tomando $t \to 0$, o resultado segue. $\square$

Corolário 34. *Uma matriz quadrada A é inversível se, e somente se,* $\det(A) \neq 0$.

Demonstração. Se A é inversível, então existe B com $AB = \mathrm{Id}$, assim $\det(A)\det(B) = \det(\mathrm{Id}) = 1$, portanto é impossível que $\det(A)$ seja nulo. Por outro lado, se A não é inversível, então A não é sobrejetora, portanto sua imagem, o espaço gerado pelas suas colunas, é próprio, da onde segue que suas colunas são linearmente dependentes e assim o seu determinante é nulo. $\square$

Deixo aqui mais duas fórmulas que podem ser úteis para o cálculo de determinantes e de inversas, sem as demonstrações. se A é uma matriz $n \times n$, denotamos por A^i_j a matriz $(n-1)\times(n-1)$ dada pela remoção da i-ésima linha e da j-ésima coluna de A. Valem as seguintes identidades:

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a^i_j \det(A^i_j) \quad \text{e} \quad A^{-1} = \left[(-1)^{i+j}\frac{\det(A_{ji})}{\det(A)}\right]_{n\times n}. \tag{113}$$

Para terminarmos esse capítulo, vamos falar do traço. Se A é uma matriz $n \times n$, seu **traço** é definido por

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a^i_i. \tag{114}$$

Fica claro da definição que o traço é linear.

Proposição 35. *Se A e B são matrizes $n \times n$, $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.*

Demonstração. Temos

$$\mathrm{tr}(AB) = \sum_{i=1}^{n}\sum_{k=1}^{n} a^i_k b^k_i = \sum_{k=1}^{n}\sum_{i=1}^{n} a^k_i b^i_k = \mathrm{tr}(BA) \tag{115}$$

$\square$

Proposição 36. *Se A é uma matriz $n \times n$, então*

$$\mathrm{tr}(AA^\top) = \sum_{i=1}^{n}\sum_{j=1}^{n} (a^i_j)^2 \tag{116}$$

Demonstração. De fato,

$$\mathrm{tr}(AA^\top) = \sum_{i=1}^{n}\sum_{j=1}^{n} a^i_j a^i_j = \sum_{i=1}^{n}\sum_{j=1}^{n} (a^i_j)^2 \tag{117}$$

$\square$

Agora recordamos a noção de similaridade, mas para matrizes. Se M e T são matrizes $n \times n$, dizemos que M é **similar** a T se $M = STS^{-1}$ para alguma matriz S inversível.

Proposição 37. *Matrizes similares possuem mesmo traço e determinante.*

Demonstração. Se $M = STS^{-1}$, então

$$\det(M) = \det(STS^{-1}) = \det(S^{-1}ST) = \det(T) \quad \text{e} \quad \mathrm{tr}(M) = \mathrm{tr}(STS^{-1}) = \mathrm{tr}(S^{-1}ST) = \mathrm{tr}(T). \tag{118}$$

$\square$

O próximo passo agora é falar sobre mudanças de coordenadas. Se $e = (e_1, \ldots, e_n)$ e $f = (f_1, \ldots, f_n)$ são bases de V, então podemos escrever

$$e_j = \sum_{i=1}^{n} a^i_j f_i. \tag{119}$$

A matriz $A = [a^i_j]_{n\times n}$ leva a base f na base e (escrevemos $e = Af$). Note que A é inversível, afinal, caso seu determinante fosse nulo, os vetores e_j seriam linearmente dependentes. Agora, se $v[e] = (v^1, \ldots, v^n)$ e $v[f] = (u^1, \ldots, u^n)$, então

$$v = \sum_{j=1}^{n} v^j e_j = \sum_{j=1}^{n} v^j \sum_{i=1}^{n} a^i_j f_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a^i_j v^j\right) f_i, \tag{120}$$

portanto $v[f] = Av[e]$ e assim $v[e] = A^{-1}v[f]$. A matriz A portanto é chamada de **matriz de mudança da base e para a base f**. Note que a mesma matriz que leva a base f na base e, leva as coordenadas de v na base e para as coordenadas de v na base f, isso é:

"as coordenadas de um vetor mudam contra a mudança de base".

Por esse motivo, dizemos que vetores são **quantidades contravariantes**. Por outro lado, considerando as bases duais ε e φ de e e f, se $\mu[\varepsilon] = (\mu_1, \ldots, \mu_n)$ então

$$\mu(v) = \sum_{i=1}^{n} \mu_i \varphi^i(v) = \sum_{i=1}^{n} \mu_i u^i = \sum_{i=1}^{n} \mu_i \sum_{j=1}^{n} a^i_j v^j = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a^i_j \mu_i \right) \varepsilon^j(v), \tag{121}$$

da onde segue que $\mu[\varepsilon] = A\mu[\varphi]$, ou seja,

"as coordenadas de um covetor mudam a favor da mudança de base".

Por esse motivo, dizemos que vetores são **quantidades covariantes**. Índices em quantidades contravariantes sempre são colocados embaixo, e os índices em quantidades covariantes sempre são colocados em cima. Índices em coordenadas sempre são colocados ao contrário (coordenadas de vetores tem índice em cima, e coordenadas de covetores tem índice em baixo).

Se $T\colon V \to W$ é linear, então fixe e e f bases de V e x e y bases de W. Vamos entender como se da a transformação de $[T]_{f,y}$ em $[T]_{e,x}$. Se

$$Te_j = \sum_{i=1}^{m} \lambda^i_j x_i \quad \text{e} \quad Tf_j = \sum_{i=1}^{m} \mu^i_j y_i, \tag{122}$$

então seja $A = [a^i_j]_{n\times n}$ a matriz de mudança de base de e para f e $B = [b^i_j]_{m\times m}$ a matriz de mudança de base de y para x, então

$$Tf_i = \sum_{r=1}^{n} \mu^r_i y_r = \sum_{k=1}^{n} \sum_{r=1}^{n} \mu^r_i b^k_r x_k \tag{123}$$

e assim

$$Te_j = \sum_{i=1}^{n} a^i_j Tf_i = \sum_{k=1}^{n} \sum_{r=1}^{n} \sum_{i=1}^{n} b^k_r \mu^r_i a^i_j x_k. \tag{124}$$

Isso mostra que $[T]_{e,x} = B[T]_{f,y}A$. Tomando $W = V$, $x = e$ e $y = f$, então $B = A^{-1}$, assim $[T]_{e,e}$ é similar a $[T]_{f,f}$. Portanto, podemos definir o **determinante** de T como sendo o determinante da matriz de T em alguma base, e o conceito está bem definido, pois matrizes similares possuem o mesmo determinante. O mesmo vale para o traço.

1.3 Estrutura Euclidiana, Formas Bilineares e Quadráticas

Estrutura Euclidiana

A partir de agora, tomamos $\mathbb{K} = \mathbb{R}$, ou seja, todo escalar é real. Uma **estrutura Euclidiana** em um espaço vetorial V é um mapa

$$\begin{aligned} \langle \cdot, \cdot \rangle \colon V \times V &\to \mathbb{R} \\ (v, u) &\mapsto \langle v, u \rangle \end{aligned} \tag{125}$$

que satisfaz:

- para todo $v \in V$ com $v \neq 0$, $\langle v, v \rangle > 0$ - **positividade** (dizemos que a função é **positivo-definida**);
- para todos $u, v \in V$, $\langle v, u \rangle = \langle u, v \rangle$ - **simetria**;
- para todos $u, v, w \in V$ e $\lambda \in \mathbb{R}$, $\langle \lambda v + u, w \rangle = \lambda \langle v, w \rangle + \langle u, w \rangle$ - **linearidade** na primeira entrada, que junto com a simetria se torna **bilinearidade**.

Um mapa dessa forma é chamado de **produto interno**, e um espaço com uma estrutura Euclidiana é um **espaço Euclidiano**. A **norma** de um vetor v é o número $||v|| = \sqrt{\langle v, v \rangle}$. A norma representa a distância de v a 0, portanto a **distância** entre u e v é definida por $||u - v||$. Os próximos dois teoremas são chamados de desigualdade de Cauchy-Schwarz e desigualdade triangular.

Teorema 38. *Dados $u, v \in V$, temos $|\langle v, u \rangle| \leq ||v|| \cdot ||u||$.*

Demonstração. Se $u = 0$, então a desigualdade é trivialmente verdadeira. Considere o mapa $q(t) = ||v+tu||^2$. Usando a bilinearidade, temos que

$$q(t) = ||v||^2 + 2t\langle v, u\rangle + t^2||u||^2. \tag{126}$$

Tome $t = -\langle v, u\rangle/||u||^2$ e temos

$$q(t) = ||v||^2 - \frac{\langle v, u\rangle^2}{||u||^2} \geq 0, \tag{127}$$

da onde segue o resultado. □

Teorema 39. *Dados $u, v \in V$, temos $||v + u|| \leq ||v|| + ||u||$.*

Demonstração. Temos

$$||v + u||^2 = ||v||^2 + 2\langle v, u\rangle + ||u||^2 \leq ||v||^2 + 2||v|| \cdot ||u|| + ||u||^2 = (||v|| + ||u||)^2. \tag{128}$$

Tirando a raíz dos dois lados, a desigualdade segue. □

Dois vetores u e v são **perpendiculares** ou **ortogonais** se $\langle v, u\rangle = 0$. Fica claro que, nesse caso, vale o teorema de pitágoras: $||v - u||^2 = ||v||^2 + ||u||^2$. Se $e_1, \ldots, e_n$ é uma base de V, dizemos que ela é **ortonormmal** se

$$\langle e_i, e_j\rangle = \delta^i_j = \begin{cases} 0, & \text{se } i \neq j, \\ 1, & \text{se } i = j. \end{cases} \tag{129}$$

Note que $||e_i|| = \sqrt{\langle e_i, e_i\rangle} = 1$. O principal resultado é que toda base pode ser transformada numa base ortonormal.

Teorema 40. *O teorema a seguir é chamado de ortonormalização de Gram-Schmidt. Dada uma base $f_1, \ldots, f_n$ de V, existe uma outra base $e_1, \ldots, e_n$ com as seguintes propriedades:*

1. *$e_1, \ldots, e_n$ é ortonormal;*
2. *e_k é uma combinação linear de $f_1, \ldots, f_k$ para todo k.*

Demonstração. Definimos $e_1 = f_1/||f_1||$. Se $e_1, \ldots, e_{k-1}$ já estiverem definidos, definimos

$$e_k = c\left(f_k - \sum_{j=1}^{k-1} c_j e_j\right) \tag{130}$$

com $c_j = \langle f_k, e_j\rangle$ e c escolhido de tal forma que $||e_k|| = 1$. Se $l < k$, então

$$\langle e_k, e_l\rangle = c\langle f_k, e_l\rangle - c\sum_{j=1}^{k-1}\langle f_k, e_j\rangle\langle e_j, e_l\rangle = (c - c)\langle f_k, e_l\rangle = 0. \tag{131}$$

O caso $l > k$ é consequência do caso $l < k$ por simetria. □

Se $e = (e_1, \ldots, e_n)$ é uma base ortonormal, $x[e] = (x^1, \ldots, x^n)$ e $y[e] = (y^1, \ldots, y^n)$, então fica claro que

$$\langle x, y\rangle = \sum_{i=1}^{n} x^i y^i \quad \text{e} \quad ||x||^2 = \sum_{i=1}^{n}(x^i)^2. \tag{132}$$

O resultado que segue é chamado de teorema de representação de Riesz (versão de dimensão finita).

Teorema 41. *Dado $\varphi\colon V \to \mathbb{R}$ um covetor, existe um $v \in V$ tal que $\varphi(u) = \langle u, v\rangle$. A associação $\varphi \mapsto u$ é um isomorfismo linear entre V e V^*.*

Demonstração. Seja $e = (e_1, \ldots, e_n)$ uma base ortonormal. Se $v[e] = (v^1, \ldots, v^n)$, então

$$\varphi(u) = \sum_{i=1}^{n} v^i \varphi(e_i) = \sum_{i=1}^{n} \langle v, e_i \rangle \varphi(e_i) = \left\langle v, \sum_{i=1}^{n} \varphi(e_i) e_i \right\rangle. \tag{133}$$

Como o segundo vetor não depende de v, fica provada a primeira parte do resultado. Se $\varphi(v) = \langle v, u \rangle$ e $\iota(v) = \langle v, w \rangle$, então

$$(\varphi + \iota)(v) = \langle v, u \rangle + \langle v, w \rangle = \langle v, u + w \rangle \quad \text{e} \quad (\lambda \varphi)(v) = \lambda \langle v, u \rangle = \langle v, \lambda u \rangle, \tag{134}$$

assim a associação $\varphi \mapsto u$ é linear. Se $u = 0$, então claramente $\varphi = 0$, assim o mapa é injetor e, como $\dim V = \dim V^*$, é um isomorfismo. □

Dado um subespaço U de V, o **complemento ortogonal** de U em V é o subespaço

$$U^{\perp} = \{ v \in V \mid \langle v, u \rangle = 0, \text{ para todo } u \in U \}. \tag{135}$$

Essa notação pode ser ambígua com a notação para o aniquilador de U, mas isso é por que eles são o mesmo espaço, através da representação de Riesz.

Proposição 42. *Se U é um subespaço de V, então $U \oplus U^{\perp} = V$.*

Demonstração. Se U' é o aniquilador, sabemos que $\dim U' = \dim U^{\perp}$, assim $\dim U^{\perp} + \dim U = \dim V$, portanto $\dim U \cap U^{\perp} = 0$, assim a soma é direta e o resultado está provado. □

Pelo resultado anterior, cada $v \in V$ se quebra unicamente em $v = x + y$ com $x \in U$ e $y \in U^{\perp}$. A **projeção** em U é o mapa $Pv = x$.

Proposição 43. *O mapa P é linear e $P^2 = P$.*

Demonstração. Se $v = x + y$ e $w = r + s$ com $x, r \in U$ e $y, s \in U^{\perp}$, então $v + w = (x + r) + (y + s)$ e, pela unicidade da decomposição, $P(v + w) = x + r = Pv + Pw$. Além disso, $\lambda v = \lambda x + \lambda y$ e, pela unicidade da decomposição $P(\lambda v) = \lambda x = \lambda P v$. Por fim, $P^2 v = P(Pv) = Px = x$, pois $x \in U$, assim $P^2 = P$. □

Teorema 44. *O vetor Pv minimiza, em U, a distância até x.*

Demonstração. Se $v = x + y$ com $x \in U$ e $y \in U^{\perp}$, dado $w \in U$ temos

$$v - w = x - w + y \tag{136}$$

e, como $x - w \in U$, então $||v - w||^2 = ||x - w||^2 + ||y||^2$. Claramente o valor $||v - w||^2$ é mínimo quando $w = x$. □

Se $T \colon V \to W$ é um mapa entre espaços Euclidianos, então considere o transposto $T^{\top} \colon W^* \to V^*$. Fazendo a identificação de W^* e V^* com W e V, temos um mapa $T^* \colon W \to V$, chamado de **adjunto** de T.

Proposição 45. *Se $v \in V$, $w \in W$ e $T \colon V \to W$ é linear, então $\langle Tv, w \rangle = \langle v, T^* w \rangle$. De fato, se $S \colon W \to V$ é linear e satisfaz $\langle Tv, w \rangle = \langle v, Sw \rangle$, então $S = T^*$.*

Demonstração. Denotamos $\varphi_x(v) = \langle v, x \rangle$. Se $u = T^* w$, então $\varphi_u = T^{\top}(\varphi_w)$, assim

$$\langle v, T^* w \rangle = \varphi_u(v) = T^{\top}(\varphi_w) v = (\varphi_w \circ T) v = \varphi_w(Tv) = \langle Tv, w \rangle. \tag{137}$$

Por fim, se S satisfaz a igualdade, então se $u = Sw$, temos $\varphi_u(v) = \varphi_w(Tv) = T^{\top}(\varphi_w) v$, assim $\varphi_u = T^{\top}(\varphi_w)$ da onde segue que $S = T^*$. □

O próximo passo é demonstrar algumas propriedades sobre o adjunto.

Proposição 46. *Se $T, S \colon V \to W$ e $R \colon W \to U$ são lineares, então*

$$(T + S)^* = T^* + S^*, , \quad (T^*)^* = T \quad \text{e} \quad (RT)^* = T^* R^*. \tag{138}$$

Além disso, se T for um isomorfismo, então $(T^{-1})^* = (T^*)^{-1}$.

Demonstração. Note que

$$\langle (T+S)v, u\rangle = \langle Tv, u\rangle + \langle Sv, u\rangle = \langle v, T^*u\rangle + \langle v, S^*u\rangle = \langle v, (T^*+S^*)u\rangle. \tag{139}$$

Além disso,

$$\langle T^*u, v\rangle = \langle v, T^*u\rangle = \langle Tv, u\rangle = \langle u, Tv\rangle \tag{140}$$

e também temos

$$\langle RTv, u\rangle = \langle TvR^*u\rangle = \langle v, T^*R^*u\rangle. \tag{141}$$

Por fim,

$$TT^{-1} = \mathrm{Id}_W \implies (T^{-1})^*T^* = \mathrm{Id}_W, \tag{142}$$

concluindo o resultado desejado. □

Vamos agora achar a matriz da transformação de Riesz. Se e é uma base ortonormal de V e ε é sua base dual, então se R_V é a representação de Riesz, temos que

$$R_V\varepsilon^j = e_j \tag{143}$$

assim a matriz de R_V (com respeito a essas bases) é a identidade! Segue que a matriz do mapa adjunto, com respeito a duas bases ortogonais, assim como do mapa transposto, é a matriz transposta do mapa original.

Se V e W são espaços Euclidianos, um mapa $T\colon V \to W$ **preserva distâncias** se

$$||T(v) - T(u)|| = ||v-u|| \tag{144}$$

para todos $v, u \in V$. Mapas que preserva distâncias são sempre injetores. Um mapa bijetor que preserva distâncias é uma **isometria** entre V e W, e nesse caso dizemos que V e W são **isométricos**. Uma **translação** é um mapa da forma $T_a(v) = v + a$ para algum $a \in V$ fixado. Se S preserva distâncias, então se $a = -S(0)$ temos que $T = T_a \circ S$ também preserva distâncias e $T(0) = 0$.

Proposição 47. *Seja $T\colon V \to V$ um mapa que preserva distâncias e tal que $T(0) = 0$. Então, T é linear, $T^*T = \mathrm{Id}_V$, T é inversível, sua inversa é uma isometria e $\det(T) = \pm 1$.*

Demonstração. Sabemos que $||T(v) - T(u)|| = ||v - u||$ para todos $u, v \in V$, assim tomando $u = 0$ temos que $||T(v)|| = ||v||$. Assim,

$$||T(v)||^2 - 2\langle T(v), T(u)\rangle + ||T(u)||^2 = ||v||^2 - 2\langle v, u\rangle + ||u||^2 \implies \langle T(v), T(u)\rangle = \langle v, u\rangle. \tag{145}$$

Dessa maneira,

$$\begin{aligned}
||T(v+u) - T(v) - T(u)|| &= ||T(v+u)||^2 + ||T(v)||^2 + ||T(u)||^2 \\
&\quad - 2\langle T(v+u), T(v)\rangle - 2\langle T(v+u), T(u)\rangle + 2\langle T(v), T(u)\rangle && (146)\\
&= ||v+u||^2 + ||v||^2 + ||u||^2 - 2\langle v+u, v\rangle - 2\langle v+u, u\rangle + 2\langle v, u\rangle && (147)\\
&= ||v+u-v-u|| = 0, && (148)
\end{aligned}$$

assim $T(v+u) = T(v) + T(u)$. Além disso,

$$\begin{aligned}
||T(\lambda v) - \lambda T(v)|| &= ||T(\lambda v)||^2 - 2\langle T(\lambda v), \lambda T(v)\rangle + ||\lambda T(v)||^2 && (149)\\
&= ||\lambda v||^2 - 2\lambda\langle T(\lambda v), T(v)\rangle + \lambda^2||T(v)||^2 && (150)\\
&= ||\lambda v||^2 - 2\langle \lambda v, \lambda v\rangle + ||\lambda v||^2 = 2||\lambda v||^2 - 2||\lambda v||^2 = 0, && (151)
\end{aligned}$$

assim $T(\lambda v) = \lambda T(v)$. Segue que T é linear. Como é injetora, é um isomorfismo. Além disso,

$$||T^{-1}(v) - T^{-1}(u)|| = ||T(T^{-1}(v)) - T(T^{-1}(u))|| = ||v-u||, \tag{152}$$

portanto T^{-1} é uma isometria. Por fim,

$$\langle v, u\rangle = \langle Tv, Tu\rangle = \langle v, T^*Tu\rangle. \tag{153}$$

Como v é arbitrário, $T^*Tu = u$, assim $T^*T = \mathrm{Id}_V$. Por fim, seja e uma base ortonormal. Então,

$$\det(T)^2 = \det([T]_{e,e})^2 = \det([T]_{e,e}^\top[T]_{e,e}) = \det([T^*]_{e,e}[T]_{e,e}) = \det([T^*T]_{e,e}) = \det(\mathrm{Id}_{n\times n}) = 1, \tag{154}$$

assim $\det(T)^2 = 1$, portanto $\det(T) = \pm 1$. □

Proposição 48. *Se $T\colon V \to V$ é linear e $T^*T = \mathrm{Id}$, então T é uma isometria.*

Demonstração. Temos que

$$||Tv - Tu||^2 = \langle Tv - Tu, Tv - Tu\rangle = \langle T(v-u), T(v-u)\rangle \tag{155}$$
$$= \langle v-u, T^*T(v-u)\rangle = \langle v-u, v-u\rangle = ||v-u||^2, \tag{156}$$

portanto $||Tv - Tu|| = ||v-u||$. □

Se M é uma matriz $n \times n$ que satisfaz $M^\top M = \mathrm{Id}$, dizemos que M é ortogonal. Fica claro que toda isometria pode ser representada, em uma base ortonormal, por uma matriz ortogonal. Por outro lado, a Proposição anterior mostra que toda matriz ortogonal define uma isometria $\mathbb{R}^n \to \mathbb{R}^n$. O conjunto de todas as matrizes ortogonais $n \times n$ é denotado por $O(n)$ e chamade de **grupo ortogonal**. As matrizes de $O(n)$ com determinante 1 formam o **grupo ortogonal especial**, denotado por $SO(n)$. O nome ortogonal é bem justificado.

Proposição 49. *As colunas de uma matriz ortogonal são ortonormais (como vetores de $\mathbb{R}^n$). O mesmo vale para suas linhas.*

Demonstração. Se $e_1, \dots, e_n$ são as colunas de M, então são as linhas de $M^\top$. Se c^i_j é uma entrada de $M^\top M$, sabemos que $c^i_j = \langle e_i, e_j\rangle$ (aqui, o produto interno é o usual de $\mathbb{R}^n$, em que a base canônica é ortonormal). Porém, como $c^i_j = \delta^i_j$, segue que $e_1, \dots, e_n$ são ortonormais. □

Formas Bilineares e Quadráticas

Uma **forma bilineear** é um mapa $A\colon V \times V \to \mathbb{K}$ que é linear em cada entrada, isso é,

$$A(\lambda v + u, w) = \lambda A(v,w) + A(u,w) \quad \text{e} \quad A(v, \lambda w + u) = \lambda A(v,w) + A(v,u). \tag{157}$$

Se $v, u \in \mathbb{R}^n$ e M é uma matriz $n \times n$, fica claro que, considerando v, u como matrizes $n \times 1$, o mapa $A(v,u) = v^\top M u$ é uma forma bilinear. Agora, se $v = (v^1, \dots, v^n)$ e $u = (u^1, \dots, u^n)$ e e é a base canônica de $\mathbb{R}^n$, então

$$A(v,u) = A\left(\sum_{i=1}^n v^i e_i, \sum_{j=1}^n u^j e_j\right) = \sum_{i=1}^n \sum_{j=1}^n v^i A(e_i, e_j) u^j. \tag{158}$$

A matriz $G = [A(e_i,e_j)]_{n\times n}$ é a **matriz da forma bilinear** A e fica claro que $A(v,u) = v^\top G u$. Se essa matriz for simétrica, a forma bilinear é simétrica, afinal $A(u,v) = u^\top G v = u^\top G^\top v = v^\top G u = A(v,u)$. Por fim, dizemos que uma matriz G é positivo-definida se a forma bilinear resultante é positivo-definida. Assim, segue o resultado.

Proposição 50. *Uma forma bilinear em $\mathbb{R}^n$ é um produto interno se, e somente se, sua matriz é simétrica e positivo-definida.*

O caso geral, quando a forma bilinear está definida sobre um espaço vetorial qualquer, depende de base. Se e é uma base de V, $v[e] = (v^1, \dots, v^n)$ e $u[e] = (u^1, \dots, u^n)$, então a matriz de A, dada por $G_e = [A(e_i,e_j)]_{n\times n}$ depende da base e $A(v,u) = v[e]^\top G_e u[e]$. O resultado análogo vale no caso geral.

Proposição 51. *Uma forma bilinear é um produto interno se, e somente se, uma de suas matrizes (e por consequência todas elas), é simétrica e positivo-definida.*

Se f é outra base de V, seja M a matriz de mudança da base e para a base f. Temos $v[f] = Mv[e]$ e $u[f] = Mu[e]$. Assim,

$$A(v,u) = v[f]^\top G_f u[f] = v[e]^\top M^\top G_f M u[e], \tag{159}$$

assim $G_e = M^\top G_f M$.

Dada A uma forma bilinear simétrica, a associação $Q(v) = A(v,v)$ é uma **forma quadrática**. Dizemos que A é a **forma polar** de Q.

Proposição 52. *A forma polar está unicamente determinada pela forma quadrática.*

Demonstração. Como A é bilinear e simétrica, então $A(x+y, x+y) = A(x,x) + A(y,y) + 2A(x,y)$, então

$$A(x,y) = \frac{Q(x+y) - Q(x) - Q(y)}{2}. \tag{160}$$

□

Uma forma quadrática é **positivo-definida** se $Q(x) > 0$ sempre que $x \neq 0$. Fica claro que uma forma quadrática é positivo-definida se, e somente se, sua forma polar é positivo-definida.

Proposição 53. *Se e é uma base de V e M é uma matriz inversível, então existe uma base f de V tal que M é a matriz de mudança da base e para a base f.*

Demonstração. Se $M = [m^i_j]_{n\times n}$, então definimos

$$f_j = m^1_j e_1 + \cdots + m^n_j e_n. \tag{161}$$

Vamos verificar que é uma base, verificando a independência linear. Considere a combinação linear

$$\lambda^1 f_1 + \cdots + \lambda^n f_n = 0. \tag{162}$$

Podemos reescrever essa expressão como

$$(m^1_1 \lambda^1 + \cdots + m^1_n \lambda^n) e_1 + \cdots + (m^n_1 \lambda^1 + \cdots + m^n_n \lambda^n) e_n = 0. \tag{163}$$

Pela independência linear, isso é equivalente a um sistema homogêneo cujas incógnitas são $\lambda^1, \ldots, \lambda^n$ e M é matriz dos coeficientes. Como M é inversível, a solução é única e portanto $\lambda^1 = \cdots = \lambda^n = 0$, mostrando que os f_i formam base. □

Corolário 54. *Se $T\colon V \to \mathbb{R}$ é linear e $T(v^1, \ldots, v^n)$ é a representação de T usando uma base e de V, se definimos novas coordenadas $w^1, \ldots, w^n$ em função linear e inversível de $v^1, \ldots, v^n$, então existe uma base f de V tal que $T(w^1, \ldots, w^n)$ é a representação de T usando as cooordenadas dadas por f.*

Demonstração. Como os w^j estão em função linear dos v^i, podemos escrever

$$w^j = a^j_1 v^1 + \cdots + a^j_n v^n. \tag{164}$$

Dizer que a mudança é inversível é equivalente a dizer que a matriz $A = [a^i_j]$ é inversível, e assim pela Proposição anterior existe uma base f tal que A é a mudança da base e para a base f. Ou seja, dado $v \in V$, $v[f] = Av[e]$, isso é, $v[f] = (w^1, \ldots, w^n)$ e o resultado está provado. □

Antes de prosseguirmos, definimos que uma matriz $A = [a^i_j]_{n\times n}$ é **diagonal** se $a^i_j = 0$ sempre que $i \neq j$.

Teorema 55. *Se Q é uma forma quadrática em V (de dimensão n), então existe uma base de V tal que, nas coordenadas dessa base,*

$$Q(v) = \lambda_1 (v^1)^2 + \cdots + \lambda_n (v^n)^2 \tag{165}$$

para $\lambda_i \in \mathbb{K}$.

Demonstração. Considere uma representação em coordenadas de Q:

$$Q(v) = \sum_{r=1}^{n} \sum_{s=1}^{n} v_r g^r_s v^s \tag{166}$$

(aqui, $v_i = v^i$ e a mudança de lugar no índice é feita pois em coordenadas $Q(v) = v^\top G v$). Podemos supor que existe k tal que $g^k_k \neq 0$. Afinal, se não existe, então necessariamente algum $g^i_j \neq 0$, e assim podemos considerar a mudança de coordenadas

$$v^i = w^i + w^j, \quad v^j = w^i - w^j \quad \text{e} \quad v^k = w^k, k \neq i, j, \tag{167}$$

que é claramente dada por uma mudança de base, já que é linear e inversível. Então, temos

$$2 v_i g^i_j v^j = 2 g^i_j ((v_i)^2 - (w^j)^2) \tag{168}$$

e, como esses são os únicos termos em que aparecem $(v_i)^2$ e $(v^j)^2$ (já que $g_i^i = g_j^j = 0$), então seus coeficientes são não nulos e chegamos na conclusão desejada.

Agora assumindo que algum g_k^k é não nulo, após uma reordenação da base podemos assumir que $g_1^1 \neq 0$. Agora, juntamos os termos em que aparece v^1 ou v^1:

$$v_1 g_1^1 v^1 + 2v_1 g_2^1 v^2 + \cdots + 2v_1 g_n^1 v^n \tag{169}$$

e, em seguida, completamos o quadrado para obter

$$v_1 g_1^1 v^1 + 2v_1 g_2^1 v^2 + \cdots + 2v_1 g_n^1 v^n = \frac{1}{g_1^1}(g_1^1 v^1 + \cdots + g_n^1 v^n)^2 - B, \tag{170}$$

onde B é formado por somas e multiplicações dos termos $g_2^1 v^2, \ldots, g_n^1 v^n$. Assim, podemos fazer essa substiuição em Q para obtermos

$$Q(v) = \frac{1}{g_1^1}(g_1^1 v^1 + \cdots g_n^1 v^n)^2 + \sum_{r=2}^{n} \sum_{s=2}^{n} v_r h_s^r v^s, \tag{171}$$

em que os coeficientes h_s^r são os coeficientes g_s^r modificados pela adição de B. Agora, fazemos a mudança de coordenadas

$$w^1 = g_1^1 v^1 + \cdots + g_n^1 v^n \quad \text{e} \quad w^k = v^k, k \neq 1. \tag{172}$$

Fica claro que essa mudança é linear e inversível pois $g_1^1 \neq 0$, assim vem de uma mudança de bases e portanto podemos realizá-la sem problemas. Nessa nova mudança, temos

$$Q(v) = \frac{1}{g_1^1}(w^1)^2 + \sum_{r=2}^{n} \sum_{s=2}^{n} w_r h_s^r w^s \tag{173}$$

Se a soma restante em Q for nula, o resultado está provado. Caso contrário, repetimos o processo anterior (sempre mantendo o primeiro vetor da base fixado durante as mudanças de variáveis, para garantir que o trabalho anterior não seja alterado), e eventualmente chegaremos no formato desejado. $\square$

Fica claro que, após uma mudança de coordenadas razoável, podemos supor que os coeficientes λ_i no Teorema anterior são todos ± 1 ou 0. Dessa forma, fica evidente a próxima pergunta: o número de coeficiennte positivos, negativos e nulo, é sempre fixo independente da mudança de base?

Lema 56. *Se U e W são subespaços de V e $\dim U + \dim W > \dim V$, então existe $v \neq 0$ com $v \in U \cap W$.*

Demonstração. De fato,

$$\dim U \cap W = \dim U + \dim W - \dim V > 0. \tag{174}$$

$\square$

O teorema a seguir é chamado de lei da inércia de Sylvester.

Teorema 57. *Os números de coeficientes positivos e negativos em duas formas canônicas de uma forma quadrática sempre são os mesmos.*

Demonstração. Considere duas bases e e f de V e as representações em coordenadas Q nessas duas bases, respectivamente:

$$\begin{aligned} Q(v) = (v_1^1)^2 + \cdots + (v_1^p)^2 - (v_1^{p+1})^2 + \cdots + (v_1^{p+q})^2 \quad \text{e} \\ Q(v) = (w^1)^2 + \cdots + (w^r)^2 - (w^{r+1})^2 + \cdots + (w^{r+s})^2. \end{aligned} \tag{175}$$

Queremos mostrar que $p = r$ e $q = s$. Suponha, por absurdo, que $p > r$. Sejam U o subespaço gerado por $e_1, \ldots, e_p$ e W o subespaço gerado por $f_{r+1}, \ldots, f_n$. Temos

$$\dim U + \dim W = p + n - r > n, \tag{176}$$

assim existe $x \in U \cap W$. Porém, fica claro que $Q(x) > 0$ (quando olhamos para a representação na base e) e $Q(x) \leq 0$ (quando olhamos para a representação na base e). Concluímos então que $p \leq r$. Porém, o argumento contrário diz que $r \leq p$, assim $p = r$. Um argumento similar mostra que $q = s$ e o resultado está provado. $\square$

O **posto** de uma forma quadrática é o número de coeficientes não nulos em uma forma canônica. Pelo teorema anterior, esse número está bem definido. O **núcleo** de uma forma bilinear A é o conjunto de todos os vetores perpendiculares a todos os outros vetores, isso é,

$$\ker A = \{w \in V \mid A(v, w) = 0 \text{ para todo } v \in V\}. \tag{177}$$

Fica claro que o posto de uma forma quadrática pode ser definido como o posto de uma de suas matrizes na forma canônica, afinal, elas são diagonais e o número de linhas não nulas é precisamente o número de coeficientes não nulos na forma canônica.

Bônus: Espaços Complexos

A partir de agora, $\mathbb{K} = \mathbb{C}$. Um mapa $T: V \to W$ é **antilinear** se $T(\lambda v + w) = \overline{\lambda}T(v) + T(w)$. Uma **forma sesquilinear** é um mapa

$$\begin{aligned} A: V \times V &\to \mathbb{K} \\ (v, u) &\mapsto \langle v, u\rangle \end{aligned} \tag{178}$$

que é antilinear na primeira entrada e linear na segunda entrada. Um **produto interno** em V é uma forma bilinear $\langle -, - \rangle$ positivo-definida e que satisfaz a **condição Hermitiana**, isso é,

$$\langle u, v\rangle = \overline{\langle v, u\rangle}. \tag{179}$$

Perceba que se estamos sobre os reais, essa condição é precisamente a simetria. Além disso, temos $\langle u, u\rangle = \overline{\langle u, u\rangle}$, assim $\langle u, u\rangle \in \mathbb{R}$ e portanto a condição de positividade faz sentido. Assim como no caso real, um **espaço Euclidiano** é um espaço vetorial equipado com um produto interno. As noções de ortogonalidade e de base ortonormal são análogas ao caso real. Além disso, o processo de ortonormalização de Gram-Schmidt funciona também em espaços complexos. Se $e_1, \ldots, e_n$ é uma base ortonormal, $v[e] = (v^1, \ldots, v^n)$ e $u[e] = (u^1, \ldots, u^n)$, então

$$\langle v, u\rangle = \overline{v_1}u^1 + \cdots + \overline{v_n}u^n \tag{180}$$

(trocamos a direção dos índices para o vetor v, como é padrão para falar de formas bilineares, quadráticas e produtos internos).

Considerando agora uma forma sesquilinear qualquer A, em coordenadas podemos, assim como no caso real, expressá-la como $A(v, u) = v^*Mu$ (aqui, v^* é a matriz linha obtida ao transpor v e conjugar todas as suas entradas) e dizemos que M é a **matriz de** A na base que determina as coordenadas. Definimos similarmente uma **forma quadrática** a partir de uma forma sesquilinear por $Q(v) = A(v, v)$. Da mesma maneira que antes, chamamos A de **forma polar** de Q, e esta está unicamente determinada por Q através da fórmula

$$A(v, u) = \frac{Q(v + u) + iQ(v + iu) - Q(v - u) - iQ(v - iu)}{4} \tag{181}$$

Se a forma polar A é Hermitiana, então Q só possui valores reais. Lembramos que, no caso real, se e e f são bases do espaço, então

$$A(v, u) = v[e]^\top G_e v[e] = v[f]^\top M^\top G_e M v[f] \tag{182}$$

e assim $G_f = M^\top G_e M$ em que M é a matriz de mudança da base f para a base e. Um procedimento análogo funciona no caso complexo, porém, nesse caso temos

$$A(v, u) = v[e]^* G_e u[e] = v[f]^* M^* G_e M u[f] \tag{183}$$

e assim $G_f = M^* G_e M$. Assim como no caso real, valem os teoremas de forma canônica e da inércia de Sylvester, com demonstrações muito parecidas.

Teorema 58. *Seja A uma forma sesquilinear Hermitiana e considere sua forma quadrática Q. Então existe uma base de V na qual, em coordenadas, a forma biilinear é dada por*

$$Q(v) = \lambda_1|v^1|^2 + \cdots + \lambda_n|v^n|^2, \tag{184}$$

com $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

Teorema 59. *O número de coeficientes positivos e negativos em uma forma canônica de uma forma quadrática independe da base.*

1.4 Teoria Espectral

Teoria Espectral

Nessa seção, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, a menos que seja dito o contrário em alguma situação específica. Seja $T\colon V \to V$ um mapa linear. Um **autovetor** de T é uma direção fixada, isso é, um vetor $v \in V$ não nulo tal que $Tv = \lambda v$ para algum $\lambda \in \mathbb{K}$. O valor λ é chamado de **autovalor** de T correspondente a v.

Proposição 60. *Todo operador linear em um espaço complexo possui autovetor.*

Demonstração. Seja $v \in V$ com $v \neq 0$ e considere os vetores

$$v, Tv, T^2v, \dots, T^nv \tag{185}$$

onde $n = \dim V$. Como $n+1$ vetores em um espaço de dimensão n sempre são linearmente dependentes, existe uma combinação linear não trivial

$$\sum_{j=0}^{n} c_j T^j v = 0. \tag{186}$$

Se $p(t) = c_0 + c_1 t + \cdots + c_n t^n$, reescrevemos a combinação linear por $p(T)v = 0$. Se $k = \deg(p(t))$, então podemos fatorar

$$p(t) = c_k(t - a_1)\cdots(t - a_k) \tag{187}$$

e assim

$$p(T)v = c_k(T - a_1\,\mathrm{Id})\cdots(T - a_k\,\mathrm{Id})v = 0. \tag{188}$$

Como $v \neq 0$, então $p(T)$ não é injetor, assim existe algum j com $T - a_j\,\mathrm{Id}$ não injetor, da onde segue que existe w não nulo tal que $(T - a_j\,\mathrm{Id})w = 0$. Fica claro que w é autovetor de T com autovalor a_j. $\square$

Vamos agora demonstrar o mesmo fato utilizando um resultado muito mais poderoso.

Proposição 61. *Seja $T\colon V \to V$ linear (V pode ser real ou complexo). Então, $\lambda \in \mathbb{K}$ é autovalor se, e somente se, for raíz do polinômio*

$$p(\lambda) = \det(T - \lambda\,\mathrm{Id}). \tag{189}$$

*Chamamos p de **polinômio característico** de T.*

Demonstração. Se λ é autovalor, então existe $v \neq 0$ tal que $(T - \lambda\,\mathrm{Id})v = 0$, assim $T - \lambda\,\mathrm{Id}$ não é injetor, portanto não é invertível e assim $\det(T - \lambda\,\mathrm{Id}) = 0$. Por outro lado, se $\det(T - \lambda\,\mathrm{Id}) = 0$, então $T - \lambda\,\mathrm{Id}$ não é injetor e assim existe $v \in V$ não nulo com $(T - \lambda\,\mathrm{Id})v = 0$, assim λ é autovalor. $\square$

Fica claro então que, no caso real, podem não existir autovalores em dimensão par, já que polinômios reais de grau par não necessariamente possuem raízes reais. No caso complexo, sempre existirão n autovalores, não necessariamente distintos.

Da mesma maneira que toda matriz real pode ser tratada como uma matriz complexa, todo espaço vetorial real da origem a um espaço vetorial complexo pelo processo de **complexificação**. Se V é um espaço vetorial real, então consideramos o espaço vetorial $V(\mathbb{C})$ que, como conjunto, é V^2, mas que é munido das operações

$$(u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2) \quad \text{e} \quad (a + bi)(u, v) = (au - bv, bu + av). \tag{190}$$

Se $e_1, \dots, e_n$ é base de V, então $(e_1, 0), \dots, (e_n, 0)$ é base de $V(\mathbb{C})$. De fato, se

$$(a^1 + ib^1)(e_1, 0) + \cdots + (a^n + ib^n)(e_n, 0) = (0, 0) \tag{191}$$

então

$$(a^1 e_1, b^1 e_1) + \cdots + (a^n e_n, b^n e_n) = (0, 0) \tag{192}$$

e assim

$$(a^1 e_1 + \cdots + a^n e_n, b^1 e_1 + \cdots + b^n e_n) = (0, 0) \tag{193}$$

portanto $a^i = b^i = 0$ para todo i pela independência linear dos e_i. Se $w = (u, v) \in \mathbb{C}$, então podemos escrever

$$u = \lambda^1 e_1 + \cdots + \lambda^n e_n \quad \text{e} \quad v = \eta^1 e_1 + \cdots + \eta^n e_n, \tag{194}$$

assim

$$(\lambda^1 + i\eta^1)(e_1, 0) + \cdots + (\lambda^n + i\eta^n)(e_n, 0) = (u, v), \tag{195}$$

portanto os vetores $(e_i, 0)$ formam base. Isso da origem ao principal resultado, em termos de teoria espectral, da complexificação:

Proposição 62. *Se $T\colon V \to V$ é linear, então a extensão natural*

$$\begin{aligned} T(\mathbb{C})\colon V(\mathbb{C}) &\to V(\mathbb{C}) \\ (u,v) &\mapsto (Tu, Tv) \end{aligned} \tag{196}$$

é linear. Mais que isso, se $e = \{e_1, \ldots, e_n\}$ é uma base e $e(\mathbb{C}) = \{(e_1, 0), \ldots, (e_n, 0)\}$, então $[T]_{e,e} = [T(\mathbb{C})]_{e(\mathbb{C}),e(\mathbb{C})}$.

Demonstração. A linearidade é óbvia. A parte importante aqui é a segunda. Seja

$$Te_j = \sum_{i=1}^{n} a^i_j e_i. \tag{197}$$

Então

$$T(\mathbb{C})(e_j, 0) = (Te_j, 0) = \left(\sum_{i=1}^{n} a^i_j e_i, 0\right) = \sum_{i=1}^{n} (a^i_j e_i, 0) = \sum_{i=1}^{n} a^i_j (e_i, 0), \tag{198}$$

da onde segue o resultado. □

O teorema acima é a versão "sem coordenadas" do fato de que toda matriz real pode ser considerada como uma matriz complexa. Se V é um espaço complexo, definimos $V(\mathbb{C}) = V$.

Lema 63. *Os autovalores reais de T são autovalores de $T(\mathbb{C})$, com as mesmas multiplicidades algébricas.*

Demonstração. Os polinômios característicos coincidem, afinal, o determinante não depende da escolha da base e, como vimos, existem bases em que as matrizes de T e $T(\mathbb{C})$ coincidem. □

Proposição 64. *Se $T\colon V \to V$ é linear (V pode ser real ou complexo), então $\det(T)$ é o produto dos autovalores de $T(\mathbb{C})$ e $\operatorname{tr} T$ é a soma dos autovalores de $T(\mathbb{C})$.*

Demonstração. Seja $p(\lambda)$ o polinômio característico de $T(\mathbb{C})$ (e também de T), então se $A = [a^i_j]$ é a matriz de T para alguma base, temos

$$p(\lambda) = \sum_{\sigma \in S_n} (-1)^\sigma (a^1_{\sigma(1)} - \lambda\delta^1_{\sigma(1)}) \cdots (a^n_{\sigma(n)} - \lambda\delta^n_{\sigma(n)}). \tag{199}$$

Dessa maneira, o termo livre de $p(\lambda)$ é

$$p(0) = \sum_{\sigma \in S_n} (-1)^\sigma a^1_{\sigma(1)} \cdots a^n_{\sigma(n)} = \det(A). \tag{200}$$

Mais ainda, como é impossível que apenas um $\delta^i_{\sigma(i)}$ seja nulo, então o único termo que contribui para o coeficiente de t^{n-1} é

$$(a^1_1 - \lambda) \cdots (a^n_n - \lambda) \tag{201}$$

Dessa maneira, os elementos da diagonal de A são as raízes desse polinômio e, pelas relações de Girard, segue que o coeficiente de t^{n-1} é $(-1)^{n-1}\operatorname{tr}(A)$. □

Agora, começamos com o primeiro resultado relevante para diagonalizar transformações lineares:

Proposição 65. *Autoespaços correspondentes a autovalores diferentes estão em soma direta.*

Demonstração. Vamos proceder por indução na quantidade de autovalores. Com 1 autoespaço, não há nada a provar. Se $V_1, \ldots, V_k$ são autoespaços dos autovalores $\lambda^1, \ldots, \lambda^k$ (todos diferentes), então basta mostrar que se $v_i \in V_i$ são tais que $v_1 + \cdots + v_k = 0$, então todo $v_i = 0$. De fato, se $v_1 + \cdots + v_k = 0$, então $v_1 + \cdots + v_{k-1} \in V_k$. Dessa maneira, podemos aplicar T de duas maneiras diferentes nessa soma, para obtermos

$$T(v_1 + \cdots + v_{k-1}) = \lambda^1 v_1 + \cdots + \lambda^{k-1} v_{k-1} \quad \text{e} \quad T(v_1 + \cdots + v_{k-1}) = \lambda^k v_1 + \cdots + \lambda^k v_{k-1}, \tag{202}$$

assim segue que

$$(\lambda^k - \lambda^1)v_1 + \cdots + (\lambda^k - \lambda^{k-1})v_{k-1} = 0. \tag{203}$$

Pela hipótese de indução, isso significa que $(\lambda^k - \lambda^i)v_i = 0$ para todo $i = 1, \ldots, k-1$. Commo $\lambda^k \neq \lambda^i$, temos $v_i = 0$ Dessa forma, $v_k = 0$ e o resultado segue. □

Corolário 66. *Se um mapa $T\colon V \to V$ possui $n = \dim V$ autovalores distintos, então existe uma base de autovetores.*

Proposição 67. *A matriz de $T\colon V \to V$ com respeito a uma base de autovetores é diagonal.*

Demonstração. De fato, se $e_1, \ldots, e_n$ é uma base de autovetores, então

$$Te_j = \lambda_j^j e_j \tag{204}$$

e assim a matriz de T é diagonal, com a entrada (i, i) sendo o autovalor de e_i. □

Por conta do resultado acima, dizemos que uma transformação $T\colon V \to V$ é **diagonalizável** se existe uma base de autovetores. As matrizes de T em bases diferentes são semelhantes. Por conta disso, definimos que uma matriz M é **diagonalizável** se for semelhante a uma matriz diagonal.

Com o resultado anterior, fica claro o processo de diagonalizar uma matriz: se M é uma matriz que possui base de autovetores, então o resultado anterior diz que a matriz da transformação $\mathbb{C}^n \to \mathbb{C}^n$ dada por $x \mapsto Mx$, na base de autovetores de M, é diagonal. Porém, M é a matriz dessa transformação na base canônica, portanto temos

$$P^{-1}MP = D \tag{205}$$

, em que D é diagonal (a matriz de $x \mapsto Mx$ na base de autovetores) e P é a matriz de mudança da base de autovetores para a base canônica (suas colunas são os autovetores escritos em coordenadas canônicas). O próximo passo é falar de três resultados centrais para a teoria de transformações lineares. O primeiro deles se chama Teorema do Mapeamento Espectral.

Teorema 68. *Se $q(t)$ é um polinômio e $T\colon V \to V$ é linear, então λ é um autovalor de T se, e somente se, $q(\lambda)$ é autovalor de $q(T)$. Além disso, os autoespaços de λ para T e $q(\lambda)$ para $q(T)$ coincidem.*

Demonstração. Seja $q(t) = c_n t^n + \cdots + c_1 t + c_0$. Se λ é autovalor de T, então dado $v \in V$ autovetor de T com autovalor λ temos

$$q(T)v = (c_n T^n + \cdots + c_1 T + c_0 \operatorname{Id})v = c_n T^n v + \cdots + c_1 Tv + c_0 v \tag{206}$$
$$= c_n \lambda^n v + \cdots + c_1 \lambda v + c_0 v = (c_n \lambda^n + \cdots + c_1 \lambda + c_0)v = q(\lambda)v. \tag{207}$$

Por outro lado, se λ é autovalor de $q(T)$, vamos mostrar que existe η autovalor de T com $\lambda = q(\eta)$. De fato, a transformação linear $q(T) - \lambda \operatorname{Id}$ não é invertível. Fatorando $q(t) - b$ sobre os complexos, temos

$$q(t) - \lambda = c(t - r_1) \cdots (t - r_n) \tag{208}$$

onde $n = \dim V$. Dessa maneeira, podemos aplicar essa versão fatorada em T para obtermos

$$q(T) - \lambda \operatorname{Id} = c(T - r_1 \operatorname{Id}) \cdots (T - r_n \operatorname{Id}). \tag{209}$$

Segue então que pelo menos uma das transformações $T - r_i \operatorname{Id}$ não é invertível, ou seja, que algum r_i é autovalor de T. Como r_i é raíz de $q(t) - \lambda$, segue que $q(r_i) = \lambda$ e a prova está completa. □

O próximo Teorema é conhecido como Teorema de Cayley-Hamilton. Antes de prová-lo, precisamos de um Lema.

Lema 69. *Sejam P e Q dois polinômios com coeficientes sendo transformações lineares:*

$$P(t) = \sum_{j=1}^{n} P_j t^j \quad e \quad Q(t) = \sum_{k=1}^{m} Q_k t^k. \tag{210}$$

O produto $R = PQ$ é então dado por

$$R(t) = \sum_{l=1}^{n+m} R_l t^l \quad com \quad R_l = \sum_{j+k=l} P_j Q_k. \tag{211}$$

Então, se T comuta com todos os cooeficientes de Q, temos $P(T)Q(T) = R(T)$.

Demonstração. De fato, temos

$$R(T) = \sum_{l=1}^{n+m} R_l T^l = \sum_{l=1}^{n+m} \sum_{j+k=l} P_j Q_k T^l = \sum_{l=1}^{n+m} \sum_{j+k=l} P_j T^j Q_k T^k = P(T)Q(T). \tag{212}$$

□

Teorema 70. *Se $T\colon V \to V$ é linear e $q(t)$ é seu polinômio característico, então $q(T) = 0$.*

Demonstração. Primeiro assuma que V é complexo e que T possui autovalores distintos. Então, T possui base de autovetores $e_1, \dots, e_n$ com autovalores $\eta^1, \dots, \eta^n$. Se $v = \lambda^1 e_1 + \cdots + \lambda^n e_n$, então

$$q(T)v = \lambda^1 q(\eta^1) e_1 + \cdots + \lambda^n q(\eta^n) e_n = 0, \tag{213}$$

pois $q(\eta^i) = 0$ para todo autovalor η^i. O próximo passo é generalizar o resultado para um mapa linear qualquer. Utilizando o Lema anterior, tome

$$Q(t) = T - t\,\mathrm{Id} \quad \text{e} \quad P(t) = q(t)(T - t\,\mathrm{Id})^{-1}. \tag{214}$$

Não é difícil observar que $P(t)$ é uma matriz com entradas polinomiais (de fato, $P(t)$ é uma matriz que, em suas entradas, possui múltiplos dos determinantes de algumas submatrizes de $T - t\,\mathrm{Id}$, portanto suas entradas são polinomiais, já que determinantes são polinômios). Mais ainda, T comuta com T e $-\mathrm{Id}$, assim do Lema anterior temos que

$$P(t)Q(t) = q(t)\,\mathrm{Id} \implies P(T)Q(T) = q(T), \tag{215}$$

mas $Q(T) = 0$ da onde segue que $q(T) = 0$.

Por fim, se V é um espaço real, então considere $T(\mathbb{C})\colon V(\mathbb{C}) \to V(\mathbb{C})$ e note que $T(\mathbb{C})$ é raíz do polinômio característico de T, pelo caso anterior. Assim, $q(T(\mathbb{C})) = 0$. Se $q(t) = c_n t^n + \cdots + c_1 t + c_0$ então

$$\begin{aligned} (0,0) = q(T(\mathbb{C}))(v,0) &= c_n T(\mathbb{C})^n (v,0) + \cdots + c_1 T(\mathbb{C})(v,0) + c_0(v,0) &(216)\\ &= c_n(T^n v, 0) + \cdots + c_1(Tv, 0) + c_0(v,0) = (c_n T^n v + \cdots + c_1 Tv + c_0 v, 0) = (q(T)v, 0) &(217)\end{aligned}$$

Dessa forma, $q(T)v = 0$ e assim $q(T) = 0$ pela arbitrariedade de v. □

O terceiro resultado principal, o Teorema Espectral, que encerra essa seção, depende da definição do que chamamos de **autovetor generalizado**, que é um vetor $v \neq 0$ tal que, para algum $m \in \mathbb{Z}_{>1}$, $(T - \lambda\,\mathrm{Id})^m v = 0$. O número λ é o **autovalor generalizado** de v. Os três lemas que seguem serão utilizados na demonstração do Teorema Espectral.

Lema 71. *Se p e q são polinômios com coeficientes complexos sem um zero em comum, então existem polinômios a e b tais que $ap + bq = 1$.*

Demonstração. Seja $\mathcal{I}$ o conjunto de todos os polinômios da forma $ap + bq$. Entre eles existe um polinômio não nulo de menor grau, que chamaremos de d. Note que d divide p e q, pois caso contrário, se r é o resto da divisão de p por d, então $r = p - md$ é da forma $ap + bq$, e tem grau menor que d, o que é uma contradição. Como d divide p e q, e eles não tem zeros comuns, então d tem grau 0. Se $d(t) = k \neq 0$, tomamos $d' = d/k \in \mathcal{I}$ e o resultado está provado. □

Lema 72. *Sejam p e q polinômios com coeficientes complexos sem um zeero em comum, então se $T\colon V \to V$ é uma operador linear em um espaço complexo V, temos*

$$\ker p(T) \oplus \ker q(T) = \ker p(T)q(T). \tag{218}$$

Demonstração. Sejam a e b polinômios com $ap + bq = 1$. Temos que $a(T)p(T) - b(T)q(T) = \mathrm{Id}$. Deixando ambos os lados agirem em um vetor v, obtemos

$$a(T)p(T)v - b(T)q(T)v = v. \tag{219}$$

Se $v \in \ker p(T)q(T)$, então

$$q(T)a(T)p(T)v = a(T)q(T)p(T)v = a(T)p(T)q(T)v = 0 \quad \text{e} \quad p(T)b(T)q(T)v = b(T)p(T)q(T)v = 0, \tag{220}$$

assim v se decompõe em uma soma de um elemento em $\ker q(T)$ e outro em $\ker p(T)$. Por outro lado, se $v \in \ker p(T) + \ker q(T)$, então escrevemos $v = u + w$ com $p(T)u = 0$ e $q(T)w = 0$ e temos

$$p(T)q(T)v = p(T)q(T)(u + w) = p(T)q(T)u = q(T)p(T)u = 0, \tag{221}$$

assim $v \in \ker p(T)q(T)$.

Por fim, provamos que a soma é direta. Se $u \in \ker p(T)$ e $w \in \ker q(T)$ são tais que $u + w = 0$, então $u = -w$ e

$$a(A)p(A)u + b(A)q(A)w = u \implies u = 0 \implies w = 0 \tag{222}$$

e o resultado segue. □

Lema 73. *Sejam $p_1, \ldots, p_k$ polinômios com coeficientes complexos tais que p_i e p_j não tem raíz em comum para todo i e j. Então, se $T\colon V \to V$ é um operador em um espaço complexo V, temos*

$$\ker p_1(T) \cdots p_k(T) = \ker p_1(T) \oplus \cdots \oplus \ker p_k(T). \tag{223}$$

Demonstração. Seguimos por indução em k. Se $k = 2$, o resultado está provado. Assumindo que o resultado vale para k, provemos para $k + 1$. Isso porém, segue do caso $k = 2$, afinal,

$$\ker p_1(T) \cdots p_k(T)p_{k+1}(T) = \ker p_1(T) \cdots p_k(T) \oplus \ker p_{k+1}(T) \tag{224}$$

$$= \ker p_1(T) \oplus \cdots \ker p_k(T) \oplus \ker p_{k+1}(T). \tag{225}$$

□

Temos, finalmente, o Teorema Espectral.

Teorema 74. *Se $T\colon V \to V$ é um operador em um espaço complexo, então todo vetor de v pode ser escrito como uma soma de autovetores de T, sejam eles generalizados ou não.*

Demonstração. Se $v = 0$, então dado u autovetor de T temos $v = u - u$ e não há nada mais a provar. Se $v \neq 0$ então os vetores $v, Tv, \ldots, T^n v$ (com $n = \dim V$) são linearmente dependentes, portanto existe uma combinação linear nula deles, que se escreve como um polinômio em T agindo em v: $p(T)v = 0$. Prosseguimos fatorando p

$$(T - r_1 \operatorname{Id})^{m_1} \cdots (T - r_k \operatorname{Id})^{m_k} v = 0. \tag{226}$$

Como polinômios em T comutam, podemmos remover todos os fatores invertíveis desse produto e assumir que todo r_i é autovalor de T (seja generalizado ou não). Denotando $p_i(t) = (t-r_i)^{m_i}$ temos que $p_1(T) \cdots p_k(T)v = 0$, ou seja, $v \in \ker p_1(T) \cdots p_k(T)$, portanto v se decompõe unicamente em uma soma de elementos de $\ker p_1(T), \ldots, \ker p_k(T)$. Porém, $w \in \ker p_i(T)$ se, e somente se, w é autovetor (possivelmente generalizado) de T, encerrando assim a prova. □

O teorema não vale para operadores em espaços reais já que podem não existir autovetores, nem generalizados.

Teoria Espectral em Espaços Euclidianos

A partir de agora V é um espaço Euclidiano complexo, a menos que o contrário seja dito. Vamos relembrar um resultado sobre formas quadráticas:

Teorema 75. *Se A é uma forma sesquilinear e Hermitiana (isso é, linear na primeira coordenada e tal que $A(v, u) = \overline{A(u, v)}$), então existe uma base tal que sua forma quadrática, nas coordenadas dessa base, se escreve como*

$$Q(v) = \xi^1 |v_1|^2 + \cdots + \xi^n |v_n|^2 \tag{227}$$

com $\xi^1, \ldots, \xi^n \in \mathbb{R}$.

Em particular, se H é uma matriz Hermitiana ($H = H^*$, onde H^* é a transposta conjugada), então a forma bilinear $A(u, v) = u^* H v$ é Hermitiana e portanto existe uma base em que A está na forma canônica. Como a mudança de base é dada, na matriz, por $M^* H M$, então temos a primeira forma do Teorema Espectral para Espaços Euclidianos.

Teorema 76. *Se H é Hermitiana, então existe uma matriz M invertível tal que $M^* H M$ é diagonal.*

Isso não implica, a priori, que H é diagonalizável, já que pra isso precisaríamos de $M^* = M^{-1}$, e portanto que a base que da a forma canônica de A fosse ortonormal. Será que isso é possível de conseguir?

Teorema 77. *Um mapa $H\colon V \to V$ auto-adjunto (isso é, tal que $H = H^*$) possui todos os autovalores reais e uma base ortonormal da autovetores.*

Demonstração. Sabemos, pelo Teorema Espectral, que todo vetor de V se escreve como soma de autovetores (possivelmente generalizados) de H. Basta mostrarmmos que todos esses autovetores são genuínos (não generalizados), que seus autovalores são reais e que autovetores com autovalores diferentes são ortogonais.

Note que não precisamos tratar de autovalores generalizados, afinal, se $(H - \lambda \operatorname{Id})^m v = 0$, então λ é autovalor genuíno de H com autovetor $(H - a \operatorname{Id})^{m-1} v$. Se $a + ib$ é um autovalor de H então existe $v \neq 0$ tal que $Hv = (a + ib)v$ e assim $(H - a \operatorname{Id})v = ibv$, portanto ib é autovalor de $H - a \operatorname{Id}$, que também é auto-adjunto. Porém, isso significa que

$$\langle (H - a \operatorname{Id})v, v\rangle = \langle ibv, v\rangle = ib\langle v, v\rangle. \tag{228}$$

Da definição de produto interno, sabemos que $ib\langle v, v\rangle$ é imaginário puro, mas como $H - a \operatorname{Id}$ é auto-adjunto, $\langle (H - a \operatorname{Id})v, v\rangle$ é real, o que é uma contradição a menos que $b = 0$, portanto H só possui autovalores reais.

Seja v um autovetor generalizado, ou seja, existem $\lambda \in \mathbb{R}$ e $m \in \mathbb{Z}_{>1}$ com $(H - a \operatorname{Id})^m v = 0$. Proseguimos por indução em m. Se $m = 2$, então

$$(H - a \operatorname{Id})^2 v = 0 \implies 0 = \langle (H - a \operatorname{Id})^2 v, v\rangle = \langle (H - a \operatorname{Id})v, (H - a \operatorname{Id})v\rangle, \tag{229}$$

portanto $(H - a \operatorname{Id})v = 0$ e v é autovetor genuíno de H. Agora, supondo que vale o resultado para m, provemos para $m + 1$. Se $(H - a \operatorname{Id})^{m+1} v = 0$, então

$$(H - a \operatorname{Id})^2 (H - a \operatorname{Id})^{m-1} v = 0, \tag{230}$$

portanto pelo caso anterior temos $(H - a \operatorname{Id})^m v = 0$, assim $(H - a \operatorname{Id})v = 0$ pela hipótese de indução.

Por fim, resta a ortonormalidade. Se $\lambda \neq \eta$ são autovalores de H com autovetores u e v, respectivamente, então

$$\lambda\langle u, v\rangle = \langle \lambda u, v\rangle = \langle Hu, v\rangle = \langle u, Hv\rangle = \langle u, \eta v\rangle = \eta\langle u, v\rangle \tag{231}$$

e, como $\lambda \neq \eta$, obrigatoriamente temos $\langle u, v\rangle = 0$. Dessa forma, basta tomar autovetores de norma 1 e conseguimos a base desejada. □

Corolário 78. *Toda matriz Hermitiana é diagonalizável para uma matriz real.*

Demonstração. Seja H uma matriz Hermitiana. O mapa $x \mapsto Hx$ é linear e auto-adjunto, então possui base ortonormal de autovetores. A matriz M de mudança dessa base para a base canônica é ortonormal, portanto $M^* = M^{-1}$. Mais ainda, a matriz $M^{-1}HM$ é a matriz do mapa $x \mapsto Hx$ correspondente a uma base de autovetores, e portanto é diagonal. Além disso, os elementos da diagonal são os autovalores de H, que são reais. □

O último objetivo é apresentar a versão mais geral do Teorema Espectral para Espaços Euclidianos. Uma **resolução da identidade** é uma coleção $P_1, \ldots, P_k$ de mapas tais que cada P_i é uma projeção auto-adjunta,

$$\operatorname{Id} = \sum_{i=1}^{k} P_i \quad \text{e} \quad P_i P_j = 0, \text{ se } i \neq j. \tag{232}$$

Se, além disso, para algum mapa linear $H\colon V \to V$ temos uma combinação linear

$$H = \sum_{i=1}^{k} a^i P_i \tag{233}$$

então a coleção dos mapas P_i são uma **resolução espectral** de H. Perceba que, se H for auto-adjunto, podemos pegar cada P_i como as projeções nos autoespaços e cada a^i como o autovalor correspondente.

Lema 79. *Se H possui uma resolução espectral, possui base ortonormal de autovetores.*

Demonstração. Considere uma resolução de H

$$H = \sum_{i=1}^{k} a^i P_i. \tag{234}$$

Então, dado $x \in V$, denotando $x_i = P_i x$ temos $Hx_i = a^i x_i$ pois $P_j P_i x = 0$, assim $\operatorname{im} P_i$ é autoespaço de H com autovalor a^i. Além disso, como

$$\mathrm{Id} = \sum_{i=1}^{k} P_i, \tag{235}$$

então $V = \operatorname{im} P_1 + \cdots + \operatorname{im} P_k$. Mais ainda, a soma é direta pois, se $x_i \in \operatorname{im} P_i$ e $x_1 + \cdots + x_k = 0$, aplicando P_i de ambos os lados temos $x_i = 0$. Por fim, se $v \in \operatorname{im} P_i$ e $u \in \operatorname{im} P_j$, então

$$\langle v, u\rangle = \langle P_i v, P_j u\rangle = \langle v, P_i P_j u\rangle = \langle v, 0\rangle = 0, \tag{236}$$

assim podemos construir bases ortonormais em cada $\operatorname{im} P_i$ e juntá-las para obter uma base ortonormal de autovetores para V. □

Teorema 80. *Se $H, K\colon V \to V$ são auto-adjuntos que comutam, então existe uma resolução espectral comum a H e K.*

Demonstração. Sabemos que H possui resolução espectral, pois são auto-adjuntos. Denote-as por

$$H = \sum_{i=1}^{k} a^i P_i \tag{237}$$

Seja V_i um autoespaço de H com autovalor a^i. Então, se $v \in V_i$, temos $Hv = a^i v$ e, aplicando K dos dois lados e usando a comutatividade, chegamos em $HKv = a^i Kv$, portanto $Kv \in V_i$. Dessa maneira, podemos construir a restrição $K|_{V_i}\colon V_i \to V_i$, que também é auto-adjunta e portanto possui resolução espectral. Juntando as resoluções de K restrito a cada autoespaço de H, temos uma resolução comum a H e K. □

Um mapa $H\colon V \to V$ é **anti-auto-adjunto** se $H^* = -H$. Note que nesse caso o mapa iH é auto-adjunto e portanto seus autovalores são reais e existe base ortonormal de autovetores de iH, da onde segue o teorema a seguir.

Teorema 81. *Se H é anti-auto-adjunto, seus autovalores são imaginários puros e existe base ortonormal de V formada por autovetores de H.*

Termminamos o conteúdo de álgebra linear com a última versão do Teorema Espectral para Espaços Euclidianos, que depende da noção de normalidade. Dizemos que $N\colon V \to V$ é **normal** se $NN^* = N^*N$. Perceba que todo auto-adjunto é normal.

Teorema 82. *Todo mapa normal possui base ortonormal de autovetores.*

Demonstração. Definimos

$$H = \frac{N + N^*}{2} \quad \text{e} \quad A = \frac{N - N^*}{2}. \tag{238}$$

Fica claro que H é auto-adjunto e A é anti-auto-adjunto. Mais ainda, H e $K = iA$ comutam e K é auto-adjunto, portanto existe uma resolução espectral comum a H e K, então como $N = H - iK$, essa mesma resolução espectral também é resolução de N e assim o resultado está provado. □

1.5 Referências

- Peter Lax - Linear Algebra and Its Applications - Seções 1.1, 1.2, 1.3 e 1.4;
- Israel Gelfand - Lectures on Linear Algebra - Seção 1.3;
- Ernest Borisovich Vinberg - A Course in Algebra - Seção 1.4.

2 Análise Real

2.1 Sequências e Séries em $\mathbb{R}$

Supremos e Ínfimos

Sequências

Uma **sequência** em um conjunto X é uma função $x\colon \mathbb{Z}_{>0} \to X$. É costumeiro denotar $x(n)$ por x_n. Estamos interessados em sequências em $\mathbb{R}$. Costumamos denotar o mapa x por x_n, (x_n) ou $(x_n)_{n=1}^{\infty}$. A nomenclatura usual de funções (limitada, monótona, estritamente crescente/decrescente, crescente/decrescente, constante e etc.) se aplicam também no contexto de sequências.

Corolário 1. *Como consequência das definições, x_n é limitada se, e somente se, $|x_n|$ é limitada.*

Uma **subsequência** de x_n é uma sequência y_k de maneira que para todo k existe um $n_k \in \mathbb{Z}_{>0}$ tal que $y_k = x_{n_k}$. Por conta disso, ignoramos y_k e denotamos a subsequência apenas por x_{n_k}. Fica claro que toda subsequência de uma sequência monótona (respectivamente limitada) também é monótona (respectivamente limitada).

Proposição 2. *Uma sequência x_n monótona é limitada se, e somente se, possui uma subsequência limitada.*

Demonstração. Vamos assumir que x_n é crescente, pois os outros casos são análogos. Como já observamos, toda subsequência de uma sequência limitada, é limitada. Se x_n é monótona e possui uma subsequência limitada, então $|x_{n_k}| \leq M$ para todo k em algum $M > 0$. Note que, se $m \in \mathbb{Z}_{>0}$, então existe $k \in \mathbb{Z}_{>1}$ com $n_k > m$, ou seja, pela monotonicidade, temos

$$|x_m| \leq |x_{n_k}| < M \tag{1}$$

e assim x_n é limitada. □

Uma sequência satisfaz uma propriedade **eventualmente** se a propriedade é satisfeita para todo x_n com n **suficientemente grande**, isso é, para todo x_n com $n \geq m$ para algum m fixado. Por exemplo, x_n é eventualmente crescente se existe m tal que $x_n \leq x_{n+1}$ para todo $n \geq m$. Dizemos que x_n **converge** para $x \in \mathbb{R}$ se x_n fica arbitrariamente próximo de x para n suficientemente grande, isso é, se para todo $\varepsilon > 0$ bem pequeno existe um m tal que $|x_n - x| < \varepsilon$ sempre que $n \geq m$. Nesse caso, dizemos que x é o **limite** de x_n, e denotamos isso por

$$x_n \to x, \quad \lim_{n\to\infty} x_n = x \quad \text{ou ainda} \quad \lim x_n = x. \tag{2}$$

Quando, para n arbitrariamente grande, x_n se torna arbitrariamente grande, dizemos que x_n **converge para o infinito** e escrevemos

$$x_n \to \infty, \quad \lim_{n\to\infty} x_n = \infty \quad \text{ou ainda} \quad \lim x_n = \infty. \tag{3}$$

Se a sequência $-x_n$ converge para o infinito, então escrevemos

$$x_n \to -\infty, \quad \lim_{n\to\infty} x_n = -\infty \quad \text{ou ainda} \quad \lim x_n = -\infty \tag{4}$$

e dizemos que x_n **converge para o infinito negativo**. Se existe $x \in \mathbb{R}$ com $x_n \to x$, ou x_n converge para o infinito, seja ele positivo ou negativo, dizemos que x_n **converge** ou que x_n é **convergente**.

Teorema 3. *Se $x_n \to a$ e $x_n \to b$, então $a = b$.*

Demonstração. Se $x_n \to a$, então para n suficientemente grande temos

$$|x_n - a| < |a - b|/3 \quad \text{e} \quad |x_n - b| < |a - b|/3. \tag{5}$$

Dessa maneira, vale que

$$|a - b| = |a - x_n + x_n - b| \leq |x_n - a| + |x_n - b| < 2|a - b|/3 < |a - b|, \tag{6}$$

o que é um absurdo. □

Teorema 4. *Se $x_n \to x$, então toda subsequência x_{n_k} de x_n converge para x.*

Demonstração. Dado $\varepsilon > 0$, para n suficietemente grande temos $|x_n - x| < \varepsilon$. Porém, existe k tal que $n_k > n$ e assim $n_s > n$ para todo $s \geq k$, da onde segue que $|x_{n_s} - x| < \varepsilon$. □

Corolário 5. *Se $x_n \to x$ e $k \in \mathbb{Z}_{>1}$, então $x_{n+k} \to x$.*

Demonstração. Basta notar que $x_{n_k} = x_{n+k}$ é uma subsequência de x_n. □

Teorema 6. *Toda sequência que converge a um número real é limitada.*

Demonstração. Se $x_n \to x$, dado $\varepsilon > 0$, para n suficientemente grande temos $|x_n - x| < \varepsilon$. Dessa maneira, $x_n \in (x - \varepsilon, x + \varepsilon)$. Isso significa que existe um $M > 0$ tal que, para n suficientemente grande, $|x_n| < M$. Porém, isso significa que apenas finitos x_n estouram essa cota, assim podemos tomar

$$R = \max\left\{\max_{|x_k| \geq M} |x_k|, M\right\}. \tag{7}$$

e assim $|x_n| < R + 1$ para todo n. □

Teorema 7. *Toda sequência monótona e limitada converge para um número real.*

Demonstração. Vamos assumir que x_n é crescente (os outros casos são análogos). Seja $a = \sup x_n$. Por definição, dado $\varepsilon > 0$, existe $n \in \mathbb{Z}_{>1}$ com $x_n > a - \varepsilon$. Como a sequência é crescente, então para todo $m \geq n$ temos $x_m \geq x_n > a - \varepsilon$, assim $a - x_m < \varepsilon$. Porém, como $a = \sup x_n$, $|a - x_m| = a - x_m < \varepsilon$, da onde segue que $x_n \to a$. Fica claro que, se x_n fosse decrescente (estritamente ou não) teríamos $x_n \to \inf x_n$. □

Corolário 8. *Se uma sequência monótona possui uma subsequência convergente, então converge.*

Demonstração. Se x_{n_k} converge, é limitada. Como x_n é monótona e possui uma subsequência limitada, então x_n é limitada. Como é monótona e limitada, converge. □

Teorema 9. *Se $\lim x_n = 0$ e y_n é limitada, então $\lim x_n y_n = 0$.*

Demonstração. Como y_n é limitada, existe $M > 0$ tal que $|y_n| < M$ para todo n. Dado $\varepsilon > 0$, para todo n suficientemente grande temos $|x_n| < \varepsilon/M$, assim

$$|x_n y_n| = |x_n||y_n| < M\varepsilon/M = \varepsilon, \tag{8}$$

da onde sai que $x_n y_n \to 0$. □

Proposição 10. *Se $x_n \to x$ e $y_n \to y$, então:*

- *se $x_n \leq y_n$ para todo n, então $x \leq y$;*
- *$x_n + y_n \to x + y$;*
- *$x_n y_n \to xy$;*
- *se $y_n \neq 0$ para todo n e $y \neq 0$, então $x_n/y_n \to x/y$.*

Demonstração.

- Suponha que $x > y$. Então, tome $\varepsilon = (x - y)/2 > 0$. Para todo n suficientemente grande, temos
$$|x_n - x| < (x - y)/2 \quad \text{e} \quad |y_n - y| < (x - y)/2 \tag{9}$$
para n suficientemente grande. Porém, isso significa que
- Temos
$$|(x_n + y_n) - (x + y)| = |x_n - x + y_n - y| \leq |x_n - x| + |y_n - y| \tag{10}$$
e, portanto, como ambas as parcelas da soma ficam arbitraiamente pequenas, $|x_n + y_n - x - y|$ também fica;

- Temos

$$|x_ny_n - xy| = |x_ny_n - x_ny + x_ny - xy| \tag{11}$$
$$= |x_n(y_n - y) + y(x_n - x)| \tag{12}$$
$$\leq |x_n||y_n - y| + |y||x_n - x|. \tag{13}$$

Como x_n converge, é limitada, assim ambas as parcelas ficam arbitrariamente pequenas e portanto $|x_ny_n - xy|$ também fica;

- Temos

$$\left|\frac{x_n}{y_n} - \frac{x}{y}\right| = \frac{|x_ny - xy_n|}{|y_ny|} = \frac{|x_ny - xy + xy - xy_n|}{|y_ny|} \leq \frac{|y||x_n - x| + |x||y_n - y|}{|y_n||y|} \tag{14}$$
$$= \tag{15}$$

□

3 Anéis

4 Módulos, Grupos e Corpos

5 Análise Complexa

6 Teoria da Medida

7 Análise Funcional

8 Teoria Quantitativa de EDOs

9 Teoria Básica de EDPs

10 Probabilidade

11 Topologia

12 Topologia Algébrica

13 Topologia Diferencial

14 Teoria Qualitativa de EDOs

15 Teoria Geral de EDPs

16 Análise em Variedades

17 Riemanniana

18 Álgebra Comutativa

19 Geometria Algébrica

20 Teoria de Lie

21 Métodos Numéricos

22 Otimização Linear

23 Aprendizado de Máquina