Álgebra Linear

Notas da aula[1]
MA327 2021-2

versão preliminar

Joa Weber[2]

9 de setembro de 2021

[1] versão final vai aparecer la:
www.math.stonybrook.edu/∼joa/PUBLICATIONS/00-publications.html

[2] joa@math.uni-bielefeld.de

Extensão

Totalitarismo defronte de Democracia[1]

A ferramenta chave do totalitarismo é medo; suportado de pânica e histeria.
Meias palavras, difamações, e denunciações formam a estratégia.
Não tem discussões livres, nem diferenças de opinião honestas.
Argumentos não tem mais valor – o pensamento baixo já levou a melhor.
O pensamento mesmo é o inimigo.

Democracia é o domínio da dignidade do humano
e o direito, a pensar propriamente, o direito para uma opinião própria,
ainda mais, o direito, afirmar explícito a opinião própria e
proteger-se contra a intrusão na sua psique e contra constrangimento psíquico.

Democracia solicita uma alta atividade intelectual dos seus membros.
Democracia rege a sociedade por os seus erros sem intimidação.

[1] Em parte tirado do sumário [Hir21] do livro [Mee56].

Prefácio

Este texto oriunda de notas da aula para o curso MA327 Álgebra Linear dado múltiplas vezes, isto é, nos semestres 2013-2, 2014-1, 2015-1, 2016-2, 2019-2, 2020-2, e 2021-2. Quando cheguei em agosto 2013 fui assinado ensinar este curso e como meu Português foi (ainda mais) fraco, na realidade quase inexistente, eu busquei um livro em Português mesmo. Encontrei um texto excelente, o livro do Elon Lages Lima [Lim11], no qual eu baseei meu manuscrito em 2013-2. Nos semestres na seguida eu usei como base meu manuscrito de 2013, adicionando e melhorando uma ou outra coisa. Só em 2020-2 quando a aula foi mandado ser remota eu digitei meus manuscritos no computador resultando no texto presente o qual em grandes partes ainda é um condensado do Lima. Um outro texto excelente o qual eu recomendo particularmente para alunos de disciplinas de estudo não matemáticas é o Pulino [Pul12].

O que é diferente neste texto é na seção 5 a abordagem a matrizes de transformações lineares como ilustrado no diagrama (5.0.1).
O texto na forma atual ainda precisa ser estendido ca e la, o que vai acontecer no semestre 2021-2 e no futuro.

Agradecimentos

É um prazer agradecer os pagadores de imposto do Brasil para as condições excelentes de pesquisar e até 2019 de ensinar.

Espero que o Deus protege todos os cidadãos de todas as nações
e leve a responsabilidade todos fomentando medo e gerando pânica.

Campinas, *Joa Weber*
Semestre 2021-2

Sumário

Introdução **1**
Notações . . . 3
Convenções . . . 4

I Teoria dos espaços vetoriais **5**

1 Espaços vetoriais **7**
1.1 Axiomas . . . 7
1.1.1 Grupo . . . 8
1.1.2 Corpo . . . 9
1.1.3 Espaço vetorial . . . 13
1.2 Exemplos . . . 14
1.2.1 Listas ordenadas . . . 14
1.2.2 Matrizes . . . 16
1.2.3 Funções e polinômios . . . 17
1.3 Independência linear . . . 18
1.3.1 Combinação linear . . . 18
1.3.2 Linearmente independente . . . 18

2 Subespaços **23**
2.1 Definição e exemplos . . . 23
2.2 Sistemas de geradores . . . 25
2.3 Soma direta . . . 27

3 Bases **29**
3.1 Aplicações . . . 32
3.1.1 Coordenadas de um vetor . . . 32
3.1.2 Dimensão de um espaço vetorial . . . 33
3.1.3 Escalonamento: Dimensão de um subespaço gerado . . . 36
3.1.4 Complexificação e realificação . . . 37
3.2 Existência e extensão . . . 37

II Teoria das transformações lineares 41

4 Transformações lineares 43
4.1 Exemplos e construção . . . 43
4.1.1 O espaço vetorial das transformações lineares . . . 45
4.1.2 Isomorfismos . . . 46
4.1.3 Construção de transformações lineares . . . 46
4.1.4 O espaço dual . . . 48
4.2 Matrizes . . . 49
4.3 Dimensão dois – o plano . . . 53
4.3.1 Rotações . . . 54
4.3.2 Projeção ortogonal sobre uma reta . . . 57
4.3.3 Reflexão em torno de uma reta . . . 58

5 Matrizes de transformações lineares 61
5.1 Bases induzem isomorfismos . . . 62
5.2 A matriz em respeito a uma base . . . 63
5.3 Mudança de base – comutatividade da diagrama . . . 65
5.3.1 Como o vetor coordenada muda . . . 65
5.3.2 Como a matriz de uma TL muda . . . 66
5.3.3 Determinante de uma transformação linear . . . 68
5.4 Exercícios e umas soluções . . . 68

6 Núcleo e imagem 73
6.1 Escalonamento: Cálculo do posto . . . 75
6.2 Sobrejetividade – inversa à direita . . . 76
6.3 Injetividade – inversa à esquerda . . . 77
6.4 Bijetividade – inversa . . . 79
6.4.1 Isomorfismos . . . 79
6.5 Teorema de núcleo e imagem . . . 82

7 Soma direta e projeções 85
7.1 Projeções . . . 87
7.2 Involuções . . . 88
7.3 Exercícios . . . 91

8 Eliminação e aplicações (repetição de MA141) 93
8.1 Resolução de sistemas lineares . . . 93
8.2 Exercícios e umas soluções . . . 95

9 Subespaços invariantes – autovalores/vetores 97
9.1 Autovalores e autovetores . . . 98
9.2 Polinômio característico . . . 102
9.3 Existência – caso real . . . 104
9.4 Exercícios . . . 105

III Estruturas adicionais e operadores especiais 107

10 Produto interno 109
10.1 Produto interno, norma, distância . . . 109
10.1.1 Produto interno e espaço dual – dualidade . . . 112
10.1.2 Produto interno e matrizes . . . 114
10.2 Ortogonalidade . . . 116
10.2.1 Projeção ortogonal sobre uma reta . . . 117
10.3 Ângulos e cumprimentos em $(\mathbb{R}^2, \langle\cdot,\cdot\rangle_0)$ . . . 117
10.4 Desigualdades . . . 118
10.5 Ortonormalização – processo de Gram-Schmidt . . . 118
10.5.1 Existência e extensão de bases ortogonais . . . 120
10.5.2 Projeção ortogonal sobre um subespaço . . . 120
10.6 Complemento ortogonal . . . 121
10.7 Exercícios e umas soluções . . . 122

11 A adjunta 125
11.1 Definição e propriedades . . . 125
11.2 Fórmula para inversa à direita/esquerda . . . 129
11.3 Traço – produto interno em $\mathcal{L}(E, F)$ . . . 129
11.4 Operadores normais . . . 130
11.5 Exercícios . . . 130

12 Operadores auto-adjuntos 133
12.1 Auto-adjunto e ortogonalidade . . . 134
12.2 Matrizes simétricas . . . 134
12.3 Teorema espectral . . . 136
12.4 Operadores não-negativos . . . 138
12.5 Teorema dos valores singulares (operadores gerais) . . . 141
12.6 Exercícios . . . 141

13 Operadores ortogonais 143
13.1 Matrizes ortogonais . . . 143
13.2 Operadores ortogonais . . . 145
13.2.1 Forma normal . . . 146
13.3 Decomposição polar . . . 146
13.4 Exercícios . . . 147

14 Produto hermitiano 149
14.1 Definições . . . 149
14.2 Adjunta complexa $A^\dagger$ . . . 151
14.2.1 Operadores normais $A^\dagger A = AA^\dagger$ . . . 151
14.2.2 Operadores hermitianos $A^\dagger = A$ . . . 152
14.2.3 Operadores unitários $A^\dagger = A^{-1}$ . . . 152

IV Apêndices **153**

A MA141 – Revisão de matrizes e sistemas lineares **155**
A.1 Matrizes . . . 155
A.2 Escalonamento de matrizes segundo Gauss . . . 157
A.2.1 Aplicações de escalonamento . . . 159
A.3 Sistemas Lineares . . . 159
A.4 Cálculo da matriz inversa – Gauss-Jordan . . . 162
A.4.1 O determinante de matrizes quadradas . . . 162
A.5 Exercícios . . . 163

B Demonstrações restantes **165**
B.1 Espaços vetoriais . . . 165
B.2 Subespaços . . . 167
B.3 Bases – SLH . . . 167
B.4 Transformações lineares . . . 168
B.5 Existência de subespaço invariante ($\mathbb{K} = \mathbb{R}$) . . . 170
B.6 Operadores ortogonais . . . 172

Referências Bibliográficas **173**

Índice Remissivo **175**

Introdução

Àlgebra Linear

é o estudo dos espaços lineares e das transformações lineares.

Uma outra palavra para espaço linear é espaço vetorial.

O próximo exemplo descreve como os conceitos de vetor e espaço vetorial foram descobertos - como flechas no plano onde se considera igual todos da mesma direção e do mesmo comprimento. A identificação "considerar igual" lida a uma enorme importância e poder do conceito na física (campos de forças).

Exemplo 0.0.1 (O espaço vetorial F das flechas equipolentes no plano)**.** Seja F o conjunto das flechas no plano Π, onde *consideramos iguais* duas flechas se

(i) as duas flechas são paralelas,

(ii) têm o mesmo sentido de percurso, e

(iii) têm o mesmo comprimento.

Chama-se duas flechas satisfazendo (i-iii) de **flechas equipolentes**, símbolo $\sim$. Chamamos de **vetores** os elementos de F. Assim um vetor é uma flecha v onde consideramos iguais todas flechas equipolente a v, todas as quais denotamos de v também, e cada uma chamamos de um **representante** do vetor.[2]

Seja F munido de duas operações, primeiro, multiplicar uma flecha v com um número real $\alpha \in \mathbb{R}$ e, segundo, adicionar duas flechas v e w. Veja Figura 1.

Multiplicação (escalar). Pela definição αv é a flecha na direção de v cujo comprimento é α vezes aquele de v (muda-se a direção caso o número α é negativo).

Adição (vetorial). Pela definição $v + w$ é a flecha cujo ponto inicial é aquela de v e cujo ponto termino p é obtido depois fazer uma translação de w movendo o ponto inicial de w no ponto termino de v. Então p é definido como o ponto termino do novo w.

O conjunto F munido das duas operações é um espaço vetorial sobre o corpo $\mathbb{R}$. Um exemplo de uma transformação linear em F é uma rotação $r_\theta : F \to F$ que vira cada uma flecha v pelo angulo θ em torno do ponto inicial.

[2] Esta definição dos elementos de F é informal e só serve para nossa introducao. Uma excelente referência básica para a definição de F como um conjunto composto de coleções de flechas equipolentes é [Lim05].

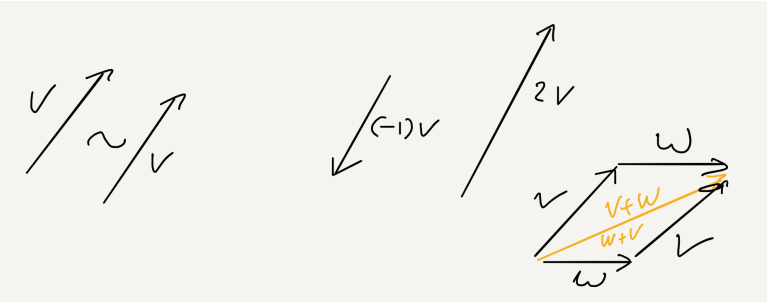


Figura 1: Flechas consideradas iguais, multiplicação escalar, e adição

Exemplo 0.0.2 (Pares de números reais)**.** Seja $\mathbb{R}^2 := \{(x, y) \mid x, y \in \mathbb{R}\}$ o conjunto de todas listas ordenadas de dois membros reais munido de adição membro-por-membro e de multiplicação com um número real $\alpha \in \mathbb{R}$, também membro-por-membro. Então $\mathbb{R}^2$ é um espaço vetorial sobre o corpo $\mathbb{R}$.

Comentário 0.0.3 (Identificação dos conjuntos e operações – isomorfismo)**.** Os dois exemplos anteriores são "iguais" no sentido seguinte. Suponhamos que na reta podemos medir a distância 1. No plano Π escolha um **eixo** OX, ou seja uma reta com dois pontos diferentes O e X da distância 1, e um segundo eixo OY cujo primeiro ponto O é aquele do OX e qual intersecta OX no ponto O só (equivalentemente $OX \neq OY$). Uma tal escolha de dois eixos é chamado um **sistema de coordenadas** no plano, símbolo OXY. Veja Figura 2.

Observe-se que um eixo OX chega com uma direção (de O para X) e com um comprimento unitário (o comprimento do segmento entre O e X). Uma escolha de coordenadas OXY no plano Π nos da uma aplicação

$$F \to \mathbb{R}^2, \quad v \mapsto (x, y) \tag{0.0.1}$$

a qual identifica os elementos de F com os elementos de $\mathbb{R}^2$ unicamente (bijetora) – e ainda é **linear**: compatível com as duas operações no domínio e

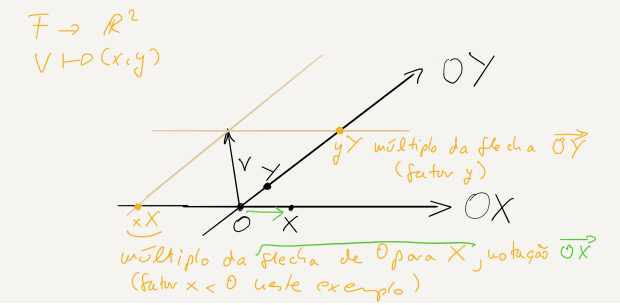


Figura 2: Sistema de coordenadas OXY composto de dois eixos OX e OY

as duas no contra-domínio. Uma tal aplicação (bijetora linear) é chamado um **isomorfismo** entre espaços vetoriais. Deixamos ao leitor definir esta aplicação.[3]

Escolhendo dois eixos *ortogonais* René Descartes (1596-1650) introduziu em 1637 a identificação (0.0.1) a qual é assim chamado de **sistema de coordenadas Cartesianas**; veja [Koe85, p. 7].

Exemplo 0.0.4 (Funções contínuas e integração)**.** Sejam $a < b$ dois números reais. Então o quadruplo $V = (C^0([a,b],\mathbb{R}),+,\cdot,\mathbb{R})$ que é composto do conjunto das funções contínuas $f : [a,b] \to \mathbb{C}$ munido com as duas operações de adicionar $f+g$ duas funções e multiplicar αf uma função com um número real $\alpha \in \mathbb{R}$ é um espaço vetorial sobre o corpo $\mathbb{R}$.

Também $W = (\mathbb{R},+,\cdot,\mathbb{R})$ composto dos números reais $\mathbb{R}$ e munido das operações óbvias é um espaço vetorial sobre o corpo $\mathbb{R}$.

Integração $T : V \to W$, $f \mapsto \int_a^b f(x)\,dx$, *é compatível com*, as duas adições e as duas multiplicações (em V e em W) no sentido que

$$T(f+g) = Tf + Tg, \qquad T(\alpha f) = \alpha T f$$

para todos os vetores $f, g \in V$ e escalares α do corpo $\mathbb{R}$. Uma aplicação T entre espaços vetoriais qual respeita as duas operações no domínio e no contra-domínio é chamada de **transformação linear**.

Notações

Para uma lista extensiva dos símbolos usados veja o Índice Remissivo B.8.

Comentário 0.0.5 (Números)**.** Vamos trabalhar com os seguintes **números**

$\mathbb{N} := \{1,2,3,\dots\}$, $\mathbb{N}_0 := \{0,1,2,\dots\}$	naturais
$\mathbb{Z} := \{\dots,-2,-1,0,1,2,\dots\}$	inteiros
$\mathbb{Q} := \{\frac{p}{q} \mid p \in \mathbb{Z},\, q \in \mathbb{N}\}$	racionais
$\mathbb{R} := (-\infty,\infty)$ "a reta real"	reais
$\mathbb{C} := \{a+ib \mid a,b \in \mathbb{R}\}$ "o plano complexo"	complexos

Com $|\alpha|$ denotamos o absoluto de um número real α. Denotamos **intervalos** fechados de $[a,b] \subset \mathbb{R}$ e abertos de $(a,b) \subset \mathbb{R}$. Usamos os símbolos

$$\forall \text{ "para todos os"} \qquad \exists \text{ "existe um"} \qquad \exists! \text{ "existe um único"}$$

A notação $A := B$ significa que A **é definido** pelo lado direito B. Escrevendo $\dim E = n$ ou $\{\xi_1,\dots,\xi_k\}$ indica sem ser mencionado explicitamente que n e k são números naturais, particularmente têm valor **finito**.

"Sejam $x_1,\dots,x_\ell$ elementos de um conjunto X" é uma frase encontrada frequentemente e depois quer-se trabalhar com o conjunto composto destes elementos. Um ponto útil é que não é proibido que uns dos elementos, ainda

[3] Dica: Os pontos O, X e O, Y dão duas flechas. Represente um elemento v de F por uma flecha equipolente com ponto inicial O. Pensa num paralelogramo tal que O e o ponto termino da flecha equivalente são dois vértices opostos.

todos, são iguais. O jeito certo de escrever o conjunto correspondente é assim $\{x_1\}\cup\cdots\cup\{x_\ell\}$. Para este conjunto usa-se também a notação $\{x_i \mid i = 1, \ldots, \ell\}$. Veja Definição 1.1.2.

Convenções

É comum usar para o mesmo conceito às vezes terminologia diferente, por exemplo **transformação linear** e **operador linear**, ou ainda só **operador**, denota todo o mesmo conceito.

Cor cinza. Parágrafos e maiores partes de texto em cinza indicam matéria avançada direcionado às turmas A e B do "cursão", mas não às outras turmas. Palavras individuais em cinza geralmente são nomes ou informações complementares.

Parte I

Teoria dos espaços vetoriais

Capítulo 1

Espaços vetoriais

1.1 Axiomas

Definição 1.1.1. Um **conjunto** X é composto de elementos os quais são dois-a-dois diferentes. Então não faz sentido escrever expressões da forma ~~$\{2,3,2\}$~~. Um conjunto não é ordenado, por exemplo $\{1,2\} = \{2,1\}$. A **união** de dois conjuntos A e B é o conjunto $A \cup B$ cujos elementos pertencem a A ou a B. Por exemplo

$$\{2,3\} \cup \{2\} = \{2,3\} = \{3,2\} \tag{1.1.1}$$

A **interseção** de dois conjuntos A e B é o conjunto $A \cap B$ composto de todos os elementos que pertencem a ambos A e B. Um conjunto chama-se **ordenado** se seus elementos são enumerados, por exemplo $X = \{x_1, \ldots, x_n\}$. O conjunto que não contem nenhum elemento é chamado **o conjunto vazio**, símbolo $\emptyset$. Usamos a notação $A \dot{\cup} B$ para transferir a informação adicional que os dois conjuntos A e B são **disjuntos**, ou seja não tem nenhum elemento comum, em símbolos $A \cap B = \emptyset$. Denotamos de $|X|$ o **número de elementos de um conjunto** quando o número é finito. Neste caso X é chamado de **conjunto finito**.

Um **subconjunto** de um conjunto X é um conjunto A tal que cada um elemento de A é elemento de X, notação $A \subset X$. Observe que conforme esta definição, o conjunto vazio $\emptyset$ é subconjunto de todos conjuntos: para todo conjunto X temos $\emptyset \subset X$.

Definição 1.1.2. *"Sejam $x_1, \ldots, x_\ell$ elementos de um conjunto X"* é uma frase encontrada frequentemente e depois quer-se trabalhar com o conjunto composto destes elementos. Um ponto sutil é que não é proibido que uns dos elementos, ainda todos, são iguais. Mas conforme nossa convenção para denotar conjuntos, veja Definição 1.1.1, a notação $\{x_1, \ldots, x_\ell\}$ só faz sentido, e é permitida, quando os elementos são dois-a-dois diferente. A notação certa, junta com sua abreviação, para **o conjunto contendo os elementos $x_1, \ldots, x_\ell$ de X** é

$$\{x_1\} \cup \ldots \{x_\ell\} =: \{v_i \mid i = 1, \ldots, \ell\}$$

Uma escolha arbitraria forma o conjunto $\cup_{\lambda \in \Lambda} \{x_\lambda\} =: \{x_\lambda \mid \lambda \in \Lambda\}$.

Definição 1.1.3. O **produto cartesiano** $X \times Y$ de dois conjuntos X e Y é o conjunto de todas listas ordenadas (x, y) dos elementos $x \in X$ e $y \in Y$, ou seja

$$X \times Y := \{(x, y) \mid x \in X,\, y \in Y\}$$

Observe que se um fator fica vazio, ou seja $X = \emptyset$ ou $Y = \emptyset$, então $X \times Y = \emptyset$. Abreviamos

$$Y^{\times k} := Y \times \cdots \times Y \tag{1.1.2}$$

se na direita temos k fatores.

1.1.1 Grupo

Definição 1.1.4. Um conjunto não-vazio $G \neq \emptyset$ munido de uma operação

$$* : G \times G \to G, \quad (f, g) \mapsto f * g$$

é chamado um **grupo**, notação $(G, *)$, se valem os três axiomas

1. $f*(g*h) = (f*g)*h$ para todos os elementos $f, g, h \in G$ (associatividade)
2. existe um elemento $e \in G$ tal que (elemento neutro)

   $$e * g = g, \qquad g * e = g$$

   para todos os elementos $g \in G$.
3. para todo $g \in G$ existe um elemento $h \in G$ t.q. (inverso)

   $$g * h = e, \quad h * g = e$$

   Chama-se h de **inverso** do elemento g e denota-se h com o símbolo $\bar{g}$.

Em palavras,

> *um grupo é um conjunto não-vazio munido de uma operação associativa, contendo um elemento neutro, e tal que cada um elemento admite um inverso.*

O seguinte lema diz que um grupo G tem exatamente um elemento neutro, notação comum e, e cada um elemento g de G tem exatamente um inverso, notação $\bar{g}$. Ás vezes é comum e útil escrever o elemento neutro na forma 0 ou 1 e os inversos na forma $-g$ ou g^{-1}; veja os dois exemplos em Exercício 1.1.7 a).

Lema 1.1.5. *Seja $(G, *)$ um grupo. Então vale o seguinte.*
1) *O elemento neutro é único.*
2) *Os elementos inversos são únicos.*
3) *Dado elementos $f, g, h \in G$, então vale:*
   a) $f * g = f * h \;\Rightarrow\; g = h$ *(lei da corte)*
   b) $f * g = f \;\Rightarrow\; g = e$
   c) $f * g = e \;\Rightarrow\; g = \bar{f}$

Note que b) e c) são consequências imediatas de a).

Demonstração. Lema B.1.1. □

Definição 1.1.6. Um grupo $(G, *)$ é chamado de **abeliano** se a ordem dos dois elementos na operação não importa, em símbolos $f*g = g*f$. (comutatividade)

Exercício 1.1.7. Mostre que
a) são grupos (ainda abelianos): $(\mathbb{Z}, +)$ e $(\mathbb{R}, \cdot)$
b) não são grupos: $(\mathbb{N}, +)$ e $(\mathbb{N}_0, +)$ e $(\mathbb{Z}, \cdot)$
c) não são grupos abelianos: – as matrizes 3×3 – as rotações em $\mathbb{R}^3$.

1.1.2 Corpo

Definição 1.1.8. Um conjunto $\mathbb{K}$ munido de duas operações[1]

$$+ : \mathbb{K} \times \mathbb{K} \mapsto \mathbb{K} \qquad \cdot : \mathbb{K} \times \mathbb{K} \mapsto \mathbb{K}$$

é chamado um **corpo** se valem os três axiomas

1. $(\mathbb{K}, +)$ é um grupo abeliano.
   (O *elemento neutro* seja denotado 0 e $-\alpha$ denota o *inverso* de $\alpha \in \mathbb{K}$.)
2. $(\mathbb{K} \setminus \{0\}, \cdot)$ é um grupo abeliano.
   (O *elemento neutro* seja denotado 1 e α^{-1} denota o *inverso* de $\alpha \in \mathbb{K}\setminus\{0\}$.)
3. Distributividade: $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$ para todos $\alpha, \beta, \gamma \in \mathbb{K}$.
   (É costume escrever $\alpha\beta$ em vez de $\alpha \cdot \beta$.)

Para distinguir chamamos o elemento neutro da primeira operação – para a qual temos usado o símbolo "+" ainda que geralmente não tem nada ver com adição de números – o **elemento neutro aditivo**. Chamamos o elemento neutro da segunda operação – motivado pelo uso do símbolo "$\cdot$" – o **elemento neutro multiplicativo**. Como é feio escrever $\alpha + (-\beta)$ para a soma de um elemento com um elemento inverso aditivo definimos $\alpha - \beta := \alpha + (-\beta)$. Isso é uma abreviação só, não é, nem tem diferença. Analogamente simplificamos a notação escrevendo α/β em vez de $\alpha\beta^{-1}$.

Corolário 1.1.9. *Um corpo contem pelo menos dois elementos.*

Demonstração. Pelos axiomas 1. e 2. cada uma operação tem um elemento neutro as quais não podem ser iguais por causa de 2. □

Lema 1.1.10. *Seja $\mathbb{K}$ um corpo e 0 o elemento neutro da adição. Então $0\beta = 0$ e $\beta 0 = 0$ para todos os elementos $\beta \in \mathbb{K}$.*

Demonstração. Lema B.1.2. □

[1] as quais vamos batizar aos nomes "+" e "$\cdot$" – ainda que *geralmente não tem nada ver com adição e multiplicação de números*, mas esta escolha é motivada pelos exemplos principais (Exemplo 1.1.12) nos quais "+" e "$\cdot$" são adição e multiplicação de números

Definição 1.1.11. Seja $\mathbb{K}$ um corpo e 0 e 1 os elementos neutros da adição e multiplicação. O menor número natural $k \geq 2$, caso existisse, tal que a soma seguinte se anula

$$\underbrace{1 + \cdots + 1}_{k \text{ vezes}} = 0$$

é chamado de **caraterística** do corpo. Caso não existe nenhum tal k chamamos o corpo de **caraterística nula**.

Exemplos de corpos

Exemplo 1.1.12 (Os corpos $\mathbb{R}$ e $\mathbb{Q}$)**.** São corpos $\mathbb{R} = (\mathbb{R}, +, \cdot)$ e $\mathbb{Q} = (\mathbb{Q}, +, \cdot)$.

Exemplo 1.1.13 (O corpo dos números complexos $\mathbb{C}$)**.** Vamos chamar a letra i de **unidade imaginária**. Um **número complexo** é uma soma formal da forma $a + ib$ onde a e b são números reais. O conjunto $\mathbb{C}$ de todos os números complexos munido das duas operações definidas assim

$$\begin{aligned}(a + ib) + (c + id) &:= (a + c) + i(b + d)\\ (a + ib) \cdot (c + id) &:= (ac - bd) + i(ad + bc)\end{aligned}$$

é um corpo, denotado $(\mathbb{C}, +, \cdot)$, ou simplesmente $\mathbb{C}$.

Enquanto a regra da adição é fácil para memorizar, você consegue memorizar a multiplicação? Eu não. Deixa ver se tem um jeito simples para chegar na fórmula certa: Fazemos multiplicação formal como fossem números, ou seja

$$\begin{aligned}(a + ib) \cdot (c + id) &= ac + aid + ibc + i^2bd\\ &= ac + i^2bd + i(ad + bc)\\ &= (ac - bd) + i(ad + bc)\end{aligned}$$

Assim só precisamos memorizar que $i^2 = -1$, mas isso é uma fórmula famosa.

Considere um número complexo $z = a + ib$. Chama-se a a **parte real** de z, notação $\mathrm{Re}(z) = a$, e chama-se b a **parte imaginária** de z, notação $\mathrm{Im}(z) = b$.

O **complexo conjugado** de $z = a + ib$ é definido por $\bar{z} := a - ib$ e o **absoluto**, ou módulo, por $|z| := \sqrt{a^2 + b^2}$.

Exercício 1.1.14. Os números inteiros $\mathbb{Z} = (\mathbb{Z}, +, \cdot)$ não formam um corpo.

Exemplo 1.1.15 (Adição e multiplicação modulo n)**.** Dado um número natural $n \geq 2$,[2] defina no conjunto $\mathbb{Z}_n := \{0, 1, \ldots, n-1\}$ as duas operações

$$a +_n b := a + b \pmod{n}, \qquad a \cdot_n b := ab \pmod{n}$$

para todos os elementos $a, b \in \mathbb{Z}_n$.[3]

Fato. $(\mathbb{Z}_n, +_n, \cdot_n)$ é um corpo $\iff$ n é um número primo.

[2] para $n = 1$ obtemos o grupo aditivo trivial $\mathbb{Z}_1 = \{0\}$ o qual nunca pode ser um corpo, porque contem 1 elemento só; veja Corolário 1.1.9

[3] Dado $n \in \mathbb{N}$, seja $\ell \in \mathbb{Z}$ um número inteiro. Pela definição o elemento $\ell \pmod{n} \in \mathbb{Z}_n$ é o resto $r \in \{0, 1, \ldots, n-1\}$ que falta depois você "enche" ℓ com múltiplos de n. Em símbolos, $\ell \pmod{n} := r$ onde $r \in \{0, 1, \ldots, n-1\}$ é o único elemento tal que $\ell = kn + r$ para um $k \in \mathbb{Z}$.

Um primo p é a caraterística de $\mathbb{Z}_p$. Para valores pequenos de n pode-se checar de mão se $\mathbb{Z}_n$ é um corpo ou não. Só precisa-se calcular as tabelas de adição e de multiplicação. Vamos ilustrar isso num exemplo.

Exemplo 1.1.16 ($\mathbb{Z}_4$ não é um corpo.)**.** Para checar se $(\mathbb{Z}_4, +_4)$ e $(\mathbb{Z}_4 \backslash \{e_{+_4}\}, \cdot_4)$ são grupos abelianos é útil calcular as tabelas de adição e de multiplicação.

• $(\mathbb{Z}_4, +_4)$ é um grupo abeliano? Para responder calculamos os valores na tabela

$+_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

São 4 passos:

1. Determinar o elemento neutro de $+_4$: Checamos se a linha em cima da linha solida horizontal, ou seja a linha 0 1 2 3, tem uma cópia nas linhas embaixo. Sim, tem 0 1 2 3. Neste caso o elemento em frente da cópia é o elemento neutro de $+_4$, certo? No nosso caso $e_{+_4} = 0$. Se não tem copia, não tem elemento neutro, e assim não seria um grupo.
2. Inversos: Na cada dos (neste caso 4) linhas de valores na tabela localiza o elemento neutro 0 (se existir). Então o elemento g em frente da linha de 0 e o elemento em cima da coluna de 0, notação $\bar{g}$, são inversos um do outro. Caso uma linha não contem 0, então este g não tem inverso, e assim não seria um grupo. No nosso caso todo elemento g tem um inverso:

g	$\bar{g}$ (denotado $-g$)
0	0
1	3
2	2
3	1

3. Associatividade: Calculando caso por caso temos que checar se $f +_4 (g +_4 h) = (f +_4 g) +_4 h$ para todas as possibilidades. No nosso caso vale.
4. Grupo abeliano (comutatividade): Vale se a tabela é simétrica em respeito à diagonal. No nosso caso vale.

Na verdade temos esquecido um passo: No início de tudo temos que checar se a operação é bem definida, ou seja os valores da operação (os valores na tabela) realmente são elementos do conjunto, ou não. Olhamos a tabela - sim.
Nosso resultado é que $(\mathbb{Z}_4, +_4)$ é um grupo abeliano.

• $(\mathbb{Z}_4 \setminus \{0\}, \cdot_4)$ é um grupo abeliano? Para responder calculamos a tabela

$\cdot_4$	1	2	3
1	1	2	3
2	2	0	2
3	3	2	1

Como o valor 0 não é elemento de $\mathbb{Z}_4 \setminus \{0\}$ a multiplicado $\cdot_4$ não é uma operação em $\mathbb{Z}_4 \setminus \{0\}$, então não pode ser um grupo.

Ainda assim vamos repetir os 4 passos para $\cdot_4$ (em vez de $+_4$) para ver se tem outras falhas ainda. As respostas são:

1. Elemento neutro de $\cdot_4$: Tem, é o elemento $e_{\cdot_4} = 1$.
2. Inversos: Na cada dos (neste caso 3) linhas de valores na tabela localizamos o elemento neutro 1 (se existir). No nosso caso

g	$\bar{g}$ (denotado g^{-1})
1	1
2	não tem!
3	3

   o elemento 2 não tem um inverso e já por isso não temos um grupo.
3. Associatividade: Ainda que a fórmula $f +_4 (g +_4 h) = (f +_4 g) +_4 h$ vale, os valores não são todos em $Z_4 \setminus \{0\}$.
4. Grupo abeliano (comutatividade): A tabela é simétrica em respeito à diagonal, mas os valores não são todos em $Z_4 \setminus \{0\}$.

Nosso resultado é que $(\mathbb{Z}_4 \setminus \{0\}, \cdot_4)$ não é um grupo abeliano.

Exercício 1.1.17. Seja $n = 6$:

1. Calcule a tabela da adição e da multiplicação no caso $\mathbb{Z}_6$.
2. Identifique os elementos neutros da adição e multiplicação em $\mathbb{Z}_6$. Eles sempre existem?
3. Para todo $a \in \mathbb{Z}_6$ identifique o elemento inverso aditivo.
4. Para todo $a \in \mathbb{Z}_6 \setminus \{0\}$ identifique o elemento inverso multiplicativo, se existir.
5. Cheque que $\mathbb{Z}_6$ não é um corpo. Quais dos axiomas não valem?

Matéria avançada

Motivado pelas perguntas da Turma C na 1ª aula 2016-2 vamos dar um exemplo de um corpo onde a primeira operação não está relacionada à adição de números nem a segunda à multiplicação de números.

Exercício 1.1.18 (Corpo $(P, \cdot, \circ, \mathbb{R})$ onde $\cdot$ não é adição e $\circ$ não é multiplicação)**.** Para $\alpha \in \mathbb{R}$ considere a função $p_\alpha : (0, \infty) \to (0, \infty)$, $x \mapsto x^\alpha$. Seja o conjunto

$$P := \{p_\alpha \mid \alpha \in \mathbb{R}\}$$

composto de todas funções $p_\alpha(x) = x^\alpha$ com $\alpha \in \mathbb{R}$ e munido das operações

$$\begin{aligned} \cdot : P \times P &\to P \\ (p_\alpha, p_\beta) &\mapsto p_\alpha \cdot p_\beta \end{aligned} \qquad \begin{aligned} \circ : P \times P &\to P \\ (p_\alpha, p_\beta) &\mapsto p_\alpha \circ p_\beta \end{aligned}$$

chamado de **multiplicação**[4] e **composição**[5] de funções, respectivamente. Mostre que:

1. As duas operações são bem definidas: $p_\alpha \cdot p_\beta \in P$ e $p_\alpha \circ p_\beta \in P$, de fato
$$p_\alpha \cdot p_\beta = p_{\alpha+\beta}, \qquad p_\alpha \circ p_\beta = p_{\alpha\beta}$$
2. $(P, \cdot)$ é um grupo abeliano com elemento neutro $p_0 \equiv 1$.
3. $(P \setminus \{p_0\}, \circ)$ é um grupo abeliano com elemento neutro $p_1(x) = x$.
4. Distributividade: $(p_\alpha \cdot p_\beta) \circ p_\gamma = (p_\alpha \circ p_\gamma) \cdot (p_\beta \circ p_\gamma),\ \ \forall p_\alpha, p_\beta, p_\gamma \in P$.

1.1.3 Espaço vetorial

O conceito de espaço vetorial reportar-se a um livro de 1844 do suiço Hermann Günther Grassmann (1809-1877), professor num ginásio, veja [Koe85, p. 10-15]. A forma abstrata a qual introduzimos no seguinte só apareceu em torno de 1900.

Definição 1.1.19. Um **espaço vetorial E sobre um corpo** [6] $\mathbb{K}$ é um quádruplo $(E, +, \cdot, \mathbb{K})$ composto de um conjunto E, um corpo $\mathbb{K}$, e duas operações

$$\begin{aligned} + : E \times E &\to E & \cdot : \mathbb{K} \times E &\to E \\ (v, w) &\mapsto v + w & (\alpha, v) &\mapsto \alpha v \end{aligned}$$

chamadas de **adição** e **multiplicação escalar**, respectivamente, tal que vale

1. $(E, +)$ é um grupo abeliano.
(O *elemento neutro* é denotado $\mathcal{O}$ e chamado o **vetor nulo**.)
2. Distributividade: $\begin{cases} (\alpha + \beta)v = \alpha v + \beta v \\ \alpha(v + w) = \alpha v + \alpha w \end{cases}$
3. Compatibilidade: $\begin{cases} (\alpha\beta)v = \alpha(\beta v) \\ 1v = v \end{cases}$

Onde as identidades tem que ser válidas para todos $\alpha, \beta \in \mathbb{K}$ e todos $v, w \in E$. Chama-se **escalares** os elementos do corpo $\mathbb{K}$ e **vetores** os elementos de E.

Lema 1.1.20. *Seja $(E, +, \cdot, \mathbb{K})$ um espaço vetorial e $0 \in \mathbb{K}$ e $\mathcal{O} \in E$, então:*

(i) $\alpha\mathcal{O} = \mathcal{O}$ *para todos os escalares* $\alpha \in \mathbb{K}$.

(ii) $0v = \mathcal{O}$ *para todos os vetores* $v \in E$.

[4] $(p_\alpha \cdot p_\beta)(x) := p_\alpha(x) \cdot p_\beta(x)$
[5] $(p_\alpha \circ p_\beta)(x) := p_\alpha(p_\beta(x))$
[6] fala-se abreviando "*E é um espaço vetorial sobre* $\mathbb{K}$" ou ainda "*E é um espaço vetorial*".

(iii) *Para todo o escalar $\alpha \in \mathbb{K}$ e todo o vetor $w \in E$ são equivalentes:*

$$\alpha w = \mathcal{O} \quad \Leftrightarrow \quad \alpha = 0 \text{ ou } w = \mathcal{O} \tag{1.1.3}$$

Demonstração. Lema B.1.3. □

Corolário 1.1.21 (Compatibilidade dos inversos aditivos com multiplicação)**.** *Para todo o escalar $\alpha \in \mathbb{K}$ e todo o vetor $w \in E$ vale:*

a) $(-\alpha)w = -(\alpha w)$
b) $\alpha(-w) = -(\alpha w)$

Demonstração. Corolário B.1.4. □

Corolário 1.1.22. *Seja $(E, +, \cdot, \mathbb{K})$ um espaço vetorial sobre um corpo $\mathbb{K}$ no qual $1 + 1 \neq 0$. Neste caso para $v \in E$ temos*

$$v + v = \mathcal{O} \qquad \Rightarrow \qquad v = \mathcal{O}$$

Demonstração. Como $\mathcal{O} = v + v = 1v + 1v = (1 + 1)v$ segue de (1.1.3) que ou $1 + 1 = 0$ no corpo $\mathbb{K}$ ou $v = \mathcal{O}$. (Lembre-se que $1 \in \mathbb{K}$, assim 1 geralmente não é um número e $1+1$ não tem nada ver com 2... Veja nota de rodapé no Lema 7.2.5.) □

1.2 Exemplos de espaços vetoriais

Exemplo 1.2.1 (O espaço vetorial trivial $\{\mathcal{O}\}$)**.** Seja E um conjunto com 1 elemento só. Vamos já denotar aquele elemento com o símbolo $\mathcal{O}$ (porque?). Então $E = \{\mathcal{O}\}$. Seja $\mathbb{K}$ um corpo qualquer. Não tem escolha nenhuma para definir as duas operações

$$\begin{aligned} + : E \times E &\to E & \qquad \cdot : \mathbb{K} \times E &\to E \\ (\mathcal{O}, \mathcal{O}) &\mapsto \mathcal{O} & \qquad (\alpha, \mathcal{O}) &\mapsto \mathcal{O} \end{aligned}$$

Então $(E, +, \cdot, \mathbb{K})$ satisfaz os axiomas de um espaço vetorial, denotado simplesmente $E = \{\mathcal{O}\}$ e chamado de **espaço vetorial trivial**.

Exemplo 1.2.2 (Um corpo $\mathbb{K}$ como um espaço vetorial sobre $\mathbb{K}$)**.** Usa-se as duas operações chegando com o corpo $(\mathbb{K}, +, \cdot)$ como as duas operações necessárias para tornar um conjunto (escolhemos $E := \mathbb{K}$) num espaço vetorial sobre um corpo (escolhemos $\mathbb{K}$). Com efeito $(\mathbb{K}, +, \cdot, \mathbb{K})$ é um espaço vetorial sobre $\mathbb{K}$.

1.2.1 Listas ordenadas

Números reais

Exemplo 1.2.3 (O espaço vetorial $\mathbb{R}^n$ sobre $\mathbb{R}$)**.** Seja

$$\mathbb{R}^n := \{u = (\alpha_1, \ldots, \alpha_n) \mid \alpha_1, \ldots \alpha_n \in \mathbb{R}\}$$

o conjunto de todas as listas ordenadas de n números reais. Chamamos α_i o i-ésimo membro da lista. As duas operações

$$+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \qquad \cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$$

são definidas como adição membro-por-membro e multiplicação de todos membros com um escalar $\beta \in \mathbb{R}$. Checando todos axiomas vê-se que $\mathbb{R}^n$ é um espaco vetorial sobre o corpo dos números reais, notação $(\mathbb{R}^n, +, \cdot, \mathbb{R})$ ou $\mathbb{R}^n$ só. O vetor nulo, também chamado de **origem**, é a lista

$$\mathcal{O} = (0, \dots, 0)$$

e o inverso de um elemento $u = (\alpha_1, \dots, \alpha_n)$ é a lista $(-\alpha_1, \dots, -\alpha_n)$ a qual denotamos com o símbolo $-u$.

O ***i*-ésimo vetor canônico** é a lista de n membros

$$e_i = (0, \dots, 0, 1, 0, \dots, 0)$$

cujo i-ésimo membro é o número 1 e todos outros sao nulo 0. O conjunto

$$\mathcal{E}^n := \{e_1, \dots, e_n\} \tag{1.2.1}$$

de todos os vetores canônicos é chamado de **base canônica** de $\mathbb{R}^n$.

Exemplo 1.2.4 (O espaço vetorial $\mathbb{R}^\infty$ sobre $\mathbb{R}$)**.** O conjunto

$$\mathbb{R}^\infty := \{u = (\alpha_1, \alpha_2, \alpha_3, \dots) \mid \alpha_1, \alpha_2, \alpha_3, \dots \in \mathbb{R}\}$$

de todas as sequências reais é um espaço vetorial sobre $\mathbb{R}$ sob adição e multiplicação membro-por-membro, notação $(\mathbb{R}^\infty, +, \cdot, \mathbb{R})$.

Exemplo 1.2.5 (O espaço vetorial $\mathbb{R}_0^\infty$ sobre $\mathbb{R}$)**.** O conjunto

$$\mathbb{R}_0^\infty := \{u \in \mathbb{R}^\infty \mid \text{só um número finito de membros são não-nulos}\}$$

é um espaço vetorial sobre $\mathbb{R}$ sob adição e multiplicação membro-por-membro como no exemplo prévio, notação $(\mathbb{R}_0^\infty, +, \cdot, \mathbb{R})$.
Dado $i \in \mathbb{N}$, a sequência com todos membros nulos exceto o i-ésimo qual é 1 denotamos também de e_i. O conjunto de todos os e_i's é denotado de

$$\mathcal{E}^\infty := \{e_1, e_2, \dots\} \tag{1.2.2}$$

e chamado de **base canônica** de $\mathbb{R}_0^\infty$.

Corpos gerais

Comentário 1.2.6 ($\mathbb{K}^n$ e $\mathbb{K}^\infty$)**.** Os espaços vetoriais $\mathbb{K}^n$ e $\mathbb{K}^\infty$ sobre qualquer corpo $\mathbb{K}$ são definidos analogamente Exemplos 1.2.3 e 1.2.4.

1.2.2 Matrizes

No Apêndice A.1 introduzimos o conjunto $\mathrm{M}(m \times n)$ das matrizes $m \times n$ cujas entradas a_{ij} são números reais, munido de duas operações, adição e multiplicação escalar. O conteúdo do Apêndice A.1, de fato do Apêndice A inteiro, também vale para matrizes com entradas num corpo $\mathbb{K}$. Repetimos uns fatos.

Exemplo 1.2.7 (Espaço vetorial das matrizes $m \times n$)**.** O **espaço vetorial das matrizes $m \times n$ sobre um corpo** $\mathbb{K}$ é o conjunto

$$\mathrm{M}(m \times n; \mathbb{K}) := \left\{ \mathbf{a} = (a_{ij}) \;\middle|\; a_{ij} \in \mathbb{K},\, i = 1, \dots m,\, j = 1, \dots, n \right\}$$

onde a matriz $\mathbf{a} = (a_{ij})$ é o quadro de escalares com m linhas e n colunas [7]

$$\mathbf{a} = (a_{ij}) := \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

munido da adição entrada por entrada e da multiplicação escalar com escalares $\beta \in \mathbb{K}$ também entrada por entrada.

O vetor nulo é a matriz nula $\mathbf{0}$ cujas entradas são todas o escalar nulo $0 \in \mathbb{K}$. Se na matriz nula $n \times n$ colocamos o escalar $1 \in \mathbb{K}$ ao longo da diagonal obtemos a **matriz identidade** $\mathbb{1} = \mathbb{1}_n$. O elemento inverso aditivo, notação $-\mathbf{a}$, de uma matriz $\mathbf{a} = (a_{ij})$ tem como entradas os inversos aditivos dos a_{ij}, notação $-a_{ij}$.

Definição 1.2.8 (Linhas e colunas de matrizes)**.** Seja $\mathbf{a} = (a_{ij}) \in \mathrm{M}(m \times n; \mathbb{K})$ uma matriz $m \times n$. Note-se que o primeiro índice i de uma entrada a_{ij} indica a linha e o segundo j a coluna dela. Tendo isso na vista vamos denotar a ***k*-ésima coluna**, respectivamente a ***ℓ*-ésima linha**, de uma matriz $\mathbf{a}$ com os símbolos

$$\mathbf{a}_{\bullet k} = \begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix}, \qquad \mathbf{a}_{\ell \bullet} = \begin{bmatrix} a_{\ell 1} & \dots & a_{\ell n} \end{bmatrix} \tag{1.2.3}$$

Temos escolhido o símbolo $\bullet$ para sugerir "este índice é aberto" – ele corre e assim gera uma lista, ou vertical ou horizontal dependente se $\bullet$ fica no primeiro ou no segundo lugar. Assim podemos escrever a matriz $\mathbf{a}$ nas formas seguintes

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}_{1\bullet} \\ \vdots \\ \mathbf{a}_{m\bullet} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_{\bullet 1} & \dots & \mathbf{a}_{\bullet n} \end{bmatrix}$$

Definição 1.2.9 (Espaço-coluna e espaço-linha)**.** Seja $\mathbf{a} = (a_{ij}) \in \mathrm{M}(m \times n; \mathbb{K})$ uma matriz $m \times n$. O **espaço-coluna** é o conjunto de todas as somas das colunas $\mathbf{a}_{\bullet k}$ da matriz decorado com fatores escalares α_k, em símbolos

$$\text{Esp-col}(\mathbf{a}) := \{\alpha_1 \mathbf{a}_{\bullet 1} + \dots + \alpha_n \mathbf{a}_{\bullet n} \mid \alpha_1, \dots, \alpha_n \in \mathbb{K}\} \subset \mathrm{M}(m \times 1; \mathbb{K})$$

[7] Os escalares a_{ij} são chamadas as **entradas da matriz**. Por definição que a entrada a_{ij} está localizada na i-ésima linha e j-ésima coluna.

Analogamente no **espaço-linha** usa-se as linhas da matriz $\mathbf{a}$, ou seja

$$\text{Esp-lin}(\mathbf{a}) := \{\alpha_1 \mathbf{a}_{1\bullet} + \cdots + \alpha_n \mathbf{a}_{n\bullet} \mid \alpha_1, \ldots, \alpha_n \in \mathbb{K}\} \subset \mathrm{M}(1 \times n; \mathbb{K})$$

Produto matriz

O produto entre duas matrizes de tipos adequadas é definido no apêndice na formula (A.1.2) e umas propriedades do produto encontram-se no Lema A.1.5.

Para matrizes quadradas faz sentido investigar se admitem uma inversa ou não. Uma matriz quadrada $\mathbf{a}$ é chamada de **invertível** se existe uma matriz quadrada $\mathbf{b}$ tal que $\mathbf{ab} = \mathbb{1}$.

1.2.3 Funções e polinômios

Exercício 1.2.10. Dado um conjunto não-vazio $X \neq \emptyset$ e um corpo $\mathbb{K}$, seja

$$\mathcal{F}(X, \mathbb{K}) := \{f \mid f : X \to \mathbb{K} \text{ função}\}$$

o conjunto de todas as funções $f : X \to \mathbb{K}$. Adição de funções e multiplicação com um escalar são definidas assim

$$(f + g)(x) := f(x) + g(x), \qquad (\alpha f)(x) := \alpha f(x)$$

para todos os $x \in X$, $\alpha \in \mathbb{K}$. Mostre que $\mathcal{F}(X, \mathbb{K})$ é um espaço vetorial sobre $\mathbb{K}$.

Comentário 1.2.11. A próxima observação ilustra o poder da matemática e um *ponto fundamental* dela - economizar através de abstração e *encontrar o certo ponto da vista*.

Observação 1.2.12.

a) Se $X = \{1, \ldots, n\}$, então $\mathcal{F}(X, \mathbb{R}) = \mathbb{R}^n$.

b) Se $X = \mathbb{N}$, então $\mathcal{F}(X, \mathbb{R}) = \mathbb{R}^\infty$.

c) Se $X = \{1, \ldots, m\} \times \{1, \ldots, n\}$, então $\mathcal{F}(X, \mathbb{R}) = \mathrm{M}(m \times n)$.

Exercício 1.2.13 (Polinômios $\mathcal{P}(\mathbb{K})$)**.** Dados escalares $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{K}$, então chama-se uma soma finita

$$p = p(x) := \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n$$

de **polinômio** na variável $x \in \mathbb{K}$, no caso $\alpha_n \neq 0$ de **polinômio de grau n**, e no caso $\alpha_n = 1$ de **polinômio mónico**. Forneça o conjunto dos polinômios com uma estrutura de um espaço vetorial $(\mathcal{P}(\mathbb{K}), +, \cdot, \mathbb{K})$.

1.3 Independência linear

1.3.1 Combinação linear

Definição 1.3.1. Uma soma *finita* da forma

$$\underbrace{\alpha_1 v_1 + \ldots + \alpha_\ell v_\ell}_{=:w} \qquad \text{com } \alpha_j \in \mathbb{K},\, v_j \in E \tag{1.3.1}$$

é chamado de **combinação linear (CL)**. Uma tal CL chama-se de **trivial** se todos os coeficientes $\alpha_1, \ldots, \alpha_\ell$ são nulos. Dizemos que

"os vetores $v_1, \ldots, v_\ell$ representam o vetor w"

ou

"o vetor w é CL dos vetores $v_1, \ldots, v_\ell$"

Um problema central na teoria dos espaços vetoriais: Dado vetores $v_1, \ldots, v_\ell$, existe uma CL não-trivial deles representando o vetor nulo?
Nas outras palavras, existem escalares $\alpha_1, \ldots, \alpha_\ell$ não todos nulos, tal que a CL correspondente dos vetores representa o vetor nulo? Em símbolos,

"existem coeficientes $(\alpha_1, \ldots, \alpha_\ell) \neq (0, \ldots, 0)$ tal que $\alpha_1 v_1 + \ldots + \alpha_\ell v_\ell = \mathcal{O}$?"

Exercício 1.3.2. a) Caso possível escreva o vetor $b = (1, -3, 10) \in \mathbb{R}^3$ como combinação linear dos vetores $u = (2, -3, 5)$, $v = (1, 1, 0)$, e $w = (1, 0, 0)$.
b) Sejam $u = (1, 1)$, $v = (1, 2)$ e $w = (2, 1)$. Encontre números a, b, c e α, β, γ todos não-nulos, tais que

$$au + bv + cw = \alpha u + \beta v + \gamma w$$

com $a \neq \alpha$, $b \neq \beta$ e $c \neq \gamma$.

[Dica: a) Determinar os coeficientes α, β, γ na CL de u, v, w a qual representa b lida a um SL. Escalonamento.[8]
b) Defina $x = a - \alpha$, $y = b - \beta$, e $z = c - \gamma$ para obter um SLH. Resolve.][9]

1.3.2 Linearmente independente

Definição 1.3.3 (Independência linear)**. Vetores** $v_1, \ldots, v_\ell$ são chamados de

- **linearmente dependente (LD)** se representam o vetor nulo através de uma CL não-trivial;

[8] encontre "o certo ponto da vista" (Comentário 1.2.11) e o SL vai chegar já escalonada..
[9] Respostas para seu controle: a) $(\alpha, \beta, \gamma) = (2, 3, -6)$. b) $(x, y, z) = z(-3, 1, 1)$. Escolha um $z \neq 0$, por exemplo $z = 1$. Então $(a, b, c) = (\alpha - 3, 1 + \beta, 1 + \gamma)$. Toda escolha de reais $\alpha \neq 0, 3$ e $\beta, \gamma \neq 0, -1$ da uma solução. A escolha $\alpha = 5$ e $\beta = \gamma = 1$ resulta em $a = b = c = 2$.

- **linearmente independente (LI)** se não são linearmente dependente, em símbolos

$$\alpha_1 v_1 + \ldots + \alpha_\ell v_\ell = \mathcal{O} \quad \Rightarrow \quad \alpha_1 = \cdots = \alpha_\ell = 0 \tag{1.3.2}$$

ou em palavras

> *"não existe nenhuma CL não-trivial dos vetores $v_1, \ldots, v_\ell$ representando o vetor nulo"*

Um **subconjunto** X de um espaço vetorial E é dito de

- **linearmente independente (conjunto LI)** se cada uma escolha finita de elementos dois-a-dois diferentes[10] é linearmente independente;

- **linearmente dependente (conjunto LD)** se não é linearmente independente, isto é se o vetor nulo é CL não -trivial de uma escolha finita de elementos dois-a-dois diferentes.

Ainda que se vetores $v_1, \ldots, v_\ell$ são elementos de um subconjunto X de E, uma CL deles não necessariamente encontra-se mais em X. Encontra-se sim, quando o subconjunto X de E é um chamado "subespaço" (Lema 2.1.2).

Comentário 1.3.4.

(i) O conjunto vazio $\emptyset$ é LI: Com efeito, com os elementos de $\emptyset$ não pode-se representar o vetor nulo nunca - simplesmente não tem elementos. Chama-se tal argumentação de **verdade vazia**.

(ii) Um conjunto $X = \{v\}$, contendo só um vetor, é LI se e somente se $v \neq \mathcal{O}$.

(iii) Se (1.3.2) vale para uma escolha $v_1, \ldots, v_\ell$, então vale para qualquer subescolha destes vetores. [Use os coeficientes $\alpha_i = 0$ nos restantes.]

Para provar a afirmação (ii), lembra (1.1.3).

Lema 1.3.5. *Sejam A, B, X subconjuntos de um espaço vetorial E.*

a) $\mathcal{O} \in X \Rightarrow X$ *LD.* *(O vetor nulo rende conjuntos LD)*

b) *Todo conjunto A contido num conjunto LI X é LI.*

c) *Todo conjunto B contendo um conjunto LD X é LD.*

Demonstração. a) O termo $1\mathcal{O}$ é uma CL não-trivial em X representando o vetor nulo (segundo Lema 1.1.20). b) Os elementos de A são elementos de X, agora usa a definição da independência linear de X. c) Cada uma CL não-trivial em X representando o vetor nulo é também em B. □

[10] Para que precisa-se a condição *dois-a-dois diferentes*?

Exemplo 1.3.6. Para saber se o subconjunto $X := \{(1,0),(2,1)\}$ de $\mathbb{R}^2$ é LI temos que checar (1.3.2) para cada uma escolha finita de elementos dois-a-dois diferentes de X. Como X é um conjunto finito, e tendo em vista Comentário 1.3.4 (iii), começamos com a escolha máxima, ou seja todos os (dois) elementos. Sejam $\alpha, \beta \in \mathbb{R}$. Conforme (1.3.2) suponhamos que

$$(0,0) = \alpha(1,0) + \beta(2,1) = (\alpha + 2\beta, \beta)$$

onde o termo cinza resulta através das regras de multiplicação escalar e adição em $\mathbb{R}^2$. Comparando os segundos membros vemos que $0 = \beta$ o qual usamos na comparação dos primeiros membros: recebemos $0 = \alpha + 2 \cdot 0 = \alpha$. Assim temos provado que ambos os coeficientes α e β são nulos. Então X é LI.

Exercício 1.3.7. Quais dos seguintes conjuntos X_i de vetores de $\mathbb{R}^2$ são ou não são conjuntos linearmente independentes (LI)? Explique porque são ou não são.

1. Elementos de X_1: os vetores $(1,1)$ e $(-1,-1)$.
2. $X_2 := \{(2, \frac{1}{2}), (\frac{1}{2}, 2)\}$.
3. Escolha dois vetores diferentes $u, v \in \mathbb{R}^2$ e defina $X_3 := \{u, v\} \cup \{(1,1)\}$.

Exercício 1.3.8. Prove as afirmações seguintes.

1. A base canônica $\mathcal{E}^n$ em (1.2.1) é um conjunto LI no $\mathbb{R}^n$ para $n \in \mathbb{N}$.
2. A base canônica $\mathcal{E}^\infty$ em (1.2.2) é um conjunto LI no $\mathbb{R}_0^\infty$ e no $\mathbb{R}^\infty$.
3. Suponha $u, v \in \mathbb{K}^2$ não são múltiplos um do outro. Prove que o conjunto $\{u, v\}$ é LI.

   [Dica: Seja $\alpha u + \beta v = \mathcal{O}$. Considere $\beta \neq 0$ e, lembrando (1.1.3), $\beta = 0$.]
4. Sejam $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ vetores de $\mathbb{R}^n$. Prove que um deles é múltiplo do outro se, e somente se, para todo $i, j = 1, \ldots, n$ temos $x_i y_j = x_j y_i$.
5. O subconjunto $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\} \subset \mathrm{M}(2 \times 2)$ composto das matrizes

   $$\mathbf{a} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{c} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

   é um conjunto LI.
6. O conjunto X composto dos três polinômios

   $$\begin{aligned} p &= p(x) = x^3 - 5x^2 + 1 \\ q &= q(x) = 2x^4 + 5x - 6 \\ r &= r(x) = x^2 - 5x + 2 \end{aligned}$$

   é um conjunto LI no espaço vetorial $\mathcal{P}(\mathbb{R})$ dos polinômios reais.
7. Se o conjunto de vetores $\{v_1, \ldots, v_m\}$ é LI, prove que o mesmo se dá com o conjunto $\{v_1, v_2 - v_1, \ldots, v_m - v_1\}$. Vale a recíproca?

Exercício 1.3.9 (Funções com valores complexos)**.** Seja $E = \mathcal{F}(\mathbb{R}, \mathbb{C})$ o espaço vetorial sobre o corpo $\mathbb{C}$ composto de todas as funções $f : \mathbb{R} \to \mathbb{C}$ como introduzido em Exercício 1.2.10. Considere os subconjuntos $X = \{f_1, f_2, f_3\}$ e $Y = \{g_1, g_2, g_3\}$ de E composto das funções

$$f_1(t) = \cos t - i \sin t, \qquad f_2(t) = \cos t + i \sin t, \qquad f_3(t) = 1$$

e

$$g_1(t) = \sin t, \qquad g_2(t) = 1, \qquad g_3(t) = \cos t$$

(a) Prove que cada um dos conjuntos X e Y é LI.

(b) Escreva cada g_i como combinação linear dos f_i's.

(c) Escreva cada f_i como combinação linear dos g_i's.

Este exercício vai ter continuação em Exercício 5.4.7.

Capítulo 2

Subespaços

Um subespaços de um espaço vetorial E é um subconjuntos F qual é invariante sob as duas operações de E. Assim faz sentido restringir as duas operações a F. Munido das operações restritas F torna-se um espaço vetorial mesmo.

2.1 Definição e exemplos

Definição 2.1.1. Um subconjunto $F \subset E$ de um espaço vetorial $(E, +, \cdot, \mathbb{K})$ é chamado de **subespaço** se é **fechado sob as duas operações**, ou seja

(i) $u, v \in F \Rightarrow u + v \in F$ (F é fechado sob adição)

(ii) $\alpha \in \mathbb{K}, u \in F \Rightarrow \alpha u \in F$ (F é fechado sob multiplicação escalar)

Lema 2.1.2. *Seja F um subespaço de um espaço vetorial $(E, +, \cdot, \mathbb{K})$. Então*

a) $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$, $v_1, \ldots, v_k \in F \ \Rightarrow \ \sum_{i=1}^{k} \alpha_i v_i \in F$ *(fechado sob CL)*

b) *F é um espaço vetorial sobre o corpo $\mathbb{K}$ onde as duas operações são aquelas de E restrito ao subconjunto $F \subset E$.* *(subespaços são espaços vetoriais)*

Demonstração. a) Indução. b) As restrições tomam valores em F segundo parte a) e as axiomas valem como os elementos de F são elementos de E para as quais os axiomas valem pela hipótese que E é um espaço vetorial. □

Exercício 2.1.3 (Vetor nulo)**.** O vetor nulo de um subespaço F é o vetor nulo $\mathcal{O}$ do espaço vetorial ambiente. [Dica: Mostre $\mathcal{O} \in F$. O vetor nulo de F é único.]

Checar se um subconjunto $F \subset E$ é um espaço vetorial é bastante trabalhoso dado os muitos axiomas. Isso mostra o valor alto da parte b) do lema dizendo que é suficiente checar "fechado sob as duas operações" – tarefa rapidinha.

Exercício 2.1.4. Mostre que o espaço vetorial $\mathbb{R}$ só tem dois subespaços, isto é $\{0\}$ e $\mathbb{R}$.

Exercício 2.1.5. Mostre que são subespaços de um espaço vetorial $(E, +, \cdot, \mathbb{K})$:

a) $F := \{\mathcal{O}\}$ (o subespaço mínimo / trivial)

a) $F := E$ (o subespaço máximo)

b) $\mathbb{K}v := \{\alpha v \mid \alpha \in \mathbb{K}\}$ (a reta passando v e a origem $\mathcal{O}$)
onde v é um vetor não-nulo de E. Observe que $\mathbb{K}\mathcal{O} = \{\mathcal{O}\}$ é um ponto só.

Exemplo 2.1.6 (O subespaço $\mathbb{R}_0^\infty$ de $\mathbb{R}^\infty$)**.** O subconjunto $\mathbb{R}_0^\infty \subset \mathbb{R}^\infty$, composto de todas sequências reais tal que só um número finito de membros são nao-nulos, é um subespaço: Se a lista u tem k membros não-nulos e v tem ℓ, então (i) $u + v$ tem no máximo $k + \ell$ e (ii) αu tem no máximo k.

Exercício 2.1.7 (Espaços vetoriais de funções)**.** O conjunto $\mathcal{F}(\mathbb{R}) := \mathcal{F}(\mathbb{R}, \mathbb{R})$ das funções reais é um espaço vetorial sob adição e multiplicação com constantes $\alpha \in \mathbb{R}$, veja Exercício 1.2.10. Para $n \in \mathbb{N}_0$ seja

$$\mathcal{P}_n(\mathbb{R}) := \{\alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n \mid \alpha_1, \ldots, \alpha_n \in \mathbb{R}\}$$

o conjunto dos **polinômios reais do grau menor ou igual n** e $\mathcal{P}(\mathbb{R}) := \cup_{n=0}^\infty \mathcal{P}_n(\mathbb{R})$ o conjunto de todos os **polinômios reais**. Seja

$$C^0(\mathbb{R}) := C^0(\mathbb{R}, \mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ é contínua}\}$$

o conjunto das **funções contínuas**. Para $k \in \mathbb{N}$ seja $C^k(\mathbb{R}) := C^k(\mathbb{R}, \mathbb{R})$ o conjunto das **funções k vezes continuamente diferenciáveis**. Chama-se

$$C^\infty(\mathbb{R}) := C^\infty(\mathbb{R}, \mathbb{R}) := \bigcap_{k=0}^{\infty} C^k(\mathbb{R})$$

o conjunto das **funções suaves**. Sejam $n \in \mathbb{N}_0$ e $k \in \mathbb{N}$. Mostre que

$$\mathcal{P}_n(\mathbb{R}) \subset \mathcal{P}(\mathbb{R}) \subset C^\infty(\mathbb{R}) \subset C^k(\mathbb{R}) \subset C^0(\mathbb{R}) \subset \mathcal{F}(\mathbb{R})$$

são subespaços do espaço vetorial $\mathcal{F}(\mathbb{R})$ do Exercício 1.2.10. Segundo parte b) do Lema 2.1.2 todos estes conjuntos são espaços vetoriais sob adição de funções e multiplicação com constantes.

Exemplo 2.1.8 (Hiperplanos no $\mathbb{R}^n$)**.** Dada uma lista $\alpha \in \mathbb{R}^n$, o subconjunto

$$\mathrm{H}_\alpha := \{x \in \mathbb{R}^n \mid \alpha_1 x_1 + \cdots + \alpha_n x_n = 0\}$$

é um subespaço de $\mathbb{R}^n$. O vetor nulo lida ao subespaço máximo $\mathrm{H}_\mathcal{O} = \mathbb{R}^n$. No caso não-nulo $\alpha \neq \mathcal{O}$ chama-se H_α de **hiperplano** no $\mathbb{R}^n$ passando a origem $\mathcal{O}$.

Lema 2.1.9 (Conjunto de subespaços é fechado sob interseções)**.** *Cada interseção $F := \cap_{\lambda\in\Lambda} F_\lambda$ de subespaços F_λ de um espaço vetorial E é um subespaço.*

Demonstração. Dado $u, v \in F := \cap_\lambda F_\lambda$, ou seja $u, v \in F_\lambda$ $\forall\lambda$. Como subespaço cada um F_λ é fechado sob adição, ou seja $u + v \in F_\lambda$ para todos os $\lambda \in \Lambda$. Em símbolos $u + v \in \cap_\lambda F_\lambda =: F$. Analogamente F é fechado sob mult. escalar. □

Exemplo 2.1.10. Dada uma matriz $\mathbf{a} = (a_{ij}) \in \mathrm{M}(m \times n)$, então o conjunto

$$F_{\mathbf{a}} := \{x \in \mathbb{R}^n \mid \mathbf{a}x = \mathcal{O}\}$$

é um subespaço de $\mathbb{R}^n$. Para ver isso lembramos de (A.3.1) que $\mathbf{a}x = \mathcal{O}$ é o SLH

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \quad \vdots \qquad\qquad \vdots \qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = 0 \end{cases}$$

para n incógnitas $x_1, \ldots, x_n \in \mathbb{R}$, notação $x := (x_1, \ldots, x_n)$. Note-se que as soluções x da primeira linha formam o hiperplano $\mathrm{H}_1 := \mathrm{H}_{\mathbf{a}_{1\bullet}}$ associado à primeira linha $\mathbf{a}_{1\bullet}$ da matriz $\mathbf{a}$. Isso é o certo ponto da vista, com efeito assim

$$F_{\mathbf{a}} = \mathrm{H}_1 \cap \cdots \cap \mathrm{H}_m$$

é uma interseção de subespaços e por isso é um subespaço segundo Lema 2.1.9.

Exercício 2.1.11.

1. Quais dos seguintes subconjuntos X_j são subespaços de $\mathbb{R}^n$? Em cada caso faça um desenho e explique porque é subespaço ou não é.
   (a) $X_1 := \{(\alpha, \alpha) \mid \alpha \in \mathbb{R}\} \subset \mathbb{R}^2$;
   (b) $X_2 := \{(\alpha + 1, \alpha) \mid \alpha \in \mathbb{R}\} \subset \mathbb{R}^2$;
   (c) $X_3 := \{(\alpha, \beta) \mid \alpha, \beta \text{ reais não-negativos}\} \subset \mathbb{R}^2$.
2. (LI transfere-se a espaços vetoriais *ambientes*). Seja F um subespaço de um espaço vetorial E. Mostre que se um subconjunto de F é LI em respeito ao espaço vetorial F então o também é LI em respeito ao espaço vetorial ambiente E.

2.2 Sistemas de geradores

Definição 2.2.1 (Subespaço gerado por um subconjunto)**.** Seja E um espaço vetorial e X um subconjunto. O **subespaço de E gerado por X** é o conjunto[1]

$$\langle X \rangle := \{\text{todas as combinações lineares em } X\} \cup \{\mathcal{O}\}$$

Dizemos que X é um **sistema de geradores** do subespaço $\langle X \rangle$. Observa que o conjunto vazio gera o subespaço trivial, ou seja $\langle \emptyset \rangle = \{\mathcal{O}\}$.

Se $\langle X \rangle = E$ dizemos que **o conjunto X gera E**. Neste caso cada um elemento de E é uma CL de elementos de X. Se $v_1, \ldots, v_\ell$ são vetores de E vamos usar a notação curta[2]

$$\langle v_1, \ldots, v_\ell \rangle := \langle \{v_1\} \cup \cdots \cup \{v_\ell\} \rangle \subset E$$

para o subespaço gerado e chamamos $v_1, \ldots, v_\ell$ um **sistema de geradores**.

[1] Lembre-se da nossa convenção (1.1.1) para conjuntos, por exemplo $\{1, 2\} \cup \{2\} = \{1, 2\}$.
[2] O ponto sutil é que uns dos vetores podem ser iguais, veja Definição 1.1.2.

Exercício 2.2.2. Mostre que $\langle X\rangle$ é um subespaço de E e que $\mathbb{K}v = \langle v\rangle$.

Lema 2.2.3. *Seja X um subconjunto de um espaço vetorial $(E, +, \cdot, \mathbb{K})$. Então*

(i) $X \subset \langle X\rangle$ (contido no subespaço gerado)

(ii) $Y \subset X \Rightarrow \langle Y\rangle \subset \langle X\rangle$ (naturalidade sob inclusão)

(iii) $F \subset E$ *subespaço* $\Rightarrow \langle F\rangle = F$ (não muda subespaços)

(iv) *Um subespaço $F \subset E$ contendo X contem $\langle X\rangle$.* (respeita subespaços)

Demonstração. (i) Seja $v \in X$, então $v \overset{(\text{comp.})}{=} 1v \in \langle X\rangle$. (ii) Como $Y \subset X$, CL's em Y são CL's em X. (iii) Igualdade é consequência das duas inclusões $F \subset \langle F\rangle \subset F$, onde a primeira é (i) e para a segunda usamos que os elementos de $\langle F\rangle$ são CL's em F, mas um subespaço é fechado sob CL's segundo Lema 2.1.2 a). (iv) Com efeito $F \overset{\text{(iii)}}{=} \langle F\rangle \overset{\text{(ii)}}{\supset} \langle X\rangle$. □

Lema 2.2.4. *Todo subconjunto LI $\{u, v\} \subset \mathbb{R}^2$ de dois elementos já gera $\mathbb{R}^2$.*

Demonstração. Lema B.2.1. □

Lema 2.2.5 (Os subespaços de $\mathbb{R}^2$)**.** *$\{\mathcal{O}\}$, $\mathbb{R}^2$, e as retas passando a origem.*

Demonstração. '$\supset$' Exercício 2.1.5. '$\subset$' Seja F um subespaço de $\mathbb{R}^2$. Caso $F = \{\mathcal{O}\}$, pronto. Caso contrario existe $u \in F$ não-nulo. Se os demais $f \in F$ são múltiplos de u temos $F = \mathbb{R}u$, pronto. Caso contrario existe um $v \in F$, não múltiplo de u. Então $\{u, v\}$ é LI segundo Exercício 1.3.8 parte 2. Mas neste caso segundo Lema 2.2.4 e Lema 2.2.3 (iv) obtemos $\mathbb{R}^2 = \langle u, v\rangle \subset F \subset \mathbb{R}^2$. □

Exemplo 2.2.6 (Os espaços $\mathbb{R}^n$, $\mathbb{R}^\infty_0$, $\mathbb{R}^\infty$)**.**

a) A base canônica $\mathcal{E}^n = \{e_1, \dots, e_n\}$, veja (1.2.1), gera $\mathbb{R}^n$. Com efeito

$$\mathbb{R}^n \ni v = \begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix} = \alpha_1 \begin{bmatrix}1\\ 0\\ \vdots\\ 0\end{bmatrix} + \cdots + \alpha_n \begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix}$$

A base canônica $\mathcal{E}^0 := \emptyset$ gera o subespaço vetorial trivial $\{0\} =: \mathbb{R}^0 \subset \mathbb{R}$.

b) Dado $i \in \mathbb{N}$, a sequência com todos membros nulos exceto o i-ésimo qual é 1 denotamos também de e_i. A **base canônica** $\mathcal{E}^\infty := \{e_1, e_2, \dots\}$ gera $\mathbb{R}^\infty_0$.

c) A base canônica $\mathcal{E}^\infty$ não gera $\mathbb{R}^\infty$: Uma CLe deve ser uma soma *finita*, tente escrever a sequência cujos membros são todos 1 como uma CL.

Exemplo 2.2.7 (Polinômios)**.** O conjunto de **monômios** $\{x^0, x, x^2, \dots, x^n\}$ onde $x^0 := 1$ gera $\mathcal{P}_n(\mathbb{R})$. Todos os monômios $\{1, x, x^2, \dots\}$ geram $\mathcal{P}(\mathbb{R})$.

Exemplo 2.2.8 (Sistemas lineares)**.** Dado um sistema linear $[\mathbf{a} : b]$ onde $\mathbf{a}$ é uma matriz $m \times n$. Sabemos de (A.3.2) que existe uma solução x se e somente se a lista b é CL das colunas da matriz $\mathbf{a}$. Consequentemente se as colunas de $\mathbf{a}$ formam um conjunto de geradores de $\mathbb{R}^m$, então para cada uma inomogeneidade $b \in \mathbb{R}^m$ o SL admite uma solução.

2.3 Soma direta

Definição 2.3.1 (Soma de subconjuntos)**.** A **soma de subconjuntos X e Y** de um espaço vetorial E é o conjunto de todas as somas

$$X + Y := \{x + y \mid x \in X,\ y \in Y\} \subset E$$

Em vez de $\{u\} + Y$ escreve-se $u + Y$ e chama-se **a translação de Y por u**.

Lema 2.3.2. *A soma de dois subespaços é gerado da união deles, em símbolos*

$$F, G \subset E \text{ subespaços} \quad \Rightarrow \quad F + G = \langle F \cup G \rangle$$

Particularmente, a soma de dois subespaços é um subespaço mesmo.

Demonstração. Para provar igualdade de dois conjuntos prova-se as duas inclusões. '$\subset$' Os elementos de $F + G$ são CL's da forma especial $f + g$ enquanto $\langle F \cup G \rangle$ contem todas as CL's em $F \cup G$.

'$\supset$' Pegue um elemento h de $\langle F \cup G \rangle$ e use comutatividade para re-escrever a soma finita com os somandos em F no frente e depois aqueles em G. Assim recebemos um elemento, igual h, em $F + G$. □

Definição 2.3.3 (Soma direta de subespaços)**.** Sejam $F_1, F_2 \subset E$ subespaços de um espaço vetorial E. No caso da interseção trivial $F_1 \cap F_2 = \{\mathcal{O}\}$ escreve-se $F_1 \oplus F_2$ em vez de $F_1 + F_2$ e chama-se **soma direta dos subespaços F_1 e F_2.**

O símbolo $F \oplus G$ é simplesmente uma abreviação para duas informações, com efeito

$$F \oplus G = H \quad \Leftrightarrow \quad \begin{cases} F \cap G = \{\mathcal{O}\} \\ F + G = H \end{cases}$$

Use-se a soma direta para decompor um vetor unicamente em componentes.

Teorema 2.3.4. *Sejam $F_1, F_2 \subset F$ três subespaços de um espaço vetorial E:*

$$F = F_1 \oplus F_2 \quad \Leftrightarrow \quad \forall f \in F,\ \exists!\ f_1 \in F_1,\ f_2 \in F_2 \ \text{ tal que } \ f = f_1 + f_2$$

Demonstração. Teorema B.2.2. □

Exercício 2.3.5. No espaço vetorial $\mathcal{F}(\mathbb{R})$ das funções $\mathbb{R} \to \mathbb{R}$ sejam

$F_1 = \{f : \mathbb{R} \to \mathbb{R}$ que se anulam em todos os pontos do intervalo [0,1]$\}$
$F_2 = \{g : \mathbb{R} \to \mathbb{R}$ que se anulam em todos os pontos do intervalo [2,3]$\}$

Mostre que F_1 e F_2 são subespaços de $\mathcal{F}(\mathbb{R})$, que $\mathcal{F}(\mathbb{R}) = F_1 + F_2$, mas não se tem $\mathcal{F}(\mathbb{R}) = F_1 \oplus F_2$.

Exercício 2.3.6. Verdadeiro ou falso? Para todos subconjuntos $X, Y \subset E$ vale

(i) $\langle X \cup Y \rangle = \langle X \rangle + \langle Y \rangle$

(ii) $\langle X \cap Y \rangle = \langle X \rangle \cap \langle Y \rangle$

Exercício 2.3.7. Uma matriz quadrada $\mathbf{a} = (a_{ij})_{i,j=1,\dots,n}$ chama-se

$$\textbf{simétrica} \text{ se } a_{ij} = a_{ji}\ \forall i,j \qquad \textbf{anti-simétrica} \text{ se } a_{ij} = -a_{ji}\ \forall i,j$$

Então as matrizes simétricas sao aquelas iguais às suas transpostas $\mathbf{a}^t = \mathbf{a}$ e as anti-simétricas aquelas com $\mathbf{a}^t = -\mathbf{a}$.

Prove que a) o conjunto $\mathcal{S} = \mathcal{S}(n)$ das matrizes simétricas $n \times n$ e o conjunto $\mathcal{A} = \mathcal{A}(n)$ das anti-simétricas são subespaços de $\mathrm{M}(n \times n; \mathbb{K})$ e b)

$$\mathrm{M}(n \times n; \mathbb{K}) = \mathcal{S} \oplus \mathcal{A}.$$

[Dica: b) Considere as duas matrizes $\mathbf{a}^{\pm} := \frac{1}{2}(\mathbf{a} \pm \mathbf{a}^t)$.]

Capítulo 3

Bases

Durante o Capítulo 3 denotamos de E um espaço vetorial

$$E = (E, +, \cdot, \mathbb{K})$$

sobre um corpo $\mathbb{K}$.

Bases de um espaço vetorial E sao subconjuntos LI as quais geram E no sentido que todo vetor de E pode ser escrito como combinação linear (CL) de elementos da base. Os coeficientes escalares na CL são únicos (propriedade LI) e chamados de coordenadas de um vetor em respeito à base. Quando E admite uma base finita de n elementos chama-se n a dimensão de E. Se escolhemos uma outra base, recebemos uma outra dimensão? Veremos na Seção 3.1.2 que não: Se E admite uma base finita todas as bases tem o mesmo número de elementos.

Então bases são LI, contem suficientemente muitos elementos para que todo vetor pode ser escrito como CL deles, e na dimensão finita bases ainda são conjuntos máximos no sentido que adicionando mais um outro vetor só já recebe-se um conjunto LD.

Definição 3.0.8 (Base)**.** Para um subconjunto $\mathcal{B}$ de E definimos

$$\mathcal{B} \textbf{ base de } E \quad :\Leftrightarrow \quad \begin{cases} \mathcal{B} \text{ gera } E \\ \mathcal{B} \text{ é LI} \end{cases}$$

O número de elementos de $\mathcal{B}$ pode ser finito ou infinito. Uma **base ordenada** é uma base $\mathcal{B} = \{\xi_1, \xi_2, \dots\}$ cujos elementos são enumerados. Equivalentemente podemos escrever uma base ordenada na forma de uma lista ordenada $(\xi_1, \xi_2, \dots)$, finita ou infinita.

A observação chave é essa: Se uma base $\mathcal{B}$ de um espaço vetorial contem exatamente m elementos, então todas as bases contêm exatamente m elementos. Isso será o conteúdo da Proposição 3.1.12. Vamos aproveitar do resultado já:

Definição 3.0.9 (Dimensão)**.** Se um espaço vetorial E admite uma base finita, dizemos $\mathcal{B}$, então o número dos elementos é dito a **dimensão de E**, em símbolos

$$\dim E := |\mathcal{B}|$$

Se E não admite uma base finita, então dizemos que E é de **dimensão infinita** e escrevemos $\dim E = \infty$.

Comentário 3.0.10 (Dimensão do espaço vetorial trivial)**.** O conjunto vazio é uma base do espaço vetorial trivial $E = \{\mathcal{O}\}$, Exemplo 1.3.4, assim $\dim\{\mathcal{O}\} = 0$.

Exemplo 3.0.11. Sejam $u = (1, 1)$ e $v = (2, 0)$. Os conjuntos $\{e_1, u\}$ e $\{u, v\}$ são bases de $\mathbb{R}^2$. Ambos conjuntos são LI segundo Exercício 1.3.8 3. (os elementos não são múltiplos um do outro, veja Teorema 3.1.1) e por isso geram $\mathbb{R}^2$ (Lema B.2.1). Um exemplo para LD é o conjunto $\{e_1, v\}$, no qual um elemento é múltiplo do outro.

Exercício 3.0.12 (Para soma direta a união de bases é base)**.** Seja $E = F_1 \oplus F_2$. a) Mostre que uma união $\mathcal{B}_1 \cup \mathcal{B}_2$ de bases de F_1 e F_2 é uma base de E.

[Dica: a) União LI – ideia de (B.2.1).]

Exemplo 3.0.13 (Bases canônicas e dimensões)**.** Análogo para corpos gerais $\mathbb{K}$.

a) **Listas.** A base canônica $\mathcal{E}^n := \{e_1, \ldots, e_n\}$ é base de $\mathbb{R}^n$ e assim

$$\dim \mathbb{R}^n := |\mathcal{E}^n| = n, \quad n \in \mathbb{N}_0$$

Caso $n \geq 1$: Exercício 1.3.8 confirma LI, Exercício 2.2.6 diz que gera $\mathbb{R}^n$. Caso $n = 0$: Note que $\mathbb{R}^0 = \{0\}$ é o espaço vetorial trivial. O conjunto vazio $\mathcal{E}^0 = \emptyset$ é LI (Comentário 1.3.4) e gera o espaço trivial (Definição 2.2.1).

b) **Sequências.** A base canônica $\mathcal{E}^\infty := \bigcup_{j=1}^{\infty}\{e_j\}$ é base de $\mathbb{R}_0^\infty$ e assim

$$\dim \mathbb{R}_0^\infty := |\mathcal{E}^\infty| = \infty$$

Base de $\mathbb{R}_0^\infty$, não de $\mathbb{R}^\infty$: O conjunto $\mathcal{E}^\infty$ é LI em $\mathbb{R}_0^\infty$ e em $\mathbb{R}^\infty$ (Exercício 1.3.8). Mas enquanto $\mathcal{E}^\infty$ gera $\mathbb{R}_0^\infty$, não gera $\mathbb{R}^\infty$, veja Exemplo 2.2.6.

Ainda que não temos na mão uma base de $\mathbb{R}^\infty$ no Lema 3.1.16 mostramos

$$\dim \mathbb{R}^\infty = \infty$$

Isso é baseado no fato que $\mathbb{R}^\infty$ possui um subespaço de dimensão infinita.

c) **Matrizes.** Seja $\mathbf{e^{ij}} \in \mathrm{M}(m \times n)$ a matriz com todas entradas nulas exceto a ij-ésima entrada a qual é $(\mathbf{e^{ij}})_{ij} = 1$. A base canônica de $\mathrm{M}(m \times n)$ é

$$\mathcal{E}^{m\times n} := \{\mathbf{e^{ij}} \mid i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\} \subset \mathrm{M}(m \times n)$$

Como a base canônica tem $|\mathcal{E}^{m\times n}| = mn$ elementos obtemos

$$\dim \mathrm{M}(m \times n) = mn$$

d) **Polinômios.** A base canônica são os monômios $\{x^n \mid n \in \mathbb{N}_0\} \subset \mathcal{P}(\mathbb{R})$.

Gerando: Por definição de $\mathcal{P}(\mathbb{R})$. LI: Uma CL de monômios é um polinômio $p \not\equiv 0$ (não constantemente nulo). Se p representa o vetor nulo (o polinômio constantemente nulo) todos coeficientes devem se anular porque polinômios de grau $\ell \geq 1$ tem um número finito de raízes. Caso a p é de grau zero, ele é da forma $p(x) = \alpha_0 x^0$ e para se anular α_0 deve-se anular.

Analogamente $\{1, x, \ldots, x^n\}$ é uma base de $\mathcal{P}_n(\mathbb{R})$. Consequentemente

$$\dim \mathcal{P}(\mathbb{R}) = \infty, \qquad \dim \mathcal{P}_n(\mathbb{R}) = n+1, \quad n \in \mathbb{N}_0 \tag{3.0.1}$$

e) **Hiperplanos.** Dado uma lista $\alpha \in \mathbb{R}^n$ com $\alpha_n \neq 0$, o hiperplano

$$\mathrm{H}_\alpha := \{x \in \mathbb{R}^n \mid \alpha_1 x_1 + \cdots + \alpha_n x_n = 0\}$$

tem como base o conjunto $\mathcal{B}_\alpha := \{\xi_1, \ldots \xi_{n-1}\}$ no qual a lista

$$\xi_i := \left(0, \ldots, 0, 1, 0, \ldots, 0, -\tfrac{\alpha_i}{\alpha_n}\right) \in \mathbb{R}^n, \quad i = 1, \ldots, n-1$$

tem todos membros nulos exceto o i-ésimo e o último. É óbvio que $\mathcal{B}_\alpha$ é LI, que gera $\mathbb{R}^n$ podemos ver assim: Para $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ vale

$$\begin{aligned}
& x \in \mathrm{H}_\alpha \\
\Leftrightarrow \quad & 0 = \alpha_1 x_1 + \cdots + \alpha_n x_n \\
\Leftrightarrow \quad & x_n = -\tfrac{\alpha_1}{\alpha_n} x_1 - \cdots - \tfrac{\alpha_{n-1}}{\alpha_n} x_{n-1} \\
\Leftrightarrow \quad & x = \left(x_1, \ldots, x_{n-1}, -\tfrac{\alpha_1}{\alpha_n} x_1 - \cdots - \tfrac{\alpha_{n-1}}{\alpha_n} x_{n-1}\right) \in \mathbb{R}^n \\
\Leftrightarrow \quad & x = x_1 \xi_1 + \cdots + x_{n-1} \xi_{n-1}
\end{aligned}$$

Consequentemente

$$\dim \mathrm{H}_\alpha = n-1, \quad \alpha \in \mathbb{R}^n, \, \alpha \neq \mathcal{O}$$

Note que $\dim \mathrm{H}_\mathcal{O} = n$.

A parte **hiper** em hiperplano refere-se ao fato do que na dimensão falta 1 para a dimensão do espaço vetorial ambiente, no caso presente $\mathrm{H}_\alpha \subset \mathbb{R}^n$.

Comentário 3.0.14 (Corpo geral $\mathbb{K}$)**.** Todas afirmações no Exemplo 3.0.13 ficam verdadeiro se em vez do corpo $\mathbb{R}$ usa-se um corpo geral $\mathbb{K}$.

Exercício 3.0.15 (Produto cartesiano)**.** Seja $n \in \mathbb{N}_0$ e seja F um espaço vetorial de dimensão m. Mostre que o produto cartesiano $F^{\times n}$, veja (1.1.2), é um espaço vetorial sob as operações de adição e multiplicação escalar, ambas componente-por-componente, e que $\dim F^{\times n} = mn$.

[Dica: Dimensão – escolha uma base de F e use para definir uma base de $F^{\times n}$.]

3.1 Aplicações

3.1.1 Coordenadas de um vetor

Teorema 3.1.1. *Seja $X \subset E$ um subconjunto tal que $|X| \geq 2$. Então*

a) *X é LI $\Leftrightarrow$ nenhum elemento de X é CL de outros elementos de X*

b) *X é LD $\Leftrightarrow$ existe um elemento de X que é CL de outros elementos de X*

Demonstração. a) '$\Rightarrow$' Seja X LI, suponha por absurdo que um elemento $u \in X$ fosse CL $u = \alpha_1 v_1 + \cdots + \alpha_k v_k$ de outros elementos v_j (tem outros como $|X| \geq 2$). Adicionando $-u$ em ambos lados obtemos $\mathcal{O} = (-1)u + \alpha_1 v_1 + \cdots + \alpha_k v_k$. Como $-1 \neq 0$ trata-se de uma CL não-trivial em X representando o vetor nulo. Assim X é LD. Contradição.
'$\Leftarrow$' Suponha por absurdo X fosse LD. Então existe uma CL não-trivial em X

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = \mathcal{O}$$

representando o vetor nulo. Pelo menos um dos α_i é não-nulo. Renomeando podemos supor $\alpha_1 \neq 0$. Caso $k = 1$. Então $\alpha_1 v_1 = \mathcal{O}$ e assim $v_1 = \alpha_1^{-1}\mathcal{O} = \mathcal{O}$. Contradição. Caso $k \geq 2$. Então $v_1 = -\alpha_1^{-1}\alpha_2 v_2 - \cdots - \alpha_1^{-1}\alpha_k v_k$ é CL de outros elementos de X. Contradição. b) é equivalente à parte a). □

Corolário 3.1.2 (Unicidade dos coeficientes de CL's em conjuntos LI)**.** *Seja $\{v_1, \ldots, v_k\}$ um subconjunto LI de E, então*

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = \beta_1 v_1 + \cdots + \beta_k v_k \quad \Rightarrow \quad \alpha_1 = \beta_1, \ldots, \alpha_k = \beta_k$$

Em palavras, se duas CL's num conjunto LI representam o mesmo vetor, então os coeficientes escalares coincidem.

Demonstração. $\alpha_1 - \beta_1 = 0$: Suponha por absurdo $\alpha_1 - \beta_1 \neq 0$. Então o vetor

$$v_1 = (\alpha_1 - \beta_1)^{-1}\left((\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_k - \beta_k)v_k\right)$$

é CL de outros elementos. Contradição (Teorema 3.1.1 a)). Análogo para os outros $\alpha_j - \beta_j$. Outro argumento (usando LI): Suponha $(\alpha_1 - \beta_1)v_1 + \cdots + (\alpha_k - \beta_k)v_k = \mathcal{O}$. LI diz que todos coeficientes são nulos. □

Lema 3.1.3.

a) Um subconjunto Y de um conjunto LI X é LI. *(Subconjuntos herdam LI)*

b) Um conjunto X contendo um Y LD é LD. *(Superconjuntos herdam LD)*

c) Um subconjunto LI X num subespaço $F \subset E$, também é LI em E. *(LI transfere-se para superespaços)*

Demonstração. a) Como X é LI, toda CL em X representando $\mathcal{O}$ tem todos coeficientes nulos. Como $Y \subset X$, toda tal CL em Y é uma em X e assim tem todos coeficientes nulos. b) Como $Y \subset X$, uma CL não-trivial em Y representando $\mathcal{O}$ é uma tal em X. c) Isso resta no fato que o vetor nulo de um subespaço é o vetor nulo do espaço vetorial ambiente, veja Exercício 2.1.3. □

Comentário 3.1.4 (Consequências das duas propriedades de ser base $\mathcal{B}$ de E)**.**
$\langle B\rangle = E$: Assim todo $v \in E$ pode ser escrito como CL em $\mathcal{B}$, com efeito

$$v = \alpha_1\xi_1 + \cdots + \alpha_k\xi_k \tag{3.1.1}$$

para escalares $\alpha_i \in \mathbb{K}$ e vetores $\xi_j \in \mathcal{B}$ da base.
$\mathcal{B}$ é LI: Assim os coeficientes α_j em cima são únicos (Corolário 3.1.2).

> *Todo vetor $v \in E$ admite coordenadas únicas $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ em respeito a uma base ordenada $\mathcal{B}$ de E.*

Definição 3.1.5 (Coordenadas)**.** Suponha $\mathcal{B} = \{\xi_1, \ldots, \xi_n\}$ é uma base ordenada de um espaco vetorial E. As **coordenadas** de um vetor $v \in E$ em respeito à base $\mathcal{B}$ são os coeficientes $\alpha_1, \ldots, \alpha_n$ em (3.1.1). A matriz coluna $n \times 1$ das coordenadas

$$[v]_{\mathcal{B}} := \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{K}^n \tag{3.1.2}$$

é chamado de **vetor coordenada** de v em respeito à base $\mathcal{B}$. Abreviamos $[v] := [v]_{\mathcal{E}^m}$ no caso de $E = \mathbb{K}^m$ munido da base canônica $\mathcal{E}^m$.

Lema 3.1.6. *Duas bases ordenadas $\mathcal{B} = \{\xi_1, \ldots, \xi_n\}$ e $\widetilde{\mathcal{B}}$ de E são iguais se e somente se cada um elemento de E tem o mesmo vetor coordenada em respeito a $\mathcal{B}$ e a $\widetilde{\mathcal{B}}$. Em símbolos*

$$[v]_{\mathcal{B}} = [v]_{\widetilde{\mathcal{B}}} \quad \forall v \in E \qquad \Leftrightarrow \qquad \mathcal{B} = \widetilde{\mathcal{B}}$$

Demonstração. "⇒" A hipótese para $v := \xi_1$ diz que $[\xi_1]_{\mathcal{B}} = [\xi_1]_{\widetilde{\mathcal{B}}}$. Note-se que $[\xi_1]_{\mathcal{B}} = (1, 0, \ldots, 0)$. E assim $[\xi_1]_{\widetilde{\mathcal{B}}} = (1, 0, \ldots, 0)$. Mas isso significa que $\xi_1 = 1 \cdot \tilde{\xi}_1 + 0 \cdot \tilde{\xi}_2 + \cdots + 0 \cdot \tilde{\xi}_n = \tilde{\xi}_1$. Repita para $v = \xi_2, \ldots, \xi_n$. "⇐" óbvio. □

Exercício 3.1.7. Seja $E = \mathbb{R}^2$ munido da base canônica $\mathcal{E} = \{e_1, e_2\}$ e da base ordenada $\mathcal{B} = \{\xi_1, \xi_2\}$ onde $\xi_1 = (1, 1)$ e $\xi_2 = (-1, 1)$. Determine $[e_1]_{\mathcal{B}}$, $[e_2]_{\mathcal{B}}$, $[e_1]$, $[e_2]$ e também $[\xi_1]_{\mathcal{B}}$, $[\xi_2]_{\mathcal{B}}$, $[\xi_1]$, $[\xi_2]$.

Exercício 3.1.8. Mostre que os polinômios 1, $x-1$, e x^2-3x+1 formam uma base de $\mathcal{P}_2(\mathbb{R})$. Exprima o polinômio $2x^2 - 5x + 6$ como CL nessa base.

3.1.2 Dimensão de um espaço vetorial

Teorema 3.1.9. *Se um conjunto finito gera E, então qualquer conjunto $Y \subset E$ com mais elementos é LD.*

Corolário 3.1.10. *Suponha um conjunto finito X gera E, então*

$$Y \subset E \text{ LI} \qquad \Rightarrow \qquad |Y| \leq |X|.$$

Para provar Teorema 3.1.9 vamos usar o seguinte resultado sobre SLH's.

Teorema 3.1.11 (Existência de soluções não-triviais de um SLH)**.** *Dado uma matriz* $\mathbf{a} \in \mathrm{M}(m \times n; \mathbb{K})$*. Se a matriz tem menos linhas como colunas* $(m < n)$*, então o SLH* $\mathbf{a}x = \mathcal{O}$*, compare (A.3.1), admite soluções não-triviais* $x = (x_1, \ldots, x_n) \neq (0, \ldots, 0)$*.*

Demonstração. Indução sobre o número m de linhas. Veja Teorema B.3.1. □

Demonstração de Teorema 3.1.9. Suponha que o conjunto $X = \{v_1, \ldots, v_m\}$ gera o espaço vetorial E e seja $Y \subset E$ um outro subconjunto com mais elementos, ou seja $|Y| > m$. Para mostrar que Y é LD, basta mostrar segundo Lema 3.1.3 b) que um subconjunto $U = \{u_1, \ldots, u_{m+1}\} \subset Y$ de $m+1$ elementos é LD. Como X gera E e cada um u_j pertence a E existem escalares a_{ij} tal que

$$u_j = a_{1j}v_1 + \cdots + a_{mj}v_m \tag{$*_j$}$$

para $j = 1, \ldots, m+1$.

Para U é LD resta mostrar: existem escalares não-nulos $x_1, \ldots, x_{m+1}$ tal que

$$x_1u_1 + \cdots + x_{m+1}u_{m+1} = \mathcal{O} \tag{3.1.3}$$

Para este fim considere o SLH de m equações de $n = m+1$ incógnitas x_j

$$\begin{cases} a_{11}x_1 + \cdots + a_{1,m+1}x_{m+1} = 0 \\ \quad\vdots \qquad\qquad\qquad\quad \vdots \qquad \vdots \\ a_{m1}x_1 + \cdots + a_{m,m+1}x_{m+1} = 0 \end{cases} \tag{SLH}$$

o qual tem uma solução não-trivial $x = (x_1, \ldots, x_{m+1}) \neq (0, \ldots, 0)$ segundo Teorema 3.1.11 como $m < n$. Obtemos (3.1.3) assim: usando $(*_1 - *_{m+1})$ temos

$$\begin{aligned} & x_1u_1 + \cdots + x_{m+1}u_{m+1} \\ &= x_1\left(a_{11}v_1 + \cdots + a_{m1}v_m\right) \\ &\quad + x_2\left(a_{12}v_1 + \cdots + a_{m2}v_m\right) \\ &\quad \vdots \\ &\quad + x_{m+1}\left(\alpha_{1,m+1}v_1 + \cdots + \alpha_{m,m+1}v_m\right) \\ &= v_1 \underbrace{\sum_{j=1}^{m+1} a_{1j}x_j}_{=0\ (\mathrm{SLH})_1} + \cdots + v_m \underbrace{\sum_{j=1}^{m+1} a_{mj}x_j}_{=0\ (\mathrm{SLH})_m} \\ &= \mathcal{O} \end{aligned}$$

□

Proposição 3.1.12. *Se uma base* $\mathcal{B}$ *de um espaço vetorial contem exatamente* m *elementos, então todas bases contêm exatamente* m *elementos.*

Demonstração. Sejam $\mathcal{B} = \{\xi_1, \dots, \xi_m\}$ e $\tilde{\mathcal{B}}$ bases de E.
1) $\langle \mathcal{B} \rangle = E$ e $Y := \tilde{\mathcal{B}}$ LI implicam (Corolário 3.1.10) $\ell := |\tilde{\mathcal{B}}| \leq |\mathcal{B}| = m < \infty$.
2) Analogamente como $\langle \tilde{\mathcal{B}} \rangle = E$ e $Y := \mathcal{B}$ é LI, temos que $m = |\mathcal{B}| \leq |\tilde{\mathcal{B}}| = \ell$. □

A nocao de dimensão é baseada nessa proposição: Se um espaço vetorial E admite uma base finita, dizemos $\mathcal{B}$, então o número dos elementos é dito a **dimensão de E**, em símbolos

$$\dim E := |\mathcal{B}|$$

Caso E não admite nenhuma base finita dizemos que E é de **dimensão infinita** e escrevemos $\dim E = \infty$.

Lema 3.1.13 (Aumentando conjuntos LI)**.** *Seja $X = \{v_1, \dots, v_k\}$ um subconjunto LI e seja $u \in E \setminus \langle X \rangle$ um vetor de E mas não em $\langle X \rangle$. Então o conjunto extendido $\{v_1, \dots, v_k, u\}$ também é LI.*

Demonstração. Se $X = \emptyset$, então $u \notin \langle \emptyset \rangle = \{\mathcal{O}\}$, assim $u \neq \mathcal{O}$ e $\{u\}$ é LI segundo Corolário 1.3.4. Se $X \neq \emptyset$, suponha por absurdo que $\{v_1, \dots, v_k, u\}$ fosse LD. Assim existe uma CL não-trivial $\alpha_1 v_1 + \cdots + \alpha_k v_k + \beta u = \mathcal{O}$. Caso $\beta = 0$, entao $(\alpha_1, \dots, \alpha_k) \neq (0, \dots, 0)$ e $\alpha_1 v_1 + \cdots + \alpha_k v_k = \mathcal{O}$. Assim $\{v_1, \dots, v_k\}$ é LD. Contradição. Caso $\beta \neq 0$, então $u = -\beta^{-1}(\alpha_1 v_1 + \cdots + \alpha_k v_k) \in \langle X \rangle$. Contradição. □

Exercício 3.1.14. Sejam $X_1, X_2, \dots$ subconjuntos LI de um espaço vetorial E.

1. Caso são encaixados $X_1 \subset X_2 \subset \dots$, prove que $X = \bigcup X_n$ é LI.
2. Se cada X_n tem n elementos, prove que existe um conjunto LI $\tilde{X} = \{x_1, x_2, \dots\}$ com $x_j \in X_j$, para todo $j \in \mathbb{N}$.
3. Supondo $E = \mathbb{R}^\infty$ e as hipóteses em 1. e 2., é verdadeiro que $X = \bigcup X_n$ seja uma base de E?

Corolários do Teorema 3.1.9

Nos corolários seguintes $n \in N_0$, particularmente é um número, assim finito.

Corolário 3.1.15. *$Y \subset E$, $|Y| > n := \dim E \;\Rightarrow\; Y$ LD*

Demonstração. Como $\dim E = n$ existe uma base $\mathcal{B}$ de E com n elementos. □

Lema 3.1.16. *A dimensão de $\mathbb{R}^\infty$ é infinito.*

Demonstração. Suponha por absurdo que é finita a dimensão $k := \dim \mathbb{R}^\infty$. Segundo Corolário 3.1.15 para $Y = \mathcal{E}^\infty$ e $E = \mathbb{R}^\infty$, como $|Y| = |\mathcal{E}^\infty| = \infty > k = \dim \mathbb{R}^\infty$, segue que $\mathcal{E}^\infty$ é LD em $\mathbb{R}^\infty$. Mas $\mathcal{E}^\infty$ é LI em $\mathbb{R}^\infty$ segundo Exercício 1.3.8. (Alternativamente, como $\mathcal{E}^\infty$ é LI em $\mathbb{R}_0^\infty$, $\mathcal{E}^\infty$ deve ser LI em $\mathbb{R}^\infty$ segundo parte c) do Lema 3.1.3.) Contradição. □

Corolário 3.1.17. *Se um conjunto Y é LI em E, então $|Y| \leq \dim E$.*

Demonstração. Caso $\dim E = \infty$: verdadeiro trivialmente. Caso $\dim E < \infty$: escolha para X em Corolário 3.1.10 uma base de E para obter $|Y| \leq \dim E$. □

Corolário 3.1.18. *Suponha $X \subset E$ tem $n := \dim E$ elementos, então*

$$X \text{ gera } E \qquad \Leftrightarrow \qquad X \text{ é LI}$$

Demonstração. $\boldsymbol{n = 0}$**.** Assim $X = \emptyset$, ambos lados valem automaticamente.
$\boldsymbol{n = 1}$**.** Assim $X = \{v\}$ onde $v \in E$, ambos lados são equivalentes a $v \neq \mathcal{O}$.
$\boldsymbol{n = 2}$**.** '$\Rightarrow$' Suponha que $X = \{v_1, \ldots, v_n\}$ gera E. Por absurdo suponha que X é LD. Segundo Teorema 3.1.1 b) um elemento de X, dizemos v_n, é CL de outros elementos de X. Então $E = \langle X \rangle = \langle v_1, \ldots, v_n \rangle = \langle v_1, \ldots, v_{n-1} \rangle$. Qualquer base $\mathcal{B}$ de E tem n elementos pela hipótese $n = \dim E$ – mais elementos como o conjunto $\{v_1, \ldots, v_{n-1}\}$ gerando E. Então $\mathcal{B}$ é LD segundo Teorema 3.1.9. Contradição. '$\Leftarrow$' Suponha que $X = \{v_1, \ldots, v_n\}$ é LI. Por absurdo suponha que X não gera E. Então existe $u \in E$ não elemento de $\langle v_1, \ldots, v_n \rangle$. Assim o conjunto aumentado $\{v_1, \ldots, v_n, u\}$ é LI segundo Lema 3.1.13. Mas um subconjunto com mais elementos $(n + 1)$ como a dimensão (n) é LD segundo Corolário 3.1.15. Contradição. □

Corolário 3.1.19. *Um subconjunto LI com $n = \dim E$ elementos é uma base.*

Demonstração. Tal subconjunto LI gera E segundo Corolário 3.1.18 '$\Leftarrow$'. □

Lema 3.1.20. *Se um conjunto finito X gera E, então $|X| \geq \dim E$.*

Demonstração. Suponha que $X = \{v_1, \ldots, v_m\}$ gera E.
Caso $X = \{v_1, \ldots, v_m\}$ é LI: Então X é base e assim $|X| = \dim E$.
Caso $X = \{v_1, \ldots, v_m\}$ é LD: Assim $X \neq \emptyset$.
Subcaso $m = 1$: Então $v_1 = \mathcal{O}$ e $E = \langle v_1 \rangle = \{\mathcal{O}\}$. Assim $|X| = 1 > 0 = \dim E$.
Subcaso $m \geq 2$: Como $\{v_1, \ldots, v_m\}$ é LD, pelo menos um elemento, dizemos v_m, deve ser CL de outros. Iterando até chegamos num conjunto LI obtemos que

$$E = \langle v_1, \ldots, v_m \rangle = \langle v_1, \ldots, v_{m-1} \rangle = \cdots = \langle v_1, \ldots, v_\ell \rangle$$

onde $\{v_1, \ldots, v_\ell\}$ é LI e $\ell \geq 1$. Então $\{v_1, \ldots, v_\ell\} =: \mathcal{B}$ é base de E e assim $|X| > \ell = |\mathcal{B}| =: \dim E$. □

3.1.3 Escalonamento: Dimensão de um subespaço gerado

Aplicamos o processo de escalonar uma matriz **a** – conteúdo do curso MA141 e revisado no Apêndice A.2 – para calcular a dimensão do subespaço gerado por m vetores.

Consideramos m vetores $v_1, \ldots, v_m$ do espaço vetorial $\mathbb{K}^n$. (No caso geral de um espaço vetorial E de dimensão n use uma base para chegar em $\mathbb{K}^n \simeq E$.)

Escreve as m listas $v_1, \dots, v_m \in \mathbb{K}^n$ como linhas de uma matriz $m \times n$, ou seja

$$\mathbf{a} = (a_{ij}) := \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{matrix} \leftarrow v_1 \\ \vdots \\ \leftarrow v_m \end{matrix}$$

Escalonamento da matriz $\mathbf{a}$, veja Seção A.2, lida à matriz escalonada $\mathbf{a}_{\text{esc}}$. Enumere as linhas não-nulas de $\mathbf{a}_{\text{esc}}$ de cima para baixo, dizemos $\ell_1, \dots, \ell_d$.

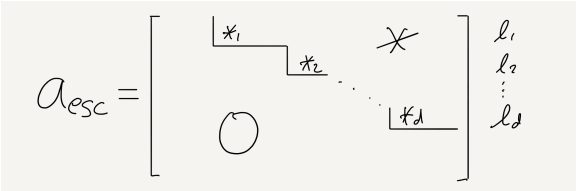


Figura 3.1: Linhas não-nulas $\ell_1, \dots, \ell_d$ da matriz escalonada $\mathbf{a}_{\text{esc}}$

É fácil checar que estas linhas formam um conjunto LI, com efeito

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \alpha_1\ell_1 + \alpha_2\ell_2 + \dots + \alpha_d\ell_d = \begin{bmatrix} \alpha_1 * _1 \\ \alpha_1 * _1 + \alpha_2 * _2 \\ \vdots \\ \dots \alpha_{d-1} * _{d-1} + \alpha_d * _d \end{bmatrix} \begin{matrix} \Rightarrow \alpha_1 = 0 \\ \Rightarrow \alpha_2 = 0 \\ \vdots \\ \Rightarrow \alpha_d = 0 \end{matrix}$$

Então $\{\ell_1, \dots, \ell_d\}$ é uma base do Esp-lin($\mathbf{a}_{\text{esc}}$), qual íguale Esp-lin($\mathbf{a}$) porque operações elementares não mudam o espaço linha, veja Teorema A.2.2. Então

$$\langle v_1, \dots, v_m \rangle = \text{Esp-lin}(\mathbf{a}) = \text{Esp-lin}(\mathbf{a}_{\text{esc}}) \subset \mathbb{K}^n$$

é um subespaço com base as linhas não-nulas $\{\ell_1, \dots, \ell_d\}$ da matriz $\mathbf{a}_{\text{esc}}$.

Exercício 3.1.21. Obtenha uma base para o subespaço F de $\mathbb{R}^4$ gerado por

$$v_1 = (1, 2, 3, 4), \qquad v_2 = (3, 4, 7, 10), \qquad v_3 = (2, 1, 3, 5)$$

Determine a dimensão de F.

3.1.4 Complexificação e realificação

3.2 Existência e extensão

Teorema 3.2.1. *Seja E da dimensão finita $n \in \mathbb{N}_0$.*

(a) *Todo conjunto gerando E contem uma base de E.* *(Existência de bases)*

(b) *Todo subconjunto LI é contido numa base de E.* *(Extensão de bases)*

(c) *A dimensão de qualquer subespaço de E é $\leq n$.* *(Dimensão)*

(d) *Um subespaço F de E da mesma dimensão n é igual a E.*

Demonstração. LI refere-se a E se não especificado diferente. **(a)** Suponha X gera E. Seja $B \subset X$ qualquer subconjunto LI (existe como $B = \emptyset$ mostra), então $|B| \leq \dim E =: n$ segundo Corolário 3.1.17. Para $k \in \mathbb{N}_0$ seja

$$\mathcal{C}_k := \{B \subset X \mid B \text{ é LI e } |B| = k\}$$

a família de todos os subconjuntos $B \subset X$ os quais são LI e composto de k elementos. Seja $B_* \subset X$ um subconjunto LI com o número máximo de elementos. Então $B_* \in \mathcal{C}_\ell$ para um $\ell \in \{0, 1, \ldots, n\}$. Seja $B_* = \{\xi_1, \ldots, \xi_\ell\}$. Considere as quatro inclusões (dois deles sendo igualdades)

$$E = \langle X \rangle \subset \langle \langle B_* \rangle \rangle = \langle B_* \rangle \subset E$$

Consequentemente o conjunto LI B_* gera E, ou seja B_* é uma base. Resta justificar as quatro inclusões. Inclusão 1. Pela hipótese X gera E.
Inclusão 2. Parte (ii) de Lema 2.2.3 aplica como $X \subset \langle B_* \rangle$: Suponha por absurdo que existe um vetor $v \in X$ o qual não é CL em B_*, ou seja $v \notin \langle B_* \rangle$. Segundo Lema 3.1.13 o subconjunto aumentado $\{\xi_1, \ldots, \xi_\ell, v\}$ de X ainda é LI, mas contem $\ell + 1$ elementos, então mais como B_*. Contradição.
Inclusão 3. Como $\langle B_* \rangle$ é um subespaço parte (iii) de Lema 2.2.3 aplica.
Inclusão 4. Como $B_* \subset X \subset E$ parte (ii) de Lema 2.2.3 aplica.

(b) Suponha $X \subset E$ é um subconjunto LI. Como $k := |X| \in \{0, \ldots, n\}$, veja Corolário 3.1.17, trata-se de um conjunto finito, ou seja $X = \{v_1, \ldots, v_k\}$. Subconjuntos $B \subset E$ LI e contendo X (existem como $B := X$ mostra) são compostos de ℓ elementos para um $\ell \in \{k, \ldots, n\}$. Seja B_* um tal subconjunto com o número máximo ℓ_* de elementos. Então B_* é LI e contem $X \subset B_*$. Para B_* é uma base de E, resta mostrar que gera E, ou seja $\langle B_* \rangle = E$:
'$\subset$' trivial como $B_* \subset E$. '$\supset$' Suponha por absurdo que existe um vetor $u \in E$ o qual não pertence a $\langle B_* \rangle$, então o conjunto aumentado $B_* \cup \{u\}$ é LI segundo Lema 3.1.13, contem X porque B_* contem X – mas tem mais elementos como B_*. Contradição.

(c) Suponha F é um subespaço de E. Seja $B \subset F$ qualquer subconjunto LI em respeito a F (existe como $B = \emptyset$ mostra). Note que B é LI em respeito a E segundo Lema 3.1.3 c). Assim $|B| \leq n := \dim E$ segundo Corolário 3.1.17. Agora escolha um subconjunto $B_* \subset F$ LI em respeito a F com o número máximo de elementos. Como temos visto $k := |B_*| \leq \dim E =: n$. Resta mostrar que B_* é uma base de F (neste caso $\dim F = |B_*|$). Pela escolha B_* é LI em F, então basta mostrar $\langle B_* \rangle = F$:
'$\subset$' trivial como $B_* \subset F$. '$\supset$' Suponha por absurdo que existe um vetor $u \in F$ o qual não pertence a $\langle B_* \rangle$, então o conjunto aumentado $B_* \cup \{u\}$ é LI em F segundo Lema 3.1.13 – mas tem mais elementos como B_*. Contradição.

(d) Seja $F \subset E$ um subespaço de dimensão $n := \dim E$. Pela definição de dimensão existe uma base $\mathcal{B}$ de F com n elementos. Como $\mathcal{B}$ é LI em

respeito a F, é LI em respeito a E segundo Lema 3.1.3 c). Como além disso $|\mathcal{B}| = n := \dim E$ o Corolário 3.1.18 diz que $\mathcal{B}$ gera E. Então $E = \langle \mathcal{B} \rangle = F$, onde a segunda igualdade segue porque $\mathcal{B}$ é base de F, então gera F. □

Proposição 3.2.2. *Seja F um espaço vetorial e F_1, F_2 subespaços de dimensões finitas k, ℓ. Então existe uma base finita $\mathcal{B}$ do subespaço F_1+F_2 de F que contem uma base $\mathcal{B}_1$ de F_1, uma base $\mathcal{B}_2$ de F_2, e uma base $\mathcal{B}_{12}$ de $F_1 \cap F_2$. Vale que*

$$\dim(F_1 + F_2) = \dim F_1 + \dim F_2 - \dim(F_1 \cap F_2) \tag{3.2.1}$$

Demonstração. Vamos denotar de (b),(c) as partes correspondentes do Teorema 3.2.1. O subespaço $F_1 \cap F_2 \subset F_1$ tem dimensão finita m (segundo (c) para $E = F_1$) e assim admite uma base finita $\mathcal{B}_{12} = \{\zeta_1, \dots, \zeta_m\}$ (segundo a definição de dimensão). Segundo (b) para $E = F_1$ o conjunto $\mathcal{B}_{12}$ – LI em $F_1 \cap F_2$ e segundo Lema 3.1.3 LI no superespaço F_1 – é contido numa base $\mathcal{B}_1$ de F_1. Analogamente $\mathcal{B}_{12}$ é contido numa base $\mathcal{B}_2$ de F_2. Uma base de $F_1 + F_2$ contendo as bases desejadas é

$$\mathcal{B} := (\mathcal{B}_1 \setminus \mathcal{B}_{12}) \dot\cup \overbrace{\mathcal{B}_{12} \dot\cup (\mathcal{B}_2 \setminus \mathcal{B}_{12})}^{\mathcal{B}_2} = \mathcal{B}_1 \dot\cup (\mathcal{B}_2 \setminus \mathcal{B}_{12}) = \mathcal{B}_1 \cup \mathcal{B}_2$$

Contando elementos obtemos

$$\dim(F_1 + F_2) := |\mathcal{B}| = (k - m) + m + (\ell - m) = k + \ell - m.$$

Resta checar as duas propriedades de uma base. $\mathcal{B}$ gera F_1+F_2: Os elementos de $F_1 + F_2$ são da forma $f_1 + f_2$ onde $f_1 \in F_1$ (assim é CL em $\mathcal{B}_1$) e $f_2 \in F_2$ (assim é CL em $\mathcal{B}_2$). Consequentemente $f_1 + f_2$ é CL em $\mathcal{B}_1 \cup \mathcal{B}_2 = \mathcal{B}$.
$\mathcal{B}$ é LI em F: Seja $\mathcal{B}_1 = \{\xi_1, \dots, \xi_k\}$ e $\mathcal{B}_2 \setminus \mathcal{B}_{12} = \{\eta_1, \dots, \eta_{\ell-m}\}$. Suponha por absurdo que $\mathcal{B}$ é LD, ou seja existem escalares α_i, β_i não todos nulos tal que

$$\underbrace{\alpha_1\xi_1 + \cdots + \alpha_k\xi_k}_{=-v_1} + \underbrace{\beta_1\eta_1 + \cdots + \beta_{\ell-m}\eta_{\ell-m}}_{=:v_1} = \mathcal{O}$$

Não todos β_i's são nulos (caso contrário $\mathcal{B}_1$ é LD, contradição). Assim $v_1 \in F_2 \setminus (F_1 \cap F_2)$. De outro lado $-v_1$, então v_1, é elemento do subespaço F_1. Assim $v_1 \in (F_1 \cap F_2)$ e $v_1 \notin (F_1 \cap F_2)$. Contradição. □

Corolário 3.2.3. *Sejam $F, G \subset E$ subespaços de dimensões finitas, então:*

$$F \oplus G = E \qquad \Leftrightarrow \qquad \begin{cases} \dim F + \dim G = \dim E \\ F \cap G = \{\mathcal{O}\} \end{cases}$$

Demonstração. '⇒' Fórmula (3.2.1) usando que a interseção tem dimensão zero. '⇐' Suponha que $F \cap G = \{\mathcal{O}\}$ e que as dimensões de F e G adicionam à dimensão de E. Segundo Lema 2.3.2 a soma $F + G$ é um subespaço de E. Então $F + G = E$ segundo Teorema 3.2.1 (d). □

Exercício 3.2.4 (Subespaços do espaço $\mathrm{M}(n\times n)$ das matrizes quadradas)**.** [1]

1. Sejam $\mathcal{A},\mathcal{S}\subset \mathrm{M}(n\times n)$ os subespaços das matrizes anti-/simétricas.

   (a) Para cada par $(i,j)\in\{1,\dots,n\}\times\{1,\dots,n\}$ seja $\mathbf{e}^{\mathbf{ij}}_{+}$ a matriz $n\times n$ cujos elementos nas posições ij e ji são iguais a 1 e os demais são zero. Prove que estas matrizes constituem uma base $\{\mathbf{e}^{\mathbf{ij}}_{+}\}$ para $\mathcal{S}$.

   (b) De modo análogo, obtenha uma base $\{\mathbf{e}^{\mathbf{ij}}_{-}\}$ para $\mathcal{A}$.

   (c) Conclua que

   $$\dim\mathcal{S}=\frac{n(n+1)}{2}\qquad\qquad \dim\mathcal{A}=\frac{n(n-1)}{2} \tag{3.2.2}$$

   Calcule $\dim\mathcal{S}+\dim\mathcal{A}$ e lembre-se que $\dim\mathrm{M}(n\times n)=n^2$. Conclua que

   $$\mathrm{M}(n\times n)=\mathcal{S}\oplus\mathcal{A}$$

   Antes, no Exercício 2.3.7, tenhamos obtido uma prova alternativa desse.

2. As matrizes quadradas $\mathbf{t}=(t_{ij})\in\mathrm{M}(n\times n)$ tal que $t_{ij}=0$ quando $i<j$ chama-se **triangular inferior**. Prove que elas constituem um subespaço $\mathcal{T}\subset\mathrm{M}(n\times n)$. Obtenha uma base para $\mathcal{T}$ e determine a sua dimensão.

3. Obtenha uma base e consequentemente determine a dimensão de cada um dos seguintes subespaços de $\mathrm{M}(n\times n)$ as quais são composto de

   (a) as matrizes $\mathbf{a}=(a_{ij})$ de **traço** (a soma dos elementos da diagonal)

   $$\mathrm{tr}\,:\mathrm{M}(n\times n)\to\mathbb{R},\quad \mathbf{a}\mapsto\mathrm{tr}\,\mathbf{a}:=\sum_{i=1}^{n}a_{ii}$$

   nulo, ou seja $\mathrm{tr}\,\mathbf{a}=0$.

   (b) as matrizes cuja primeira e última linha são iguais

   (c) as matrizes cuja primeira linha e primeira coluna são iguais

[1] As dimensões para seu controle: 2. $\dim\mathcal{T}=n(n+1)/2$
3. (a) $2\frac{(n-1)n}{2}+(n-1)=n^2-1$ (b) $n(n-2)+n=n(n-1)$ (c) $(n-1)^2+n=n^2-(n-1)$

Parte II

Teoria das transformações lineares

Capítulo 4

Transformações lineares

No Capítulo 4 denotamos de E, F espaços vetoriais

$$E = (E, +, \cdot, \mathbb{K}), \qquad F = (F, +, \cdot, \mathbb{K})$$

ambos sobre o mesmo corpo $\mathbb{K}$. Na primeira leitura pense em $\mathbb{K} = \mathbb{R}$. As letras m, n denotam números naturais ou zero.

4.1 Exemplos e construção

Definição 4.1.1. Uma **transformação linear (TL)**, também chamado de **homomorfismo de espaços vetoriais** ou **operador linear**, é uma aplicação

$$A : E \to F, \quad v \mapsto A(v) =: Av$$

a qual preserva as operações em E e F, ou seja

$$\text{(Linearidade)} \qquad \begin{cases} A(\alpha v) = \alpha Av \\ A(v + w) = Av + Aw \end{cases}$$

para todos os escalares $\alpha \in \mathbb{K}$ e todos os vetores $v, w \in E$. Note-se que nos lados esquerdos aparecem as operações em E e nos lados direitos aquelas em F.

Como indicado acima vamos escrever no caso de aplicações lineares geralmente Av em vez de $A(v)$. Assim Av já sinaliza que A é linear. Note-se que

$$\text{(Linearidade)} \quad \Leftrightarrow \quad A(\alpha u + \beta v) = \alpha Au + \beta Av, \quad \forall \alpha, \beta \in \mathbb{K}, \forall u, v \in E$$

Lema 4.1.2. *Seja $A : E \to F$ uma TL. Então*

(i) $A\mathcal{O} = \mathcal{O}$ *(leva o vetor nulo de E no vetor nulo de F)*

(ii) $A(-v) = -(Av)$ *(leva inversos em inversos)*

(iii) $A(u - v) = Au - Av$

(iv) $A(\alpha_1 v_1 + \cdots + \alpha_k v_k) = \alpha_1 Av_1 + \cdots + \alpha_k Av_k$ *(leva CLs em CLs)*

para todos os vetores $u, v, v_i \in E$ *e escalares* $\alpha_i \in \mathbb{K}$.

Demonstração.
(i) $A\mathcal{O} = A(\mathcal{O} + \mathcal{O}) \overset{\text{linear}}{=} A\mathcal{O} + A\mathcal{O}$, então $A\mathcal{O} = \mathcal{O}$ segundo Lema 1.1.5 3b)
(ii) $A(-v) + Av \overset{\text{linear}}{=} A(-v + v) = A\mathcal{O} \overset{\text{(i)}}{=} \mathcal{O}$.
(iii) $A(u - v) \overset{\text{linear}}{=} Au + A(-v) \overset{\text{(ii)}}{=} Au - Av$.
(iv) Indução sobre k baseado na linearidade. □

Exercício 4.1.3. Considere os elementos de $\mathbb{R}^2$ dados por

$$u_1 = (2, -1), \qquad u_2 = (1, 1), \qquad u_3 = (-1, -4),$$

e

$$v_1 = (1, 3), \qquad v_2 = (2, 3), \qquad v_3 = (5, 6).$$

Decida se existe ou não uma transformação linear $A : \mathbb{R}^2 \to \mathbb{R}^2$ tal que

$$Au_1 = v_1, \qquad Au_2 = v_2, \qquad Au_3 = v_3$$

Solução. Suponha que A é linear. Então escrevendo u_3 como CL de u_1 e u_2, ou seja $u_3 = \alpha u_1 + \beta u_2$, o elemento $v_3 = Au_3$ deve ser CL de $v_1 = Au_1$ e $v_2 = Au_2$ com os mesmos coeficientes. Com efeito

$$v_3 = Au_3 = A(\alpha u_1 + \beta u_2) \overset{\text{lin.}}{=} \alpha Au_1 + \beta Au_2 = \alpha v_1 + \beta v_2$$

Então vamos checar se é verdadeiro isso: Determinamos α e β primeiro

$$\begin{bmatrix} -1 \\ -4 \end{bmatrix} = u_3 = \alpha u_1 + \beta u_2 = \begin{bmatrix} 2\alpha \\ -\alpha \end{bmatrix} + \begin{bmatrix} \beta \\ \beta \end{bmatrix} = \begin{bmatrix} 2\alpha + \beta \\ -\alpha + \beta \end{bmatrix}$$

Comparando os primeiros membros obtemos $\beta = -1 - 2\alpha$. Use isso na comparação dos segundos membros para obter $\alpha = \beta + 4 = -1 - 2\alpha + 4 = 3 - 2\alpha$. Assim $\alpha = 1$ e $\beta = -3$. Basta calcular

$$\alpha v_1 + \beta v_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} - 3 \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -5 \\ -6 \end{bmatrix} \neq \begin{bmatrix} 5 \\ 6 \end{bmatrix} = v_3$$

Então não existe uma tal transformação linear A.

Exercício 4.1.4. Mesma pergunta como no Exercício 4.1.3 mas a) com $v_3 = (-5, -6)$ e b) com $v_3 = (5, -6)$.

Exemplo 4.1.5 (Derivação e convolução)**.**

(Derivação) Seja $k \in \mathbb{N}$ ou $k = \infty$, entao é linear o operador derivada

$$D \colon C^k(\mathbb{R}, \mathbb{R}) \to C^k(\mathbb{R}, \mathbb{R}), \quad f \mapsto f' := \frac{d}{dx} f$$

(Convolução) Dada uma função contínua $k\colon [a,b]\times[a,b]\to\mathbb{R}$, Seja $k\in\mathbb{N}$ ou $k=\infty$, entao é linear o operador definido por

$$K\colon C^0([a,b],\mathbb{R})\to C^0([a,b],\mathbb{R}),\quad f\mapsto\int_a^b k(\cdot,y)f(y)\,dy$$

No caso particular $k(x,y)=g(x-y)$ onde $g\colon [a,b]\to\mathbb{R}$ é uma dada função contínua o operador K_g definido por

$$(K_g f)(x) := \int_a^b g(x-y)f(y)\,dy$$

é chamado de **convolução** das funções f e g, notação $f*g := K_g f$.

4.1.1 O espaço vetorial das transformações lineares

Definição 4.1.6 (O espaço vetorial $\mathcal{L}(E,F)$)**.** O conjunto

$$\mathcal{L}(E,F) := \{A \mid A : E\to F \text{ transformação linear}\}$$

de todas as transformações lineares entre E e F seja munido das operações

$$\begin{aligned} +\,:\mathcal{L}(E,F)\times\mathcal{L}(E,F)&\to\mathcal{L}(E,F) & \cdot\,:\mathbb{K}\times\mathcal{L}(E,F)&\to\mathcal{L}(E,F)\\ (A,B)&\mapsto A+B & (\alpha,A)&\mapsto \alpha A\end{aligned}$$

definidas assim $(A+B)v := Av+Bv$ e $(\alpha A)v := \alpha(Av)$.

Note que $A+B, \alpha A : E\to F$ realmente são lineares. Por exemplo, vale

$$(\alpha A)(v+w) \overset{\text{def.}}{=} \alpha(A(v+w)) \overset{\text{lin.}}{=} \alpha(Av+Aw) \overset{\text{distr.}}{=} \alpha(Av)+\alpha(Aw)$$

e o lado direito é $(\alpha A)v+(\alpha A)w$ pela definição de αA.

Lema 4.1.7. *O conjunto $\mathcal{L}(E,F)$ das transformações lineares de E para F munido das operações '$+$' e '$\cdot$' forma um espaço vetorial*

$$\mathcal{L}(E,F) = (\mathcal{L}(E,F),+,\cdot,\mathbb{K})$$

sobre o corpo $\mathbb{K}$. O vetor nulo de $\mathcal{L}(E,F)$ é a TL nula $\mathcal{O}: E\to F$, $v\mapsto\mathcal{O}$.[1]

Demonstração. Deixamos ao leitor verificar os axiomas na Definição 1.1.19. □

Definição 4.1.8 (Operadores lineares em E)**.** No caso $F=E$ os elementos de $\mathcal{L}(E) := \mathcal{L}(E,E)$ são chamados de **operadores lineares em E** e o operador

$$I = I_E : E\to E,\quad v\mapsto v \tag{4.1.1}$$

é chamado de **operador identidade** em E.

[1] o primeiro $\mathcal{O}$ é $\mathcal{O}_{\mathcal{L}(E,F)}$ e o outro $\mathcal{O}_F$; para legibilidade não escrevemos demais subscritos

4.1.2 Isomorfismos

Isomorfismo e inversa serão tratados com mais detalhes na Seção 6.4.

Definição 4.1.9 (Isomorfismo)**.** Um **isomorfismo** entre espaços vetoriais E e F é uma transformação linear (homomorfismo) $T : E \to F$ tal que a aplicação

$$T : E \to F \text{ é } \textbf{bijetiva} \quad :\Leftrightarrow \quad \begin{cases} \textbf{injetiva} & :\Leftrightarrow \;\; Tu = Tv \Rightarrow u = v \\ \text{e} & \\ \textbf{sobrejetiva} & :\Leftrightarrow \;\; \forall v \in F \; \exists u \in E \colon \; Tu = v \end{cases}$$

Se existe um isomorfismo entre E e F dizemos "E e F são **isomorfos**" e escrevemos $E \simeq F$, ou ainda $E \overset{T}{\simeq} F$ para destacar quem é o isomorfismo.

Definição 4.1.10 (Inversa)**.** Uma transformação linear $A \colon E \to F$ é chamado de **invertível** caso existe uma transformação linear $B \colon F \to E$ tal que valem as duas condições $AB = I_F$ e $BA = I_E$. Neste caso B é único e chamado **a inversa** de A, símbolo $A^{-1} := B$.

Comentário 4.1.11. Dado um isomorfismo $T : E \to F$, definimos a aplicação

$$S : F \to E, \quad f \mapsto v$$

onde $v \in E$ é o único vetor tal que $Tv = f$; encontraremos em (6.4.1). Pode checar que S é linear e bijetiva, ou seja um isomorfismo, e que S é a inversa de T. A composição BA de dois isomorfismos é um isomorfismo e sua inversa é a composição das inversas – mas na ordem oposta

$$(BA)^{-1} = A^{-1}B^{-1} \tag{4.1.2}$$

4.1.3 Construção de transformações lineares

Uma base **ordenada** é uma base $\mathcal{B} = \{\xi_1, \ldots, \xi_n\}$ cujos elementos são enumerados, alternativamente escreve-se na forma de uma lista ordenada $(\xi_1, \ldots, \xi_n)$.

Proposição 4.1.12. *A fim de definir um homomorfismo $A \in \mathcal{L}(E, F)$ basta escolher as imagens de uma base (ordenada) $\mathcal{B} = \{\xi_1, \ldots, \xi_n\}$ de E:*

Existência. *Escolha uma lista $f := (f_1, \ldots, f_n)$ de $n := \dim E$ elementos f_j do contra-domínio F, repetições não excluídas, e defina*

$$A_f \xi_j := f_j, \qquad j = 1, \ldots, n \tag{$*_f$}$$

Então extende A_f ao E inteiro usando (Linearidade): Dado $u \in E$, exprime u em respeito à base $\mathcal{B}$ na forma $u = \sum_{j=1}^{n} \alpha_j \xi_j$ onde os escalares α_j são únicas – são as chamadas coordenadas do vetor u, veja (3.1.1). Defina

$$A_f u := \sum_{j=1}^{n} \alpha_j f_j \overset{(*_f)}{=} \sum_{j=1}^{n} \alpha_j A_f \xi_j \tag{4.1.3}$$

Unicidade. *Se $B \in \mathcal{L}(E, F)$ satisfaz $(*_f)$, levando os ξ_j nos f_j, então $B = A_f$.*

Demonstração. Deixamos ao leitor a tarefa simples de checar que A_f definido acima é linear, ou seja $A_f \in \mathcal{L}(E,F)$, e é unicamente determinado por $(*_f)$. □

Note-se que A_f não só depende da escolha dos elementos f_j de F, mas também da escolha da base $\mathcal{B}$ de E. Por isso às vezes escrevemos

$$A_f^{\mathcal{B}} = A_f$$

Exercício 4.1.13. Mostre: os membros da lista $f = (f_1, \dots, f_n) \in F^{\times n}$

a) são vetores LI $\quad\Leftrightarrow\quad$ A_f é injetivo

b) geram o espaço vetorial F $\quad\Leftrightarrow\quad$ A_f é sobrejetivo

c) formam uma base de F $\quad\Leftrightarrow\quad$ A_f é um isomorfismo (e $\dim E = \dim F$)

Teorema 4.1.14. *Seja* $\dim F$ *finita e* $\mathcal{B} = \{\xi_1, \dots, \xi_n\}$ *uma base de* E*, então*

$$\begin{aligned} \Psi = \Psi_{\mathcal{B}} : F^{\times n} &\to \mathcal{L}(E,F) \\ f &\mapsto A_f \end{aligned} \tag{4.1.4}$$

é um isomorfismo. Lembre-se de $(*_f)$ *que* A_f *é determinado por* $A_f\xi_j := f_j$*.*

Demonstração. Segundo Proposição 4.1.12 é suficiente avaliar TLs numa base.
Linear. Segue de $A_{\alpha f+\beta g}\xi_j := (\alpha f + \beta g)_j = \alpha f_j + \beta g_j =: \alpha A_f\xi_j + \beta A_g\xi_j$.
Injetivo. Suponha $A_f = A_g$. Então $f_j =: A_f\xi_j = A_g\xi_j := g_j$ para todos os j.
Sobrejetivo. Dado $B \in \mathcal{L}(E,F)$, defina $f_j := B\xi_j$, $\forall j$. Assim $A_f = B$. □

Lembre-se do Exercício 3.0.15 que $\dim F^{\times n} = n \dim F$. Vamos ver no futuro, em Corolário 6.4.9, que isomorfismos preservam dimensões – o que implica

Corolário 4.1.15. $\dim \mathcal{L}(E,F) = \dim F^{\times n} = \dim E \cdot \dim F$

Exercício 4.1.16. Mostre que $\dim \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) = \dim \mathrm{M}(m \times n)$. Dado um corpo $\mathbb{K}$, então vale analogamente que $\dim \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m) = \dim \mathrm{M}(m \times n; \mathbb{K})$.

Comentário 4.1.17 (Extensão de TLs)**.** Suponha que em vez de uma base de E só temos um subconjunto LI $\mathcal{U} \subset E$ de k elementos $\mathcal{U} = \{\xi_1, \dots, \xi_k\}$, em particular $k = |\mathcal{U}| \le \dim E =: n$. Note-se que $\mathcal{U}$ é uma base do subespaço $\langle\mathcal{U}\rangle$ gerado por $\mathcal{U}$. Seja $f = (f_1, \dots, f_k) \in F^{\times k}$ uma lista com $k := \dim\langle\mathcal{U}\rangle = |\mathcal{U}|$ membros. Proposição 4.1.12 diz que isso determina unicamente uma TL injetiva

$$A_f^{\mathcal{U}} : E \supset \langle\mathcal{U}\rangle \to F, \quad A_f^{\mathcal{U}}\xi_j := f_j \qquad (j = 1, \dots, k)$$

Então existe uma transformação linear

$$A_{\tilde{f}}^{\widetilde{\mathcal{U}}} : E \to F$$

extendendo $A_f^{\mathcal{U}}$, ou seja $A_{\tilde{f}}^{\widetilde{\mathcal{U}}}$ restrito a $\langle\mathcal{U}\rangle$ é $A_f^{\mathcal{U}}$. Para construir a extensão 1) estende-se o conjunto LI $\mathcal{U}$ a uma base $\mathcal{B}$ de E usando o Teorema 3.2.1 (b) e 2) apensa-se à lista f mais $n-k$ membros.

4.1.4 O espaço dual

Definição 4.1.18 (O espaço dual E^*)**.** No caso $F = \mathbb{K}$ o espaço $E^* := \mathcal{L}(E, \mathbb{K})$ é chamado de **espaço dual** de E. Chama-se os elementos $\phi \in E^*$ de **funcionais $\mathbb{K}$-lineares** em E ou, no caso $\mathbb{K} = \mathbb{R}$, **funcionais lineares**.

Definição 4.1.19 (A base dual $\mathcal{B}^*$)**.** Na dimensão finita uma base ordenada $\mathcal{B} = \{\xi_1, \ldots, \xi_n\}$ de E induz uma base de E^*,[2] a chamada **base dual** $\mathcal{B}^* := \{\phi_1, \ldots, \phi_n\}$. Como transformação linear, cada um membro $\phi_i : E \to \mathbb{K}$ é determinado pelos valores numa base (Proposição 4.1.12) e a escolha seja essa

$$\phi_i(\xi_j) := \delta_{ij} := \begin{cases} 1 & , i = j \\ 0 & , i \neq j \end{cases} \tag{4.1.5}$$

onde chama-se δ_{ij} o **símbolo de Kronecker**. Equivalentemente

$$\phi_i(\alpha_1\xi_1 + \cdots + \alpha_n\xi_n) := \alpha_i \tag{4.1.6}$$

para $i = 1, \ldots, n$. Note-se que $\dim E = n = \dim E^*$.

Lema 4.1.20. *A base dual $\mathcal{B}^* = \{\phi_1, \ldots, \phi_n\}$ é base de E^* e $\dim E^* = \dim E$.*

Demonstração. Bem definida: Deixamos ao leitor verificar que os membros $\phi_i \colon E \to \mathbb{R}$ definidos por (4.1.6) são lineares. Gera: Dado $\psi \in E^*$, denotamos as imagens dos membros ξ_i da base $\mathcal{B}$ de $\beta_i := \psi\xi_i$, então $\beta_1\phi_1 + \ldots \beta_n\phi_n = \psi$. Com efeito, escrevendo $E \ni v = \alpha_1\xi_1 + \ldots, \alpha_n\xi_n$ como CL na base $\mathcal{B}$, obtemos

$$\begin{aligned} \psi(v) &= \psi(\alpha_1\xi_1 + \cdots + \alpha_n\xi_n) \\ &= \alpha_1\psi(\xi_1) + \cdots + \alpha_n\psi(\xi_n) \\ &= \alpha_1\beta_1 + \cdots + \alpha_n\beta_n \\ &= \phi_1(v)\beta_1 + \cdots + \phi_n(v)\beta_n \\ &= (\beta_1\phi_1 + \cdots + \beta_n\phi_n)\, v \end{aligned}$$

LI: Suponha que $\beta_1\phi_1 + \cdots + \beta_n\phi_n = \mathcal{O}$. Segundo (4.1.5), avaliando no vetor ξ_1 obtemos $\beta_1 = 0$, avaliando em $\xi_2, \ldots, \xi_n$ obtemos $\beta_2 = 0, \ldots, \beta_n = 0$. □

Exercício 4.1.21. A expressão geral de um funcional linear $\phi : \mathbb{R}^3 \to \mathbb{R}$ é

$$\phi(x, y, z) = ax + by + cz$$

onde a, b, c são números reais determinando ϕ. Dados os elementos

$$u = (1, 2, 3), \qquad v = (-1, 2, 3), \qquad w = (1, -2, 3),$$

de $\mathbb{R}^3$ determine a, b, c de tal modo que se tenha $\phi u = 1$, $\phi v = 0$ e $\phi w = 0$.[3]

[2] Errado na dimensão infinita. Porque? Relembre-se que CL's são somas *finitas*.

[3] A resposta para seu controle: $a = -\frac{1}{2}$, $b = \frac{1}{4}$, $c = 0$.

Exemplo 4.1.22 (Funcionais lineares $\varphi, \psi \in E^*$)**.** Seja $E = C^0([a,b])$ o espaço vetorial real[4] das funções contínuas $f\colon [a,b] \to \mathbb{R}$ neste intervalo.

(Integração) A função $\varphi : E \to \mathbb{R}$ definida por

$$\varphi(f) := \int_a^b f(x)\,dx$$

é linear e assim $\varphi \in E^*$

(Avaliação) Dado um ponto $x_0 \in [a,b]$, a função $\psi : E \to \mathbb{R}$ definida por

$$\psi(f) := f(x_0)$$

é linear e assim $\psi \in E^*$.

4.2 Matrizes

Matrizes são transformações lineares

Para ver isso escolhemos uma matriz $\mathbf{a} \in \mathrm{M}(m \times n; \mathbb{K})$ e consideramos a aplicação

$$\mathbf{a}\colon \mathbb{K}^n \to \mathbb{K}^m, \quad x \mapsto \mathbf{a}x$$

a qual leva uma lista $x \in \mathbb{K}^m$ para a lista definida pelo produto matriz

$$\mathbf{a}x = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix}$$

Lembrando a notação $\mathbf{a}_{\bullet j}$ para colunas introduzido em (1.2.3), continuamos

$$= \underbrace{\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}}_{\mathbf{a}_{\bullet 1}} x_1 + \cdots + \underbrace{\begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}}_{\mathbf{a}_{\bullet n}} x_n =: (\mathbf{a}_{\bullet 1}, \dots, \mathbf{a}_{\bullet n}) \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} \tag{4.2.1}$$

No último passo definimos uma nova notação a qual vai ser bem útil. O símbolo que usamos quer lembrar o produto matriz, para não precisamos memorizar mais uma fórmula. Na nova notação é facil ver que $\mathbf{a}(\alpha x + \beta y) = \alpha \mathbf{a}x + \beta \mathbf{a}y$ mostrando que $\mathbf{a}\colon \mathbb{K}^n \to \mathbb{K}^m$ é linear, então $\mathbf{a} \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$.

[4] 'real' indica que o corpo são os números reais $\mathbb{R}$

Comentário 4.2.1 (Espaço-coluna e -linha)**.** A imagem de uma matriz

$$\mathrm{Im}(\mathbf{a}) := \{\mathbf{a}x \mid x \in \mathbb{K}^n\} = \text{Esp-col}(\mathbf{a}) \subset \mathbb{K}^m, \qquad \mathbf{a} \in \mathrm{M}(m \times n; \mathbb{K}) \tag{4.2.2}$$

é igual ao espaço-coluna como (4.2.1) mostra. Como Esp-col($\mathbf{a}$) e Esp-lin($\mathbf{a}$) são fechados sob adição e multiplicação, são subespaços de $\mathbb{K}^m$ e $\mathbb{K}^n$. As dimensões

$$\mathrm{pc}(\mathbf{a}) := \dim \text{Esp-col}(\mathbf{a}) = \dim \mathrm{Im}(\mathbf{a}), \qquad \mathrm{pl}(\mathbf{a}) := \dim \text{Esp-lin}(\mathbf{a}) \tag{4.2.3}$$

são chamadas de **posto-coluna** e **posto-linha** da matriz $\mathbf{a}$.

Teorema 4.2.2 (Postos linha e coluna são iguais)**.** $\mathrm{pl}(\mathbf{a}) = \mathrm{pc}(\mathbf{a}) = \dim \mathrm{Im}(\mathbf{a})$

Demonstração. Seja $\mathbf{a}$ uma matriz $m \times n$. '$\leq$' Seja $p := \mathrm{pc}(a)$ a dimensão do espaço coluna e $\mathcal{X} = \{\xi_1, \ldots, \xi_p\}$ uma base ordenada dele. Usamos a notação

$$\xi_\ell = \begin{bmatrix} b_{1\ell} \\ \vdots \\ b_{m\ell} \end{bmatrix}$$

Assim cada uma coluna $\mathbf{a}_{\bullet j}$ é CL em $\mathcal{X}$ com coeficientes únicos $c_{ij} \in \mathbb{K}$, ou seja

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} =: \mathbf{a}_{\bullet j} = \xi_1 c_{1j} + \cdots + \xi_p c_{pj} = \sum_{\ell=1}^{p} \xi_\ell c_{\ell j} = \begin{bmatrix} \sum_{\ell=1}^{p} b_{1\ell} c_{\ell j} \\ \vdots \\ \sum_{\ell=1}^{p} b_{m\ell} c_{\ell j} \end{bmatrix}$$

Isso mostra que a ij-ésima entrada da matriz $\mathbf{a}$ é dada por

$$a_{ij} = \sum_{\ell=1}^{p} b_{i\ell} c_{\ell j}$$

Usamos esta fórmula para ver que a i-ésima linha

$$\begin{aligned} \mathbf{a}_{i\bullet} = \begin{bmatrix} a_{i1} & \ldots & a_{in} \end{bmatrix} &= \begin{bmatrix} \sum_{\ell=1}^{p} b_{i\ell} c_{\ell 1} & \ldots & \sum_{\ell=1}^{p} b_{i\ell} c_{\ell n} \end{bmatrix} \\ &= \sum_{\ell=1}^{p} b_{i\ell} \underbrace{\begin{bmatrix} c_{\ell 1} & \ldots & c_{\ell n} \end{bmatrix}}_{=:\eta_\ell \in \text{Esp-lin}(\mathbf{a})} \end{aligned}$$

é CL das listas $\eta_1, \ldots, \eta_p$ de n escalares cada uma. Assim Esp-lin($\mathbf{a}$) é contido no subespaco Y gerado pelo conjunto $\mathcal{Y} := \{\eta_k \mid k = 1, \ldots, p\}$. Note que $\mathcal{Y}$ contém no máximo p elementos ($< p$ no caso de dobros). Assim

$$\mathrm{pl}(\mathbf{a}) := \dim \text{Esp-lin}(\mathbf{a}) \leq \dim Y \leq |\mathcal{Y}| \leq p =: \mathrm{pc}(\mathbf{a})$$

A primeira desigualdade segue de Teorema 3.2.1 (c) e a segunda de Lema 3.1.20. '$\geq$' Usando '$\leq$' para a transposta obtemos $\mathrm{pc}(\mathbf{a}) = \mathrm{pl}(\mathbf{a}^t) \leq \mathrm{pc}(\mathbf{a}^t) = \mathrm{pl}(\mathbf{a})$. □

Acima temos verificado que cada uma matriz $m \times n$ é uma transformação linear $\mathbb{K}^n \to \mathbb{K}^m$. E vice versa?
Escreve as colunas de uma matriz $\mathbf{a}$ como lista, ou seja $f_{\mathbf{a}} = (\mathbf{a}_{\bullet 1}, \ldots, \mathbf{a}_{\bullet n})$. Agora considere o operador linear $A^{\mathcal{E}^n}_{f_{\mathbf{a}}}$ definido em (4.1.3) e defina a aplicação

$$\mathrm{M}(m \times n; \mathbb{K}) \to \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m), \quad \mathbf{a} \mapsto A^{\mathcal{E}^n}_{f_{\mathbf{a}}} \tag{4.2.4}$$

onde $\mathcal{E}^n = \{e_1, \ldots, e_n\}$ é a base canónica de $\mathbb{K}^n$. Note que $A^{\mathcal{E}^n}_{f_{\mathbf{a}}} = \mathbf{a}$, com efeito

$$A^{\mathcal{E}^n}_{f_{\mathbf{a}}} e_i \overset{(*_{f_{\mathbf{a}}})}{=} \mathbf{a}_{\bullet i} = \mathbf{a} e_i, \qquad i = 1, \ldots, n$$

e então lembre-se de UNICIDADE em Proposição 4.1.12.
Como a aplicação $\mathbf{a} \mapsto \mathbf{a}$ é obviamente linear e injetivo, só falta sobrejetivo para ser um isomorfismo. Dado $A \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$, coloque as listas $Ae_1, \ldots, Ae_n \in \mathbb{K}^m$ como colunas de uma matriz, notação $[A]$.[5] O leitor pode verificar que esta matriz $[A]$ é levado ao operador A, em símbolos $A^{\mathcal{E}^n}_{f_{[A]}} = A$. Isso prova sobrejetividade.[6] Deixa nos formalizar esta ideia no seguinte.

Transformações lineares representadas como matrizes

Como temos visto acima uma transformação linear $A : \mathbb{K}^n \to \mathbb{K}^m$ corresponde naturalmente, utilizando as bases canônicas

$$\mathcal{E}^n = \{e_1, \ldots, e_n\}, \qquad \mathcal{E}^m = \{E_1, \ldots, E_m\}$$

a uma matriz $[A] \in \mathrm{M}(m \times n; \mathbb{K})$. Com efeito, seja $[Ae_i]_{\mathcal{E}^m} \in \mathrm{M}(m \times 1; \mathbb{K})$ o vetor coordenada do elemento $Ae_i \in \mathbb{K}^m$, veja (3.1.2). Usando estes vetores coordenadas como colunas de uma matriz obtém-se

$$[A] = [A]_{\mathcal{E}^n, \mathcal{E}^m} := [[Ae_1]_{\mathcal{E}^m} \ldots [Ae_n]_{\mathcal{E}^m}] \quad \in \mathrm{M}(m \times n; \mathbb{K})$$

chamado de **matriz da transformação linear $A\colon \mathbb{K}^n \to \mathbb{K}^m$ em respeito às bases canônicas**.

O caso geral de associar uma matriz $\mathbf{a}$ a uma transformação linear $A\colon E \to F$ entre espaços vetoriais munidos de bases ordenadas $\mathcal{U} = \{\xi_1, \ldots, \xi_n\}$ e $\mathcal{V} = \{\eta_1, \ldots, \eta_m\}$ vai ser investigado em grande detalhe na Seção 5 depois tratar isomorfismos e inversas na Seção 6.4.1. Sim, as colunas de esta matriz serão os vetores coordenadas (3.1.2), com efeito defina-se

$$\mathbf{a} = [A]_{\mathcal{U}, \mathcal{V}} := [[A\xi_1]_{\mathcal{V}} \ldots [A\xi_n]_{\mathcal{V}}] \quad \in \mathrm{M}(m \times n; \mathbb{K})$$

Proposição 4.2.3. *A aplicação entre espaços vetoriais definido por*

$$\begin{aligned} [\cdot] = [\cdot]_{\mathcal{E}^n, \mathcal{E}^m} : \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m) &\to \mathrm{M}(m \times n; \mathbb{K}) \\ A &\mapsto [[Ae_1]_{\mathcal{E}^m} \ldots [Ae_n]_{\mathcal{E}^m}] \end{aligned} \tag{4.2.5}$$

é um isomorfismo entre espaços vetoriais.

[5] Verifique que os membros da lista Ae_i são as entradas do vetor coordenada $[Ae_i]_{\mathcal{E}^m}$.
[6] Alternativamente, vamos ver em mais em frente que, como as dimensões são iguais e finito, injetividade de uma TL é equivalente a sobrejetividade (Corolário 6.5.2).

Demonstração. Checar linearidade é rotina. *Injetivo.* Se as matrizes $[A] = [B]$ são iguais, os vetores coordenadas $[Ae_i]_{\mathcal{E}^m} = [Be_i]_{\mathcal{E}^m}$ são iguais, e assim as imagens $Ae_i = Be_i$ dos elementos da base são iguais. Assim $A = B$ segundo unicidade em Proposição 4.1.12. *Sobrejetivo.* Dado uma matriz $\mathbf{a}$, então a matriz do operador $A^{\mathcal{E}^n}_{f_{\mathbf{a}}}$, veja (4.2.4), é $\mathbf{a}$. □

Exercício 4.2.4. Mostre que as entradas a_{ij} da matriz $\mathbf{a} := [A]$ satisfazem

$$Ae_i = E_1 a_{1i} + \cdots + E_m a_{mi} =: \mathcal{E}^m \mathbf{a}_{\bullet i}$$

para cada um elemento e_i da base $\mathcal{E}^n$ e onde $\mathcal{E}^m = \{E_1, \ldots, E_m\}$.

Exemplo 4.2.5. Seja $A \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$ determinado por

$$A(1,1) = (1,2,3) \quad \text{e} \quad A(-1,1) = (1,1,1)$$

Pede-se a matriz $\mathbf{a}$ de A relativamente às bases canônicas.

Uma solução. Denotamos de $\mathcal{E}^2 = \{e_1, e_2\}$ e $\mathcal{E}^3 = \{E_1, E_2, E_3\}$ as bases canônicas. Precisamos escrever Ae_1 e Ae_2 como CL's dos vetores E_1, E_2, E_3 e colocar os coeficientes como colunas da matriz desejada. Sabemos que

$$\begin{aligned} A(1,1) &= (1,2,3) = (1,0,0) + (0,2,0) + (0,0,3) = E_1 + 2E_2 + 3E_3 \\ A(1,1) &= A((1,0) + (0,1)) = A(e_1 + e_2) = Ae_1 + Ae_2 \end{aligned}$$

e

$$\begin{aligned} A(-1,1) &= (1,1,1) = (1,0,0) + (0,1,0) + (0,0,1) = E_1 + E_2 + E_3 \\ A(-1,1) &= A((-1,0) + (0,1)) = A(-e_1 + e_2) = -Ae_1 + Ae_2 \end{aligned}$$

Assim temos 2 equações lineares inomogeneas para as 2 incógnitas $x := Ae_1$ e $y := Ae_2$, com efeito

$$\begin{cases} x + y = E_1 + 2E_2 + 3E_3 \\ -x + y = E_1 + E_2 + E_3 \end{cases}$$

Aplicamos escalonamento, adicionando a primeira equação para a segunda obtemos $2y = 2E_1 + 3E_2 + 4E_3$ e assim a CL

$$y = E_1 + \frac{3}{2}E_2 + 2E_3$$

cujas coeficientes formam a segunda ($y = Ae_2$) coluna da matriz $\mathbf{a} := [A]$. Use na primeira equação para receber os coeficientes da primeira coluna, ou seja

$$x = -y + (E_1 + 2E_2 + 4E_3) = 0E_1 + \frac{1}{2}E_2 + 1E_3$$

Então a matriz é a seguinte

$$\mathbf{a} = [A] = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{3}{2} \\ 1 & 2 \end{bmatrix}$$

Com certeza, vai ter outros caminhos como resolver. Acima vemos um.

Exercício 4.2.6. Tem-se uma transformação linear $A : \mathbb{R}^2 \to \mathbb{R}^4$ tal que

$$A(1,2) = (1,1,1,-1) \quad \text{e} \quad A(3,4) = (1,1,1,1)$$

Pede-se a matriz **a** de A relativamente às bases canônicas.

Exercício 4.2.7 (Vetores linha e coluna)**.** a) Mostre que a matriz de um funcional linear $\varphi \in (\mathbb{R}^n)^* := \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ é uma linha (matriz $1 \times n$) da forma

$$[\varphi] = \begin{bmatrix} \varphi e_1 & \dots & \varphi e_n \end{bmatrix}$$

b) Mostre que a matriz de uma reta no $\mathbb{R}^n$ passando a origem, ou seja $R \in \mathcal{L}(\mathbb{R}, \mathbb{R}^n)$, é uma coluna (matriz $n \times 1$) da forma

$$[R] = \begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix}$$

onde de $R_1, \dots, R_n$ denotamos os n membros da lista $Re_1 = R(1) \in \mathbb{R}^n$.

Lema 4.2.8 (Na dimensão 1 operadores correspondem a escalares)**.** *Seja* $\dim E = 1$ *e* $A \in \mathcal{L}(E)$*, então existe um único escalar* $\alpha \in \mathbb{K}$ *tal que o operador corresponde a multiplicação com* α*, em símbolos* $A = \alpha I_E$*.*

Demonstração. Pegue um elemento não-nulo $\xi \in E$. Então $\mathcal{B} := \{\xi\}$ é uma base de E: Com efeito é LI como $\xi \neq \mathcal{O}$, segundo Comentário 1.3.4 (ii), mas LI é equivalente a gera segundo Corolário 3.1.18. Como $\mathcal{B}$ é base com um elemento só, todo elemento de E é uma CL em $\{\xi\}$, assim um múltiplo escalar de ξ com coeficiente único. Então $E \ni A\xi = \alpha\xi$ para um único $\alpha \in \mathbb{K}$.
Analogamente todo $w \in E$ é da forma $w = \lambda\xi$ para um único $\lambda \in \mathbb{K}$. Segue que

$$Aw = A(\lambda\xi) = \lambda A\xi = \lambda(\alpha\xi) = (\lambda\alpha)\xi = (\alpha\lambda)\xi = \alpha(\lambda\xi) = \alpha w = \alpha I_E w$$

onde usamos vários axiomas do espaço vetorial. Isso prova que $A = \alpha I_E$. □

4.3 Dimensão dois – o plano

No plano Π queremos estudar três tipos elementares de transformações lineares, nomeadamente

- rotação R_θ por um ângulo θ em torno de um centro O no plano
- projeção ortogonal P_L sobre uma reta L no plano
- reflexão S_L em torno de uma reta L no plano

Mas o plano Π é composto de pontos... Como pode-se dar Π a estrutura de um espaço vetorial sobre o corpo dos números reais $\mathbb{R}$? Como pode-se adicionar pontos ou multiplicar por números? Não da.

Mas pode-se adicionar flechas v no plano se consideramos iguais todas as flechas do mesmo comprimento e direção, veja Exemplo 0.0.1. Mais detalhado duas flechas são consideradas iguais se formam os lados opostos de um paralelogramo no qual os outros dois lados conectam, respectivamente, os dois pontos iniciais e os dois pontos terminais. Multiplicação de uma flecha v com um número real α muda o comprimento pelo fator α, trocando a direção caso $\alpha < 0$ é negativo. Adicionamos duas flechas pondo no ponto termino da primeira flecha o ponto inicial da segunda, veja Figura 1 na introdução do manuscrito.

Definição 4.3.1 (O espaço vetorial Π_O das flechas no plano de ponto início O)**.** Para eliminar a complicação que, dado uma flecha v, todo ponto $p \in \Pi$ nos da uma flecha equivalente (escolhendo p como ponto início), vamos fixar um ponto do plano, notação $O \in \Pi$. Neste caso todo ponto $p \in \Pi$ representa uma flecha só: por definição a flecha correndo de O a p. Vice versa, cada uma flecha em Π é equivalente a uma iniciando no ponto O. Seja

$$\Pi_O := (\Pi, O)$$

o conjunto das flechas no plano Π com ponto início O. Identificamos tal flecha com seu ponto termino $p \in \Pi$. Escrevendo $p \in \Pi$ significa que p é um ponto do plano, escrevendo $p \in \Pi_O$ significa que p é a flecha correndo de O a p. Para Π_O pode-se verificar os axiomas de um espaço vetorial real sob multiplicação escalar αp definida como mudando o comprimento da flecha com ponto termino p pelo fator $\alpha \in \mathbb{R}$ e adição $p + q$ definida pelo paralelogramo gerado, veja Figura 4.3.

Comentário 4.3.2. Note-se que são em bijeção o conjunto Π_O das flechas no plano Π iniciando no ponto O e o conjunto F no Exemplo 0.0.1 cujos elementos são flechas v no plano junto com todas flechas equivalentes a v. Adição e multiplicação escalar coincidem. Assim os espaços vetoriais F e Π_O são isomorfos.

Comentário 4.3.3 (Sistema de coordenadas ortogonal OXY)**.** Escolhendo no plano Π dois eixos OX e OY ortogonal um ao outro, notação OXY, recebemos uma bijeção linear

$$\Pi_O \to \mathbb{R}^2, \quad p \mapsto (x, y)$$

como definida em (0.0.1) e ilustrada na Figura 4.1.

4.3.1 Rotações

Seja Π_O o plano Π junto com um ponto $O \in \Pi$ fixado. Suponha que podemos medir distância no plano, assim ângulos entre semi-retas do mesmo ponto inicial. Para os elementos $p \in \Pi_O$ (pontos do plano interpretado simultaneamente como flecha de O ao ponto), denotamos de

- C_p o círculo com centro O e passando p, veja Figura 4.2
- ℓ_p a semi-reta iniciando em O e passando p (se $p = O$ seja $\ell_O := \{O\}$)
- $\ell_p(\theta)$ a semi-reta obtida pela rotação de ℓ_p em torno O pelo ângulo θ no sentido contra-horário

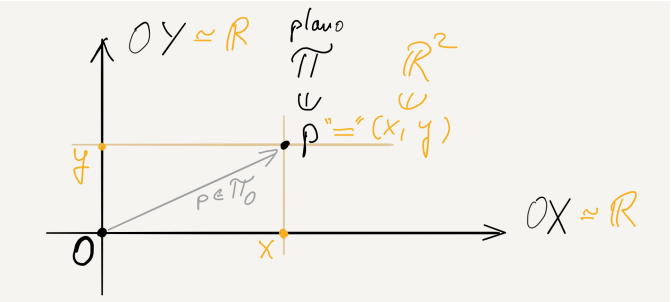


Figura 4.1: Sistema ortogonal de coordenadas OXY no plano Π

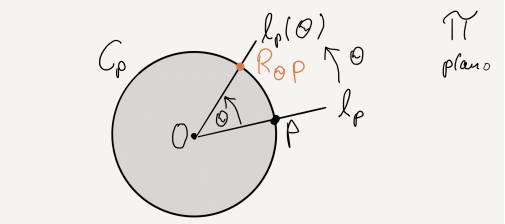


Figura 4.2: Rotação R_θ no plano Π em torno do ponto O por um ângulo θ

Definição 4.3.4 (Rotação)**.** A aplicação definida por

$$R_\theta : \Pi_O \to \Pi_O, \quad p \mapsto \ell_p(\theta) \cap C_p$$

é chamado de **rotação** no plano Π em torno de O por o ângulo θ.

Comentário 4.3.5 (Preservação de comprimento e ângulos)**.** Como o resultado da rotação é localizado no mesmo círculo a distância de O fica constante. O ângulo φ entre duas flechas $p, q \in \Pi_O$ fica constante se aplicamos a rotação R_θ pelo *mesmo* ângulo θ em ambas flechas. aplicar a rotação R_θ pelo mesmo ângulo θ em ambas flechas, veja Figura 4.3.

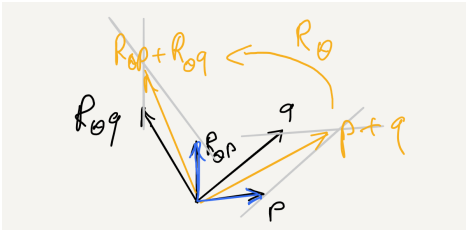


Figura 4.3: Preservação de ângulos

Lema 4.3.6. *Dado um ângulo θ, a rotação $R_\theta : \Pi_O \to \Pi_O$ é linear.*

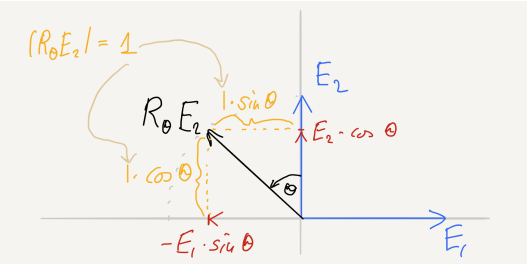


Figura 4.4: Rotação do vetor unitário – coeficientes formam coluna 2 da matriz

Demonstração. Lembre que os elementos $p \in \Pi_O$ são pontos do plano visto como flechas de O a p. O comprimento da flecha é a distância dos ponto p e O. Dado p, denotamos ambos, **comprimento da flecha** e **distância de $\mathcal{O}$**, com o símbolo $|\boldsymbol{p}|$.

Preservação de comprimento $\Rightarrow$ multiplicativo. Dado um ponto p e um escalar $\alpha \in \mathbb{R}$. Se $p = O$ ou $\alpha = 0$ temos $R_\theta(\alpha p) = R_\theta(O) = O = \alpha R_\theta(p)$. Seja então $p \neq O$ e $\alpha \neq 0$. Segundo preservação de comprimento obtemos

$$\frac{|R_\theta(\alpha p)|}{|R_\theta(p)|} = \frac{|\alpha p|}{|p|} = \pm\alpha \quad \text{então} \quad |R_\theta(\alpha p)| = \pm\alpha |R_\theta(p)|$$

onde o sinal em $+/-\alpha$ depende se α é positivo/negativo.

Resta eliminar os absolutos. No caso $\alpha > 0$ os pontos αp e p estão na mesma semi-reta, mas rotação preserva esta propriedade, assim $R_\theta(\alpha p) = \alpha R_\theta(p)$. No caso $\alpha < 0$ os pontos αp e p estão em semi-retas opostas. Rotação também preserva esta propriedade, assim $R_\theta(\alpha p)$ e $R_\theta(p)$ são múltiplos negativos um do outro. Como $|R_\theta(\alpha p)| = -\alpha |R_\theta(p)|$ segue que $R_\theta(\alpha p) = \alpha R_\theta(p)$.

Preservação de ângulos $\Rightarrow$ aditivo. Dado pontos p, q, considere o paralelogramo definindo a soma $p + q$. Aplicando a rotação sabemos que $R_\theta p$ e $R_\theta q$ formam o mesmo ângulo como p e q. Então o paralelogramo gerado por $R_\theta p$ e $R_\theta q$ resulta daquele gerado por p e q através de aplicar R_θ. Mas assim a diagonal $R_\theta p + R_\theta q$ resulta de aplicar R_θ aà diagonal $p + q$ do paralelogramo original, em símbolos $R_\theta p + R_\theta q = R_\theta(p + q)$. □

A matriz da rotação num sistema ortogonal de coordenadas

Conforme a definição de eixo, veja Comentário 0.0.3, o vetor unitário no eixo OX é a flecha correndo de O ao ponto X, notação E_1. Analogamente denotamos de E_2 o vetor unitário no eixo OY

Por definição a matriz $\mathbf{r}_\theta$ da rotação R_θ em respeito à base $\mathcal{E} := \{E_1, E_2\}$ contem como colunas os coeficientes (veja Figura 4.4) de

$$R_\theta E_1 = E_1 \cos\theta + E_2 \sin\theta, \qquad R_\theta E_2 = -E_1 \sin\theta + E_2 \cos\theta$$

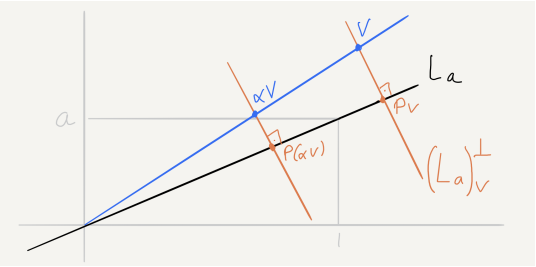


Figura 4.5: Projeção ortogonal P sobre a reta L_a

Lema 4.3.7. *A **matriz da rotação** pelo ângulo θ é a matriz real*

$$\mathbf{r}_\theta := [R_\theta]_{\mathcal{E},\mathcal{E}} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{4.3.1}$$

Lembramos que o sistema ortogonal de coordenadas disponibiliza uma correspondência $\Pi_O \simeq \mathbb{R}^2$ na qual a base $\{E_1, E_2\}$ corresponde à base canônica $\mathcal{E}^2 = \{e_1, e_2\}$. Obviamente é mais confortável trabalhar com as listas de $\mathbb{R}^2$ como com as flechas de Π_O. Assim vamos trabalhar no futuro com $\mathbb{R}^2$.

4.3.2 Projeção ortogonal sobre uma reta

Trabalhamos no plano identificado com $\mathbb{R}^2$ mediante um sistema ortogonal de coordenadas.

Definição 4.3.8 (Projeção ortogonal)**.** Seja $a \in \mathbb{R}$ uma constante e seja $L_a := \mathbb{R}(1, a)$ a reta no $\mathbb{R}^2$ passando a origem $\mathcal{O} = (0, 0)$ e o ponto $(1, a)$ como ilustrado na Figura 4.5. Para um elemento $v \in \mathbb{R}^2$ seja $(L_a)^\perp_v$ a reta ortogonal a L_a e passando o ponto v. Então a aplicação que leva v à interseção das duas retas

$$P = P_{L_a} : \mathbb{R}^2 \to \mathbb{R}^2, \quad v \mapsto L_a \cap (L_a)^\perp_v \tag{4.3.2}$$

é chamado de **projeção ortogonal** sobre a reta L_a.

Lema 4.3.9. *A projeção ortogonal $P = P_{L_a} : \mathbb{R}^2 \to \mathbb{R}^2$ é linear.*

Demonstração. Multiplicativo. Similarmente como na prova de Lema 4.3.4 discrimina-se três casos $\alpha < 0$, ($\alpha = 0$ ou $v = \mathcal{O}$), e $\alpha > 0$. Vamos tratar o caso $\alpha > 0$ e deixar os outros ao leitor. Para $\alpha > 0$ e $v \neq \mathcal{O}$ obtemos

$$\frac{|v|}{|Pv|} = \frac{|\alpha v|}{|P\alpha v|} = \frac{\alpha|v|}{|P\alpha v|}$$

onde temos usado o Teorema do Raio na primeira igualdade. Como $|v| \neq 0$ segue, cortando $|v|$, que $|P\alpha v| = \alpha|Pv|$. Como $P\alpha v$ e Pv são elementos da mesma ($\alpha > 0$) semi-reta de L_a, obtém-se $P\alpha v = \alpha Pv$.

Aditivo. A identidade $P(v + w) = Pv + Pw$ resulta da Figura 4.6. □

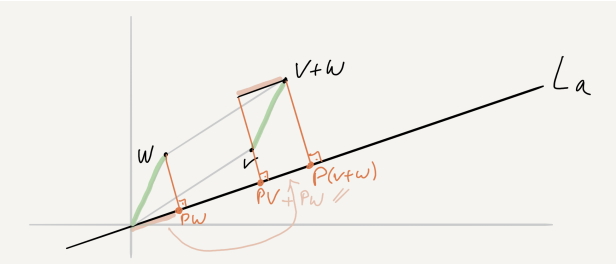


Figura 4.6: A identidade $P(v+w) = Pv + Pw$

A matriz da projeção ortogonal

Trabalhamos no plano identificado com $\mathbb{R}^2$ mediante um sistema ortogonal de coordenadas.

Lema 4.3.10. *A **matriz da projeção ortogonal** sobre a reta L_a é dada por*

$$\mathbf{p}_a := [P_{L_a}] = \frac{1}{1+a^2}\begin{bmatrix} 1 & a \\ a & a^2 \end{bmatrix}$$

Demonstração. Lema B.4.1 □

4.3.3 Reflexão em torno de uma reta

Trabalhamos no plano identificado com $\mathbb{R}^2$ mediante um sistema ortogonal de coordenadas.

Definição 4.3.11 (Reflexão)**.** Dado $a \in \mathbb{R}$, a aplicação definida assim

$$\begin{aligned} S = S_{L_a} \colon \mathbb{R}^2 &\to \mathbb{R}^2 \\ v &\mapsto v + 2(P_{L_a}v - v) = (2P_{L_a} - I)v \end{aligned} \tag{4.3.3}$$

é chamado de **reflexão** em torno da reta L_a. Note que $S = 2P - I$ é linear.

Use Proposição 4.2.3 e a matriz de P para obter a **matriz da reflexão** em torno da reta L_a, com efeito

$$\mathbf{s}_a := [S_{L_a}] = 2\mathbf{p}_a - \mathbb{1} = \frac{1}{1+a^2}\begin{bmatrix} 1-a^2 & 2a \\ 2a & -(1-a^2) \end{bmatrix}$$

Exercício 4.3.12 (Rotação, projeção, reflexão)**.**

1. Sejam $R, P, S \in \mathcal{L}(\mathbb{R}^2)$ respectivamente a rotação de 30° em torno da origem, a projeção ortogonal sobre a reta $y = \frac{1}{3}x$ (notação $L_{\frac{1}{3}}$) e a reflexão em torno da mesma reta.

   Dado o vetor $v = (2, 5)$, determine suas imagens Rv, Pv, Sv.

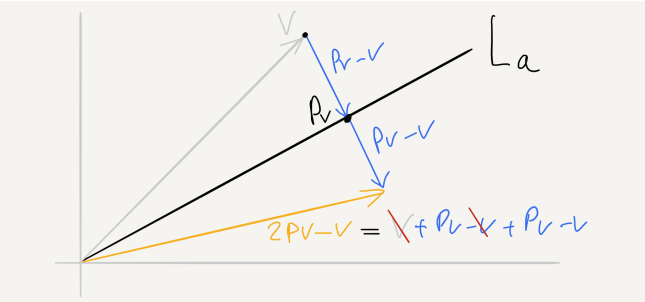


Figura 4.7: Reflexão $S = 2P - I$ em torno da reta L_a

2. Considere os operadores lineares $\mathbb{R}^2 \to \mathbb{R}^2$ dado por

$$R = R_{30^\circ}, \qquad S = S_{L_2}, \qquad P = P_{L_2}.$$

   (a) Mostre que se tem $PS = SP = P$.
   (b) Verifique a igualdade $RSR = S$.
   (c) Mostre que R não comuta com S nem com P.
   (d) Determine todos os vetores v tais que $RPv = 0$ e $RPv \neq 0$.

3. Encontre $a, b, c, d \in \mathbb{R}$ tais que o operador

$$\begin{aligned} A : \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\mapsto (ax + by, cx + dy) \end{aligned}$$

   tenha como núcleo a reta $y = 3x$.

Capítulo 5

Matrizes de transformações lineares

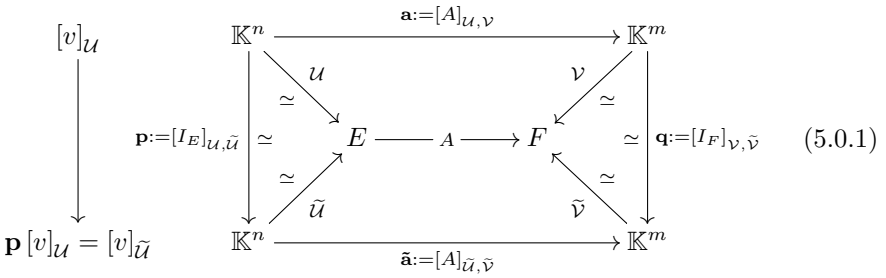
Na dimensão *finita* consideramos uma transformação linear

$$A : E \to F$$

entre espaços vetoriais sobre um corpo $\mathbb{K}$. Denotamos o operador identidade de

$$I_E : E \to E, \quad v \mapsto v$$

e o em F de I_F, veja (4.1.1). Agora será muito útil escrever uma base ordenada na forma de uma lista ordenada. Sejam $\mathcal{U} = (\xi_1, \ldots, \xi_n)$ e $\tilde{\mathcal{U}}$ bases ordenadas de E e $\mathcal{V} = (\eta_1, \ldots, \eta_m)$ e $\tilde{\mathcal{V}}$ de F. Nas seguintes seções vamos estabelecer e provar os detalhes da seguinte diagrama comutativa[1]



(5.0.1)

Na diagrama $\mathbf{p}$ é a chamada **matriz de passagem** da base $\mathcal{U}$ de E para $\tilde{\mathcal{U}}$. Ela leva, dado $v \in E$, o vetor coordenada $[v]_{\mathcal{U}}$ em respeito à base $\mathcal{U}$ ao vetor coordenada $[v]_{\tilde{\mathcal{U}}}$ em respeito à base $\tilde{\mathcal{U}}$. Além disso $\mathbf{a}$ é a **matriz da transformação linear** $A : E \to F$ em respeito às bases $\mathcal{U}$ do domínio e $\mathcal{V}$ do contra-domínio.

[1] **Comutatividade** significa que caso entre dois espaços vetoriais no diagrama tem dois caminhos de flechas, então não importa o qual usamos. Note que a flecha de um isomorfismo '$\simeq$' também existe na direção reversa (no diagrama só mostramos uma flecha para simplicidade).

Comentário 5.0.13 (Interpretação da parte triangular esquerda da diagrama)**.** Sejam $\mathcal{U}$,$\tilde{\mathcal{U}}$, e $\mathcal{W}$ bases do espaço vetorial E da dimensão n. Sejam

$$\begin{aligned} \mathbf{p}&: \quad \text{a matriz de passagem de } \mathcal{U} \text{ para } \tilde{\mathcal{U}} \\ \mathbf{r}&: \quad \text{a matriz de passagem de } \tilde{\mathcal{U}} \text{ para } \mathcal{W} \end{aligned}$$

Vamos entender nesta seção que sob estas hipóteses vale o seguinte

$$\begin{aligned} \mathbf{rp} &\quad \text{é a matriz de passagem de } \mathcal{U} \text{ para } \mathcal{W} \\ \mathbf{p^{-1}} &\quad \text{é a matriz de passagem de } \tilde{\mathcal{U}} \text{ para } \mathcal{U} \end{aligned}$$

5.1 Bases induzem isomorfismos

Definição 5.1.1 (Um símbolo com duas significados)**.** Usamos *o mesmo símbolo* para a base e para o isomorfismo determinado pela: Escrevemos

$$\mathcal{U} : \mathbb{K}^n \to E, \quad x \mapsto \mathcal{U}x$$

para a transformação linear determinada pela base $\mathcal{U} = (\xi_1, \dots, \xi_n)$, ou seja

$$\mathcal{U}x := (\xi_1, \dots, \xi_n) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} := \underbrace{\xi_1 x_1 + \cdots + \xi_n x_n}_{=v} \in E \tag{5.1.1}$$

Obviamente a aplicação $\mathcal{U} : \mathbb{K}^n \to E$ é linear. Ela é injetiva como a base $\mathcal{U}$ é LI (Corolário 3.1.2) e sobrejetiva como $\mathcal{B}$ gera E.

Portanto $\mathcal{U} : \mathbb{K}^n \to E$ é um isomorfismo (símbolo $\simeq$).

No caso $E = \mathbb{K}^n$ a base canônica $\mathcal{E}^n$ produz o operador identidade $\mathcal{E}^n = I_{\mathbb{K}^n}$.

Seja $\mathcal{U} = (\xi_1, \dots, \xi_n)$ uma base ordenada de E. Dado $v \in E$, então $x := \mathcal{U}^{-1}v \in \mathbb{K}^n$ é o vetor coordenada $[v]_{\mathcal{U}}$ de v em respeito à base $\mathcal{U}$ introduzido em (3.1.2): Com efeito como $\mathcal{B}$ é base exprime-se v como CL dos elementos de $\mathcal{B}$ com coeficientes únicos x_i, ou seja $v = \xi_1 x_1 + \cdots + \xi_n x_n =: \mathcal{U}x$. Assim

$$\mathcal{U}^{-1}v = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = [v]_{\mathcal{U}} \in \mathbb{K}^n \tag{5.1.2}$$

O isomorfismo $\mathcal{U}^{-1} : E \to \mathbb{K}^n$ é chamado de **sistema de coordenadas** em E. No caso de $E = \mathbb{K}^n$ com base canônica $\mathcal{U} = \mathcal{E}$ abreviamos $[v] := [v]_{\mathcal{E}}$.

5.2 A matriz em respeito a uma base

Dado uma transformação linear $A \in \mathcal{L}(E,F)$ e bases $\mathcal{U} = (\xi_1, \dots, \xi_n)$ de E e $\mathcal{V} = (\eta_1, \dots, \eta_m)$ de F, então podemos representar os elementos $A\xi_j \in F$ como combinação linear na base $\mathcal{V}$ com coeficientes $a_{ij} \in \mathbb{K}$ únicos. Com efeito

$$\boxed{A\xi_j = \eta_1 a_{1j} + \cdots + \eta_m a_{mj}} \tag{5.2.1}$$

Note como pela nossa definição o índice do η_i coincide com o primeiro (mais perto) índice do escalar a_{ij} o qual deve ser escrito *atrás*.

Os escalares a_{ij} formam uma matriz $m \times n$ chamada de **matriz de A** em respeito às bases $\mathcal{U}$ e $\mathcal{V}$, símbolo

$$[A]_{\mathcal{U},\mathcal{V}} := \mathbf{a} = [a_{ij}] := \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \tag{5.2.2}$$

Note-se que a matriz $\mathbf{a}$ tem como colunas

$$\mathbf{a}_{\bullet j} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} = [A\xi_j]_{\mathcal{V}}$$

os vetores coordenadas dos imagens $A\xi_j$, ou seja

$$[A]_{\mathcal{U},\mathcal{V}} = \big[[A\xi_1]_{\mathcal{V}} \dots [A\xi_n]_{\mathcal{V}}\big] = [\mathbf{a}_{\bullet 1} \dots \mathbf{a}_{\bullet n}] = \mathbf{a}$$

No caso $E = F$ e $\mathcal{U} = \mathcal{V}$ abreviamos $[A]_{\mathcal{U}} := [A]_{\mathcal{U},\mathcal{U}}$. No caso $E = F = \mathbb{K}^n$ e $\mathcal{U} = \mathcal{V} = \mathcal{E}$ abreviamos $[A] := [A]_{\mathcal{E},\mathcal{E}}$.

Exercício 5.2.1. Seja $\mathcal{E} \subset \mathbb{R}^3$ a base canônica e $A\colon \mathbb{R}^3 \to \mathbb{R}^3$ determinado por

$$\begin{aligned} Ae_1 &= \ 2e_1 - e_2 - e_3 \\ Ae_2 &= -e_1 + e_2 \\ Ae_3 &= -e_1 \qquad + e_3 \end{aligned} \tag{5.2.3}$$

Considere a base ordenada $\mathcal{V} := (e_1, e_1 + e_3, e_1 + e_2)$ e determine a matriz $\mathbf{a} := [A]_{\mathcal{E},\mathcal{V}}$. (Vamos reencontrar A nos Exercícios 5.3.10 e 8.1.3.)

Exercício 5.2.2 (Identidade $I = I_E$)**.** Mostre que a matriz da identidade

$$[I]_{\mathcal{U}} := [I]_{\mathcal{U},\mathcal{U}} = \mathbb{1} \tag{5.2.4}$$

sempre é a matriz identidade se usamos a mesma base $\mathcal{U}$ para o domínio e o contra-domínio de $I\colon E \to E$.

Exercício 5.2.3 (Homotetias αI)**.** Seja E um espaço vetorial sobre $\mathbb{R}$ de dimensão $n < \infty$. Suponha que $A \in \mathcal{L}(E)$ não seja um múltiplo do operador identidade: $A \neq \alpha I$, para todo $\alpha \in \mathbb{R}$.

1. Mostre que existem bases de E do tipo $\mathcal{U} = (u, Au, \dots)$ e $\mathcal{V} = (v, 2Av, \dots)$ tais que as matrizes $[A]_{\mathcal{U}}$ e $[A]_{\mathcal{V}}$ de A são diferentes.
2. Conclua que as **homotetias** (múltiplos αI do operador identidade) são os únicos operadores cuja matriz não depende da base escolhida.
3. Conclua que as matrizes do tipo $\alpha \mathbb{1}_n$ são os únicos que comutam ($\mathbf{ab} = \mathbf{ba}$) com todas matrizes invertíveis $n \times n$.

[ad 1.: Conclua $n \geq 2$. Mostre que existe conjunto LI da forma $X = \{v, Av\}$. Depois estenda X para receber uma base $(\xi_1 = v, \xi_2 = Av, \dots, \xi_n)$ de E.]

Exercício 5.2.4. Suponha que $E = F \oplus G$ e $n = \dim E$ é finita. Mostre que existe uma base ordenada $\mathcal{X}$ de E tal que

$$[P_{F,G}]_{\mathcal{X}} = \begin{bmatrix} \mathbb{1}_k & \mathcal{O}_{k,\ell} \\ \mathcal{O}_{\ell,k} & \mathcal{O}_\ell \end{bmatrix}, \qquad [S_{F,G}]_{\mathcal{X}} = \begin{bmatrix} \mathbb{1}_k & \mathcal{O} \\ \mathcal{O} & -\mathbb{1}_\ell \end{bmatrix}$$

Teorema 5.2.5. *Levando transformações lineares às suas matrizes*

$$\begin{aligned} \Phi = \Phi_{\mathcal{U},\mathcal{V}} : \mathcal{L}(E,F) &\xrightarrow{\simeq} \mathrm{M}(n \times m; \mathbb{K}) \\ A &\mapsto [A]_{\mathcal{U},\mathcal{V}} \end{aligned}$$

é um isomorfismo.

Demonstração. Seja $\mathcal{U} = (\xi_1, \dots, \xi_n)$ base de E e $\mathcal{V} = (\eta_1, \dots, \eta_m)$ de F.
Linear. Escreve (5.2.2) para A, para B, e depois adiciona as duas equações e use $(A+B)\xi_j = A\xi_j + B\xi_j$.
Injetivo. Suponha $(a_{ij}) := [A]_{\mathcal{U},\mathcal{V}} = [B]_{\mathcal{U},\mathcal{V}} =: (b_{ij})$. Então A e B coincidem

$$A\xi_j = \eta_1 a_{1j} + \cdots + \eta_m a_{mj} = \eta_1 b_{1j} + \cdots + \eta_m b_{mj} = B\xi_j$$

nos elementos de uma base e linearidade implica que coincidem em todo $v \in E$.
Sobrejetivo. Dado $\mathbf{a} = (a_{ij}) \in \mathrm{M}(m \times n; \mathbb{K})$, para cada um j defina

$$A\xi_j := \eta_1 a_{1j} + \cdots + \eta_m a_{mj}$$

Isso determina A unicamente (Prop. 4.1.12). Então $\Psi(A) := [A]_{\mathcal{U},\mathcal{V}} = \mathbf{a}$. □

Lembre-se de Exemplo 3.0.13 (c) e de Comentário 3.0.14 que $\dim \mathrm{M}(n \times m; \mathbb{K}) = nm$. Segundo Corolário 6.4.9 isomorfismos preservam dimensões, assim obtemos $\dim \mathcal{L}(E,F)$.

Corolário 5.2.6. $\dim \mathcal{L}(E,F) = \dim \mathrm{M}(n \times m; \mathbb{K}) = nm = \dim E \cdot \dim F$

Teorema 5.2.7. *Considere duas transformações lineares entre espaços vetoriais*

$$E \xrightarrow{A} F \xrightarrow{B} G$$

com bases respectivas $\mathcal{U} = (\xi_1, \dots, \xi_n)$, $\mathcal{V} = (\eta_1, \dots, \eta_m)$, e $\mathcal{W} = (\nu_1, \dots, \nu_p)$. Então a matriz da composição

$$[BA]_{\mathcal{U},\mathcal{W}} = [B]_{\mathcal{V},\mathcal{W}}\,[A]_{\mathcal{U},\mathcal{V}} \tag{5.2.5}$$

é o produto das matrizes.

Demonstração. Sejam **a** e **b** as matrizes de A e B, seja **c** aquela de BA. Assim

$$A\xi_j = \sum_{k=1}^{m} \eta_k a_{kj}, \qquad B\eta_k = \sum_{i=1}^{p} \nu_i b_{ik}, \qquad BA\xi_j = \sum_{i=1}^{p} \nu_i c_{ij}$$

para $j = 1, \ldots, n$ e $k = 1, \ldots, m$. Use estas identidades para obter

$$\sum_{i=1}^{p} \nu_i c_{ij} = B(A\xi_j) = \sum_{k=1}^{m} (B\eta_k) a_{kj} = \sum_{i=1}^{p} \nu_i \sum_{k=1}^{m} b_{ik} a_{kj}$$

onde no último passo temos permutado a ordem das somas *finitas*. Mas como a base $\mathcal{W}$ é LI os coeficentes dos ν_i devem ser iguais (Corolário 3.1.2). □

5.3 Mudança de base – diagrama comutativa

Nesta seção estudamos na diagrama (5.0.1) o que acontece a) com um vetor coordenada $[v]_{\mathcal{U}} = \mathcal{U}^{-1}v$ se trocamos a base $\mathcal{U}$ de E para uma outra base $\tilde{U}$ e b) com a matriz de uma transformação linear onde adicionalmente permitimos trocar a base de F.

5.3.1 Como o vetor coordenada muda

A matriz $\mathbf{p} := [I_E]_{\mathcal{U},\widetilde{\mathcal{U}}}$ do operador identidade, por definição (5.2.1), satisfaz

$$\xi_j = I_E\, \xi_j = \tilde{\xi}_1 p_{1j} + \cdots + \tilde{\xi}_n p_{nj} \qquad j = 1, \ldots, n. \tag{5.3.1}$$

Nas outras palavras ela exprime os elementos da *base velha* $\mathcal{U} = (\xi_1, \ldots, \xi_n)$ de E como combinação linear dos elementos da *base nova* $\tilde{\mathcal{U}} = (\tilde{\xi}_1, \ldots, \tilde{\xi}_n)$ de E. Por isso **p** é chamado de **matriz de passagem de $\mathcal{U}$ para $\tilde{\mathcal{U}}$**. A matriz de passagem participa do diagrama (5.0.1) fazendo as partes triangulares comutativo (5.0.1).

Lema 5.3.1. *$\tilde{\mathcal{U}}\mathbf{p} = \mathcal{U}$ na diagrama (5.0.1), equivalentemente $[I_E]_{\mathcal{U},\widetilde{\mathcal{U}}} = \tilde{\mathcal{U}}^{-1}\mathcal{U}$.*

Demonstração. Para $x \in \mathbb{K}^n$ vale

$$\tilde{\mathcal{U}}\mathbf{p}x = \left(\tilde{\xi}_1, \ldots, \tilde{\xi}_n\right) \begin{bmatrix} \sum_k p_{1k}x_k \\ \vdots \\ \sum_k p_{nk}x_k \end{bmatrix} = \sum_{\ell} \tilde{\xi}_\ell \sum_k p_{\ell k}x_k = \sum_k \underbrace{\left(\sum_{\ell} \tilde{\xi}_\ell p_{\ell k}\right)}_{\overset{(5.3.1)}{=} \xi_k} x_k = \mathcal{U}x$$

□

Corolário 5.3.2. *Toda matriz de passagem é um isomorfismo e sua inversa é*

$$[I_E]^{-1}_{\mathcal{U},\widetilde{\mathcal{U}}} = [I_E]_{\widetilde{\mathcal{U}},\mathcal{U}} = \mathcal{U}^{-1}\widetilde{\mathcal{U}}$$

Demonstração. Aplicando $\widetilde{\mathcal{U}}^{-1}$ em ambos os lados de $\widetilde{\mathcal{U}}\mathbf{p} = \mathcal{U}$ obtem-se $\widetilde{\mathcal{U}}^{-1}\mathcal{U} = \mathbf{p} := [I_E]_{\mathcal{U},\widetilde{\mathcal{U}}}$. Assim $\mathbf{p}$ é um isomorfismo como $\mathcal{U}$ e $\widetilde{\mathcal{U}}$ são. Toma as inversas em ambos lados e use (4.1.2) para obter $[I_E]^{-1}_{\mathcal{U},\widetilde{\mathcal{U}}} = \mathbf{p}^{-1} = \mathcal{U}^{-1}\widetilde{\mathcal{U}} = [I_E]_{\widetilde{\mathcal{U}},\mathcal{U}}$. □

Comentário 5.3.3 (Trocando a base (o sistema de coordenadas))**.** Seja $v \in E$. Dado o vetor coordenada $[v]_{\mathcal{U}}$ em respeito a uma base $\mathcal{U}$ de E. Para calcular as coordenadas em respeito a uma outra base $\widetilde{\mathcal{U}}$ simplesmente aplique a matriz de passagem $\mathbf{p} = [I]_{\mathcal{U},\widetilde{\mathcal{U}}}$, ou seja

$$[I]_{\mathcal{U},\widetilde{\mathcal{U}}}\,[v]_{\mathcal{U}} = [v]_{\widetilde{\mathcal{U}}}$$

Exercício 5.3.4. Considere as bases $\mathcal{U} = (\xi, \eta)$ e $\widetilde{\mathcal{U}} = (\widetilde{\xi}, \widetilde{\eta})$ de $\mathbb{R}^2$ onde

$$\xi = \begin{bmatrix}1\\1\end{bmatrix}, \quad \eta = \begin{bmatrix}-1\\1\end{bmatrix}, \qquad \widetilde{\xi} = \begin{bmatrix}-1\\1\end{bmatrix}, \quad \widetilde{\eta} = \begin{bmatrix}-1\\-1\end{bmatrix}$$

Seja $v = (1, 1) \in \mathbb{R}^2$. Determine o vetores coordenadas $[v]_{\mathcal{U}}$ e $[v]_{\widetilde{\mathcal{U}}}$.[2]

5.3.2 Como a matriz de uma TL muda

Lema 5.3.5. *$A\mathcal{U} = \mathcal{V}\mathbf{a}$ na diagrama (5.0.1).*

Demonstração. Seja $x \in \mathbb{K}^n$, então

$$\begin{aligned} A\mathcal{U}x &= A\,(\xi_1, \ldots, \xi_n)\begin{bmatrix}x_1\\ \vdots \\ x_n\end{bmatrix} = A\,(\xi_1 x_1 + \ldots \xi_n x_n) \\ &= \underbrace{(A\xi_1)}_{\eta_1 a_{11} + \cdots + \eta_m a_{m1}} x_1 + \cdots + \underbrace{(A\xi_n)}_{\eta_1 a_{1n} + \cdots + \eta_m a_{mn}} x_n \\ &= \sum_{j=1}^{m}(\eta_j a_{j1})x_1 + \cdots + \sum_{j=1}^{m}(\eta_j a_{jm})x_n \end{aligned}$$

e assim

$$\begin{aligned} A\mathcal{U}x &= \sum_{k=1}^{n}\sum_{j=1}^{m}(\eta_j a_{jk})x_k = \sum_{j=1}^{m}\eta_j \underbrace{\sum_{k=1}^{n} a_{jk}x_k}_{\mathbf{a}_{j\bullet}x} \\ &= \eta_1 \mathbf{a}_{1\bullet}x + \cdots + \eta_m \mathbf{a}_{m\bullet}x = (\eta_1, \ldots, \eta_m)\begin{bmatrix}\mathbf{a}_{1\bullet}x\\ \vdots \\ \mathbf{a}_{m\bullet}x\end{bmatrix} = \mathcal{V}\mathbf{a}x \end{aligned}$$

□

[2] Controle: $[v]_{\widetilde{\mathcal{U}}} = [I]_{\mathcal{U},\widetilde{\mathcal{U}}}\,[v]_{\mathcal{U}} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}_{\mathcal{U},\widetilde{\mathcal{U}}}\begin{bmatrix}1\\0\end{bmatrix}_{\mathcal{U}} = \begin{bmatrix}0\\-1\end{bmatrix}_{\widetilde{\mathcal{U}}}$, $\widetilde{\mathcal{U}} = [I]_{\widetilde{\mathcal{U}},\mathcal{U}}\,\mathcal{U}$

Segundo as Lemas 5.3.5 e 5.3.1 o diagrama (5.0.1) é comutativo, e assim recebemos a relação

$$\tilde{\mathbf{a}} = \mathbf{qap}^{-1}$$

entre as matrizes de A em respeito às bases novas e velhas.

No caso $F = E$ munido das bases $\mathcal{V} = \mathcal{U}$ e $\tilde{\mathcal{V}} = \tilde{\mathcal{U}}$ recebemos

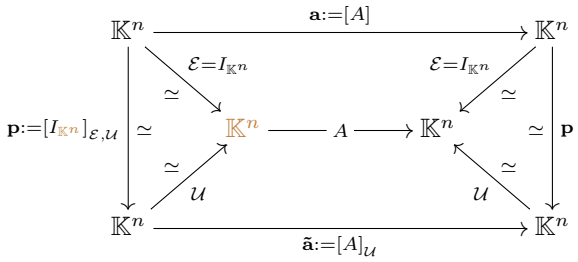
$$\tilde{\mathbf{a}} = \mathbf{pap}^{-1}, \qquad \tilde{\mathbf{a}} = [A]_{\tilde{\mathcal{U}}}, \quad \mathbf{a} = [A]_{\mathcal{U}} \tag{5.3.2}$$

Definição 5.3.6. Chama-se **semelhante** duas matrizes quadradas $\mathbf{a}$ e $\mathbf{b}$ se existe uma matriz invertível $\mathbf{p}$ tal que $\mathbf{a} = \mathbf{p}^{-1}\mathbf{bp}$.

Exercício 5.3.7. Mostre que se $\mathbf{a}$ e $\tilde{\mathbf{a}} = \mathbf{p}^{-1}\mathbf{ap}$ são matrizes $n \times n$ semelhantes, então existe $A \in \mathcal{L}(\mathbb{R}^n)$ tal que $\mathbf{a}$ e $\tilde{\mathbf{a}}$ são matrizes de A relativamente a duas bases de $\mathbb{R}^n$.

Caso especial $E = F = \mathbb{K}^n$ com a base canônica $\mathcal{E}$ e uma base $\mathcal{U}$

Considere o caso de um operador linear $A : \mathbb{K}^n \to \mathbb{K}^n$ onde $\mathbb{K}^n$ é munido originalmente da base canônica $\mathcal{E} = (e_1, \dots, e_n)$ e depois de uma base nova $\mathcal{U}$. Neste caso o diagrama (5.0.1) torna-se no diagrama comutativo



(5.3.3)

o qual disponibiliza a relação $[A]_{\mathcal{U}} = \mathbf{p}\,[A]\,\mathbf{p}^{-1}$ entre as matrizes de A quando trocar a base canônica por qualquer outra base.

Exercício 5.3.8. Suponha que $E = \mathbb{K}^n$ munido da base canônica $\mathcal{E}$ como base velha, veja diagrama (5.3.3). Então os membros da base nova $\mathcal{U} = (\xi_1 \dots, \xi_n)$ formam as colunas da matriz inversa $\mathbf{p}^{-1}$. Veja Exercício 5.4.6.

Nas outras palavras, como a inversa de $\mathbf{p} := [I_{\mathbb{K}^n}]_{\mathcal{E},\mathcal{U}}$ é $[I_{\mathbb{K}^n}]_{\mathcal{U},\mathcal{E}}$, vale a seguinte fórmula

$$\boxed{[I_{\mathbb{K}^n}]_{\mathcal{U},\mathcal{E}} = [\mathcal{U}]}$$

onde $[\mathcal{U}]$ denota a matriz cujas colunas sao os elementos da base $\mathcal{U}$ de $\mathbb{K}^n$.

Exemplo 5.3.9. Em $\mathbb{R}^3$ considere a base canônica $\mathcal{E}$ e a base $\mathcal{V} = \{\xi_1, \xi_2, \xi_3\}$

$$\xi_1 = (1, 1, 0), \qquad \xi_2 = (-1, 0, 0), \qquad \xi_3 = (0, 0, 1)$$

Determine a matriz de passagem $\mathbf{p}$ de $\mathcal{V}$ para $\mathcal{E}$, e aquela vice versa.

Uma solução. As colunas da matriz $\mathbf{p} := [I]_{\mathcal{V},\mathcal{E}} = [\mathcal{V}]$ são os ξ_i's. A outra matriz desejada $\mathbf{q} := [I]_{\mathcal{E},\mathcal{V}} = \mathbf{p}^{-1}$ é a matriz inversa de $\mathbf{p}$. Pode-se calcular com o processo de Gauss-Jordan (MA141), veja § A.4, ou seja

$$[\mathbf{p} : \mathbb{1}] = \begin{bmatrix} 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{\text{(oe)}} \dots \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} = [\mathbb{1} : \mathbf{p}^{-1}]$$

Exercício 5.3.10. Seja $\mathcal{E} \subset \mathbb{R}^3$ a base canônica e seja $A\colon \mathbb{R}^3 \to \mathbb{R}^3$ determinado por (5.2.3). Dada a base ordenada $\mathcal{V} := (e_1, e_1 + e_3, e_1 + e_2)$, determine a matriz

$$\mathbf{a} = [A]_{\mathcal{E},\mathcal{V}} = [I]_{\mathcal{E},\mathcal{V}}[A]_{\mathcal{V},\mathcal{E}}[I]_{\mathcal{E},\mathcal{V}} = \mathbf{p}\tilde{\mathbf{a}}\mathbf{p}$$

calculando $\mathbf{p}$ e $\tilde{\mathbf{a}}$, veja (5.0.1).[3] (Reencontramos A nos Exercícios 5.2.1 e 8.1.3.)

5.3.3 Determinante de uma transformação linear

Definição 5.3.11 (Determinante)**.** Para $A \in \mathcal{L}(A)$ define-se

$$\det A := \det [A]_{\mathcal{B}} \tag{5.3.4}$$

onde $\mathcal{B}$ é uma matriz ordenada de E.

Exercício 5.3.12. Verifique que $\det A$ é bem definido, o que quer dizer que é independente da escolha da base.

[Dica: Fórmula (5.3.2)]

5.4 Exercícios e umas soluções

Matriz de uma transformação linear

Exercício 5.4.1. Considere a base ordenada $\mathcal{B} = (u, v, w)$ de $\mathbb{R}^3$, onde

$$u = (1, 1, 1), \qquad v = (1, -1, 1), \qquad w = (1, 1 - 1)$$

Seja $\mathcal{B}^* = (\phi, \psi, \chi)$ a base (de $\mathbb{R}^{3^*}$) dual de $\mathcal{B}$. Calcule as matrizes $[\phi], [\psi], [\chi]$ das transformações lineares $\phi, \psi, \chi\colon \mathbb{R}^3 \to \mathbb{R}$.

Uma solução. Sejam $\mathcal{E}^3 = (e_1, e_2, e_3)$ e $\mathcal{E}^1 = (E_1)$ as bases canônicas onde o vetor unitário em $\mathbb{R}^1$ é a lista $E_1 = (1)$ de um membro com 1 no 1-ésimo lugar (e nulos nos outros lugares – as quais não têm). Seja $x_i := \phi e_i$, então usando a propriedade da base dual e linearidade obtemos

$$\begin{aligned} 1 = \phi u = \phi(1,1,1) &= \phi e_1 + \phi e_2 + \phi e_3 = x_1 + x_2 + x_3 \\ 0 = \phi v = \phi(1,-1,1) &= \phi e_1 - \phi e_2 + \phi e_3 = x_1 - x_2 + x_3 \\ 0 = \phi w = \phi(1,1,-1) &= \phi e_1 + \phi e_2 - \phi e_3 = x_1 + x_2 - x_3 \end{aligned}$$

[3] controle: inverta $\mathbf{q} = [I]_{\mathcal{V},\mathcal{E}}$ (Gauss-Jordan): $\mathbf{p} = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, $\mathbf{a} = \begin{bmatrix} 4 & -2 & -2 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$.

Usamos escalonamento para resolver o SL de 3 equações nas 3 incógnitas $x_1, x_2, x_3 \in \mathbb{R}^3$ as quais são listas de um membro só. O resultado é

$$x_1 = 1, \qquad x_2 = 0, \qquad x_3 = 0$$

Observamos que

$$\begin{aligned} 1 = x_1 = \phi e_1 = E_1\phi_{11} = 1 \cdot \phi_{11} = \phi_{11} \\ 0 = x_2 = \phi e_2 = E_1\phi_{12} = 1 \cdot \phi_{12} = \phi_{12} \\ 0 = x_3 = \phi e_3 = E_1\phi_{13} = 1 \cdot \phi_{13} = \phi_{13} \end{aligned}$$

e assim

$$[\phi]_{\mathcal{E}^3,\mathcal{E}^1} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

Analogamente obtemos

$$[\psi]_{\mathcal{E}^3,\mathcal{E}^1} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}, \qquad [\chi]_{\mathcal{E}^3,\mathcal{E}^1} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

Olha só, no caso $E = \mathbb{R}^n$ a matriz da base dual de qualquer base $\mathcal{B}$ de $\mathbb{R}^n$ tem em respeito às bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}^1$ a forma da base canônica.

Exercício 5.4.2. Considere as transformações lineares

$$A : \mathbb{R}^2 \to \mathbb{R}^3, \quad (x, y) \mapsto (x, y, x + y)$$

e $B : \mathbb{R}^3 \to \mathbb{R}^2$ definido assim

$$B(x, y, z) = (ax + (a - 1)y + (1 - a)z, -bx + (1 - b)y + bz)$$

onde $a, b \in \mathbb{R}$ são constantes. Determine o operador $BA \in \mathcal{L}(\mathbb{R}^2)$.

[Dica: Use as matrizes $[A]$ e $[B]$ que correspondem a A e B respectivamente.]

Uma solução. Denotamos de $\mathcal{E}^2 = \{e_1, e_2\}$ e $\mathcal{E}^3 = \{E_1, E_2, E_3\}$ as bases canônicas. As duas colunas da matriz $[A] \in \mathrm{M}(3 \times 2)$ são formadas das coeficientes seguintes

$$\begin{aligned} Ae_1 = A(1, 0) = (1, 0, 1) = 1E_1 + 0E_2 + 1E_3 \\ Ae_2 = A(0, 1) = (0, 1, 1) = 0E_1 + 1E_2 + 1E_3 \end{aligned}$$

As três colunas da matriz $[B] \in \mathrm{M}(2 \times 3)$ são formadas das coeficientes seguintes

$$\begin{aligned} BE_1 = B(1, 0, 0) = (a, -b) = ae_1 - be_2 \\ BE_2 = B(0, 1, 0) = (a - 1, 1 - b) = (a - 1)e_1 + (1 - b)e_2 \\ BE_3 = B(0, 0, 1) = (1 - a, b) = (1 - a)e_1 + be_2 \end{aligned}$$

Assim recebemos

$$[A] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \qquad [B] = \begin{bmatrix} a & a - 1 & 1 - a \\ -b & 1 - b & b \end{bmatrix}$$

Use (5.2.5) no primeiro passo e no segundo calcule produto matriz para obter

$$[BA] = [B]\,[A] = \begin{bmatrix} a+1-a1 & a-1+1-a \\ -b+b & 1-b+b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbb{1}_2$$

Use a relação (5.2.1) entre matriz e operador para concluir que $BA = I_{\mathbb{R}^2}$.

Exercício 5.4.3. Qual é a matriz $[A]$ do operador $A : \mathbb{R}^2 \to \mathbb{R}^2$ definido por

$$A(2,3) = (2,3) \qquad \text{e} \qquad A(-3,2) = (0,0)\ ?$$

Exercício 5.4.4. Considere as transformações lineares

$$A : \mathbb{R}^{n+1} \to \mathcal{P}_n(\mathbb{R}), \quad (\alpha_0, \alpha_1, \ldots, \alpha_n) \mapsto \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n,$$

e

$$B : \mathcal{P}_n(\mathbb{R}) \to \mathbb{R}^{n+1}, \quad p = p(x) \mapsto (p(0), p(1), \ldots, p(n))\,.$$

Determina a matriz $[BA]$ da composição $BA : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$.

Uma solução. Seja $\mathcal{E} = (e_1, \ldots, e_{n+1})$ a base canônica e seja

$$\mathcal{M} = (x^0, x, x^2, \ldots, x^n) =: (\eta_1, \ldots, \eta_{n+1}), \qquad x^0 := 1$$

a base de $\mathcal{P}_n(\mathbb{R})$ composto de monômios. Segundo a definição de A recebemos

$$Ae_i = A(0, \ldots, 0, 1, 0, \ldots, 0) = 0x^0 + \cdots + 0x^{i-1} + 1x^i + 0x^{i+1} + \cdots + 0x^n$$

para $i = 1, \ldots, n+1$. Segundo (5.2.1) obtemos $[A]_{\mathcal{E},\mathcal{M}} = \mathbb{1}_{n+1}$. Analogamente

$$B\eta_i = (\eta_i(0), \eta_i(1), \ldots, \eta_i(n)) = \left(0^{i-1}, 1^{i-1}, \ldots, n^{i-1}\right) = \sum_{j=1}^{n+1} e_j \underbrace{(j-1)^{i-1}}_{b_{ji}}$$

para cada $i = 1, \ldots, n+1$ o que nos da a i-ésima coluna da matriz

$$[B]_{\mathcal{M},\mathcal{E}} = \begin{bmatrix} 0^0 & 0^1 & \ldots & 0^n \\ 1^0 & 1^1 & \ldots & 1^n \\ 2^0 & 2^1 & \ldots & 2^n \\ \vdots & \vdots & \ldots & \vdots \\ n^0 & n^1 & \ldots & n^n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 1 & 1 & \ldots & 1 \\ 1 & 2 & \ldots & 2^n \\ \vdots & \vdots & \ldots & \vdots \\ 1 & n & \ldots & n^n \end{bmatrix}$$

Assim $[B]_{\mathcal{M},\mathcal{E}} = [B]_{\mathcal{M},\mathcal{E}}\,\mathbb{1}_{n+1} = [B]_{\mathcal{M},\mathcal{E}}\,[A]_{\mathcal{E},\mathcal{M}} = [BA]_{\mathcal{E},\mathcal{E}}$ segundo (5.2.5).

Exercício 5.4.5. Dado $w = (\alpha, \beta, \gamma) \in \mathbb{R}^3$, determine a matriz $[A]$ do operador[4]

$$A : \mathbb{R}^3 \to \mathbb{R}^3, \qquad v \mapsto v \times w$$

Descreva geometricamente o núcleo desse operador e determina sua imagem.

[4] O **produto vetorial** de dois vetores v e w de $\mathbb{R}^3$ é o vetor $v \times w$ de $\mathbb{R}^3$ definido por

$$v \times w = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \times \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} := \begin{bmatrix} y\gamma - z\beta \\ -(x\gamma - z\alpha) \\ x\beta - y\alpha \end{bmatrix} = \begin{bmatrix} y\gamma - z\beta \\ z\alpha - x\gamma \\ x\beta - y\alpha \end{bmatrix}$$

Mudança de base – vetor

Mudança de base – matriz

Exercício 5.4.6. Seja $\mathcal{E} = (e_1, \ldots, e_n)$ a base canônica de $\mathbb{R}^n$. Suponha vetores $\xi_1, \ldots, \xi_n \in \mathbb{R}^n$ e escalares $p_{ij} \in \mathbb{R}$ satisfazem

$$e_j = \xi_1 p_{1j} + \xi_2 p_{2j} + \cdots + \xi_n p_{nj}, \qquad \forall j = 1, \ldots, n$$

Mostre que

1. a lista ordenada $\mathcal{U} = (\xi_1, \ldots, \xi_n)$ é uma base de $\mathbb{R}^n$.
2. a matriz $\mathbf{p} = [p_{ij}]$ é a inversa da matriz $\mathbf{q}$ cujas colunas são por definição $\xi_1, \ldots, \xi_n$, em símbolos $\mathbf{qp} = \mathbb{1}_n$. Veja Exercício 5.3.8.

Em palavras, obtém-se a matriz $\mathbf{p} := [I_{\mathbb{R}^n}]_{\mathcal{E},\mathcal{U}}$ de mudança da base canônica $\mathcal{E}$ para uma base nova $\mathcal{U} = (\xi_1, \ldots, \xi_n)$ como a matriz inversa da matriz

$$\mathbf{q} := [\xi_1 \ldots \xi_n]$$

cujas colunas são os ξ_i's, em símbolos

$$\mathbf{p} := [I_{\mathbb{R}^n}]_{\mathcal{E},(\xi_1,\ldots,\xi_n)} = \mathbf{q}^{-1} = [\xi_1 \ldots \xi_n]^{-1}$$

Exercício 5.4.7 (Continuação de Exercício 1.3.9)**.** Seja $E = \mathcal{F}(\mathbb{R}, \mathbb{C})$ o espaço vetorial sobre o corpo $\mathbb{C}$ composto de todas as funções $f : \mathbb{R} \to \mathbb{C}$. Em Exercício 1.3.9 temos definido dois subconjuntos $\mathcal{U} = \{f_1, f_2, f_3\}$ e $\mathcal{V} = \{g_1, g_2, g_3\}$ e temos mostrado, primeiro, que cada um é LI e, segundo, que cada um elemento de um dos dois subconjuntos pode ser escrito como CL dos elementos do outro.

(a) Mostre que os subespaços gerados $F := \langle \mathcal{U} \rangle$ e $\langle \mathcal{V} \rangle$ são iguais.

(b) Conclua que $\mathcal{U}$ e $\mathcal{V}$ são bases de F.

(c) Determine a matriz de passagem $\mathbf{p} := [I_F]_{\mathcal{U},\mathcal{V}}$ da base $\mathcal{U}$ para a base $\mathcal{V}$. (Outros autores usam a notação $\mathbf{p} = I^{\mathcal{U}}_{\mathcal{V}}$.)

Exercício 5.4.8.

Seja $\mathbf{c} \in M(n \times n; \mathbb{K})$ uma matriz quadrada de posto 1.

(a) Prove que: $\mathbf{c}^2 = (\operatorname{tr} \mathbf{c})\mathbf{c}$. $(*)$

(b) Dado $n \geq 2$, generalize: $\mathbf{c}^n = (\operatorname{tr} \mathbf{c})^{n-1}\mathbf{c}$.

Poderia usar (e provar) o

Lema 5.4.9. *Para matrizes $\mathbf{a}, \mathbf{b} \in M(n \times n; \mathbb{K})$ tem-se*

$$\operatorname{tr}(\mathbf{ab}) = \operatorname{tr}(\mathbf{ba}) \tag{5.4.1}$$

[ad (a): Observe que para provar $(*)$ pode-se mudar a base de $\mathbb{K}^n$ e provar $(*)$ para a nova matriz $\tilde{\mathbf{c}}$. Lembre-se: posto$(\mathbf{c}) = 1$. Escolha base apropriada de $\mathbb{K}^n$.]

Exercício 5.4.10. Sejam $A : E \to F$ e $B : F \to G$ transformações lineares entre espaços vetoriais de dimensão finita.

1. Prove que: B injetiva $\Rightarrow$ posto$(BA) =$ posto(A).
2. Encontre uma condição sobre A a qual implica posto$(BA) =$ posto(B).

Capítulo 6

Núcleo e imagem

No Capítulo 6 denotamos de E, F espaços vetoriais

$$E = (E, +, \cdot, \mathbb{K}), \qquad F = (F, +, \cdot, \mathbb{K})$$

ambos sobre o mesmo corpo $\mathbb{K}$, e denotamos de $A\colon E \to F$ uma transformação linear. Na primeira leitura pense em $\mathbb{K} = \mathbb{R}$. As letras m, n denotam números naturais ou zero.

O primeiro objetivo no Capítulo 6 é associar a uma transformação linear $A \in \mathcal{L}(E, F)$ dois subespaços

$$\mathrm{N}(A) \subset E \xrightarrow{A} F \supset \mathrm{Im}(A)$$

e relacionar seu tamanho mínimo/máximo a injetividade/sobrejetividade de A. Outros resultados fundamentais são os seguintes: Uma transformação linear A é injetiva se e somente se leva conjuntos LI em conjuntos LI. Sobrejetividade é equivalente à existencia de uma inversa à direita e injetividade à existencia de uma inversa à esquerda.

Definição 6.0.11. Dado uma transformação linear $A \in \mathcal{L}(E, F)$ chamamos

$\mathrm{N}(A) := \{v \in E \mid Av = \mathcal{O}\}$ o **núcleo** de A e

$\mathrm{Im}(A) := \{Av \mid v \in E\}$ a **imagem** de A

Dado um subconjunto $X \subset E$, seja $AX := \{Ax \mid x \in X\}$ a imagem de X sob A.

Lema 6.0.12. *Os subconjuntos* $\mathrm{N}(A) \subset E$ *e* $\mathrm{Im}(A) \subset F$ *são subespaços.*

Demonstração. "$\mathrm{Im}(A)$ fechado sob $+$": Dado dois elementos da imagem, ou seja Av e Aw onde $v, w \in E$, então de linearidade $Av + Aw = A(v+w) \in \mathrm{Im}(A)$. "$\mathrm{Im}(A)$ fechado sob $\cdot$": Se $\alpha \in \mathbb{K}$ e $Av \in \mathrm{Im}(A)$, então $\alpha Av = A(\alpha v) \in \mathrm{Im}(A)$. Deixamos ao leitor provar que $\mathrm{N}(A)$ é fechado sob $\cdot$ e $+$. □

Definição 6.0.13 (Posto)**.** A dimensão da imagem é chamado de **posto de uma transformação linear** $A \in \mathcal{L}(E, F)$, em simbolos

$$\operatorname{posto}(A) := \dim \operatorname{Im}(A)$$

Lema 6.0.14 (Os dois subespaços naturais – mínimo e máximo)**.** *Para uma transformação linear $A \in \mathcal{L}(E, F)$ injetividade e sobrejetividade correspondem a*

(i) $\operatorname{N}(A) = \{\mathcal{O}\} \quad \Leftrightarrow \quad A$ *é injetivo*

(ii) $\operatorname{Im}(A) = F \quad \Leftrightarrow \quad A$ *é sobrejetivo*

Demonstração. (i) **"⇐"** '$\subset$' Seja $v \in \operatorname{N}(A)$, então $Av = \mathcal{O} = A\mathcal{O}$ onde temos usado linearidade no segundo passo. Então segundo injetividade como as imagens são iguais, os elementos $v = \mathcal{O}$ devem ser iguais. '$\supset$' Como subespaço $\operatorname{N}(A)$ contem o vetor nulo. **"⇒"** Suponha que são iguais as imagens $Av = Aw$ de dois elementos $v, w \in E$. Então $\mathcal{O} = Av - Aw = A(v - w)$, e assim $v - w \in \operatorname{N}(A) = \{\mathcal{O}\}$. Então $v = w$.
(ii) **"⇐"** '$\subset$' trivial. '$\supset$' Dado $f \in F$, como A é sobrejetivo existe um $v \in E$ tal que $f = Av$. Assim $f \in \operatorname{Im}(A)$. **"⇒"** Seja $f \in F = \operatorname{Im}(A)$, ou seja $f = Av$ para um $v \in E$, mostrando que A é sobrejetivo. □

Lema 6.0.15. *Seja $A \in \mathcal{L}(E, F)$ e $X \subset E$, então*

$$\langle X \rangle = E \quad \Rightarrow \quad \langle AX \rangle = \operatorname{Im}(A)$$

Demonstração. '$\subset$' Sem usar $\langle X \rangle = E$, um elemento $f \in \langle AX \rangle$ é da forma de uma soma finita $f = \sum \alpha_i A x_i = A \sum \alpha_i x_i \in \operatorname{Im}(A)$ onde $x_i \in X$ e $\alpha_i \in \mathbb{K}$.
'$\supset$' Um elemento $f \in \operatorname{Im}(A) = AE = A\langle X \rangle$ é da forma de uma soma finita $f = A \sum \alpha_i x_i = \sum \alpha_i A x_i \in \langle AX \rangle$ onde $x_i \in X$ e $\alpha_i \in \mathbb{K}$. □

Como $\operatorname{Im}(A) \subset F$ é um subespaço já sabemos de Teorema 3.2.1 que sua dimensão não é maior daquela de F. É uma surpresa que isso vale para $\dim E$ também. Este fato será utilizado no famoso Teorema 6.5.1 de nucelo e imagem.

Corolário 6.0.16. $\forall A \in \mathcal{L}(E, F)$ *vale* $\dim \operatorname{Im}(A) \leq \dim E$.

Demonstração. Se $\dim E = \infty$ não tem nada a provar. No caso $n = \dim E \in \mathbb{N}_0$ seja $\mathcal{B} = \{\xi_1, \dots, \xi_n\}$ uma base de E. Como $\langle \mathcal{B} \rangle = E$ temos $\langle A\mathcal{B} \rangle = \operatorname{Im}(A)$ segundo Lema 6.0.15. Nas outras palavras, o conjunto finito $A\mathcal{B} = \{A\xi_i \mid i = 1, \dots, n\}$ de $m \leq n$ elementos (possivelmente uns $A\xi_i$'s são iguais) gera o espaço vetorial $\operatorname{Im}(A)$. Então $\dim \operatorname{Im}(A) \leq m \leq n = \dim E$ segundo Lema 3.1.20. □

Exercício 6.0.17. Defina operadores lineares $A, B : \mathbb{R}^\infty \to \mathbb{R}^\infty$ como

$$A(x_1, x_2, x_3, \dots) := (x_1, 0, x_2, 0, x_3, 0, \dots)$$
$$B(x_1, x_2, x_3, \dots) := (x_2 - 2x_1, x_3 - 2x_2, \dots).$$

Determine o núcleo e a imagem de A e de B.

Exemplo 6.0.18 (SL)**.** Dada uma matriz $\mathbf{a} \in \mathrm{M}(m \times n; \mathbb{K})$ e uma lista $b \in \mathbb{K}^m$. São equivalente os seguintes:

O sistema linear (SL) de m equações a n incógnitas

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

admite uma solução $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$.

$\overset{(A.3.2)}{\iff}$ O vetor b é CL das colunas $\mathbf{a}_{\bullet 1}, \ldots, \mathbf{a}_{\bullet n}$ da matriz $\mathbf{a}$; veja (1.2.3)

$\overset{2}{\iff}$ $b \in \mathrm{Im}(\mathbf{a})$ considerando a matriz como transformação linear $\mathbf{a}\colon \mathbb{K}^n \to \mathbb{K}^m$

$\overset{3}{\iff}$ $\langle \mathbf{a}_{\bullet 1}, \ldots, \mathbf{a}_{\bullet n}, b\rangle = \langle \mathbf{a}_{\bullet 1}, \ldots, \mathbf{a}_{\bullet n}\rangle$

Equivalência 2: Isso é o fato que a imagem de uma matriz é o espaço-coluna, ou seja o conjunto de todas as CLs das colunas da matriz.

Equivalência 3: "⇒" Se b é CL das colunas a igualdade é trivial. Caso geral: Como as colunas $\{\mathbf{a}_{\bullet 1}, \ldots, \mathbf{a}_{\bullet n}\}$ geram $\mathrm{Im}(\mathbf{a})$, $b \in \mathrm{Im}(\mathbf{a})$, e $\mathrm{Im}(\mathbf{a})$ é um subespaço obtemos a primeira identidade

$$\langle \mathbf{a}_{\bullet 1}, \ldots, \mathbf{a}_{\bullet n}, b\rangle = \mathrm{Im}(\mathbf{a}) = \langle \mathbf{a}\mathcal{E}^n\rangle = \langle \underbrace{\mathbf{a}e_1}_{\mathbf{a}_{\bullet 1}}, \ldots, \underbrace{\mathbf{a}e_n}_{\mathbf{a}_{\bullet n}}\rangle$$

Identidade dois segue do Lema 6.0.15 com a base canônica $\mathcal{E}^n$ de $\mathbb{K}^n$.
"⇐" A identidade fala que b é CL das colunas $\mathbf{a}_{\bullet i} = \mathbf{a}e_i \in \mathrm{Im}(\mathbf{a})$, então $b \in \mathrm{Im}(\mathbf{a})$ porque $\mathrm{Im}(\mathbf{a})$ é um subespaço e assim fechado sob adição.

Exercício 6.0.19. Seja $A : \mathbb{R}^4 \to \mathbb{R}^3$ dada por:

$$(x, y, z, t) \mapsto (x + y + z + 2t, x - y + 2z, 4x + 2y + 5z + 6t)\,.$$

Encontre $b \in \mathbb{R}^3$ que não pertença à imagem de A. Com b, exiba um sistema linear de 3 equações e 4 incógnitas sem solução.

6.1 Escalonamento: Cálculo do posto

Aplicamos o processo de escalonar uma matriz $\mathbf{a}$ – conteúdo do curso MA141 e revisado no Apêndice A.2 – para calcular o posto de uma transformação linear.

Lema 6.1.1. *Dado uma transformação linear $A\colon E \to F$ entre espaços vetoriais das dimensões finitas n e m, respectivamente, então*

$$\underbrace{\mathrm{posto}(A)}_{:=\dim \mathrm{Im}(A)} = \underbrace{\mathrm{posto}(\mathbf{a})}_{:=\dim \mathrm{Im}(\mathbf{a})}, \qquad \mathbf{a} := [A]_{\mathcal{U},\mathcal{V}}$$

para qualquer escolha de bases ordenadas $\mathcal{U}$ de E e $\mathcal{V}$ de F. Repetimos que vale $\mathrm{posto}(\mathbf{a}) = \mathrm{pc}(\mathbf{a}) = \mathrm{pl}(\mathbf{a})$ *segundo Teorema 4.2.2.*

Demonstração. Temos que

$$\operatorname{posto}(A) := \dim \operatorname{Im}(A) = \dim \operatorname{Im}\big([A]_{\mathcal{U},\mathcal{V}}\big) = \dim \operatorname{Im}(\mathbf{a}) =: \operatorname{posto}(\mathbf{a})$$

onde a primeira identidade segue do isomorfismo entre as imagens das transformações lineares A e $\mathbf{a} := [A]_{\mathcal{U},\mathcal{V}}$, veja a diagrama comutativa (5.0.1). □

Exemplo 6.1.2. Dado $A \in \mathcal{L}(E,F)$, determine $\operatorname{posto}(A)$ e uma base de $\operatorname{Im}(A)$.

Uma solução. Escolhe bases e considere a matriz $\mathbf{a} := [A]_{\mathcal{U},\mathcal{V}}$ de A. Aplique escalonamento para a transposta $\mathbf{a}^t$, então as linhas não-nulas de $(\mathbf{a}^t)_{\text{esc}}$, dizemos $\ell_1, \dots, \ell_d$, formam uma base de

$$\text{Esp-lin}((\mathbf{a}^t)_{\text{esc}}) \overset{\text{Teor.A.2.2}}{=} \text{Esp-lin}(\mathbf{a}^t) = \text{Esp-col}(\mathbf{a}) \overset{(4.2.3)}{=} \operatorname{Im}(\mathbf{a})$$

e o $\operatorname{posto}(A) \overset{\text{Le.6.1.1}}{=} \operatorname{posto}(\mathbf{a}) := \dim \operatorname{Im}(\mathbf{a}) = d$ é dado pelo número d das linhas não-nulas do escalonamento da transposta $\mathbf{a}^t$. Resta traduzir a base $\{\ell_1, \dots, \ell_d\}$ de $\operatorname{Im}(\mathbf{a}) \subset \mathbb{K}^m$ numa base de $\operatorname{Im}(A) \subset F$. Usamos o isomorfismo $\mathcal{V} : \mathbb{K}^m \to F$ gerado pela base ordenada $\mathcal{V} = (\eta_1, \dots, \eta_m)$ de F. Definimos

$$\zeta_i := \mathcal{V}\ell_i \overset{\text{def.}}{=} (\eta_1, \dots, \eta_m) \begin{bmatrix} (\ell_i)_1 \\ \vdots \\ (\ell_i)_m \end{bmatrix} \overset{\text{def.}}{=} \eta_1(\ell_i)_1 + \cdots + \eta_m(\ell_i)_m \in F$$

para obter uma base $\{\zeta_1, \dots \zeta_d\}$ de $\operatorname{Im}(A)$.

Exercício 6.1.3. Encontre o posto de $\mathbf{a} := \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ e uma base da imagem.[1]

Exercício 6.1.4. Determine o posto da matriz

$$\mathbf{a} = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 2 & 0 & 1 & -1 \end{bmatrix}.$$

[Dicas: Calcule o posto-linha da matriz transposta. Escalonamento (modificando linhas) não muda o espaço-linha.]

6.2 Sobrejetividade – inversa à direita

Definição 6.2.1. Dado $A \in \mathcal{L}(E,F)$, uma transformação linear $B \in \mathcal{L}(F,E)$ é chamado de *uma* **inversa à direita de A** se a composição satisfaz $AB = I_F$.

[1] Obtém-se posto 2 e uma base é $\{(1,1,1),(0,1,2)\}$. Com efeito, um escalonamento é

$$(\mathbf{a}^t)_{\text{esc}} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Exemplo 6.2.2 (Geralmente inversas à direita não são únicas)**.** Seja $a \in \mathbb{R}$ e

$$A : \mathbb{R}^3 \to \mathbb{R}^2, \quad (x, y, z) \mapsto (x, y), \qquad B_a : \mathbb{R}^2 \to \mathbb{R}^3, \quad (x, y) \mapsto (x, y, ax).$$

Então $AB_a = I_{\mathbb{R}^2}$ para cada um $a \in \mathbb{R}$, mas $B_a \neq B_b$ caso $a \neq b$.

Teorema 6.2.3. *Suponha $A \in \mathcal{L}(E, F)$ onde $m := \dim F < \infty$. Então*

$$A \text{ admite uma inversa à direita} \qquad \Leftrightarrow \qquad A \text{ sobrejetivo}$$

Demonstração. "⇒" Dado uma inversa à direita B de A, então $\forall f \in F$ vale $ABf = I_F f = f$. Assim para todo $f \in F$ existe um $v \in E$, com efeito $v := Bf$, tal que $Av = f$. Mas isso significa que A é sobrejetivo.
"⇐" Usamos sobrejetividade de A para construir explicitamente uma inversa à direita de A. Escolha uma base ordenada $\mathcal{Y} = (\eta_1, \ldots, \eta_m)$ de F e, usando sobrejetividade, uma lista $v = (v_1, \ldots, v_m)$ de m elementos de E tal que $Av_j = \eta_j$ para $j = 1, \ldots, m$. Lembramos de (4.1.4) que a lista v nos da uma transformação linear $B_v : F \to E$ unicamente determinado pelos valores $B_v\eta_j := v_j$ nos membros da base $\mathcal{Y}$. Resta checar $AB_v = I_F$: Escrevendo $f \in F$ como CL única na base $\mathcal{Y}$, ou seja $f = \sum_j \beta_j \eta_j$, e usando linearidade de B_v e de A obtemos

$$AB_v f = AB_v \sum_j \beta_j \eta_j = A \sum_j \beta_j \underbrace{B_v \eta_j}_{v_j} = \sum_j \beta_j A v_j = \sum_j \beta_j \eta_j = f$$

Note-se a soma é finita porque temos exprimido f como uma CL. □

Exercício 6.2.4. (a) Mostre que $\{0\}$ e o próprio $\mathbb{R}$ são os únicos subespaços de $\mathbb{R}$.

(b) Seja E um espaço vetorial sobre o corpo $\mathbb{R}$. Mostre que $f \in \mathcal{L}(E, \mathbb{R})$ é sobrejetivo ou igual a zero.

(c) Mostre que a derivação $D : \mathcal{P}_n(\mathbb{R}) \to \mathcal{P}_{n-1}(\mathbb{R})$, $p(x) \mapsto \frac{d}{dx}p(x)$, é sobrejetiva.

(d) Mostre que a derivação $D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$, $f(x) \mapsto \frac{d}{dx}f(x)$, é sobrejetiva.

(e) Encontre uma inversa à direita $J : \mathcal{P}_{n-1}(\mathbb{R}) \to \mathcal{P}_n(\mathbb{R})$ para a derivação D em iii).

6.3 Injetividade – inversa à esquerda

Teorema 6.3.1. *Dada uma transformação linear $A \in \mathcal{L}(E, F)$, então*

$$A \text{ injetivo} \qquad \Leftrightarrow \qquad A \text{ leva conjuntos LI em conjuntos LI}$$

Demonstração. "⇒" Seja A injetivo e $X \subset E$ um subconjunto LI. Pegue elementos $Ax_1, \ldots, Ax_\ell \in AX$ na imagem e suponha que uma CL deles representa o vetor nulo, ou seja $\mathcal{O} = \alpha_1 Ax_1 + \cdots + \alpha_\ell Ax_\ell = A(\alpha_1 x_1 + \cdots + \alpha_\ell x_\ell)$ onde

os α_i's são escalares. Como A é injetivo, equivalentemente $\mathrm{N}(A) = \{\mathcal{O}\}$ segundo Lema 6.0.14, segue que $\alpha_1 x_1 + \cdots + \alpha_\ell x_\ell = \mathcal{O}$. Como X é LI segue que $\alpha_1 = \cdots = \alpha_\ell = 0$ e isso prova que o subconjunto $AX \subset F$ é LI.
"$\Leftarrow$" Seja $v \in E$. Se $v \neq \mathcal{O}$, então o subconjunto $\{v\} \subset E$ é LI segundo Comentário 1.3.4 (ii). Assim $\{Av\} \subset F$ é LI como A leva LI em LI segundo hipótese. Assim o vetor Av não pode ser nulo. Temos provado $v \neq \mathcal{O} \Rightarrow Av \neq \mathcal{O}$. O contra-positivo então diz que $Av = \mathcal{O} \Rightarrow v = \mathcal{O}$. Assim $\mathrm{N}(A) = \{\mathcal{O}\}$. □

Corolário 6.3.2. *Se $A \in \mathcal{L}(E, F)$ é injetivo, então* $\dim E \leq \dim F$.

Demonstração. Se $\dim F = \infty$ não tem nada a provar. Seja $m := \dim F \in \mathbb{N}_0$. Seja $B = \{v_1, \ldots, v_k\}$ um subconjunto LI de E, então como A leva LI em LI segundo Teorema 6.3.1, o conjunto $AB = \{Av_1, \ldots, Av_k\}$ é LI em F e por isso (Corolário 3.1.17) não pode conter mais elementos como $\dim F$, em símbolos $k := |B| = |AB| \leq m$. (Como A é injetivo vale $|B| = |AB|$.) Analogamente à prova da parte (a) de Teorema 3.2.1 um subconjunto $B_* \subset E$ LI com o número máximo n ($\leq m$) de elementos gera E e assim é uma base de E. Assim $\dim E := |B_*| = n \leq m := \dim F$. □

Exemplo 6.3.3 (Aplicação)**.** Não existe nenhuma transformação linear $A : \mathbb{R}^4 \to \mathbb{R}^2$ a qual é injetiva.

Definição 6.3.4. Dado $A \in \mathcal{L}(E, F)$, uma transformação linear $B \in \mathcal{L}(F, E)$ é chamado de *uma* **inversa à esquerda de A** se a composição satisfaz $BA = I_E$.

Exemplo 6.3.5 (Geralmente inversas à esquerda não são únicas)**.** Seja $a \in \mathbb{R}$ e

$$A : \mathbb{R}^2 \to \mathbb{R}^3, \quad (x, y) \mapsto (x, y, 0), \qquad B_a : \mathbb{R}^3 \to \mathbb{R}^2, \quad (x, y, z) \mapsto (x + az, y).$$

Então $B_a A = I_{\mathbb{R}^2}$ para cada um $a \in \mathbb{R}$, mas $B_a \neq B_b$ caso $a \neq b$.

Teorema 6.3.6. *Suponha $A \in \mathcal{L}(E, F)$ onde* $\dim E, \dim F < \infty$. *Então*

$$A \text{ admite uma inversa à esquerda } B \qquad \Leftrightarrow \qquad A \text{ injetivo}$$

Demonstração. "$\Rightarrow$" Se $Au = Av$, então $BAu = BAv$. Mas $BA = I$, daí $u = v$.
"$\Leftarrow$" Como a dimensão $n := \dim E$ é finita, escolha uma base $\mathcal{X} = \{\xi_1, \ldots, \xi_n\}$ de E. Baseado na injetividade de A, segundo Teorema 6.3.1, o conjunto das imagens $\{A\xi_1, \ldots, A\xi_n\}$ é LI em F e pode ser estendido, segundo Teorema 3.2.1 (b) usando $\dim F < \infty$, para obter a base $\mathcal{B} := \{A\xi_1, \ldots, A\xi_n, \eta_1, \ldots, \eta_k\}$ de F. A lista $w := (\xi_1, \ldots, \xi_n, \mathcal{O}, \ldots, \mathcal{O}) \in E^{\times(n+k)}$ determina $B_w \in \mathcal{L}(F, E)$ segundo (4.1.4), ou seja $B_w(A\xi_i) := \xi_i$ e $B_w \eta_j := \mathcal{O}$. Escreve $v \in E$ como CL única $v = \sum_i \alpha_i \xi_i$. Então usando linearidade de A e de B_w obtemos

$$B_w A v = B_w A \sum_i \alpha_i \xi_i = \sum_i \alpha_i \underbrace{B_w A \xi_i}_{\xi_i} = v = I_E v$$

para cada um $v \in E$. □

Exercício 6.3.7. Determine uma base para a imagem de cada uma das transformações lineares abaixo e indique quais são sobrejetivas.

(a) $A : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x - y, x - y)$;

(b) $B : \mathbb{R}^4 \to \mathbb{R}^4, (x, y, z, t) \mapsto (x + y, z + t, x + z, y + t)$;

(c) $C : \mathbb{R}^3 \to \mathbb{R}^3, (x, y, z) \mapsto \left(x + \frac{1}{2}y, y + \frac{1}{2}z, z + \frac{1}{2}x\right)$;

(d) $D : M(2 \times 2) \to M(2 \times 2), X \mapsto \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} X$;

(e) $E : \mathcal{P}_n(\mathbb{R}) \to \mathcal{P}_{n+1}(\mathbb{R}), p = p(x) \mapsto xp$.

6.4 Bijetividade – inversa

Definição 6.4.1 (Inversa)**.** Chama-se uma transformação linear $A : E \to F$ de **invertível** se A admite uma inversa à esquerda $B \in \mathcal{L}(F, E)$ e uma inversa à direita $C \in \mathcal{L}(F, E)$. Neste caso $B = C$,[2] denotado A^{-1}, é dito **a inversa** de A.

Exercício 6.4.2. Sejam $A \in \mathcal{L}(E, F)$ e $B \in \mathcal{L}(F, G)$ invertíveis, mostre que

a) a inversa de A (se existasse) é unica

b) $(A^{-1})^{-1} = A$

c) $(BA)^{-1} = A^{-1}B^{-1}$

d) $(\alpha A)^{-1} = \alpha^{-1}A^{-1}$ para escalares não-nulos $\alpha \in \mathbb{K} \setminus \{0\}$

6.4.1 Isomorfismos

Definição 6.4.3 (Isomorfismo)**.** Um **isomorfismo** (entre E e F) é uma transformação linear $A : E \to F$ a qual é bijetiva (injetivo e sobrejetivo). Neste caso se diz que E e F são espaços vetoriais **isomorfos**, símbolo $E \simeq F$.

Para aplicações gerais bijetividade é equivalente a existência da aplicação inversa (a qual claramente herda bijetividade). É interessante observar que se a aplicação bijetiva é linear a aplicação inversa não só existe mas herda linearidade.

Proposição 6.4.4. *Seja $A \in \mathcal{L}(E, F)$, então*

$$A \text{ isomorfismo} \qquad \Leftrightarrow \qquad A \text{ é invertível}$$

Demonstração. "⇒" Definimos o candidato B para ser a inversa de A assim

$$B : F \to E, \quad f \mapsto Bf := v \tag{6.4.1}$$

onde v é o único elemento de E com $Av = f$ (existência: A sobrejetivo, unicidade: A injetivo). A aplicação definida B é linear: Sejam $f, g \in F$, denotamos $v := Bf$ e $w := Bg$. Então $Av = f$ e $Aw = g$ e como A é linear obtemos

[2] Com efeito $B = BI_F = B(AC) = (BA)C = I_E C = C$.

$A(v+w) = Av + Aw = f + g$, então $B(f+g) = v + w = Bf + Bg$. Deixamos ao leitor verificar que $B(\alpha f) = \alpha Bf$. Também tem a propriedade de ser inversa à direita e esquerda, com efeito para $v \in E$ denota $f := Av$, então

$$ABf = Av = f, \qquad BAv = Bf = v$$

"$\Leftarrow$" Suponha que A admite a inversa $A^{-1} : F \to E$. Então A^{-1} é inversa à direita e à esquerda de A. Assim A é sobrejetivo e injetivo segundo os Teoremas 6.2.3 e 6.3.6. Mas bijetivo e linear significa isomorfismo. □

Corolário 6.4.5. *Dado isomorfismos $E \xrightarrow{A} F \xrightarrow{B} G$, então composição BA e múltiplos αA são isomorfismos para todos os escalares não-nulos $\alpha \neq 0$.*

Demonstração. Proposição 6.4.4 e Exercício 6.4.2. □

Exercício 6.4.6 (Isomorfismo é relação de equivalência)**.** Mostre que isomorfismo '$\simeq$' é uma **relação de equivalência** no conjunto de todos os espaços vetoriais: ou seja, mostre que são satisfeitos os três axiomas seguintes

$E \simeq E$ para cada um espaço vetorial (reflexividade)

$E \simeq F \quad \Rightarrow \quad F \simeq E$ (simetria)

$E \simeq F$ e $F \simeq G \quad \Rightarrow \quad E \simeq G$ (transitividade)

Teorema 6.4.7. *Dada uma transformação linear $A \in \mathcal{L}(E, F)$, então*

$$A \text{ bijetiva (isomorfismo)} \quad \Leftrightarrow \quad A \text{ leva uma base de } E \text{ numa base de } F$$

Demonstração. "$\Rightarrow$" Seja A um isomorfismo e $\mathcal{B}$ uma base de E. Resta mostrar que o conjunto $A\mathcal{B}$ é LI e gera F. LI segue de Teorema 6.3.1 (uma TL injetiva A leve LI em LI) e como $\mathcal{B}$ gera E o Lema 6.0.15 diz que $\langle A\mathcal{B}\rangle = \mathrm{Im}(A) = F$ onde a segunda identidade é a sobrejetividade de A.
"$\Leftarrow$" Dado um base $\mathcal{B}$ de E, então $A\mathcal{B}$ é uma base de F segundo a hipótese.
A INJETIVO ($\mathrm{N}(A) = \{\mathcal{O}\}$): Suponha $v \in E$ e $Av = \mathcal{O}$. Escrevemos v como CL $v = \sum_{i=1}^{k(v)} \alpha_i \xi_i$ de elementos ξ_i da base $\mathcal{B}$. Então

$$\mathcal{O} = Av = A \sum_i \alpha_i \xi_i = \sum_i \alpha_i \underbrace{A\xi_i}_{\in A\mathcal{B}}$$

Como $A\mathcal{B}$ é um conjunto LI todos os coeficientes $\alpha_i = 0$ se anulam. Assim $v = \mathcal{O}$ o que mostra que $\mathrm{N}(A) = \{\mathcal{O}\}$.
A SOBREJETIVO: Segundo hipótese $A\mathcal{B}$ é uma base de F. Dado $f \in F$, escrevemos f como CL $f = \sum_{j=1}^{\ell(f)} \beta_j A\xi_j$ de elementos $A\xi_j$ da base $A\mathcal{B}$. O elemento de E definido por $v := \sum_j \beta_j \xi_j$ satisfaz $Av = A\sum_j \beta_j \xi_j = \sum_j \beta_j A\xi_j = f$. □

Corolário 6.4.8. *Um espaço vetorial E sobre um corpo $\mathbb{K}$ e de dimensão $n \in \mathbb{N}_0$ é isomorfo a $\mathbb{K}^n$.*

Demonstração. Escolha uma base $\mathcal{B} = \{\xi_1, \ldots, \xi_n\}$ de E e defina a aplicação $A : \mathbb{K}^n \to E$ na forma $(\alpha_1, \ldots, \alpha_n) \mapsto \sum_{i=1}^n \alpha_i \xi_i$. Note-se que A é linear e que $Ae_j = A(0, \ldots, 0, 1, 0, \ldots, 0) = 1 \cdot \xi_j = \xi_j$. Assim A leva a base canônica $A\mathcal{E}^n = \mathcal{B}$ na base $\mathcal{B}$, então A é um isomorfismo segundo Teorema 6.4.7. □

Corolário 6.4.9. *Sejam E e F espaços vetoriais sobre um corpo $\mathbb{K}$ e de dimensões finitas, então*

$$E \simeq F \qquad \Leftrightarrow \qquad \dim E = \dim F$$

Demonstração. "⇒" Seja $A\colon E \to F$ um isomorfismo e $\mathcal{B}$ uma base de E. Então $A\mathcal{B}$ é base de F segundo Teorema 6.4.7. e $|\mathcal{B}| = |A\mathcal{B}|$ como A é bijetivo. Daí $\dim E := |\mathcal{B}| = |A\mathcal{B}| =: \dim F$.
"⇐" Corolário 6.4.8 da dois isomorfismos $E \simeq \mathbb{K}^{\dim E} = \mathbb{K}^{\dim F} \simeq F$. □

Exemplo 6.4.10. O espaço vetorial $\mathcal{S}(n)$ das matrizes $n \times n$ simétricas e o espaço vetorial $\mathcal{P}_{\frac{n(n+1)}{2}-1}$ dos polinômios de grau menor ou igual $\frac{n(n+1)}{2} - 1$ são isomorfos. Com efeito as dimensões são iguais – segundo as formulas (3.2.2) e (3.0.1) – e assim Corolário 6.4.9 aplica.

Exercício 6.4.11. Dado $A \in \mathcal{L}(E)$ onde $\dim E < \infty$, defina

$$\begin{aligned} T_A : \mathcal{L}(E) &\to \mathcal{L}(E) \\ X &\mapsto AX \end{aligned}$$

Prove que T_A é linear e que T_A é invertível se, e somente se A é invertível. Mesmo problema com $S_A(X) := XA$.

Exercício 6.4.12. Estabeleça um isomorfismo entre o espaço vetorial das matrizes reais simétricas $n \times n$ e o espaço das matrizes reais *triangulares inferiores* ($a_{ij} = 0$ se $i < j$).
Idem entre as matrizes anti-simétricas e as triangulares inferiores com diagonal nula.

Exercício 6.4.13. Sejam E, F espaços vetoriais tais que $\dim E \leq \dim F < \infty$. Prove que existem $A \in \mathcal{L}(E, F)$ e $B \in \mathcal{L}(F, E)$ tais que A é injetiva e B é sobrejetiva.

Exercício 6.4.14. Sejam E, F espaços vetoriais (de dimensão finita ou infinita). Sejam $A \in \mathcal{L}(E, F)$ e $B \in \mathcal{L}(F, E)$ tais que AB é invertível.

(a) Prove que A é sobrejetiva e B é injetiva.

(b) Se AB e BA são invertíveis, prove que A é invertível.

6.5 Teorema de núcleo e imagem

Teorema 6.5.1 (Teorema de núcleo e imagem)**.** *Para uma transformação linear $A\colon E \to F$ com domínio E de dimensão finita n vale*

$$\dim E = \dim \mathrm{N}(A) + \dim \mathrm{Im}(A)$$

Demonstração. Segundo Lema 6.0.12 os conjuntos $\mathrm{N}(A)$ e $\mathrm{Im}(A)$ são subespaços de E e F, respectivamente. Segundo Teorema 3.2.1 (c) e Corolário 6.0.16 temos

$$k := \dim \mathrm{N}(A) \leq \dim E =: n < \infty$$
$$\ell := \dim \mathrm{Im}(A) \leq \dim E =: n < \infty$$

Escolha uma base ordenada $(\xi_1, \ldots, \xi_k)$ de $\mathrm{N}(A)$ e uma $(A\nu_1, \ldots, A\nu_\ell)$ de $\mathrm{Im}(A)$. Resta mostrar que

$$\mathcal{B} := (\xi_1, \ldots, \xi_k, \nu_1, \ldots, \nu_\ell)$$

é uma base de E, porque neste caso $\dim E = k + \ell = \dim \mathrm{N}(A) + \dim \mathrm{Im}(A)$.

$\mathcal{B}$ é LI. Suponha que uma CL em $\mathcal{B}$ representa o vetor nulo, ou seja

$$\alpha_1\xi_1 + \cdots + \alpha_k\xi_k + \beta_1\nu_1 + \cdots + \beta_\ell\nu_\ell = \mathcal{O}$$

Resta mostrar que todos os coeficientes se anulam. Aplique A usando linearidade e que $\xi_i \in \mathrm{N}(A)$ para obter

$$\alpha_1 \underbrace{A\xi_1}_{\mathcal{O}} + \cdots + \alpha_k \underbrace{A\xi_k}_{\mathcal{O}} + \beta_1 A\nu_1 + \cdots + \beta_\ell A\nu_\ell = \mathcal{O}$$

Então $\beta_1 = \cdots = \beta_\ell = 0$ porque toda CL no conjunto LI $\{A\nu_1, \ldots, A\nu_\ell\}$ e representando o vetor nulo tem todos coeficientes nulos. Neste caso a primeira identidade simplifica-se para $\alpha_1\xi_1 + \cdots + \alpha_k\xi_k = \mathcal{O}$. Mas os ξ_i's formam uma base, então um conjunto LI, assim os coeficientes se anulam $\alpha_1 = \cdots = \alpha_k = 0$.

$\mathcal{B}$ gera E. Dado $v \in E$, temos que exprimir v como CL em $\mathcal{B}$. Como temos uma base de $\mathrm{Im}(A)$, escrevemos $Av \in \mathrm{Im}(A)$ como CL $Av = \beta_1 A\nu_1 + \cdots + \beta_\ell A\nu_\ell$ com coeficientes únicos β_j. Mas isso nos da um elemento w do núcleo

$$A\underbrace{(v - \beta_1\nu_1 - \cdots - \beta_\ell\nu_\ell)}_{=:w} = \mathcal{O}$$

Exprimindo w como CL na base do núcleo, com coeficientes únicos α_j, obtemos

$$\alpha_1\xi_1 + \cdots + \alpha_k\xi_k = w = v - \beta_1\nu_1 - \cdots - \beta_\ell\nu_\ell$$

Assim temos exprimido

$$v = \alpha_1\xi_1 + \cdots + \alpha_k\xi_k + \beta_1\nu_1 + \cdots + \beta_\ell\nu_\ell$$

como CL em $\mathcal{B}$. □

Corolário 6.5.2. *Para uma transformação linear $A\colon E \to F$ entre espaços vetoriais da mesma dimensão finita $n = \dim E = \dim F$ são equivalente*

$$A \text{ injetivo} \qquad \Leftrightarrow \qquad A \text{ sobrejetivo} \qquad (\Leftrightarrow \qquad A \text{ isomorfismo})$$

Demonstração. Da hipótese da mesma dimensão e do Teorema 6.5.1 sabemos

$$\dim F = \dim E = \dim \mathrm{N}(A) + \dim \mathrm{Im}(A)$$

Injetividade (equivalente a $\mathrm{N}(A) = \{\mathcal{O}\}$ segundo Lema 6.0.14) implica $\dim F = \dim \mathrm{Im}(A)$, então $F = \mathrm{Im}(A)$ (sobrejetividade) segundo Teorema 3.2.1 (d). Vice versa, sobrejetividade ($F = \mathrm{Im}(A)$) implica $\mathrm{N}(A) = \{\mathcal{O}\}$ (injetividade). □

Exercício 6.5.3 (Errado na dimensão infinita)**.** Considere os operadores lineares $A, B\colon \mathbb{R}^\infty \to \mathbb{R}^\infty$ dado por empurrar todos os membros por um lugar

$$\begin{aligned} A(x_1, x_2, x_3, \dots) &:= (0, x_1, x_2, x_3, \dots) \\ B(x_1, x_2, x_3, \dots) &:= (x_2, x_3, \dots) \end{aligned}$$

Mostre que A é linear e injetivo, mas não é sobrejetivo, enquanto B é linear e sobrejetivo, mas não é injetivo,

Corolário 6.5.4. *Na mesma dimensão finita $n = \dim E = \dim F$ ser inversa à esquerda é equivalente a ser inversa à direita, em símbolos para $A \in \mathcal{L}(E, F)$ e $B \in \mathcal{L}(F, E)$ são equivalentes*

$$BA = I_E \qquad \Leftrightarrow \qquad AB = I_F$$

No todo caso A é invertível com inversa $A^{-1} = B = C$.

Demonstração. Temos as três equivalências

$$\begin{aligned} &\exists B \in \mathcal{L}(F, E)\colon BA = I_E \\ &\Leftrightarrow A \text{ injetivo} \\ &\Leftrightarrow A \text{ sobrejetivo} \\ &\Leftrightarrow \exists C \in \mathcal{L}(F, E)\colon AC = I_F \end{aligned}$$

segundo respectivamente os três resultados Teorema 6.3.6, Corolário 6.5.2, e Teorema 6.2.3. Mas neste caso $C = B$ e este operador é a inversa de A como mostrado na Definição 6.4.1. □

Exemplo 6.5.5. Dado uma lista não-nula $\alpha \in \mathbb{R}^n \setminus \{\mathcal{O}\}$, o subconjunto

$$\mathrm{H}_\alpha := \{x \in \mathbb{R}^n \mid \varphi_\alpha(x) := \alpha_1 x_1 + \cdots + \alpha_n x_n = 0\} = \mathrm{N}(\varphi_\alpha) \subset \mathbb{R}^n$$

é chamado de hiperplano e foi introduzido no Exemplo 2.1.8. Já sabemos que

$$\dim \mathrm{H}_\alpha = n - 1$$

como no Exemplo 3.0.13 d) temos visto uma base composto de $n-1$ elementos.

Um camino alternativo para calcular a dimensão é do ponto da vista como núcleo do funcional linear $\varphi_\alpha : \mathbb{R}^n \to \mathbb{R}$. O Teorema 6.5.1 diz que

$$\underbrace{\dim \mathbb{R}^n}_{=n} = \dim \underbrace{\mathrm{N}(\varphi_\alpha)}_{\mathrm{H}_\alpha} + \underbrace{\dim \mathrm{Im}(\varphi_\alpha)}_{=1}$$

Resta ver que $\mathrm{Im}(\varphi_\alpha) = \mathbb{R}$. O subespaço $\mathrm{Im}(\varphi_\alpha)$ de $\mathbb{R}$ não é o trivial $\{0\}$ porque $\varphi_\alpha \alpha = \alpha_1^2 + \cdots + \alpha_n^2 > 0$ é não-nulo como $\alpha \neq (0, \ldots, 0)$. Entao $\mathrm{Im}(\varphi_\alpha)$ deve ser o outro subespaço de $\mathbb{R}$, o $\mathbb{R}$ mesmo, veja Exercício 2.1.4.

Capítulo 7

Soma direta e projeções

No Capítulo 7 denotamos de

$$F, G, H \subset E$$

subespaços de um espaço vetorial E sobre um corpo $\mathbb{K}$. Na parte das involuções precisamos às vezes que $1 + 1 \neq 0$ em $\mathbb{K}$, veja Corolário 1.1.22. (Vale para $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, não para $\mathbb{Z}_2$.) O objeto central do nosso interesse será o conjunto

$$\mathcal{SC} = \mathcal{SC}(E) := \{(F, G) \mid F \oplus G = E\}$$

composto de **pares** (F, G) de **subespaços complementares** de E no sentido que o par decompõe $E = F \oplus F$ como soma direta.

O nosso objetivo será relacionar o conjunto $\mathcal{SC}(E)$ bijetivamente com duas classes de operadores lineares em E – os subconjuntos de $\mathcal{L}(E)$ dados por

$$\mathcal{P} = \mathcal{P}(E) := \{P \mid P^2 = P\} \qquad \textit{“projeções em } E\textit{”}$$
$$\mathcal{I} = \mathcal{I}(E) := \{S \mid S^2 = I_E\} \qquad \textit{“involuções em } E\textit{”}$$

Ambas condições fazem sentido no contexto geral de uma aplicação $s : X \to X$ num conjunto X. Para nos são relevantes as **involuções** ($s^2 = \mathrm{id}$). Temos três involuções naturais (i) de trocar os membros

$$\mu : \mathcal{SC} \to \mathcal{SC}, \quad (F, G) \mapsto (G, F)$$

(ii) de mudar o sinal[1]

$$\sigma : \mathcal{I} \to \mathcal{I}, \quad S \mapsto -S$$

e (iii) de tomar diferença com o operador identidade

$$\delta : \mathcal{P} \to \mathcal{P}, \quad P \mapsto I - P$$

Com efeito $(I - P)^2 = I^2 - 2P + P^2 = I - P$, assim realmente é uma projeção. Capítulo 7 é ilustrado na Figura 7.1. É comum indicar injetividade de uma aplicação com tal flecha $f : X \rightarrowtail Y$, sobrejetividade com tal flecha $f : X \twoheadrightarrow Y$.

[1] Na verdade $-S$ é o inverso aditivo de $S \in \mathcal{L}(E)$ e o inverso do inverso é a identidade.

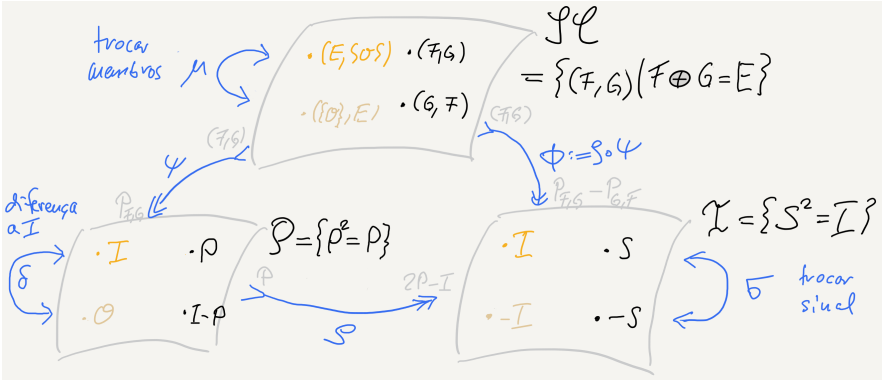


Figura 7.1: Conjuntos $\mathcal{SC}$ dos subespaços complementares, $\mathcal{P}$ das projeções, e $\mathcal{I}$ das involuções lineares – a diagrama das seis bijeções é comutativa

Preparações e lembranças

Definição 7.0.6 (Pontos fixos e anti-fixos)**.** Dado um conjunto X e uma aplicação $r : X \to X$. a) Um elemento $x \in X$ tal que $r(x) = x$ chama-se um **ponto fixo** de r. O conjunto dos pontos fixos de r satisfaz $\mathbf{Fix(}r\mathbf{)} \subset \mathrm{Im}(r)$.

b) Se X é um espaço vetorial denotamos de $\mathbf{aFix(}r\mathbf{)}$ o conjunto de todos os **pontos anti-fixos** x de r, ou seja $r(x) = -x$.

É fácil – e instrutivo – checar que para aplicações idempotentes num conjunto X os pontos fixos já formam a imagem inteira, em símbolos

$$r^2 = r \qquad \Rightarrow \qquad \mathrm{Fix}(r) = \mathrm{Im}(r) \tag{7.0.1}$$

Para aplicações idempotentes é recomendável – geralmente ajuda bastante o entendimento – trabalhar com $\mathrm{Fix}(r)$ em vez de $\mathrm{Im}(r)$.

Exercício 7.0.7. Se $B \in \mathcal{L}(E)$, então $\mathrm{Fix}(B), \mathrm{aFix}(B) \subset E$ são subespaços.

Produto cartesiano e soma

Lembre-se do Exercício 3.0.15 que o produto cartesiano $G \times H$ de dois espaços vetoriais sobre um corpo $\mathbb{K}$ é um espaço vetorial sobre $\mathbb{K}$ de dimensão

$$\dim(G \times H) = \dim G + \dim H$$

Dado dois subespaços G, H, será útil relembrar da Seção 2.3 a soma ordinária $G + H$ e a soma direta $G \oplus H$ deles. Se G, H são de dimensão finita vale a fórmula (3.2.1) a qual diz que

$$\dim(G + H) = \dim G + \dim H - \dim(G \cap H) \tag{7.0.2}$$

Exercício 7.0.8. Seja E um espaço vetorial com subespaços de intersecção trivial $G \cap H = \{\mathcal{O}\}$. Prove que $S : G \times H \to G \oplus H$, $(g, h) \mapsto g + h$, é um isomorfismo (linear, injetivo, sobrejetivo).

7.1 Projeções

Definição 7.1.1. Os operadores lineares idempotentes $P^2 = P \in \mathcal{L}(E)$ são chamados de **as projeções** de E. Um **par de subespaços complementares** de E é um par (F, G) de subespaços decompondo E no sentido que $F \oplus G = E$.

Lema 7.1.2 (Caracterização de projeção)**.** *Seja $P \in \mathcal{L}(E)$, então*

$$P \text{ projeção de } E \quad \Leftrightarrow \quad \begin{cases} a) & \forall v \in \mathrm{Im}(P)\colon Pv = v \qquad \mathrm{Im}(P) = \mathrm{Fix}(P) \\ b) & E = \mathrm{Im}(P) \oplus \mathrm{N}(P) \qquad \text{“par complementar”} \end{cases}$$

Demonstração. “⇒” Suponha $P^2 = P$. a) Já sabemos de (7.0.1) que $\mathrm{Im}(P) = \mathrm{Fix}(P)$. b) Intersecção trivial: seja $v \in \mathrm{Fix}(P) \oplus \mathrm{N}(P)$, então $\mathcal{O} = Pv = v$.
“⇐” Se $v \in E$, então $Pv \in \mathrm{Im}(P) = \mathrm{Fix}(P)$, assim $P^2v = P(Pv) = Pv$. □

Definição 7.1.3 (Projeção sobre F paralelamente G)**.** Seja (F, G) um par de subespaços complementares de E, escreva $v \in E = F \oplus G$ na forma $v = f + g$ com únicos elementos $f \in F$ e $g \in G$, veja Teorema 2.3.4. A aplicação dada por

$$P_{F,G} : E \to E, \quad v \mapsto f \tag{7.1.1}$$

é chamada de **projeção de E sobre F paralelamente G**.

Lema 7.1.4. *A aplicação $P := P_{F,G}$ definida acima é uma projeção de E. E imagem (os pontos fixos) e núcleo são dados por F e G, em símbolos*

$$F = \mathrm{Im}(P_{F,G}) = \mathrm{Fix}(P_{F,G}), \qquad G = \mathrm{N}(P_{F,G}) \tag{7.1.2}$$

Além disso $P_{G,F} = I_E - P_{F,G}$.

Demonstração. Se $v = f+g$ e $\tilde{v} = \tilde{f}+\tilde{g}$, então $v+\tilde{v} = f+g+\tilde{f}+\tilde{g} = f+\tilde{f}+g+\tilde{g}$.
Linear: Assim $P(v + \tilde{v}) = P(f + \tilde{f} + g + \tilde{g}) = f + \tilde{f} = Pv + P\tilde{v}$. Como $\alpha v = \alpha(f + g) = \alpha f + \alpha g$ obtemos $P(\alpha v) = P(\alpha f + \alpha g) = \alpha f = \alpha P v$.
Idempotente: Vale $P^2v = P(P(f + g)) = Pf = f = Pv$.
$\mathrm{Im}(P) = F$: ’$\subset$’ óbvio ’$\supset$’ dado $f \in F$, então $Pf = f$.
$\mathrm{N}(P) = G$: ’$\subset$’ para $f + g = v \in \mathrm{N}(P)$ vale $\mathcal{O} = Pv = P(f + g) = f$. Assim segue que $v = g \in G$. ’$\supset$’ para $g \in G$ vale $Pg = P(\mathcal{O} + g) = \mathcal{O}$.
Identidade: $P_{G,F}(f + g) = g = (f + g) - f = I_E(f + g) - P_{F,G}(f + g)$ □

Teorema 7.1.5. *A seguinte aplicação é uma bijeção*

$$\begin{aligned} \mathcal{SC} = \{\textit{pares de subespaços complementares de } E\} &\xrightarrow{\psi} \{\textit{projeções em } E\} = \mathcal{P} \\ (F, G) &\mapsto P_{F,G} \end{aligned}$$

com inversa $\chi\colon P \mapsto (\mathrm{Im}(P), \mathrm{N}(P))$. *Útil lembrar:* $\mathrm{Im}(P) = \mathrm{Fix}(P)$

Note-se que o subconjunto $\mathcal{P} \subset \mathcal{L}(E)$ composto das projeções P de E não é um subespaço, por exemplo $(\alpha P)^2 = \alpha^2 P^2 = \alpha^2 P \neq \alpha P$ caso $\alpha^2 \neq \alpha \in \mathbb{K}$. Então não faz sentido falar sobre linearidade da bijeção ψ.

Demonstração. Injetivo. Suponha $P_{F,G} = P_{\tilde{F},\tilde{G}}$, então aplique Lema 7.1.4 duas vezes para obter as identidades $F = \mathrm{Im}(P_{F,G}) = \mathrm{Im}(P_{\tilde{F},\tilde{G}}) = \tilde{F}$ e analogamente $G = \mathrm{N}(P_{F,G}) = \mathrm{N}(P_{\tilde{F},\tilde{G}}) = \tilde{G}$.
Sobrejetivo. Dado uma projeção P em E, Lema 7.1.2 diz que o par definido por $(F,G) := (\mathrm{Im}(P), \mathrm{N}(P))$ é um par de subespaços complementares. Resta mostrar que $P = P_{\mathrm{Im}(P),\mathrm{N}(P)} \overset{\text{def.}}{=} \psi(F,G)$. Dado $w \in E$, Lema 7.1.2 diz que $w = f + g$ para únicos elementos $f \in \mathrm{Im}(P) = \mathrm{Fix}(P)$ e $g \in \mathrm{N}(P)$. Então vale

$$P_{\mathrm{Im}(P),\mathrm{N}(P)} w \overset{\text{def.}}{=} f \overset{\text{pt.fix.}}{=} Pf = Pf + \underbrace{\mathcal{O}}_{Pg} \overset{\text{lin.}}{=} P(f+g) = Pw$$

Inversa. Dada uma projeção P em E, no item anterior temos visto que

$$P = P_{\mathrm{Im}(P),\mathrm{N}(P)} \overset{\text{def.}}{=} \psi\left(\mathrm{Im}(P), \mathrm{N}(P)\right) \overset{\text{def.}}{=} \psi\left(\chi(P)\right)$$

Vale $\chi\left(\psi(F,G)\right) \overset{\text{def.}}{=} \left(\mathrm{Im}(P_{F,G}), \mathrm{N}(P_{F,G})\right) = (F,G)$ segundo Lema 7.1.4 . □

7.2 Involuções

Definição 7.2.1. Um operador linear $S \in \mathcal{L}(E)$ cujo quadrado $S^2 = I_E$ é a identidade chama-se de **involução** de E. Involuções são isomorfismos.

Com efeito, a condição $S^2 = I_E$ para ser uma involução implica injetivo e sobrejetivo. Como o núcleo sempre é mínima $\mathrm{N}(S) = \{\mathcal{O}\}$ e a imagem sempre é máxima $\mathrm{Im}(S) = E$ estes dois subespaços não são úteis, não – em contraste ao caso de projeções. Os lugares deles como par de subespaços complementares ocupam, no caso de involuções, os subespaços dos pontos fixos e anti-fixos

$$F := \mathrm{Fix}(S), \qquad A := \mathrm{aFix}(S)$$

A vinculação entre projeções P e involuções S, além de dar decomposições

$$\mathrm{Im}(P) \oplus \mathrm{N}(P) = E = F \oplus A$$

é a igualdade $S = P_{F,A} - P_{A,F}$ baseada na identidade

$$\mathrm{Im}(P) = \mathrm{Fix}(P)$$

Nosso trajeto será assim: Suponhamos agora que $2 := 1 + 1 \neq 0$ em $\mathbb{K}$, veja Corolário 1.1.22. Primeiro mostramos que a fórmula estabelecida na dimensão 2 para a reflexão em torno de uma reta, veja (4.3.3), nos da uma bijeção

$$\rho : \mathcal{P} \to \mathcal{I}, \quad P \mapsto 2P - I_E$$

entre projeções e involuções com inversa $S \mapsto \frac{1}{2}(I_E + S)$. Caracterizamos involuções em termos de subespaços complementares com a composição de bijeções

$$\phi := \rho \circ \psi : \mathcal{SC} \to \mathcal{I}, \quad (F,G) \mapsto \rho(P_{F,G}) = P_{F,G} - P_{G,F} =: S_{F,G}$$

Todo é compatível no sentido que é comutativa a diagrama das 6 bijeções na Figura 7.1.

Lema 7.2.2 (Caracterização de involução)**.** *Seja $S \in \mathcal{L}(E)$, então*

$$S \text{ involução de } E \quad \Leftrightarrow \quad E = \underbrace{\mathrm{Fix}(S)}_{=:F} \oplus \underbrace{\mathrm{aFix}(S)}_{=:A} \quad \text{“par complementar } (F, A)\text{”}$$

Além disso uma involução S é da forma $S = P_{F,A} - P_{A,F}$.

Demonstração. “$\Rightarrow$” Cada um elemento $x \in \mathrm{Fix}(S) \cap \mathrm{aFix}(S)$ é nulo porque $x = Sx = -x$. Os elementos $v \in E$ são da forma $v = Pv + Qv$ onde $Pv := \frac{1}{2}(v + Sv)$ e $Qv := \frac{1}{2}(v - Sv)$. Mas $S^2 = I_E$ implica $S(Pv) = Pv$ e $S(Qv) = -Qv$.
“$\Leftarrow$” Como $E = \mathrm{Fix}(S) \oplus \mathrm{aFix}(S)$ os elementos $v \in E$ são da forma $v = f + a$ para únicos elementos $f \in \mathrm{Fix}(S)$ e $a \in \mathrm{aFix}(S)$, veja Teorema 2.3.4. Como S é linear obtemos

$$S^2 v = S(S(f + a)) = S(Sf + Sa) = S(f - a) = Sf - Sa = f + a = v$$

para todos os $v \in E$. “$S = S_{F,A}$” Escrevendo $v \in E$ como $v = f + a$ obtemos

$$Sv = S(f + a) = Sf + Sa = f - a = P_{F,A}v - P_{A,F}v$$

segundo Definição 7.1.3. □

Involuções e projeções

Teorema 7.2.3. *Seja $1 + 1 \neq 0$ em $\mathbb{K}$. A seguinte aplicação é uma bijeção*

$$\begin{aligned} \rho \colon \mathcal{P} = \{\text{projeções em } E\} &\to \{\text{involuções em } E\} = \mathcal{I} \\ P &\mapsto 2P - I_E =: S_P \end{aligned}$$

com inversa $\rho^{-1} = \gamma \colon S \mapsto \frac{1}{2}(I_E + S)$. As projeções $\gamma(S)$ e $\gamma(-S)$, ou seja

$$P := \tfrac{1}{2}(I_E + S), \qquad Q := \tfrac{1}{2}(I_E - S)$$

satisfazem $P + Q = I_E$ e $P - Q = S$.

Demonstração. Seja $I = I_E$. Bem definido. $(2P - I)^2 = 4P^2 - 4P + I = I$.
Injetivo. Suponha $2P - I = 2\tilde{P} - I$, adicione $-I$ para obter $2P = 2\tilde{P}$. Então $P = \tilde{P}$ segundo Corolário 1.1.22.
Sobrejetivo. Dado uma involução S em E, defina $P := \gamma(S) = \frac{1}{2}(I + S)$ para obter $\rho(P) = 2P - I = (I + S) - I = S$.
Inversa. Dada uma involução S em E, no item anterior vimos que $S = \rho(\gamma(S))$. De outro lado $\gamma(\rho(P)) = \gamma(2P - I) = \frac{1}{2}(I + (2P - I)) = P$. □

Involuções e subespaços complementares

Definição 7.2.4 (Involução/reflexão em torno de F ao longo G)**.** Seja (F, G) um par de subespaços complementares de E, escreva $v \in E = F \oplus G$ na forma $v = f + g$ com únicos elementos $f \in F$ e $g \in G$, veja Teorema 2.3.4. A aplicação

$$S_{F,G} := P_{F,G} - P_{G,F} \colon E \to E$$

é chamada de **involução (ou reflexão) de E em torno de F ao longo G**.

Vamos justificar chamar $S_{F,G}$ de involução em torno de F ao longo G:

Lema 7.2.5. *A aplicação $S_{F,G}$ definida acima é uma involução de E. Os pontos fixos e anti-fixos contém F e G, em símbolos*

$$F \subset \mathrm{Fix}(S_{F,G}), \qquad G \subset \mathrm{aFix}(S_{F,G}) \tag{7.2.1}$$

Valem igualdades nos casos $\dim E < \infty$ *ou* $1+1 \neq 0$ *em* $\mathbb{K}$.

Ter igualdades em (7.2.1) é importante para consistência: como $S_{F,G}$ euma involução o Lema 7.2.2 aplica e fala que $S_{F,G} = S_{\mathrm{Fix}(S_{F,G}),\mathrm{aFix}(S_{F,G})}$. Então espera-se igualdade dos pares $(F,G) = (\mathrm{Fix}(S_{F,G}), \mathrm{aFix}(S_{F,G}))$.

Demonstração. Dado $v \in E$, então $\exists! f \in F$ e $\exists! g \in G$ tal que $v = f + g$.
Linearidade: É óbvio como $S_{F,G}$ é soma de dois operadores lineares.
$\boldsymbol{S^2 = \mathrm{I}_E}$**:** Usamos a definição de $S := S_{F,G}$ e Lema 7.1.4 para obter

$$\begin{aligned} S^2 v &= S\left(\left(P_{F,G} - P_{G,F}\right)(f+g)\right) \\ &= S(f-g) \\ &= \left(P_{F,G} - P_{G,F}\right)(f-g) \\ &= f - (-g) \\ &= v \end{aligned}$$

$\boldsymbol{G = \mathrm{aFix}(S_{F,G})}$**:** Como $\mathrm{aFix}(S_{F,G}) = \mathrm{Fix}(S_{G,F})$ o próximo item aplica.
$\boldsymbol{F = \mathrm{Fix}(S_{F,G})}$**:** '$\subset$' Seja $f \in F$. Como $F = \mathrm{Im}(P_{F,G}) = \mathrm{Fix}(P_{F,G})$ e $F = \mathrm{N}(P_{G,F})$ obtemos $Sf = P_{F,G}f - P_{G,F}f = f$.
'$\supset$' **Caso $\boldsymbol{1+1 \neq 0}$ em $\mathbb{K}$:** Escreva $x \in \mathrm{Fix}(S) \subset E$ unicamente na forma $x = f + g$ onde $f \in F$ e $g \in G$. Então

$$f + g = x = Sx = P_{F,G}(f+g) - P_{G,F}(f+g) = f - g$$

Assim $g + g = \mathcal{O}$. Segundo Corolário 1.1.22 obtemos $g = \mathcal{O}$. Então $x = f \in F$.
'$\supset$' **Caso $\boldsymbol{\dim E < \infty}$:** Como $(F,G) \in \mathcal{SC}$ e segundo Lema 7.2.2 ($S^2 = I_E$)

$$F \oplus G = E = \mathrm{Fix}(S) \oplus \mathrm{aFix}(S)$$

Então aplicando a fórmula (7.0.2) a cada uma soma direta nos da as igualdades

$$\dim F + \dim G = \dim E = \dim \mathrm{Fix}(S) + \dim \mathrm{aFix}(S)$$

Como $0 \leq \dim F \leq \dim \mathrm{Fix}(S)$ e $0 \leq \dim G \leq \dim \mathrm{aFix}(S)$ segundo Teorema 3.2.1 (c), as dimensões devem ser iguais, ou seja

$$\dim F = \dim \mathrm{Fix}(S), \qquad \dim G = \dim \mathrm{aFix}(S)$$

Mas, segundo Teorema 3.2.1 (d), inclusão com a mesma dimensão implica igualdade, assim $F = \mathrm{Fix}(S)$ e $G = \mathrm{aFix}(S)$. □

Exercício 7.2.6. Faça um desenho de $E = \mathbb{R}^2$ com dois subespaços $F \neq G$ de dimensão 1. Ilustre para varias escolhas de $v \in E, F, G$ a imagem $S_{F,G}v$ usando os vetores (pensa em flechas) $P_{F,G}v$ e $-P_{G,F}v$.

7.3 Exercícios

Seja E um espaço vetorial sobre um corpo $\mathbb{K}$.

1. No plano $\mathbb{R}^2$, considere as retas F_1 e F_2, definidas respectivamente pelas equações $y = ax$ e $y = bx$, onde $a \neq b$ são números reais.

   (a) Exprima $v = (x, y) \in \mathbb{R}^2$ como soma de um vetor de F_1 e um de F_2.

   (b) Seja $P = P_{F_1,F_2} \in \mathcal{L}(\mathbb{R}^2)$ a projeção sobre F_1 paralelamente a F_2. Obtenha a matriz $[P]$ de P.

   (c) Encontre a matriz $[S]$ da reflexão $S = S_{F_2,F_1} : \mathbb{R}^2 \to \mathbb{R}^2$, em torno da reta F_2, paralelamente a F_1.

2. Exprima $v = (x, y, z) \in \mathbb{R}^3$ como soma de um vetor do plano F_1, cuja equação é $x + y - z = 0$, com um vetor da reta F_2, gerada pelo vetor $(1, 2, 1)$. Conclua que $\mathbb{R}^3 = F_1 \oplus F_2$. Determine a matriz $[P]$ da projeção $P : \mathbb{R}^3 \to \mathbb{R}^3$ que tem imagem F_1 e núcleo F_2.

3. Dado $P \in \mathcal{L}(E)$, *prove* ou *desprove*:

   (a) $E = \mathrm{N}(P) \oplus \mathrm{Im}(P) \quad \Rightarrow \quad P$ é projeção de E.

   (b) $E = \mathrm{N}(P) + \mathrm{Im}(P) \quad \Rightarrow \quad P$ é projeção de E.

   (c) P é projeção $\quad \Leftrightarrow \quad I - P$ é projeção.

   (d) P é projeção $\quad \Leftrightarrow \quad \mathrm{N}(P) = \mathrm{Im}(I - P) \quad \big(\Leftrightarrow \quad \mathrm{N}(I - P) = \mathrm{Im}(P)\big)$.

4. Sejam $F_1, F_2 \subset E$ subespaços com $\dim F_1 + \dim F_2 = \dim E < \infty$. Prove

$$E = F_1 \oplus F_2 \iff F_1 \cap F_2 = \{\mathcal{O}\}.$$

5. Sejam $P_1, \ldots, P_n : E \to E$ operadores lineares tais que

$$P_1 + \cdots + P_n = I \quad \text{e} \quad \forall i \neq j : \ P_i P_j = \mathcal{O}.$$

   Prove que estes operadores são projeções.

6. Sejam $P, Q \in \mathcal{L}(E)$ projeções e $1+1 \neq 0$ em $\mathbb{K}$, prove que são equivalentes:

   (a) $P + Q$ é uma projeção;

   (b) $PQ + QP = \mathcal{O}$;

   (c) $PQ = QP = \mathcal{O}$.

   [Para provar (b) $\Rightarrow$ (c), multiplique à esquerda e à direita da hipótese $PQ = -QP$ por P e conclua $\mathcal{O} = PQP$. Consequentemente $\mathcal{O} = \mathcal{O}Q = PQPQ = P(-PQ)Q = -PQ$.]

7. Seja $E = F_1 \oplus F_2$. O **gráfico** de uma transformação linear $B : F_1 \to F_2$ é o subconjunto $\mathrm{graph}(B) := \{v + Bv \mid v \in F_1\}$ de E. Prove que

   (a) $\mathrm{graph}(B)$ é um subespaço de E.

   (b) a projeção $P = P_{F_1,F_2} : E \to E$, restrita a $\mathrm{graph}(B)$, define um isomorfismo entre $\mathrm{graph}(B)$ e F_1.

Capítulo 8

Eliminação e aplicações

Agora revisamos do curso MA141 aplicações do processo de escalonar uma matriz **a**. Recomendamos recapitular os detalhes deste processo da Seção A.2.

8.1 Resolução de sistemas lineares

Sistemas lineares foram introduzido em Definição A.3.1 e tratado em Exemplo 6.0.18. Estes conteúdos são pressupostos. Seja **a** uma matriz $m \times n$ com entradas num corpo $\mathbb{K}$ e $b \in \mathbb{K}^m$ uma lista. Lembre-se de (A.3.1) que a equação $\mathbf{a}x = b$ é chamado de sistema linear de m equações a n incógnitas $(x_1, \ldots, x_n) = x$.

Existência de uma solução x é equivalente ao fato que a lista b é localizada na imagem da matriz **a**, em símbolos

$$\mathbf{a}x = b \text{ tem solução } x \quad \Leftrightarrow \quad b \in \mathrm{Im}(\mathbf{a}) \quad \Leftrightarrow \quad p := \mathrm{posto}(\mathbf{a}) = \mathrm{posto}[\mathbf{a} : b]$$

veja Exemplo 6.0.18. Mas neste caso, tem como saber quantas soluções haverá?

Lema 8.1.1. *Suponha que* $\mathbf{a}x = b$ *admite uma solução* x_0 ($b \in \mathrm{Im}(\mathbf{a})$). *Então*

a) $p = n$ ($\mathbf{a} : \mathbb{K}^n \to \mathbb{K}^m$ *é injetivo*) $\Rightarrow$ *a solução é única* $\qquad p := \mathrm{posto}(\mathbf{a})$

b) $p < n \;\Leftrightarrow$ *o número de soluções é infinito*

No caso b) o conjunto das soluções x *de* $\mathbf{a}x = b$ *é dado pela translação do núcleo*

$$x_0 + \mathrm{N}(\mathbf{a}) = \{\textit{soluções } x \textit{ de } \mathbf{a}x = b\}$$

e $\dim \mathrm{N}(\mathbf{a}) = n - p \geq 1$.

Demonstração. Como a dimensão p da imagem $\mathrm{Im}(\mathbf{a}) \subset \mathbb{K}^m$ é no máximo a dimensão n do domínio, segundo Corolário 6.0.16, temos que $p \leq \max\{n, m\}$.
a) Segundo o Teorema 6.5.1 de núcleo e imagem $n = p$ ($\dim \mathbb{K}^n = \dim \mathrm{Im}(\mathbf{a})$) é equivalente a $\mathrm{N}(\mathbf{a}) = \{\mathcal{O}\}$ o que, segundo Lema 6.0.14, significa que $\mathbf{a}\colon \mathbb{K}^n \to$

$\mathbb{K}^m$ é injetivo. Por isso $\mathbf{a}\colon \mathbb{K}^n \to \mathrm{Im}(\mathbf{a})$ é um isomorfismo, e assim $b \in \mathrm{Im}(\mathbf{a})$ corresponde a exatamente um elemento $x \in \mathbb{K}^n$ tal que $\mathbf{a}x = b$.
b) Seja $\nu \in \mathrm{N}(\mathbf{a})$, então $x := x_0 + \nu$ satisfaz $\mathbf{a}x = \mathbf{a}x_0 + \mathbf{a}\nu = b$. □

Lembra o Lema A.3.3: Uma lista x é solução do sistema linear $[\mathbf{a} : b]$ se e somente se x é solução do sistema linear associado à matriz escalonada $[\mathbf{a} : b]_{\mathrm{esc}}$.

Lembra o Comentário A.3.4: Para resolver o SL $\mathbf{a}x = b$

- escalona a matriz aumentada $[\mathbf{a} : b]$
- obtendo uma matriz escalonada $[\mathbf{a} : b]_{\mathrm{esc}} =: [\tilde{\mathbf{a}} : \tilde{b}]$, então
- resolve o SL $\tilde{\mathbf{a}}x = \tilde{b}$ "de baixo para cima", veja Exemplo A.3.5, então
- uma lista x é solução de $\tilde{\mathbf{a}}x = \tilde{b}$ se e somente se x é solução de $\mathbf{a}x = b$

Exemplo 8.1.2. Determine uma base do núcleo da transformação linear

$$A : \mathbb{R}^3 \to \mathbb{R}^3, \quad (x, y, z) \mapsto (x + 2y + z, 2x + 4y, 3x + 6y + 3z)$$

Uma solução. Para obter uma matriz de A escolhemos as bases mais simples, a base canônica $\mathcal{E}^3$. Obtemos

$$\mathbf{a} := [A] = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 0 \\ 3 & 6 & 3 \end{bmatrix}$$

Escalonamos a matriz

$$[\mathbf{a} : \mathcal{O}] = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 3 & 6 & 3 & 0 \end{bmatrix} \xrightarrow{\text{(oe)}} \cdots \xrightarrow{\text{(oe)}} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = [\mathbf{a}_{\mathrm{esc}} : \mathcal{O}]$$

Agora resolvemos o sistema escalonado $\mathbf{a}_{\mathrm{esc}}x = \mathcal{O}$, ou seja

$$\begin{cases} \alpha + 2\beta + \gamma = 0 & \Rightarrow \alpha = -2\beta,\ \beta \in \mathbb{R} \\ -2\gamma = 0 & \Rightarrow \gamma = 0 \\ 0 = 0 & \end{cases}$$

Então $\mathrm{N}(\mathbf{a}) = \mathbb{R}\xi$ onde $\xi = (-2, 1, 0)$ e $\mathcal{B} := \{\xi\}$ é uma base.

Exemplo 8.1.3. Seja $\mathcal{E} \subset \mathbb{R}^3$ a base canônica e $A\colon \mathbb{R}^3 \to \mathbb{R}^3$ determinado por

$$\begin{aligned} Ae_1 &= 2e_1 - e_2 - e_3 \\ Ae_2 &= -e_1 + e_2 \\ Ae_3 &= -e_1 + e_3 \end{aligned}$$

Determine os subespaços $\mathrm{N}(A)$, $\mathrm{Im}(A)$, as dimensões, e uma base de cada um. (Tenhamos encontrado A antes nos Exercícios 5.2.1 e 5.3.10.)

Uma solução. A definição de A já mostra que a matriz e dada por

$$\mathbf{a} := [A] = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

• Como $\mathcal{E}$ corresponde ao isomorfismo identidade, veja (5.0.1) com $\mathcal{U} = \mathcal{E}$, e usando (4.2.2) obtemos passos 1 e 2 de

$$\mathrm{Im}(A) = \mathrm{Im}(\mathbf{a}) = \text{Esp-col}(\mathbf{a}) = \text{Esp-lin}(\mathbf{a}^t)$$

Escalonamos

$$\mathbf{a}^t = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \xrightarrow{\text{(oe)}} \cdots \xrightarrow{\text{(oe)}} \begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{matrix} \ell_1 \\ \ell_2 \end{matrix}$$

Então as listas $\{\ell_1, \ell_3\}$ formam uma base da imagem e assim $\dim \mathrm{Im}(A) = 2$.
• Em respeito ao núcleo de $A = \mathbf{a}$ escalonamos o SL $\mathbf{a}x = 0$, ou seja

$$[\mathbf{a} : \mathcal{O}] = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \end{bmatrix} \xrightarrow{\text{(oe)}} \cdots \xrightarrow{\text{(oe)}} \begin{bmatrix} 2 & -1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Resolvemos o SL escalonado de baixo para cima, ou seja

$$\begin{cases} 2x - y - z = 0 & \Rightarrow x = z,\ z \in \mathbb{R} \\ \quad y - z = 0 & \Rightarrow y = z,\ z \in \mathbb{R} \\ \quad\quad 0 = 0 & \end{cases}$$

Assim $\mathrm{N}(A) = \mathbb{R}\xi$ onde $\xi = (1, 1, 1)$. Então $\{\xi\}$ é uma base e a dimensão é 1.

8.2 Exercícios e umas soluções

1. Calcule a dimensão do subespaço vetorial de $\mathbb{R}^5$ gerado pelos vetores

$$v_1 = (1, 1, 1, -1, 1) \qquad v_2 = (1, -1, -1, 0, 1)$$
$$v_3 = (0, 1, 1, -1, -1) \qquad v_4 = (-1, 1, 1, -1, 1)$$

   Decida se o vetor $b = (6, 18, 1, -9, 8)$ pertence ou não a este subespaço.

2. Encontre uma base para o núcleo da transformação linear

$$C : \mathbb{R}^4 \to \mathbb{R}^3$$
$$(x, y, z, t) \mapsto (2x + y - z + 3t, x - 4y + 2z + t, 2y + 4z - t)$$

   [Dica: Calcule a matriz de C. Escalonamento. Resolve o sistema linear homogêneo resultante.]

Capítulo 9

Subespaços invariantes

Durante o Capítulo 3 denotamos de $A \in \mathcal{L}(E)$ uma transformação linear, alternativamente chamado de operador linear

$$A\colon E \to E, \qquad E = (E, +, \cdot, \mathbb{K}), \qquad n := \dim E < \infty$$

num espaço vetorial E sobre um corpo $\mathbb{K}$ e de dimensão *finita*. No inicio do capítulo a dimensão pode ser também *infinita* no qual caso usamos a notação

$$A\colon X \to X, \qquad X = (X, +, \cdot, \mathbb{K}), \qquad \dim X \in \mathbb{N}_0 \cup \{\infty\}$$

para indicar esta generalidade maior. Ao fim do Capítulo 9 vamos despedir os corpos gerais $\mathbb{K}$ e especializar para os números reais $\mathbb{R}$. Isso é causado pelo fato que tem-se conhecimento bom sobre as raízes de polinômios reais.

Motivação. A matriz $[A]_{\mathcal{B}}$ de A depende da base $\mathcal{B}$ de E. Assim chega naturalmente o desejo de escolher uma base $\mathcal{X}$ tal que a matriz toma uma forma simples, por exemplo uma forma diagonal

$$[A]_{\mathcal{X}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_{n-1} & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{bmatrix} =: \operatorname{diag}[\lambda_1, \dots, \lambda_n] \tag{9.0.1}$$

Tal simplificação máxima, para uma diagonal, é realmente possível para a classe de operadores as quais admitem uma base composto de autovetores. Por exemplo, os operadores auto-adjuntos as quais pode-se definir depois introduzir mais uma estrutura no espaço vetorial – um chamado produto interno, veja Capítulo 10.

Definição 9.0.1. Um transformação linear $A \in \mathcal{L}(E)$ que admite uma base $\mathcal{X}$ tal que a matriz correspondente é diagonal é chamado de **diagonalizável**.

Definição 9.0.2. Um subespaço $F \subset X$ é chamado de **invariante por A** se a imagem $AF \subset F$ é contido no subespaço. Neste caso o operador linear $A|_F : F \to F$, $f \mapsto Af$, é chamado de **restrição de** A.

Exemplo 9.0.3 (Subespaços invariantes)**.** Seja $A \in \mathcal{L}(X)$.

a) $F = \{\mathcal{O}\}$ e $F = X$ (os subespaços invariantes triviais)

b) $F = \mathrm{N}(A)$, $\mathrm{Im}(A)$, $\mathrm{Fix}(A)$, $\mathrm{aFix}(A)$ (subespaços invariantes canônicos de A)

Lema 9.0.4 (Dimensão 1)**.** *Seja $A \in \mathcal{L}(X)$ e F um subespaço de $\dim F = 1$.*

$$F \text{ invariante por } A \quad \Leftrightarrow \quad \exists \lambda = \lambda(A) \in \mathbb{K} \colon Af = \lambda f \ \forall f \in F$$

Demonstração. "⇒" Fixe $\xi \in F$ não-nulo, assim $\mathcal{B} = \{\xi\}$ é base de F. Seja $f \in F$ não-nulo (para $f = \mathcal{O}$ vale $A\mathcal{O} = \lambda\mathcal{O}$ para qualquer um escalar λ). Então $f = \alpha\xi$ para um único escalar α. Como F é invariante por A temos $A\xi \in F$ e assim $A\xi = \lambda\xi$ para um único escalar $\lambda = \lambda(A, \xi)$. Vale

$$Af = A(\alpha\xi) = \alpha A\xi = \alpha(\lambda\xi) = \lambda(\alpha\xi) = \lambda f$$

O $\lambda(A, \xi)$ depende de ξ? Repetindo o argumento para $\tilde{\xi} \in F$ não-nulo obtemos

$$Af = A(\tilde{\alpha}\tilde{\xi}) = \tilde{\alpha}A\tilde{\xi} = \tilde{\alpha}(\tilde{\lambda}\tilde{\xi}) = \tilde{\lambda}(\tilde{\alpha}\tilde{\xi}) = \tilde{\lambda}f$$

Assim $\lambda f = \tilde{\lambda} f$. Daí $f \neq \mathcal{O}$ implica que $\lambda(A, \xi) = \tilde{\lambda}(A, \tilde{\xi})$. Então os escalares $\lambda = \tilde{\lambda}$ são iguais e não dependem nem de ξ nem de $\tilde{\xi}$, só de A.
"⇐" Um subespaço é fechado sob multiplicação escalar. □

Lema 9.0.5 (Dimensão 2)**.** *Para subconjuntos $\{u, v\} \subset X$ LI vale*

$$F := \langle u, v\rangle \text{ invariante por } A \quad \Leftrightarrow \quad Au, Av \in \langle u, v\rangle = F$$

Demonstração. "⇒" Invariância por A junto com o fato que $u, v \in \langle u, v\rangle$.
"⇐" Como $\mathcal{B} = \{u, v\}$ é uma base de F, todo $f \in F$ é da forma $f = \alpha u + \beta v$ para escalares $\alpha, \beta \in \mathbb{K}$. Assim $Af = \alpha Au + \beta Av$ é elemento de F, porque Au e Av são e F é fechado sob $\cdot$ e $+$. □

9.1 Autovalores e autovetores

Definição 9.1.1 (autovalor, autovetores, spectro)**.** Seja $A \in \mathcal{L}(X)$.

a) Por definição chama-se um vetor não-nulo

$$v \in X \setminus \{\mathcal{O}\} \textbf{ autovetor de } \boldsymbol{A} \qquad :\Leftrightarrow \qquad \exists \lambda \in \mathbb{K} \colon Av = \lambda v$$

em palavras, se a imagem Av é um múltiplo de v. Neste caso o escalar λ é chamado de **autovalor de A** e $v = v_\lambda$ um autovetor associado a λ.

b) O **spectro de A** é o conjunto composto de

$$\{\text{todos os autovalores de } A\} =: \operatorname{spec} A$$

Observação 9.1.2. Tendo em vista o Lema 9.0.4 podemos resumir

achar autovetores $\Leftrightarrow$ achar subespaços invariantes de dimensão 1.

Exercício 9.1.3 (Autovetores não são únicas)**.**

a) Os múltiplos não-nulos de um autovetor são autovetores .

b) Somas de autovetores associados ao mesmo autovalor λ são autovetores.

Definição 9.1.4 (Autosubespaço e multiplicidade geométrica)**.** Seja $A \in \mathcal{L}(X)$.

a) Dado um autovalor λ de A, então o conjunto

$$E_\lambda := \{\text{todos os autovetores de } A \text{ associado a } \lambda\} \cup \{\mathcal{O}\}$$

é um subespaço de X chamado de **autosubespaço** associado ao autovalor λ. Note que $E_\lambda \neq \{\mathcal{O}\}$.

b) A **multiplicidade geométrica** de um autovalor λ é a dimensão

$$g_\lambda = g_\lambda(A) := \dim E_\lambda$$

do autosubespaço.

Exemplo 9.1.5. Para $A \in \mathcal{L}(X)$ tem-se autosubespaços

$$\mathrm{N}(A) = E_0, \qquad \mathrm{Fix}A = E_1 \qquad \mathrm{aFix}A = E_{-1}$$

sempre se um dos três subespaço é não -trivial.

Exemplo 9.1.6 (Rotações, reflexões, cisalhamento)**.** Denotamos de A o operador linear na consideração e de E, F subespaços invariantes por A dos quais E tem dimensão 1 – então E é autosubespaço, enquanto F não necessariamente é.

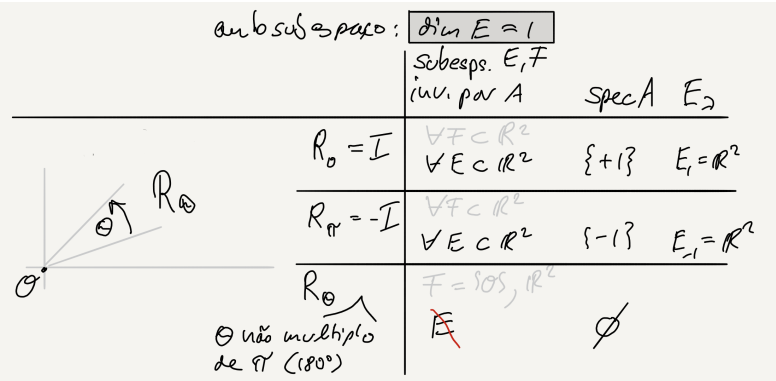


Figura 9.1: Rotação R_θ no plano pelo ângulo θ

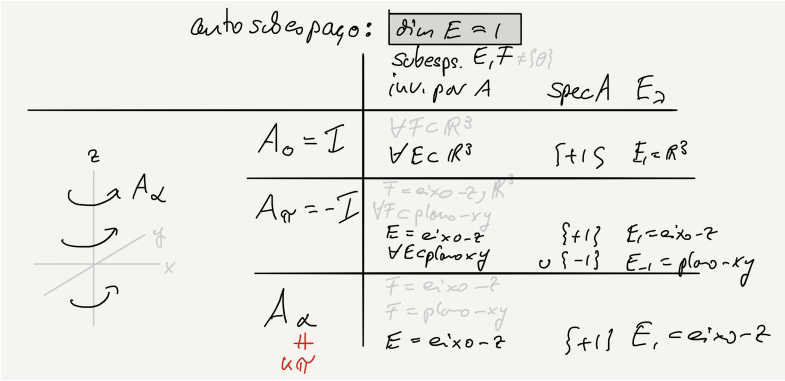


Figura 9.2: Rotação no espaço A_α em torno do eixo z pelo ângulo α

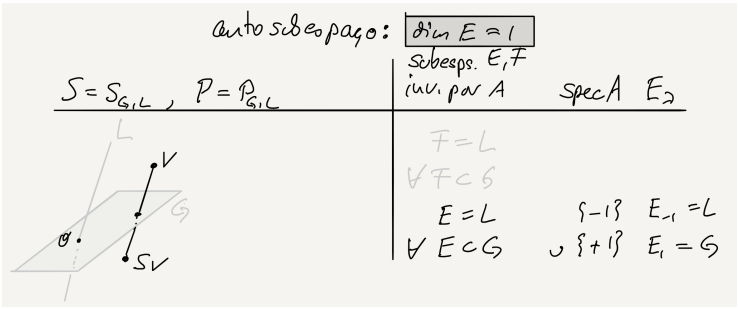


Figura 9.3: Reflexão $S = 2P - I$ sobre G paralelamente L

Teorema 9.1.7 (Autovetores associado a autovalores diferentes são LI)**.** *Se $\lambda_1, \ldots, \lambda_k$ são autovalores dois-a-dois diferentes de $A \in \mathcal{L}(X)$, então qualquer escolha $\{\xi_1, \ldots, \xi_k\}$ de autovetores associados forma um conjunto LI.*

Demonstração. Indução sobre o número k de autovalores.
$\boldsymbol{k = 1}$. Como é autovetor ξ_1 é não-nulo, assim $\{\xi_{\lambda_1}\}$ é LI.
$\boldsymbol{k - 1 \Rightarrow k}$. Dado escalares $\alpha_1, \ldots, \alpha_k$, suponhamos que

$$\alpha_1 \xi_1 + \cdots + \alpha_k \xi_k = \mathcal{O}$$

Aplicamos A para obtemos

$$\alpha_1 \underbrace{A\xi_1}_{\lambda_1 \xi_1} + \cdots + \alpha_k \underbrace{A\xi_k}_{\lambda_k \xi_k} = A\mathcal{O} = \mathcal{O}$$

Adicionamos a esta equação $-\lambda_k$ vezes a equação anterior, obtemos

$$\alpha_1 \underbrace{(\lambda_1 - \lambda_k)}_{\neq 0} \xi_1 + \cdots + \alpha_{k-1} \underbrace{(\lambda_{k-1} - \lambda_k)}_{\neq 0} \xi_{k-1} + \mathcal{O} = \mathcal{O}$$

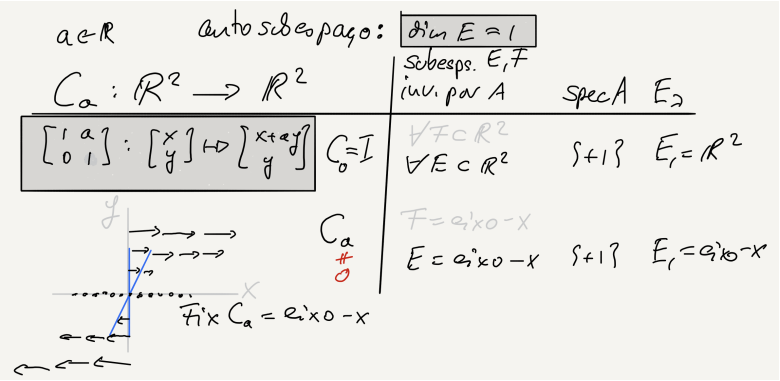


Figura 9.4: Cisalhamento $C_a\colon \mathbb{R}^2 \to \mathbb{R}^2$, dado $a \in \mathbb{R}$

Pela hipótese $k-1$ da indução $\{\xi_1, \dots, \xi_{k-1}\}$ é LI, então cada um coeficiente anula-se, consequentemente $0 = \alpha_1 = \cdots = \alpha_{k-1}$. Com estes valores a equação no inicio da indução reduz-se a $\alpha_k \xi_k = \mathcal{O}$, então como um autovetor não se anula segue que $\alpha_k = 0$ segundo (1.1.3). □

Corolário 9.1.8 (Não tem mais autovalores como $\dim E < \infty$)**.**

$$n := \dim E \qquad \Rightarrow \qquad |\operatorname{spec} A| \leq n$$

Demonstração. Se por absurdo $|\operatorname{spec} A| > n$, então segundo Teorema 9.1.7 pode-se escolher um conjunto LI composto de mais como $\dim E$ elementos. Mas isso contradiz Corolário 3.1.15. □

Corolário 9.1.9 (Diagonalizável)**.** *Se $A \in \mathcal{L}(E)$ possui $n = \dim E < \infty$ autovalores dois-a-dois diferentes $\lambda_1, \dots, \lambda_n$, então obtém-se uma matriz diagonal*

$$[A]_{\mathcal{X}} = \operatorname{diag}[\lambda_1, \dots, \lambda_n]$$

veja (9.0.1), para qualquer seleção $\mathcal{X} = \{\xi_1, \dots, \xi_n\}$ de autovetores associados.

Demonstração. A i-ésima coluna da matriz $[A]_{\mathcal{X}}$ é composto dos coeficientes de

$$A\xi_i = \lambda_i \xi_i = \xi_1 \cdot 0 + \cdots + \xi_{i-1} \cdot 0 + \xi_i \cdot \lambda_i + \xi_{i+1} \cdot 0 + \cdots + \xi_n \cdot 0$$

veja (5.2.1). □

Exercício 9.1.10. Mostre que se a matriz de um operador $A \in \mathcal{L}(E)$ em respeito a uma base $\mathcal{B} = (\xi_1, \dots, \xi_n)$ de E é uma matriz diagonal

$$[A]_{\mathcal{B}} = \operatorname{diag}[a_{11}, \dots, a_{nn}]$$

então cada um elemento a_{ii} na diagonal é autovalor de A e ξ_i é um autovetor.

Proposição 9.1.11. *Seja $A \in \mathcal{L}(E)$ ou $A \in \mathcal{L}(X)$ e $\lambda \in \operatorname{spec} A$, então vale:*

a) $E_\lambda = \mathrm{N}(\lambda I_X - A)$

b) $\lambda \in \operatorname{spec} A \Leftrightarrow \lambda I_E - A$ *não admite inversa* $\Leftrightarrow \det(\lambda I_E - A) = 0$

c) $E_\lambda \subset X$ *é subespaço invariante por* $A \in \mathcal{L}(X)$

d) $\forall \xi \in E_\lambda \setminus \{\mathcal{O}\}$: *o subespaço* $\langle \xi \rangle$ *é invariante por* $A \in \mathcal{L}(X)$

e) *Um subespaço* 1*-dimensional invariante por* A *é composto de autovetores.*

Demonstração. Abreviamos I_E e I_X de I.
a) $Av = \lambda v \Leftrightarrow (\lambda I - A)v = \mathcal{O}$.
b) $\lambda \in \operatorname{spec} A \Leftrightarrow Av = \lambda v$ para um $v \neq \mathcal{O} \Leftrightarrow (\lambda I - A)v = \mathcal{O}$ para um $v \neq \mathcal{O}$ $\Leftrightarrow \lambda I - A$ não injetivo $\Leftrightarrow$[1] $\lambda I - A$ não invertível $\Leftrightarrow$[2] $\det(\lambda I - A) = 0$.
c) Para $v \in E_\lambda$ vale $Av = \lambda v \in E_\lambda$ como o subespaço E_λ é fechado sob "$\cdot$".
d) Seja $\alpha\xi \in \langle \xi \rangle$, então $A(\alpha\xi) = \alpha A\xi = \alpha\lambda\xi \in \langle \xi \rangle$.
e) Lema 9.0.4. □

Lema 9.1.12. *Seja* $\mathcal{U} = \{\xi_1, \xi_2\} \subset E$ *LI e* $a, b, c, d \in \mathbb{K}$*, então o conjunto*

$$\{\underbrace{\alpha\xi_1 + \beta\xi_2}_{=:\eta_1}, \underbrace{\gamma\xi_1 + \delta\xi_2}_{=:\eta_2}\} \text{ é LD} \quad \Leftrightarrow \quad \alpha\delta - \beta\gamma = 0$$

Demonstração. Provamos que LI é equivalente a $\neq 0$. Seja $F := \langle \xi_1, \xi_2 \rangle$ com base $\mathcal{U}$ e seja $G := \langle \eta_1, \eta_2 \rangle$. Então LI significa que o operador linear $A : F \to G$ definido por $A\xi_1 = \eta_1$ e $A\xi_2 = \eta_2$, assim levando base em base, é um isomorfismo segundo Teorema 6.4.7. Mas isso é equivalente a sua matriz $[A]_{\mathcal{U}}$ ser invertível e isso a seu determinante não se anula $0 \neq \det [A]_{\mathcal{U}} = \alpha\delta - \beta\gamma$. □

9.2 Polinômio característico

Nesta Seção 9.2 utilizamos polinômios e o determinante e consequentemente precisamos restringir a dimensão finita ainda permitindo corpos gerais $\mathbb{K}$, ou seja $A \in \mathcal{L}(E)$ e $n := \dim E < \infty$.

Teorema 9.2.1 (Polinômio característico)**.** *Dado um operador linear* $A \in \mathcal{L}(E)$*, então as raízes*[3] *do* ***polinômio característico*** *de* A

$$p_A(\lambda) := \det(\lambda I_E - A), \qquad \lambda \in \mathbb{K}$$

são os autovalores de A*, assim*

$$\operatorname{spec} A = \{\lambda \in \mathbb{K} \mid p_A(\lambda) = 0\}$$

A ordem[4] *de uma raíz* λ *é a* ***multiplicidade algébrica*** alg_λ *do autovalor* λ*.*

Demonstração. Proposição 9.1.11 b). □

[1] Corolário 6.5.2 e Proposição 6.4.4
[2] Teorema A.4.4
[3] as **raízes** de um polinômio p são os pontos λ nos quais $p(\lambda) = 0$ anula-se
[4] só para ilustrar a ideia, a ordem de λ^2 é 2 e de $(\lambda - 5)^4 + \lambda$ é 4

Assim existência de autovalores, então autovetores (ou equivalentemente subespaços invariantes de dimensão 1), é reduzido a existência de raízes de um polinômio com coeficientes em $\mathbb{K}$ e de grau $n = \dim E$. Nos casos importantes $\mathbb{K} = \mathbb{C}$ e $\mathbb{K} = \mathbb{R}$ (veja Seção B.5) tem-se resultados sobre raízes.

Comentário 9.2.2 (Multiplicidades)**.** Para os autovalores $\lambda \in \operatorname{spec} A$ vale que

$$g_\lambda \leq \operatorname{alg}_\lambda$$

Um exemplo para uma desigualdade estrita é o cisalhamento $C_1 \in \mathcal{L}(\mathbb{R}^2)$, veja Figura 9.4. No caso C_1 o leitor pode verificar que $g_\lambda = 1 < 2 = \operatorname{alg}_\lambda$.

Caso $\dim E = 2$

Seja $A \in \mathcal{L}(E)$ um operador linear e $\mathcal{B} = \{\xi_1, \xi_2\}$ uma base ordenada de E. As colunas da matriz correspondente

$$\mathbf{a} := [A]_{\mathcal{B}} = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix}$$

são os coeficientes de $A\xi_1 = \xi_1\alpha + \xi_2\beta$ e de $A\xi_2 = \xi_1\gamma + \xi_2\delta$ segundo (5.2.1). Isso implica, dado $\lambda \in \mathbb{K}$ e $I = I_E$, as equações

$$\begin{aligned}(\lambda I - A)\xi_1 &= (\alpha - \lambda)\xi_1 + \beta\xi_2 =: v_1 \\ (\lambda I - A)\xi_2 &= \gamma\xi_1 + (\alpha - \lambda)\xi_2 =: v_2\end{aligned}$$

Usamos três equivalências chegando, respectivamente, da Proposição 9.1.11 b), do fato que $\{\xi_1, \xi_2\}$ é base, e do Lema 9.1.12, isto é

$$\begin{aligned}\lambda \in \operatorname{spec} A &\Leftrightarrow \lambda I - A \in \mathcal{L}(E) \text{ não é um isomorfismo} \\ &\Leftrightarrow \{v_1, v_2\} \text{ é LD} \\ &\Leftrightarrow 0 = \underbrace{(\alpha - \lambda)(\delta - \lambda) - \beta\gamma}_{\lambda^2 - (\underbrace{\alpha + \delta}_{\operatorname{tr}\mathbf{a}})\lambda + \underbrace{\alpha\delta - \beta\gamma}_{\det\mathbf{a}}} = \boxed{\lambda^2 - (\operatorname{tr}\mathbf{a})\lambda + \det\mathbf{a} =: p_A(\lambda)}\end{aligned}$$

Note que traço e determinante de $[A]_{\mathcal{B}}$ não dependem da escolha da base $\mathcal{B}$; veja (5.4.1) e (5.3.4). Em resumo, na dimensão dois temos mostrado

$$\lambda \in \operatorname{spec} A \ (\lambda \text{ autovalor}) \qquad \Leftrightarrow \qquad p_A(\lambda) = 0$$

Teorema 9.2.3 ($\dim E = 2$)**.** *Dado um operador linear $A \in \mathcal{L}(E)$, então as raízes do* ***polinômio característico de A***

$$p_A \colon \mathbb{K} \to \mathbb{K}, \quad \lambda \mapsto \lambda^2 - (\operatorname{tr} A)\lambda + \det A$$

onde

$$\operatorname{tr} A := \operatorname{tr}\, [A]_{\mathcal{B}}\,, \quad \det A := \det [A]_{\mathcal{B}}$$

são os autovalores de A, assim $\operatorname{spec} A = \{\lambda \in \mathbb{K} \mid p_A(\lambda) = 0\}$.

Comentário 9.2.4 (Corpo $\mathbb{K} = \mathbb{R}$)**.** As raízes de um polinômio real quadrático

$$ax^2 + bx + c = 0$$

existem no caso $b^2 \geq 4ac$ e são dados pela formula

$$x_\pm = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \qquad \textbf{caso } \boldsymbol{a = 1}\textbf{:} \;\; x_\pm = \frac{-b \pm \sqrt{b^2 - 4c}}{2} \tag{9.2.1}$$

Exemplo 9.2.5. Determine o spectro e os autosubespaços da matriz

$$\mathbf{a} = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$$

Uma solução. Passo 1 – autovalores. Temos que determinar, se existir, os autovalores as quais são as raízes do polinômio característico

$$\begin{aligned} p_{\mathbf{a}}(\lambda) :&= \lambda^2 - (\operatorname{tr}\mathbf{a})\lambda + \det\mathbf{a} \\ &= \lambda^2 - (4+2)\lambda + (4 \cdot 2 - 1 \cdot 3) \\ &= \lambda^2 - 6\lambda + 5 \end{aligned}$$

Encontramos as raízes através da fórmula (9.2.1) obtendo

$$\lambda_\pm = \frac{6 \pm \sqrt{36 - 20}}{2} = 3 \pm 2, \qquad \lambda_- = 1, \quad \lambda_+ = 5$$

Assim $\operatorname{spec}\mathbf{a} = \{1, 5\}$.

Passo 2 – autosubespaços. O núcleo da matriz $\mathbf{a} - \lambda\mathbb{1}$ para $\lambda = 1, 5$.
$E_1 = \mathrm{N}(\mathbf{a} - \mathbb{1})$: Depois calcular a matriz $\mathbf{a} - \mathbb{1}$ escalonamos ela e resolvemos de baixo para cima, veja Exemplo 8.1.2. Ou seja

$$\mathbf{a} - \mathbb{1} = \begin{bmatrix} 3 & 3 \\ 1 & 1 \end{bmatrix} \xrightarrow{\text{(oe)}} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \xrightarrow{\text{(oe)}} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

O SL obtida é $x + y = 0$ e assim $y = -x$ para $x \in \mathbb{R}$. Assim $E_1 = \mathbb{R}(1, -1)$.
$E_5 = \mathrm{N}(\mathbf{a} - 5\mathbb{1})$: Analogamente

$$\mathbf{a} - 5\mathbb{1} = \begin{bmatrix} -1 & 3 \\ 1 & -3 \end{bmatrix} \xrightarrow{\text{(oe)}} \begin{bmatrix} -1 & 3 \\ 0 & 0 \end{bmatrix}$$

O SL obtida é $-x + 3y = 0$ e assim $x = 3y$ para $y \in \mathbb{R}$. Assim $E_5 = \mathbb{R}(3, 1)$.

Exercício 9.2.6. Determine autovalores e autosubespaços do operador linear

$$A\colon \mathcal{P}_1(\mathbb{R}) \to \mathcal{P}_1(\mathbb{R}), \quad a + bx \mapsto (4a + 3b) + (a + 2b)x$$

9.3 Existência – caso real

Nesta seção restringimos a espaços vetoriais reais ($\mathbb{K} = \mathbb{R}$) de dimensão finita n e A um operador linear em E, ou seja $A \in \mathcal{L}(E)$.

Teorema 9.3.1 (Existência)**.** *Um operador linear A num espaço vetorial real E de dimensão finita n admite um subespaço invariante F de dimensão 1 ou 2.*

Demonstração. Teorema B.5.1 □

9.4 Exercícios

1. Seja $A : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (3x + y, 2x + 2y)$.
   (a) Mostre que 4 e 1 são autovalores de A.
   (b) Ache uma base ordenada $\mathcal{B} = (u, v)$ de $\mathbb{R}^2$ tal que
   $$Au = 4u \qquad \text{e} \qquad Av = v.$$
   (c) Dada a matriz $\mathbf{a} = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$, ache uma matriz invertível $\mathbf{p}$ tal que
   $$\mathbf{p}^{-1}\mathbf{a}\mathbf{p} = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}.$$

2. Dado $a \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$, determine os subespaços de $\mathbb{R}^3$ invariantes por
   $$A : \mathbb{R}^3 \to \mathbb{R}^3, \qquad v \mapsto a \times v,$$
   onde o produto vetorial $\times$ é definido no Exercício 5.4.5.

3. Sejam $A, B \in \mathcal{L}(E)$ operadores que *comutam*: $AB = BA$. Prove que
   (a) $\mathrm{N}(B)$ e $\mathrm{Im}(B)$ são subespaços invariantes por A;
   (b) Se F é um subespaço invariante por A, então $BF := \{Bf : f \in F\}$ é ainda um subespaço invariante por A.

4. Dado $A \in \mathcal{L}(E)$ e um polinômio $p = p(x)$, prove que núcleo e imagem do operador $p(A) \in \mathcal{L}(E)$ (Definição B.5.5) são subespaços invariantes por A.

5. Determine os autovetores e os autovalores do operador derivação
   $$D : \mathcal{P}_n(\mathbb{R}) \to \mathcal{P}_n(\mathbb{R}), \quad p(x) \mapsto p'(x) := \tfrac{d}{dx}p(x)$$
   no espaço vetorial dos polinômios $p = p(x)$.

6. Seja $A \in \mathcal{L}(E)$, prove que
   (a) A invertível $\iff$ A não possui autovalor 0;
   (b) Se A é invertível, então os autovetores de A e A^{-1} coincidem. E os autovalores?

7. Seja $A : \mathbb{R}^n \to \mathbb{R}^n$ o operador linear cuja matriz na base canônica tem todas as entradas iguais a 1. Prove que
   (a) $\mathrm{posto}(A) = 1$;
   (b) $\mathbb{R}^n = \mathrm{N}(A) \oplus \mathrm{Im}(A)$;
   (c) os autovalores de A são 0 e n;
   (d) os autovetores de A pertencem a $\mathrm{N}(A)$ ou a $\mathrm{Im}(A)$.

   Exiba uma base de $\mathbb{R}^n$ na qual a matriz de A tem $n^2 - 1$ zeros.

8. Mostre que todo operador $A \in \mathcal{L}(\mathbb{R}^n)$ de posto 1 possui um autovetor v cujo autovalor λ é o traco de A.

Parte III

Estruturas adicionais e operadores especiais

Capítulo 10

Produto interno

Neste Capítulo 10 consideramos exclusivamente sub/espaços vetoriais reais

$$F \subset E = (E, +, \cdot, \mathbb{R})$$

o que quer dizer que o corpo neste capítulo são os números reais. As letras $k, n \in \mathbb{N}_0$ denotam números naturais, no caso de dimensão finita elas denotam as dimensões de F e E. Além disso a letra

$$E = (E, \langle \cdot, \cdot \rangle)$$

denota um espaço vetorial munido de um produto interno.

10.1 Produto interno, norma, distância

Definição 10.1.1 (Produto interno)**.** Um **produto interno**[1] num espaço vetorial real G é uma função de duas variáveis

$$\langle \cdot, \cdot \rangle \colon G \times G \to \mathbb{R}, \quad (u, v) \mapsto \langle u, v \rangle$$

a qual satisfaz os três axiomas

(SIM) $\langle u, v \rangle = \langle v, u \rangle$ (simetria)

(BL) $\langle u + \tilde{u}, v \rangle = \langle u, v \rangle + \langle \tilde{u}, v \rangle, \quad \langle \alpha u, v \rangle = \alpha \langle u, v \rangle$ (bi-linearidade)[2]

(POS) $u \neq \mathcal{O} \Rightarrow \langle u, u \rangle > 0$ (positividade)

para todos os vetores $u, v, \tilde{u}, \tilde{v} \in G$ e escalares $\alpha \in \mathbb{R}$. Neste caso E é chamado de **espaço vetorial com produto interno**.

Lema 10.1.2. *Num espaço vetorial E com produto interno vale*

[1] Produtos internos são também chamados de **produtos escalares**.

[2] Note que simetria implica linearidade também na segunda variável, por isso o nome para o axioma dois, abreviando bi-linearidade, é justificado.

$$\langle v, \mathcal{O}\rangle = 0 \quad \forall v \in E$$

(ND) $\langle u, v\rangle = \langle \tilde{u}, v\rangle \quad \forall v \in E \quad \Rightarrow \quad u = \tilde{u}$ (não-degenerado)

(ND)′ $\langle u, v\rangle = 0 \quad \forall v \in E \quad \Rightarrow \quad u = \mathcal{O}$ (não-degenerado)′

Lema 10.1.3 (Critério para dois operadores são iguais)**.** *Dado operadores $A, B \in \mathcal{L}(E, F)$ e bases arbitrarias $\mathcal{U} = (\xi_1, \ldots, \xi_n)$ e $\mathcal{V} = (\eta_1, \ldots, \eta_m)$, então*

$$A = B \qquad \Leftrightarrow \qquad \langle \eta_i, A\xi_j\rangle = \langle \eta_i, B\xi_j\rangle \quad \forall i, j$$

Demonstração. "⇒" Trivial. "⇐" Axioma (BL) na primeira entrada de $\langle\cdot,\cdot\rangle$ e propriedade (ND) em Lema 10.1.2. □

Exemplo 10.1.4 (Produto euclidiano em $\mathbb{R}^n$)**.** A função

$$\langle\cdot,\cdot\rangle_0 \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

$$(x, y) \mapsto x_1y_1 + \cdots + x_ny_n = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

é chamado de **produto euclidiano** em $\mathbb{R}^n$. O leitor pode verificar os 3 axiomas. Caso não especificamos diferentemente o $\mathbb{R}^n$ sempre será munido do produto euclidiano.

Exemplo 10.1.5 (Produto interno mediante integração)**.** No espaço vetorial $C^0([a, b])$ das funções reais continuas num intervalo $[a, b]$ integração

$$\langle f, g\rangle := \int_a^b f(x)g(x)\, dx \tag{10.1.1}$$

define um produto interno. Deixamos ao leitor verificar os 3 axiomas.

Exemplo 10.1.6 (Integração não dando produto interno)**.** No espaço vetorial $C^0(\mathbb{R})$ das funções reais continuas no $\mathbb{R}$ inteiro, integração, nem sobre $\mathbb{R}$, nem sobre um intervalo $[a, b]$,

$$\langle f, g\rangle_\infty := \int_{-\infty}^{\infty} f(x)g(x)\, dx, \qquad \langle f, g\rangle := \int_a^b f(x)g(x)\, dx$$

define um produto interno em $C^0(\mathbb{R})$, não.

O problema no primeiro caso são valores infinitos, por exemplo $\langle 2, 3\rangle_\infty = \infty$.

O problema no segundo caso é o axioma (POS). Seja $u \colon \mathbb{R} \to \mathbb{R}$ uma função continua a qual anula-se em $[a, b]$, mas não no complemento inteiro. Por exemplo, suponha $u(b + 1) = 1$. Assim $u \neq \mathcal{O}$ não é a função nula, mas a integral $\langle u, u\rangle = \int_a^b u(x)^2\, dx = 0$ não é positivo.

Exemplo 10.1.7 (Polinômios em $\mathbb{R}$ e integração sobre $[a, b]$)**.** Em contraste ao espaço $C^0(\mathbb{R})$, no espaço vetorial dos polinômios $\mathcal{P}(\mathbb{R})$ integração sobre $[a, b]$ produz um produto interno! Deixamos ao leitor mostrar que (10.1.1) realmente satisfaz o axioma (POS).

[Dica: Em quantos pontos um polinômio de grau n pode-se anular no máximo?]

Normas – norma induzida

Definição 10.1.8. Uma **norma** num espaço vetorial real X é uma função

$$|\cdot| : X \to [0, \infty), \quad x \mapsto |x|$$

a qual satisfaz os três axiomas

(HOM) $|\alpha x| = \alpha\,|x|$ (homogeneidade) $\Rightarrow |\mathcal{O}| = 0$

(Δ) $|x + y| \leq |x| + |y|$ (desigualdade triangular)

(POS) $x \neq \mathcal{O} \Rightarrow |x| > 0$ (positividade)

para todos os vetores $x, y \in X$ e escalares $\alpha \in \mathbb{R}$. Neste caso X é chamado de **espaço vetorial normado**.

Definição 10.1.9 (Norma induzida)**.** Num espaço vetorial E com produto interno existe para cada um vetor v um número não-negativo

$$|v| = |v|_{\langle\cdot,\cdot\rangle} := \sqrt{\langle v, v\rangle} \geq 0$$

chamado de **norma induzida** de v, ou informalmente o "*comprimento*" do vetor. Um vetor de comprimento $|v| = 1$ é chamado de **vetor unitário** e às vezes denotado de $\hat{v}$ para ênfase.

Lema 10.1.10. *Num espaço vetorial E com produto interno a norma induzida é uma norma. Além disso para todo vetor não-nulo $\hat{u} := \frac{1}{|u|}u$ é um vetor unitário.*

Métricas – métrica induzida

Definição 10.1.11. Uma **métrica**[3] num conjunto M é uma função

$$\begin{aligned} d\colon M \times M &\to [0, \infty) \\ (q, p) &\mapsto d(q, p) \end{aligned}$$

a qual satisfaz os três axiomas

(SIM) $d(q, p) = d(p, q)$ (simetria)

(Δ) $d(q, r) \leq d(q, p) + d(p, r)$ (desigualdade triangular)

(POS) $d(q, q) = 0$ mas $q \neq p \Rightarrow d(q, p) > 0$ (positividade)

para todos os pontos $q, p \in M$. Neste caso M é chamado de **espaço métrico**.

Definição 10.1.12 (Métrica induzida)**.** Num espaço vetorial normado a função

$$d(x, y) = d_{|\cdot|}(x, y) := |x - y|$$

é chamado de **métrica induzida** ou **distância** entre dois pontos.

Lema 10.1.13. *A métrica induzida $d(x, y) := |x - y|$ é uma métrica.*

Então produtos internos disponibilizam normas e normas disponibilizam distâncias. As inclusões são ilustrados na Figura 10.1.

[3] Métricas são também chamadas de **funções distância** ou simplesmente **distâncias**.

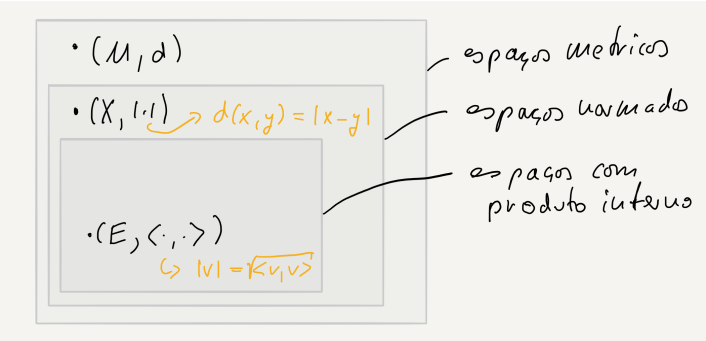


Figura 10.1: Produto interno → norma → função distância

10.1.1 Produto interno e espaço dual – dualidade

Um produto interno $\langle \cdot, \cdot \rangle$ num espaço vetorial E com $\dim E = n$ finita disponibiliza um isomorfismo canônico[4] entre E e seu espaço dual

$$\begin{aligned} D \colon E &\to E^* = \mathcal{L}(E, \mathbb{R}) \\ v &\mapsto \langle v, \cdot \rangle \end{aligned} \tag{10.1.2}$$

chamado de **dualidade** e onde $\langle v, \cdot \rangle$ é a transformação linear abreviada de

$$v^* := \langle v, \cdot \rangle \colon E \to \mathbb{R}, \quad u \mapsto \langle v, u \rangle$$

Teorema 10.1.14 (Dualidade)**.** *O operador D em (10.1.2) é um isomorfismo.*

Demonstração. Linearidade de (10.1.2) vale segundo o axioma (BL) de bilinearidade. Injetividade vale segundo o axioma (POS) na sua incarnação $(\text{ND})'$. Sobrejetividade é equivalente a injetividade segundo Corolário 6.5.2 porque as dimensões $\dim E = n = \dim E^*$ são iguais segundo Lema 4.1.20. □

Exercício 10.1.15 (Produto interno induzido no espaço dual)**.** Mostre que

$$\langle \cdot, \cdot \rangle_* := \langle D^{-1} \cdot, D^{-1} \cdot \rangle \colon E^* \times E^* \to \mathbb{R}$$

é um produto interno no espaço dual de E.

Exercício 10.1.16. Seja E um espaço vetorial com produto interno $\langle \cdot, \cdot \rangle$ e seja $\mathcal{B} = \{\xi_1, \dots, \xi_n\}$ uma base de E. Dado números $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, prove que existe um único vetor $w \in E$ tal que

$$\langle w, \xi_1 \rangle = \alpha_1, \ \dots, \ \langle w, \xi_n \rangle = \alpha_n.$$

[4] **Canônico** significa sem a necessidade de fazer escolhas das quais o objeto construído eventualmente vai depender. Por exemplo, se $\dim E = \dim G$, então E e F são isomorfos segundo Corolário 6.4.9, mas não tem um isomorfismo canônico geralmente. E de um isomorfismo com muita escolha geralmente não pode extrair informações intrínsecas.

As afirmações continuam em Lema 10.1.17.

[Dica: Proposição 4.1.12 diz que uma transformação linear $\psi\colon E \to \mathbb{R}$ é determinada por seus valores numa base, dizemos $\psi\xi_i := \alpha_i$. Defina $w := D^{-1}\psi$.]

Lema 10.1.17 (Continuando Exercício 10.1.17)**.** *Seja E um espaço vetorial com produto interno $\langle\cdot,\cdot\rangle$ e seja $\mathcal{B} = \{\xi_1, \ldots, \xi_n\}$ uma base de E. Prove que existe uma única base $\mathcal{V} = \{\eta_1, \ldots, \eta_n\}$ de E tal que*

$$\langle \eta_i, \xi_j \rangle = \delta_{ij}, \qquad i, j = 1, \ldots, n$$

Defina $a_{ij} := \langle \xi_i, \xi_j \rangle$ e $b_{ij} := \langle \eta_i, \eta_j \rangle$, onde $i, j = 1, \ldots, n$. Prove que as matrizes $\mathbf{a} = (a_{ij})$ e $\mathbf{b} = (b_{ij})$ são inversas uma da outra.

Demonstração. Dado uma base $\mathcal{B} = \{\xi_1, \ldots, \xi_n\}$ de E, seja $\mathcal{B}^* = \{\phi_1, \ldots \phi_n\}$ a base dual (4.1.5) de E^*. Isomorfismos levam base em base (Teorema 6.4.7), assim $\mathcal{V} := D^{-1}\mathcal{B}^*$ é uma base de E. Para os elementos $\eta_i := D^{-1}\phi_i$ de $\mathcal{V}$ vale

$$\langle \eta_i, \xi_j \rangle = \langle D^{-1}\phi_i, \xi_j \rangle \overset{(10.1.2)}{=} \big(D(D^{-1}\phi_i)\big)\,\xi_j = \phi_i \xi_j \overset{(4.1.5)}{=} \delta_{ij}$$

Segundo Teorema 5.2.5 $I_E = D^{-1}D$ traduz em $[I_E]_{\mathcal{B},\mathcal{B}} = \big[D^{-1}D\big]_{\mathcal{B},\mathcal{B}}$. Assim

$$\mathbb{1} \overset{(5.2.4)}{=} [I_E]_{\mathcal{B},\mathcal{B}} = \big[D^{-1}D\big]_{\mathcal{B},\mathcal{B}} \overset{(5.2.5)}{=} \big[D^{-1}\big]_{\mathcal{B}^*,\mathcal{B}}\,[D]_{\mathcal{B},\mathcal{B}^*} = \mathbf{ba}$$

onde resta provar a última identidade. (Como as matrizes são quadradas $\mathbb{1} = \mathbf{ba}$ é equivalente a $\mathbb{1} = \mathbf{ab}$.) Mais detalhado, resta provar que

$$\mathbf{c} := \big[D^{-1}\big]_{\mathcal{B}^*,\mathcal{B}} = \mathbf{b} := (\langle \eta_i, \eta_j \rangle) \quad , \quad \mathbf{d} := [D]_{\mathcal{B},\mathcal{B}^*} = \mathbf{a} := (\langle \xi_i, \xi_j \rangle)$$

Começamos com a definição de

$$\begin{aligned}
b_{ji} &:= \langle \eta_j, \eta_i \rangle \\
&= \langle \eta_i, \eta_j \rangle \\
&= \langle D^{-1}\phi_i, \eta_j \rangle \\
&= \big(D(D^{-1}\phi_i)\big)\,\eta_j \\
&= \phi_i \eta_j \\
&= \phi_i \big(D^{-1}\phi_j\big) \\
&\overset{*}{=} \phi_i \left(\xi_1 c_{1i} + \ldots \xi_n c_{ni}\right) \\
&= \phi_j \xi_j c_{ji} \\
&= c_{ji}
\end{aligned}$$

onde $*$ vale por definição (5.2.1) da matriz $\big[D^{-1}\big]_{\mathcal{B}^*,\mathcal{B}} =: (c_{ij})$. Deixamos ao leitor provar similarmente $a_{ji} = d_{ji}$. □

10.1.2 Produto interno e matrizes

Seja G um espaço vetorial de dimensão finita n. Uma vez fixado uma base ordenada $\mathcal{B}$ de G, tem uma aplicação

$$\Phi = \Phi_{\mathcal{B}} \colon \mathrm{S}^+(n) \to \{\text{produtos internos em } G\}$$

entre os conjuntos $\mathrm{S}^+(n)$ das matrizes reais $n \times n$ simétricas **positivas**[5] e dos produtos internos em G.
De outro lado, tem uma aplicação

$$\Phi = \Phi_{\mathcal{B}} \colon \{\text{bases ordenadas de } G\} \to \{\text{produtos internos em } G\}$$

entre os conjuntos das bases ordenadas de G e dos produtos internos em G.

A matriz de um produto interno

Definição 10.1.18 (Matrizes do produto interno)**.** Dado uma base ordenada $\mathcal{B} = (\xi_1, \ldots, \xi_n)$ de E, calcule todos os números reais

$$g_{ij} := \langle \xi_i, \xi_j \rangle$$

e coloque numa matriz quadrada denotada, dependente do contexto, de

$$\mathbf{g} = \mathbf{g}_{\mathcal{B}} = [g]_{\mathcal{B}} := (g_{ij})_{i,j=1}^{n} \in \mathrm{M}(n \times n)$$

Esta matriz é chamada de **matriz do produto interno** em respeito à base $\mathcal{B}$. Deixe nos simplificar também a notação dos vetores coordenadas

$$\mathbf{u} = \mathbf{u}_{\mathcal{B}} = [u]_{\mathcal{B}} := \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathrm{M}(n \times 1)$$

dos vetores $u \in E$ em respeito a $\mathcal{B}$, veja (5.1.2).

Comentário 10.1.19. Note que $\mathbf{g}_{\mathcal{B}}$ é uma matriz real simétrica positiva. Usando a base $\mathcal{B}$ e os vetores coordenadas $\mathbf{u}_{\mathcal{B}}$ pode-se escrever o produto interno em E como produto matriz ou como produto euclidiano

$$\langle u, v \rangle = \sum_{i,j=1}^{n} u_i g_{ij} v_j = \underbrace{\mathbf{u}^t}_{1\times n} \underbrace{\mathbf{g}}_{n\times n} \underbrace{\mathbf{v}}_{n\times 1} = \langle \mathbf{u}, \mathbf{g}\mathbf{v} \rangle_0 =: \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{g}} \tag{10.1.3}$$

para os vetores u, v de E.

[5] Uma matriz real quadrada $\mathbf{g}$ é chamada de **matriz positiva** se

$$\sum_{i,j=1}^{n} g_{ij} u_i u_j > 0$$

para todas as listas não nulas $u \in \mathbb{R}^n$.

Exemplo 10.1.20. Nos polinômios reais de grau menor ou igual um

$$\mathcal{P}_1(\mathbb{R}) := \{p(x) = a_0 + a_1 x \mid a_0, a_1 \in \mathbb{R}\}$$

considere a base ordenada $\mathcal{B} = (\xi_1, \xi_2) = (3, 1+x)$. Integração

$$\langle p, q\rangle := \int_{-1}^{1} p(x)q(x)\,dx$$

da um produto interno em $\mathcal{P}_1(\mathbb{R})$ cuja matriz em respeito a $\mathcal{B}$ tem entradas

$$\begin{aligned}
g_{11} &:= \langle \xi_1, \xi_1\rangle = \int_{-1}^{1} 3\cdot 3\,dx = 9x|_{-1}^{1} = 18\\
g_{22} &:= \langle \xi_2, \xi_2\rangle = \int_{-1}^{1} (1+x)\cdot(1+x)\,dx = (x + x^2 + x^3/3)\big|_{-1}^{1} = 8/3\\
g_{12} &:= \langle \xi_1, \xi_2\rangle = \int_{-1}^{1} 3(1+x)\,dx = (3x + 3x^2/2)\big|_{-1}^{1} = 6\\
g_{21} &:= \langle \xi_2, \xi_1\rangle \overset{\texttt{(SIM)}}{=} \langle \xi_1, \xi_2\rangle = g_{12} = 6
\end{aligned}$$

Exercício 10.1.21 (Continuamos Exemplo 10.1.20)**.** Determine a distância

$$d(\xi_1, \xi_2) := |\xi_1 - \xi_2| := \sqrt{\langle \xi_1 - \xi_2, \xi_1 - \xi_2\rangle}$$

dos dois membros da base de $E = \mathcal{P}_1$.

Uma solução em *E*. Inserindo na fórmula obtemos para o quadrado

$$d(\xi_1, \xi_2)^2 = \int_{-1}^{1} \underbrace{(3 - (1+x))^2}_{4+4x+x^2}\,dx = (4x + 2x^2 + \tfrac{1}{3}x^3)\big|_{x=-1}^{1} = \frac{26}{3}$$

Outra solução em coordenadas. A matriz $[g]_{\mathcal{B}}$ já conhecemos, calculamos

$$[\xi_1 - \xi_2]_{\mathcal{B}} = [\xi_1]_{\mathcal{B}} - [\xi_2]_{\mathcal{B}} = \begin{bmatrix}1\\0\end{bmatrix} - \begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}1\\-1\end{bmatrix}$$

Com isso, usando (10.1.3) no primeiro passo, obtemos

$$\begin{aligned}
d(\xi_1, \xi_2)^2 &= \langle [\xi_1 - \xi_2]_{\mathcal{B}}, [g]_{\mathcal{B}}\,[\xi_1 - \xi_2]_{\mathcal{B}}\rangle_0\\
&= \left\langle \begin{bmatrix}1\\-1\end{bmatrix}, \begin{bmatrix}18 & 6\\ 6 & 8/3\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix}\right\rangle_0\\
&= \left\langle \begin{bmatrix}1\\-1\end{bmatrix}, \begin{bmatrix}12\\10/3\end{bmatrix}\right\rangle_0\\
&= \frac{26}{3}
\end{aligned}$$

O produto interno de uma base ordenada

Proposição 10.1.22 (Existência de produtos internos)**.** *Um espaço vetorial real G de dimensão finita n admite um produto interno – um produto interno $\langle u, v\rangle_{\mathcal{B}}$ para cada uma base ordenada $\mathcal{B}$.*

Demonstração. Dado uma base ordenada $\mathcal{B} = (\xi_1, \dots, \xi_n)$ de G, defina

$$\langle u, v\rangle_{\mathcal{B}} := \langle \mathbf{u}, \mathbf{v}\rangle_0 \tag{10.1.4}$$

para $u, v \in G$ onde $\mathbf{u} = [u]_{\mathcal{B}} \in \mathbb{R}^n$ é o vetor coordenada, veja (5.1.2). □

Exercício 10.1.23. Seja $G = \mathbb{R}^n$. Mostre que o produto interno associado à base canônica $\langle \cdot, \cdot\rangle_{\mathcal{E}} = \langle \cdot, \cdot\rangle_0$ reproduz o produto euclidiano no $\mathbb{R}^n$.

10.2 Ortogonalidade

Definição 10.2.1. Seja E um espaço vetorial com produto interno.

(i) Chama-se dois vetores u e v **ortogonais**, ou **perpendiculares**, símbolo $u \perp v$, se tem produto nulo $\langle u, v\rangle = 0$. (Note $\mathcal{O} \perp v$ para todos vetores.)

(ii) Chama-se $X \subset E$ um **subconjunto ortogonal** se os vetores de X são dois-a-dois ortogonais.

(iii) Chama-se $X \subset E$ um **subconjunto ortonormal (ON)** se X é composto de vetores unitários dois-a-dois ortogonais.

(iv) Uma base $\mathcal{Z} = \{\varepsilon_1, \dots, \varepsilon_n\}$ é chamado de **base ortonormal (ON)** se

$$\langle \varepsilon_i, \varepsilon_j\rangle = \delta_{ij} := \begin{cases} 1 & , i = j \\ 0 & , i \neq j \end{cases} \tag{10.2.1}$$

onde δ_{ij} é o símbolo de Kronecker.

Teorema 10.2.2 (Conjuntos ortogonais, sem $\mathcal{O}$, são LI)**.**

$$X \subset E \setminus \{\mathcal{O}\} \text{ conjunto ortogonal} \quad \Rightarrow \quad X \text{ LI}$$

Exercício 10.2.3. Dado uma base ON $\mathcal{Z} = \{\varepsilon_1, \dots, \varepsilon_n\}$ de E, mostre que

$$v = \sum_{i=1}^{n} v_i \varepsilon_i \qquad \Leftrightarrow \qquad v_i = \langle \varepsilon_i, v\rangle, \quad i = 1, \dots, n$$

para cada um vetor $v \in E$.

Exercício 10.2.4. Seja $\mathcal{B}$ uma base ordenada de um espaço vetorial real G de dimensão finita n. Seja $\langle u, v\rangle_{\mathcal{B}}$ o produto interno correspondente (10.1.4). Mostre que $\mathcal{B}$ é uma base ON de $\langle u, v\rangle_{\mathcal{B}}$.

Exemplo 10.2.5 (Conjuntos e bases ortogonais)**.**

a) A base canônica $\mathcal{E}^n$ é ortonormal em respeito a $\langle\cdot,\cdot\rangle_0$.

b) O conjunto $\{(0,0),(-1,1)\}$ é ortogonal em $\mathbb{R}^2$.

c) O conjunto $\{(1,1),(-1,1)\}$ é uma base ortogonal de $\mathbb{R}^2$.

Teorema 10.2.6 (Teorema de Pitágoras generalizado)**.** *Sejam $u, v \in E$, então*

$$u \perp v \qquad \Leftrightarrow \qquad |u+v|^2 = |u|^2 + |v|^2$$

Exemplo 10.2.7. Para vetores não-nulos $u, v \in \mathbb{R}^2$ ter produto nulo

$$0 = \langle u, v\rangle_0 = |u| \cdot |v| \cdot \cos \angle(u, v)$$

é equivalente que o angulo entre eles é rectângulo, ou seja $\angle(u,v) \in \{\frac{\pi}{2}, \frac{3\pi}{2}\}$.

10.2.1 Projeção ortogonal sobre uma reta

Definição 10.2.8 (Projeção ortogonal sobre uma reta $\mathbb{R}\hat{u}$)**.** Seja E um espaço vetorial com produto interno. Para $u \in E$ não-nulo, seja $\hat{u} = \frac{1}{|u|}u$ o vetor unitário correspondente, definimos a transformação linear

$$\begin{aligned} \mathrm{pr}_u \colon E &\to E \\ v &\mapsto \frac{\langle u, v\rangle}{\langle u, u\rangle} u = \langle \hat{u}, v\rangle \hat{u} =: \mathrm{pr}_{\hat{u}} v \end{aligned}$$

Comentário 10.2.9 (Idempotente - projeção)**.**

a) $\mathrm{pr}_u(\alpha u) \overset{\text{def.}}{=} \alpha \frac{\langle u,u\rangle}{\langle u,u\rangle} u = \alpha u \quad \Rightarrow \quad \mathrm{pr}_u|_{\mathbb{R}u} = I_{\mathbb{R}u}$

b) $(\mathrm{pr}_u)^2 v = \underbrace{\mathrm{pr}_u}_{I \text{ em } \mathbb{R}u} \underbrace{\mathrm{pr}_u v}_{\in \mathbb{R}u} = \mathrm{pr}_u v$ e assim pr_u é uma projeção em E.

Lema 10.2.10 (Projeções não aumentam comprimento)**.** $|\mathrm{pr}_u v| \leq |v| \quad \forall v \in E$

Demonstração. □

10.3 Ângulos e cumprimentos em $(\mathbb{R}^2, \langle\cdot,\cdot\rangle_0)$

Comentário 10.3.1. Para $u = (\alpha, \beta) \in \mathbb{R}^2$ vale

$$\mathrm{dist}(u, \mathcal{O}) \overset{\text{Pit.}}{=} \sqrt{\alpha^2 + \beta^2} = \sqrt{\langle u, u\rangle_0} =: |u|_0$$

Lema 10.3.2. *Para $u, v \in \mathbb{R}^2 \setminus \{\mathcal{O}\}$ vale $\langle u, v\rangle_0 = |u|_0 \cdot |v|_0 \cdot \cos \angle(u, v)$.*

Comentário 10.3.3. São equivalentes

$$\langle u, v\rangle_0 = 0 \quad \Leftrightarrow \quad \cos\theta = 0 \quad \Leftrightarrow \quad \theta \in \{\pi/2, 3\pi/2\} \quad \Leftrightarrow \quad u \perp v$$

10.4 Desigualdades

Proposição 10.4.1 (Desigualdade de Schwarz)**.** *Para vetores $u, v \in E$ vale*

$$|\langle u, v\rangle| \leq |u| \cdot |v|$$

onde igualdade "=" é equivalente a um de u, v é múltiplo do outro.

Proposição 10.4.2 (Desigualdade triangular)**.** *Para vetores $u, v \in E$ vale*

$$|u + v| \leq |u| + |v|$$

onde igualdade "=" é equivalente a um de u, v é múltiplo não-negativo do outro.

10.5 Ortonormalização segundo Gram-Schmidt

Hipótese. Seja $\mathcal{X} = (\xi_1, \ldots, \xi_n)$ uma base ordenada de um espaço vetorial E com produto interno. Denotamos de

$$F_1 := \langle \xi_1 \rangle \subset \cdots \subset \boxed{F_k := \langle \xi_1, \ldots, \xi_k \rangle} \subset \cdots \subset F_n := \langle \xi_1, \ldots, \xi_n \rangle = E$$

os subespaços gerados pelos primeiros $1, 2, \ldots, n$ membros da base $\mathcal{X}$.

Passo 1. Vamos construir iterativamente bases ortogonais

(1) base ortogonal $\{\eta_1\}$ de F_1: escolha $\eta_1 := \xi_1$ e já pronto

(k) dado $k \geq 1$ suponha que $\{\eta_1, \ldots, \eta_k\}$ é base ortogonal de F_k e defina

$$\eta_{k+1} := \xi_{k+1} - \sum_{i=1}^{k} \mathrm{pr}_{\eta_i} \xi_{k+1} = \xi_{k+1} - \sum_{i=1}^{k} \frac{\langle \eta_i, \xi_{k+1}\rangle}{\langle \eta_i, \eta_i\rangle} \eta_i \tag{10.5.1}$$

($k+1$) então $\{\eta_1, \ldots, \eta_k, \eta_{k+1}\}$ é uma base ortogonal de F_{k+1}

o processo usa o último membro ξ_n de $\mathcal{X}$ quando $k = n - 1 \Rightarrow k + 1 = n$

(n) ao fim obtemos a base ortogonal $\{\eta_1, \ldots, \eta_k, \eta_{k+1}, \ldots, \eta_n\}$ de $F_n = E$

Demonstração (k) $\Rightarrow$ ($k+1$). Suponha (k) e defina η_{k+1}, então

a) $\eta_{k+1} \perp \eta_1, \ldots, \eta_k$ $\qquad \langle \eta_{k+1}, \eta_j\rangle \overset{(k)}{=} \langle \xi_{k+1}, \eta_j\rangle - \langle \xi_{k+1}, \eta_j\rangle = 0$

b) $\eta_{k+1} \notin F_k \overset{(k)}{=} \langle \eta_1, \ldots, \eta_k\rangle \ni \mathcal{O}$ $\qquad$ suponha por absurdo $\eta_{k+1} \in \langle \eta_1, \ldots, \eta_k\rangle$
$\Rightarrow \xi_{k+1} \in \langle \eta_1, \ldots, \eta_k\rangle = F_k := \langle \xi_1, \ldots, \xi_k\rangle$ contradição

c) $\eta_{k+1} \in F_{k+1}$ $\qquad \eta_{k+1} \in \langle \eta_1 \ldots, \eta_k, \xi_{k+1}\rangle \overset{(k)}{=} \langle \xi_1 \ldots, \xi_k, \xi_{k+1}\rangle =: F_{k+1}$

Segundo hipotese (k) o conjunto $\{\eta_1, \dots, \eta_k\}$ é LI. Além disso η_{k+1} é não-nulo segundo b) e ortogonal a $\eta_1, \dots, \eta_k$ segundo a). Sendo assim o conjunto ortogonal $\{\eta_1, \dots, \eta_k, \eta_{k+1}\}$ é LI segundo Teorema 10.2.2. Note que o subespaço

$$\langle \eta_1, \dots, \eta_{k+1} \rangle \subset \langle \xi_1, \dots, \xi_{k+1} \rangle$$

é contido num subespaço da mesma dimensão $k+1$. Então os dois são iguais segundo Teorema 3.2.1 (d). Por isso $\{\eta_1, \dots, \eta_{k+1}\}$ é base ortogonal de F_{k+1}.

Passo 2. A base $\mathcal{Z} := \{\hat{\eta}_1, \dots, \hat{\eta}_n\}$ de E é ortonormal.

Comentário 10.5.1. No caso que ξ_{k+1} já é ortogonal a $\eta_1, \dots, \eta_k$ a definição de η_{k+1} mostra que $\eta_{k+1} = \xi_{k+1}$. O processo de Gram-Schmidt não muda ξ_{k+1}.

Exercício 10.5.2 (Listas arbitrárias)**.** Seja $(\xi_1, \dots, \xi_\ell)$ uma lista arbitrária de ℓ vetores $\xi_i \in E$, dobros e o vetor nulo tudo permitido. Pode-se aplicar o processo de Gram-Schmidt com a seguinte modificação pequena da hipótese

(k) dado $k \geq 1$ suponha o conjunto $\{\eta_1, \dots, \eta_k\}$ é ortogonal e gera F_k, defina

$$\eta_{k+1} := \xi_{k+1} - \sum_{\substack{i=1 \\ \eta_i \neq \mathcal{O}}}^{k} \mathrm{pr}_{\eta_i} \xi_{k+1} = \xi_{k+1} - \sum_{\substack{i=1 \\ \eta_i \neq \mathcal{O}}}^{k} \langle \hat{\eta}_i, \xi_{k+1} \rangle \, \hat{\eta}_i$$

Obtém-se também uma lista $(\eta_1, \dots, \eta_\ell)$ cujos membros são dois-a-dois ortogonais, só agora é possível que uns são nulos. Com efeito, mostre que

$$\xi_{k+1} \in \langle \xi_1, \dots, \xi_k \rangle \overset{(k)}{=} \langle \eta_1, \dots, \eta_k \rangle \quad \Rightarrow \quad \eta_{k+1} = \mathcal{O}$$

[Dica: Note que $\langle \hat{\eta}_i, \xi_{k+1} \rangle$ é a i-ésima coordenada do vetor ξ_{k+1} na base ON composto daqueles $\hat{\eta}_i$ onde $\eta_i \neq \mathcal{O}$ é nao -nulos. Exercício 10.2.3.]

Exercício 10.5.3. Determine uma base ON do subespaço $F \subset \mathbb{R}^3$ gerado por

$$\xi_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \quad \xi_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \xi_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \langle \cdot, \cdot \rangle := \langle \cdot, \cdot \rangle_0$$

Solução com Gram-Schmidt (GS). Definição (10.5.1) dos η_{k+1} diz que

$$\eta_1 := \xi_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \quad |\eta_1|^2 := \langle \eta_1, \eta_1 \rangle = 1^2 + (-1)^2 + 1^2 = 3$$

$$\eta_2 := \xi_2 - \frac{\langle \eta_1, \xi_2 \rangle}{|\eta_1|^2} \eta_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{3} \underbrace{\left\langle \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\rangle}_{=1} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \tfrac{2}{3} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \; |\eta_2|^2 = \frac{8}{3}$$

e

$$\begin{aligned}\eta_3 := \xi_3 &\quad - \frac{\langle \eta_1, \xi_3\rangle}{|\eta_1|^2}\,\eta_1 \qquad\qquad - \frac{\langle \eta_2, \xi_3\rangle}{|\eta_2|^2}\,\eta_2 \\ &= \begin{bmatrix}1\\0\\1\end{bmatrix} - \frac{1}{3}\underbrace{\left\langle \begin{bmatrix}1\\-1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\right\rangle}_{=2} \begin{bmatrix}1\\-1\\1\end{bmatrix} - \frac{3}{8}\underbrace{\left\langle \tfrac{2}{3}\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\right\rangle}_{=\frac{2}{3}\cdot 2} \tfrac{2}{3}\begin{bmatrix}1\\2\\1\end{bmatrix} \\ &= \begin{bmatrix}1\\0\\1\end{bmatrix} - \tfrac{2}{3}\begin{bmatrix}1\\-1\\1\end{bmatrix} - \tfrac{1}{3}\begin{bmatrix}1\\2\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}\end{aligned}$$

Segundo GS e como $\eta_3 = \mathcal{O}$ sabemos que $F := \langle \xi_1, \xi_2, \xi_3\rangle = \langle \eta_1, \eta_2, \eta_3\rangle = \langle \eta_1, \eta_2\rangle$. GS diz que o conjunto $\{\eta_1, \eta_2\}$ é ortogonal, então LI segundo Teorema 10.2.2 usando que $\eta_2, \eta_2 \neq \mathcal{O}$. O conjunto ortogonal $\{\eta_1, \eta_2\}$ é uma base ortogonal de F porque é LI e gera F. Uma base ON é composto dos vetores

$$\varepsilon_1 := \tfrac{1}{|\eta_1|}\,\eta_1 = \tfrac{1}{\sqrt{3}}\begin{bmatrix}1\\-1\\1\end{bmatrix}, \quad \varepsilon_2 := \tfrac{1}{|\eta_2|}\,\eta_2 = \tfrac{\sqrt{3}}{\sqrt{8}}\tfrac{2}{3}\begin{bmatrix}1\\2\\1\end{bmatrix} = \tfrac{1}{\sqrt{6}}\begin{bmatrix}1\\2\\1\end{bmatrix} \tag{10.5.2}$$

10.5.1 Existência e extensão de bases ortogonais

Lembramos que um espaço vetorial de dimensão finita admite um produto interno – cada uma base ordenada $\mathcal{B}$ induz um, notação $\langle \cdot, \cdot\rangle_{\mathcal{B}}$, veja (10.1.4).

Teorema 10.5.4 (Existência)**.** *Um espaço vetorial E com produto interno e de dimensão finita n admite uma base ON $\mathcal{Z} = \{\varepsilon_1, \ldots, \varepsilon_n\}$.*

Demonstração. Pegue uma base ordenada $\mathcal{X} = (\xi_1, \ldots, \xi_n)$ de E e aplique o processo de ortonormalização de Gram-Schmidt. □

Proposição 10.5.5 (Extensão)**.** *Seja E um espaço vetorial com produto interno. Toda base ON $\mathcal{X}$ de um subespaço F estende-se a uma base ON de E.*

Demonstração. Segundo Teorema 3.2.1 (b) a base $\mathcal{X} = \{\xi_1, \ldots \xi_k\}$ de F é contida numa base ordenada $\mathcal{Y}$ de E, dizemos $\mathcal{Y} = (\xi_1, \ldots \xi_k, \xi_{k+1}, \ldots, \xi_n)$. Aplique Gram-Schmidt para obter $\mathcal{Z} = (\xi_1, \ldots \xi_k, \varepsilon_{k+1}, \ldots, \varepsilon_n)$. □

10.5.2 Projeção ortogonal sobre um subespaço

O processo de Gram-Schmidt prova a existência de bases ONs (pegue qualquer base e aplique o processo). É importante usar uma base ON nesta definição:

Definição 10.5.6 (Projeção ortogonal sobre um subespaço F)**.** Escolha uma base ordenada ON $\mathcal{Y} = (\varepsilon_1, \ldots, \varepsilon_k)$ de F e defina a transformação linear

$$\begin{aligned} \mathrm{pr}_F \colon E &\to E \\ v &\mapsto \sum_{i=1}^{k} \mathrm{pr}_{\varepsilon_i} v = \sum_{i=1}^{k} \langle \varepsilon_i, v \rangle \varepsilon_i \end{aligned} \tag{10.5.3}$$

Teorema 10.5.7 (Propriedades da projeção ortogonal)**.** *1.* pr_F *é linear*

2. $(\mathrm{pr}_F)^2 = \mathrm{pr}_F$

3. bem definido (independente da base ON $\mathcal{Y}$*)*

4. $\mathrm{pr}_F|_F = I_F \in \mathcal{L}(F)$

5. $\mathrm{Im}(\mathrm{pr}_F) = F$

6. $\omega := (v - \mathrm{pr}_F v) \perp f \quad \forall f \in F$

7. $\forall v \in E$ *vale*[6] $\mathrm{dist}(v, F) := \min_{f \in F} \underbrace{\mathrm{dist}(v, f)}_{=:|v-f|} = |v - \mathrm{pr}_F v|$

Demonstração. □

10.6 Complemento ortogonal

Definição 10.6.1. O **complemento ortogonal** de um subconjunto não-vazio $X \subset E$ é definido assim

$$X^\perp := \{v \in E \mid \langle v, f \rangle = 0 \quad \forall f \in F\}$$

Exercício 10.6.2. Seja $X \subset E$ um subconjunto não-vazio. Mostre que

(i) o complemento ortogonal $X^\perp$ é um subespaço de E

(ii) ou $X^\perp$ é disjunto a X, ou $X^\perp \cap X = \{\mathcal{O}\}$

(iii) $Y \subset X \Rightarrow X^\perp \subset Y^\perp$

(iv) $X^\perp = \langle X \rangle^\perp$

Proposição 10.6.3 (Relações entre F e $F^\perp$ e a projeção pr_F de (10.5.3))**.** *Para subespaços* F *de* E *vale o seguinte.*

(i) $F^\perp = \mathrm{N}(\mathrm{pr}_F)$

(ii) $F = \mathrm{Im}(\mathrm{pr}_F)$

(iii) $E = F \oplus F^\perp$ *e* $\dim E = \dim F + \dim F^\perp$

(iv) $\mathrm{pr}_F = P_{F,F^\perp}$ *veja (7.1.1),*

(v) $(F^\perp)^\perp = F$

[6] como $\dim F < \infty$ ínfimo igual mínimo

Demonstração. □

Exercício 10.6.4. No Exercício 10.5.3 temos calculado a base ON $\mathcal{Z} = \{\varepsilon_1, \varepsilon_2\}$, veja (10.5.2), do subespaço $F := \langle \xi_1, \xi_2, \xi_3 \rangle$. Determine uma base ON do complemento ortogonal

$$F^\perp := \{v \in E \mid \langle v, f\rangle = 0 \ \ \forall f \in F\} = \{v \in \mathbb{R}^3 \mid v \perp \varepsilon_1,\, v \perp \varepsilon_2\}$$

Vale a ultima igualdade porque a condição $\langle v, f\rangle = 0$ é linear em f, então é suficiente checar para os elementos f de uma base só.

Uma solução.

$$v \perp \varepsilon_1\colon \ \ 0 = \left\langle \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tfrac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\rangle = \tfrac{1}{\sqrt{3}}\,(x - y + z)$$

$$v \perp \varepsilon_2\colon \ \ 0 = \left\langle \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tfrac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\rangle = \tfrac{1}{\sqrt{6}}\,(x + 2y + z)$$

Multiplique identidade um por $\sqrt{3}$ e dois por $\sqrt{6}$ e forma a diferença das identidades resultantes para obter

$$0 - 0 = 0 - 3y - 0 \qquad \Rightarrow \qquad y = 0$$

Com isso obtemos da identidade um que

$$z = -x,\ x \in \mathbb{R} \text{ livre}, \qquad F^\perp = \mathbb{R} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

Uma base ON de $F^\perp$ é composto do vetor $\frac{1}{\sqrt{2}}(1, 0, -1)$.

10.7 Exercícios e umas soluções

Exercícios.

1. Prove que $\langle \cdot, \cdot \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ dado por

   $$((x_1, x_2), (y_1, y_2)) \mapsto 2x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2,$$

   define um produto interno em $\mathbb{R}^2$.

2. Seja E um espaço vetorial com produto interno $\langle \cdot, \cdot \rangle$. Prove que para todo $u, v \in E$:

   $$|u + v|^2 + |u - v|^2 = 2\,|u|^2 + 2\,|v|^2 \tag{10.7.1}$$

   onde $|\cdot| := \sqrt{\langle \cdot, \cdot \rangle}$ é a **norma induzida**. Interprete (10.7.1) geometricamente.

3. Considere os vetores $u = (2, -1, 2)$, $v = (1, 2, 1)$ e $w = (-2, 3, 3)$. Determine o vetor de $\mathbb{R}^3$ que é a projeção ortogonal de w sobre o plano gerado por u e v.

4. Considere a base $\mathcal{U} = (\xi_1, \xi_2, \xi_3\}$ de $\mathbb{R}^3$ onde

$$\xi_1 = (1, 1, 1), \qquad \xi_2 = (1, -1, 1), \qquad \xi_3 = (1, -1, -1).$$

   Aplique o método de Gram-Schmidt para obter uma base ortonormal $\mathcal{B} = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$. Determine a matriz $\mathbf{p}$ de passagem da base $\mathcal{U}$ para a base $\mathcal{B}$.

5. Determine as bases obtidas de $\mathcal{U} = (\xi_1, \xi_2, \xi_3)$ pelo processo de Gram-Schmidt nos casos seguintes:

   (a) $\xi_1 = (3, 0, 0), \quad \xi_2 = (-1, 3, 0), \quad \xi_3 = (2, -5, 1)$;
   (b) $\xi_1 = (-1, 1, 0), \quad \xi_2 = (5, 0, 0), \quad \xi_3 = (2, -2, 3)$.

6. Sejam $F_1, F_2 \subset E$ subespaços. Prove que

   (a) $(F_1 + F_2)^\perp = F_1^\perp \cap F_2^\perp$ (b) $F_1^\perp + F_2^\perp = (F_1 \cap F_2)^\perp$.

7. Prove que o produto vetorial $\cdot \times \cdot : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ definido no Exercício 5.4.5, satisfaz:

   (a) $u \times v = -v \times u$;
   (b) $u \times (v + \tilde{v}) = u \times v + u \times \tilde{v}$;
   (c) $u \times (\alpha v) = \alpha(u \times v)$, para todo $\alpha \in \mathbb{R}$;
   (d) $u \times v \neq 0 \iff \{u, v\}$ é um conjunto LI;
   (e) $u \times v$ é ortogonal a u e ortogonal a v;
   (f) $e_1 \times e_2 = e_3, \quad e_2 \times e_3 = e_1, \quad e_3 \times e_1 = e_2$.

Capítulo 11

A adjunta

Neste Capítulo 11 consideramos exclusivamente espaços vetoriais de dimensão finita com produtos internos

$$E = (E, \langle\cdot,\cdot\rangle_E)\,, \qquad F = (F, \langle\cdot,\cdot\rangle_F)$$

e dimensões $n := \dim E$ e $m := \dim F$. Dévido aos produtos internos o corpo sempre será $\mathbb{K} = \mathbb{R}$. Em vez de $\langle\cdot,\cdot\rangle_E$ ou $\langle\cdot,\cdot\rangle_F$ escrevemos simplesmente $\langle\cdot,\cdot\rangle$, o contexto indica do qual produto interno trata-se, aquele de E ou F.

11.1 Definição e propriedades

No teorema seguinte dimensão finita é essencial.

Teorema 11.1.1. *É um isomorfismo a transformação linear definida assim*

$$\begin{aligned} \chi = \chi_{\langle\cdot,\cdot\rangle} \colon E &\to E^* := \mathcal{L}(E, \mathbb{R}) \\ v &\mapsto \langle v, \cdot\rangle \end{aligned}$$

onde $\langle v, \cdot\rangle \colon E \to \mathbb{R}$ *é a transformação linear* $u \mapsto \langle v, u\rangle$.

Demonstração. Linear: Para $\alpha, \beta \in \mathbb{R}$ e $u, v \in E$ axioma (BL) da

$$\chi(\alpha u + \beta v) = \langle \alpha u + \beta v, \cdot\rangle = \alpha\langle u, \cdot\rangle + \beta\langle v, \cdot\rangle = \alpha\chi(u) + \alpha\chi(v)$$

Bijetivo: Segundo Corolário 5.2.6 as dimensões são iguais

$$\dim E^* = \dim \mathcal{L}(E, \mathbb{R}) = \dim E \cdot \dim \mathbb{R} = \dim E < \infty$$

Segundo Corolário 6.5.2 é suficiente mostrar injetivo: Suponha $\chi(v) = \mathcal{O} \in E^*$. Obtemos $\forall u \in E\colon\ \langle v, u\rangle = \chi(v)u = \mathcal{O}v = 0$. Então axioma (ND)′ em Lema 10.1.2 diz que $v = \mathcal{O} \in E$. Isso mostra que χ é injetivo, assim bijetivo. □

Definição 11.1.2 (Adjunta)**.** A **adjunta** de uma transformação linear $A\colon E \to F$ entre espaços vetoriais com produtos internos é num ponto w a composição

$$\begin{aligned} A^*\colon F &\to E \\ w &\mapsto \left(\chi_{\langle\cdot,\cdot\rangle_E}\right)^{-1} \langle w, A\cdot\rangle_F \end{aligned}$$

das transformações lineares $[v \mapsto \langle w, Av\rangle_F] \in E^*$ e $(\chi_{\langle\cdot,\cdot\rangle_E})^{-1}\colon E^* \to E$.

Proposição 11.1.3 (Critério para adjunta)**.** *Sejam $y \in E$ e $w \in F$, então*

$$y = A^*w \qquad \Leftrightarrow \qquad \langle y, v\rangle = \langle w, Av\rangle \ \forall v \in E$$

Demonstração. Dado $y \in E$ e $w \in F$, são equivalente

$$y = A^*w := (\chi_{\langle\cdot,\cdot\rangle_E})^{-1}\langle w, A\cdot\rangle_F \quad \Leftrightarrow \quad \langle w, A\cdot\rangle_F = \chi_{\langle\cdot,\cdot\rangle_E} y := \langle y, \cdot\rangle$$

□

Corolário 11.1.4. *Seja $A \in \mathcal{L}(E, F)$, então*

$$\langle A^*w, v\rangle = \langle w, Av\rangle \tag{11.1.1}$$

para cada um $w \in F$ e $v \in E$.

Demonstração. Proposição 11.1.3 "⇒". □

Teorema 11.1.5 (Regras básicas para a adjunta)**.**

(i) $I = I^*$
(ii) $(A + B)^* = A^* + B^*$
(iii) $(\alpha A)^* = \alpha A^*$
(iv) $(BA)^* = A^*B^*$
(v) $(A^*)^* = A$

Demonstração. Para cada um de (i-v) aplique (11.1.1) junto com Lema 10.1.3. Ilustramos o principio provando (iv) deixando os outros itens para o leitor. Vale $\langle (BA)^*w, v\rangle = \langle w, BAv\rangle = \langle B^*w, Av\rangle = \langle A^*B^*w, v\rangle$. □

Teorema 11.1.6 (Injetividade e sobrejetivade de A e A^*)**.**

(i) A *injetivo* $\Leftrightarrow$ A^* *sobrejetivo*
(ii) A *sobrejetivo* $\Leftrightarrow$ A^* *injetivo*
(iii) A *isomorfismo* $\Leftrightarrow$ A^* *isomorfismo*

Demonstração. (i) São equivalente

$$\begin{aligned} A \text{ injetivo} &\overset{1}{\Leftrightarrow} \text{existe inversa à esquerda } B \text{ de } A\colon\ BA = I_E \overset{*}{\Leftrightarrow} A^*B^* = I_E \\ &\overset{2}{\Leftrightarrow} \text{existe inversa à direita } C\ (= B^*) \text{ de } A^*\colon\ A^*C = I_E \\ &\overset{3}{\Leftrightarrow} A^* \text{ sobrejetivo} \end{aligned}$$

conforme 1 Teorema 6.3.6, 2 Teorema 11.1.5 (i,iv), e 3 Teorema 6.2.3.
(ii) Parte (i) diz que $B := A^*$ injetivo $\Leftrightarrow$ sobrejetividade de $B^* = (A^*)^* = A$.
(iii) Isomorfismo é linear e bijetivo (injetivo e sobrejetivo). Aplique (i) e (ii). □

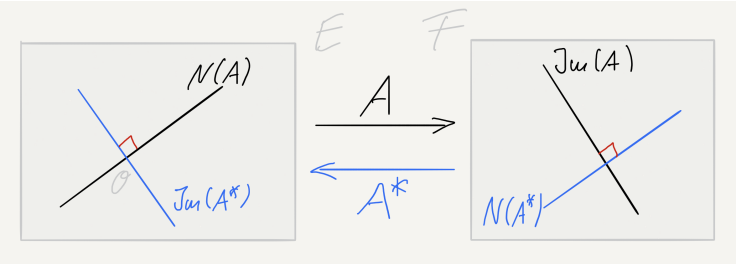


Figura 11.1: Operador A e sua adjunta A^* tem subespaços ortogonais

Adjunta e ortogonalidade

Teorema 11.1.7. *Seja $A^*\colon F \to E$ a adjunta de $A\colon E \to F$. Os dois subespaços naturais de E são complementos ortogonais, igualmente para F, ou seja*

$$\mathrm{N}(A) = \mathrm{Im}(A^*)^\perp, \qquad \mathrm{Im}(A) = \mathrm{N}(A^*)^\perp$$

Demonstração. $v \in \mathrm{N}(A) \Leftrightarrow Av = \mathcal{O} \Leftrightarrow \forall w \in F\colon 0 = \langle w, Av\rangle = \langle A^*w, v\rangle \Leftrightarrow 0 = \langle u, v\rangle\ \forall u \in \mathrm{Im}(A^*) \Leftrightarrow\colon v \in \mathrm{Im}(A^*)^\perp$. Analogamente para afirmação dois. □

Corolário 11.1.8. *São iguais* $\mathrm{posto}(A) = \mathrm{posto}(A^*)$.

Demonstração. Segundo Teorema 11.1.7 e Proposição 10.6.3 (iii) vale

$$\dim \mathrm{Im}(A^*) = \dim \mathrm{N}(A)^\perp = \dim E - \dim \mathrm{N}(A) = \dim \mathrm{Im}(A) \tag{11.1.2}$$

onde o ultimo passo é o Teorema 6.5.1 de núcleo e imagem. □

Proposição 11.1.9. *Seja $A \in \mathcal{L}(E)$ e seja $F \subset E$ um subespaço, então*

$$F \text{ subespaço invariante por } A \quad \Leftrightarrow \quad F^\perp \text{ subespaço invariante por } A^*$$

Demonstração. "⇒" Dado $g \in F^\perp$, a mostrar: $A^*g \in F^\perp$. Seja $f \in F$, então

$$\langle f, A^*g\rangle = \langle \underbrace{Af}_{\in F}, \underbrace{g}_{\in F^\perp}\rangle = 0$$

Como $f \in F$ foi arbitrário, segue que $A^*g \in F^\perp$.
"⇐" Aplique a parte já provada "⇒" para $G := F^\perp$ e $B := A^*$ usando que vale

$$G^\perp = (F^\perp)^\perp = F, \qquad B^* = (A^*)^* = A$$

segundo, respectivamente, Proposição 10.6.3 (v) e Teorema 11.1.5 (v). □

Lema 11.1.10. *Dado $A, B \in \mathcal{L}(E)$, então*

$$B^*A = \mathcal{O} \qquad \Rightarrow \qquad \forall v \in E\colon\ Av \perp Bv$$

Particularmente $A^*A = \mathcal{O} \Rightarrow A = \mathcal{O}$.

Demonstração. Seja $v \in E$. Vale $\langle Av, Bv\rangle = \langle B^*Av, v\rangle = \langle \mathcal{O}, v\rangle = 0$. Particularmente vale $\langle Av, Av\rangle = 0$. Assim $Av = \mathcal{O}$ segundo axioma (POS). Como $v \in E$ foi arbitrário o operador $A = \mathcal{O}$ é nulo. □

Matrizes

Teorema 11.1.11 (A matriz da adjunta é a matriz transposta)**.** *Seja* $\mathbf{a} = (a_{ij}) := [A]_{\mathcal{X},\mathcal{Y}}$ *a matriz de uma transformação linear* $A\colon E \to F$ *em respeito a bases ordenadas ortonormais* $\mathcal{X} = (\xi_1, \dots, \xi_n)$ *e* $\mathcal{Y} = (\eta_1, \dots, \eta_m)$*. Então*

(i) $\mathbf{a} = [A]_{\mathcal{X},\mathcal{Y}} \quad \Leftrightarrow \quad \mathbf{a}^t = [A^*]_{\mathcal{Y},\mathcal{X}}$

(ii) $a_{ij} = \langle \eta_i, A\xi_j \rangle$

Demonstração. (i) Seja $\mathbf{a} := [A]_{\mathcal{X},\mathcal{Y}}$ e $\mathbf{b} := [A^*]_{\mathcal{Y},\mathcal{X}}$. Segundo da definição das matrizes temos $A\xi_j = \sum_{\ell=1}^m \eta_\ell a_{\ell j}$ para $j = 1, \dots, n$ e $A^*\eta_i = \sum_{r=1}^n \xi_r b_{ri}$ para $i = 1, \dots, m$. Usando isso e axiomas (`BL`,`SIM`) obtemos

$$b_{ji} = \sum_{r=1}^n b_{ri} \underbrace{\delta_{jr}}_{\langle \xi_j, \xi_r \rangle} = \Big\langle \xi_j, \underbrace{\sum_{r=1}^n \xi_r b_{ri}}_{A^*\eta_i} \Big\rangle = \langle A^*\eta_i, \xi_j \rangle$$

$$\overset{4}{=} \langle \eta_i, \underbrace{A\xi_j}_{\sum_\ell \eta_\ell a_{\ell j}} \rangle = \sum_{\ell=1}^m a_{\ell j} \underbrace{\langle \eta_i, \eta_\ell \rangle}_{=\delta_{ij}} = a_{ij}$$

onde passo 4 é Proposição 11.1.3. Bases ON são essenciais. Já provamos (ii). □

O próximo resultado re-confirma o Teorema 4.2.2 dizendo que o posto de uma matriz é igual ao posto da matriz transposta.

Corolário 11.1.12. *Consideramos uma matriz real* $\mathbf{a} \in \mathrm{M}(m \times n)$ *como transformação linear* $\mathbb{R}^n \to \mathbb{R}^m$ *entre espaços cada um munido do produto euclidiano e da base canónica. Então a adjunta* $\mathbf{a}^* = \mathbf{a}^t$ *é a matriz transposta. Assim* $\mathrm{posto}(\mathbf{a}) = \mathrm{posto}(\mathbf{a}^t)$.

Demonstração. Teorema 11.1.11 (ii). □

Comentário 11.1.13. As regras básicas do Teorema 11.1.5 tomam para matrizes (visto como transformações lineares e usando $\mathbf{a}^* = \mathbf{a}^t$) a forma seguinte

$$\mathbb{1}^t = \mathbb{1}, \quad (\mathbf{a} + \mathbf{b})^t = \mathbf{a}^t + \mathbf{b}^t, \quad (\alpha\mathbf{a})^t = \alpha\mathbf{a}^t, \quad (\mathbf{ba})^t = \mathbf{a}^t\mathbf{b}^t, \quad (\mathbf{a}^t)^t = \mathbf{a}$$

Sim, estas regras prova-se mais rápido diretamente, tal-vez exceto $(\mathbf{ba})^t = \mathbf{a}^t\mathbf{b}^t$.

Comentário 11.1.14 (Injetividade e sobrejetivade de $\mathbf{a}$ e $\mathbf{a}^t$)**.** As afirmações do Teorema 11.1.6 tomam a forma seguinte para matrizes $\mathbf{a} \in \mathrm{M}(m \times n)$ – visto como transformações lineares e usando $\mathbf{a}^* = \mathbf{a}^t$.

(i) $\mathbf{a}$ injetivo $\Leftrightarrow$ $\mathbf{a}^t$ sobrejetivo

(ii) $\mathbf{a}$ sobrejetivo $\Leftrightarrow$ $\mathbf{a}^t$ injetivo

(iii) $\mathbf{a}$ isomorfismo $\Leftrightarrow$ $\mathbf{a}^t$ isomorfismo

Corolário 11.1.15. *Seja* $\mathbf{a} \in \mathrm{M}(m \times n)$ *e* $b \in \mathbb{R}^m$*, então*

$$\mathbf{a}x = b \text{ possui uma solução} \qquad \Leftrightarrow \qquad b \perp \mathrm{N}(\mathbf{a}^t)$$

Demonstração. Segundo Exemplo 6.0.18 são equivalente $\mathbf{a}x = b \Leftrightarrow b \in \mathrm{Im}(\mathbf{a})$, mas $\mathrm{Im}(\mathbf{a}) = \mathrm{N}(\mathbf{a}^t)^\perp$ segundo Teorema 11.1.7. □

11.2 Fórmula para inversa à direita/esquerda

Proposição 11.2.1 (Inversas à direita e esquerda)**.** *Seja $A \in \mathcal{L}(E,F)$.*

a) *A sobrejetivo $\Rightarrow$ $AA^* \in \mathcal{L}(F)$ é invertível e $A\,A^*(AA^*)^{-1} = I_F$*

b) *A injetivo $\Rightarrow$ $A^*A \in \mathcal{L}(E)$ é invertível e $(A^*A)^{-1}A^*A = I_E$*

Demonstração. a) Segundo Teorema 11.1.6 sobrejetividade de A significa injetividade de A^*. Isso implica[1] que $AA^*\colon F \to F$ é injetivo, assim segundo Corolário 6.5.2 (mesmas dimensões) bijetivo, então um isomorfismo.
b) Aplicando a) para $B := A^*$ usando $(A^*)^* = A$ segue que A^*A é invertível. □

De fato os postos de AA^* e de A^*A são igual ao posto de A como vamos ver no Corolário 12.4.10.

Lema 11.2.2. *Dado $A \in \mathcal{L}(E,F)$, as restrições*

$$A|\colon \operatorname{Im}(A^*) \xrightarrow{\simeq} \operatorname{Im}(A), \quad A^*|\colon \operatorname{Im}(A) \to \operatorname{Im}(A^*)$$

são isomorfismos (ainda que geralmente não são inversas um do outro).

Demonstração. É bem definido e injetivo como $\operatorname{Im}(A^*) = \operatorname{N}(A)^\perp$, então bijetivo como $\dim \operatorname{Im}(A) = \dim \operatorname{Im}(A^*)$ segundo (11.1.2). Analogamente para $A^*|$. □

Exemplo 11.2.3 (Não são inversas um do outro)**.**

$$A := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in \mathcal{L}(\mathbb{R}^2), \qquad A^* = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \in \mathcal{L}(\mathbb{R}^2)$$

São invertíveis como o determinante é não-nulo, assim sobrejetivo, ou seja $\operatorname{Im}(A) = \operatorname{Im}(A^*) = \mathbb{R}^2$, mas

$$A^*A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \neq \mathbb{1}$$

então $A^* \neq A^{-1}$.

11.3 Traço – produto interno em $\mathcal{L}(E,F)$

Exercício 11.3.1.

Considere o produto interno no espaço vetorial $M(n \times n)$ definido por

$$\langle \mathbf{a}, \mathbf{b} \rangle := \operatorname{tr}\left(\mathbf{a}^t\mathbf{b}\right) = \sum_{i,j} a_{ij}b_{ij}.$$

Mostre que o subespaço $\mathcal{A}$ das matrizes anti-simétricas é o complemento ortogonal em $M(n \times n)$ do subespaço $\mathcal{S}$ das matrizes simétricas:

$$\mathcal{A} = \mathcal{S}^\perp \quad \text{e} \quad \mathcal{S} \oplus \mathcal{A} = M(n \times n).$$

[1] Suponha $v \in \operatorname{N}(AA^*)$, ou seja $AA^*v = \mathcal{O}$, então $\operatorname{Im}(A^*) \ni A^*v \in \operatorname{N}(A) = \operatorname{Im}(A^*)^\perp$. Consequentemente $A^*v = \mathcal{O}$, então $v = \mathcal{O}$ como A^* é injetivo.

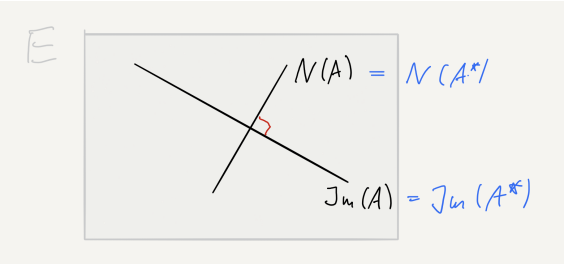


Figura 11.2: Operador normal A ($\Leftrightarrow$ A^* normal)

11.4 Operadores normais

Definição 11.4.1. Consideramos operadores lineares A em E as quais comutam com sua adjunta $AA^* = A^*A$. Tal A é chamado de **operador normal**.

Um operador A é normal se e somente sua adjunta A^* é normal, e neste caso cada um imagem Av e A^*v tem a mesma norma. Operadores normais tem a propriedade que A e a adjunta A^* tem os mesmos autovalores e autovetores associadas, o mesmo núcleo e a mesma imagem as quais são complementos um do outro como ilustrado na Figura 11.2.

Exercício 11.4.2. Seja $A \in \mathcal{L}(E)$ normal. Prove que

a) a adjunta A^* é normal também

b) $|Av| = |A^*v|$ para todos os vetores v de E

c) v autovetor de A com autovalor $\lambda \Leftrightarrow v$ autovetor de A^* com autovalor λ

d) $\mathrm{N}(A) = \mathrm{N}(A^*)$ e $\mathrm{Im}(A) = \mathrm{Im}(A^*)$ e estes são complementos ortogonais

11.5 Exercícios

Para todos os exercícios seja E um espaço vetorial de dimensão $n < \infty$, munido de um produto interno.

1. Determine uma inversa à direita para

$$A : \mathbb{R}^3 \to \mathbb{R}^2, \quad (x, y, z) \mapsto (x + 2y + 3z, 2x - y - z),$$

e uma inversa à esquerda para

$$B : \mathbb{R}^2 \to \mathbb{R}^4, \quad (x, y) \mapsto (x + 2y, 2x - y, x + 3y, 4x + y).$$

2. Dado

$$\mathbf{a} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix},$$

calcule $\mathbf{a}\mathbf{a}^t$ e, a partir daí, encontre uma matriz $\mathbf{b} \in M(3 \times 2)$ tal que $\mathbf{a}\mathbf{b} = \mathbb{1}_2$.

3. Seja P uma projeção em E ($P \in \mathcal{L}(E)$ e $P^2 = P$). Prove que a adjunta P^* também é uma projeção em E. Dê um exemplo em que $P^* \neq P$.

4. Uma matriz quadrada $\mathbf{a}$ chama-se *diagonalizável* quando é semelhante a uma matriz $\mathbf{d} = (d_{ij})$ do *tipo diagonal* ($d_{ij} = 0$ se $i \neq j$), ou seja, quando existe $\mathbf{p}$ invertível tal que $\mathbf{p}^{-1}\mathbf{a}\mathbf{p} = \mathbf{d}$. Prove que:

   (a) $\mathbf{a}$ diagonalizável $\Rightarrow \mathbf{a}^t$ diagonalizável.

   (b) Se a matriz do operador $A \in \mathcal{L}(E)$ relativamente a uma base de E é diagonalizável, então o é em relação a qualquer outra base.

5. Seja $A \in \mathcal{L}(E)$.

   (a) Seja $E = F_1 \oplus \cdots \oplus F_k$ e cada F_i é um subespaço invariante por A. Tome uma base ordenada $\mathcal{V}$ de E que seja uma união de bases das F_i. Determine a forma da matriz de A na base $\mathcal{V}$.

   (b) Se E possui uma base formada por autovetores de A, prove que existe também uma base de E formada por autovetores de $A^* : E \to E$. [Dica: (a)]

Capítulo 12

Operadores auto-adjuntos

Neste Capítulo 12 consideramos operadores $A \in \mathcal{L}(E)$ num espaço vetorial real E de dimensão finita n e com produto interno, ou seja

$$A\colon E \to E, \qquad E = (E, \langle \cdot, \cdot \rangle)$$

Lembramos que dévido ao produto interno o corpo será $\mathbb{K} = \mathbb{R}$.

Definição 12.0.1 (Auto-adjunto)**.** Chama-se um operador $A \in \mathcal{L}(E)$ de **auto-adjunto** se ele iguale a sua adjunta $A^* = A$.

Observação 12.0.2. a) O operador nulo $\mathcal{O} \in \mathcal{L}(E)$ é auto-adjunto $\mathcal{O}^* = \mathcal{O}$.
O operador identidade $I \in \mathcal{L}(E)$ é auto-adjunto $I^* = I$.
b) Operadores auto-adjuntos são operadores normais.
c) Para operadores auto-adjuntos núcleo e imagem são complementos ortogonais $\mathrm{N}(A) = \mathrm{Im}(A)^\perp$ segundo Teorema 11.1.7.

Os operadores auto-adjuntos formam um subespaço de $\mathcal{L}(E)$.

Lema 12.0.3. *Para operadores auto-adjuntos $A, B \in \mathcal{L}(E)$ vale o seguinte.*

(i) *Somas $A+B = (A+B)^*$ e múltiplos reais $\alpha A = (\alpha A)^*$ são auto-adjuntos.*

(ii) *A composição BA, igualmente AB, é auto-adjunta se e somente se A e B comutam ($BA = AB$).*

Demonstração. Teorema 11.1.5. □

Lema 12.0.4 (Restrição preserva auto-adjunto)**.** *A restrição de um operador auto-adjunto $A \in \mathcal{L}(E)$ a um subespaço F invariante por A é auto-adjunta.*

Demonstração. Para todos os $f, \tilde{f} \in F$ vale

$$\langle f, (A|_F)^* \tilde{f} \rangle = \langle (A|_F) f, \tilde{f} \rangle = \langle Af, \tilde{f} \rangle \overset{A^* = A}{=} \langle f, A\tilde{f} \rangle = \langle f, (A|_F)\tilde{f} \rangle$$

então $(A|_F)^* = (A|_F)$ segundo Lema 10.1.3. □

12.1 Auto-adjunto e ortogonalidade

Lema 12.1.1 (As projeções auto-adjuntas são as projeções ortogonais)**.** *Dado Um par de subespaços complementares $F \oplus G = E$, então são equivalente*

$$P := P_{F,G} \in \mathcal{L}(E) \text{ auto-adjunto} \qquad \Leftrightarrow \qquad F \perp G$$

Demonstração. "$\Rightarrow$" Dado $f \in F$, $g \in G$, como $F = \operatorname{Fix} P_{F,G}$ e $G = \mathrm{N}(P_{F,G})$ segundo (7.1.2) obtemos

$$\langle f, g\rangle = \langle Pf, g\rangle \overset{P=P^*}{=} \langle f, Pg\rangle = 0$$

"$\Leftarrow$" Dado $u, \tilde{u} \in E$. Como $E = F \oplus G$ escrevemos $u = f + g$ e $\tilde{u} = \tilde{f} + \tilde{g}$ para únicos $f, \tilde{f} \in F$ e $g, \tilde{g} \in G$ (Teorema 2.3.4). Como $F = \operatorname{Fix} P_{F,G}$ e $G = \mathrm{N}(P_{F,G})$

$$\langle u, P^*\tilde{u}\rangle \overset{(11.1.1)}{=} \langle P(f+g), \tilde{u}\rangle \overset{g\in\mathrm{N}(P)}{=} \langle \overbrace{Pf}^{=f}, \tilde{f} + \tilde{g}\rangle$$

$$\overset{\tilde{g}\perp f}{=} \langle f, \tilde{f}\rangle \overset{\tilde{f}\perp g}{=} \langle f + g, P\tilde{f}\rangle \overset{\tilde{g}\in\mathrm{N}(P)}{=} \langle u, P\tilde{u}\rangle$$

Então $P^* = P$ segundo Lema 10.1.3. □

Proposição 12.1.2. *Seja $A \in \mathcal{L}(E)$ auto-adjunto e $F \subset E$ subespaço, então*

$$F \text{ subespaço invariante por } A \quad \Leftrightarrow \quad F^\perp \text{ subespaço invariante por } A$$

Demonstração. Proposição 11.1.9. □

Teorema 12.1.3 (Ortogonalidade de autovetores de autovalores diferentes)**.** *Autovetores ξ_λ e ξ_μ associados a autovalores diferentes $\lambda \neq \mu$ de um operador auto-adjunto $A = A^*$ em E são ortogonais, em símbolos $\xi_\lambda \perp \xi_\mu$.*

Demonstração. Vale que o produto se anula

$$(\lambda - \mu)\langle \xi_\lambda, \xi_\mu\rangle = \langle \lambda\xi_\lambda, \xi_\mu\rangle - \langle \xi_\lambda, \mu\xi_\mu\rangle = \langle A\xi_\lambda, \xi_\mu\rangle - \langle \xi_\lambda, A\xi_\mu\rangle \overset{A^*=A}{=} 0$$

e assim, como $\lambda - \mu) \neq 0$, segue que $\langle \xi_\lambda, \xi_\mu\rangle = 0$. □

12.2 Matrizes simétricas

Como podemos checar se um operador $A \in \mathcal{L}(E)$ é auto-adjunto? Só calcular a matriz dele em respeito a uma base ON e ver se é simétrica.

Teorema 12.2.1. *Seja $\mathcal{X} = (\varepsilon_1, \ldots, \varepsilon_n)$ uma base ON de E. Então os operadores auto-adjuntos correspondem exatamente ás matrizes simétricas $n \times n$. Com efeito, a aplicação entre espaços vetoriais*

$$\Psi = \Psi_{\mathcal{X}} \colon \{\text{operadores auto-adjuntos em } E\} \to \mathcal{S}(n)$$
$$A \mapsto \mathbf{a} := [A]_{\mathcal{X}}$$

é linear e bijetivo (um isomorfismo).

Demonstração. Bem definido: Com efeito é simétrica a matriz $\mathbf{a} := [A]_{\mathcal{X}} = [A^*]_{\mathcal{X}} = \mathbf{a}^t$ onde a última igualdade é Teorema 11.1.11.
Linear e injetivo: Teorema 5.2.5.
Sobrejetivo: Dado uma matriz $n \times n$ simétrica $\mathbf{a} = (a_{ij})$, define $A \in \mathcal{L}(E)$ nos membros da base ON $\mathcal{X}$ assim $A\varepsilon_j := \varepsilon_1 a_{1j} + \cdots + \varepsilon_n a_{nj}$ e estende linearmente a E. Então A é auto-adjunto porque $[A]_{\mathcal{X}} = \mathbf{a} = \mathbf{a}^t = [A^*]_{\mathcal{X}}$ onde a última igualdade é Teorema 11.1.11. Mas se as matrizes de dois operadores são iguais os operadores são iguais porque eles tomam os mesmos valores numa base. □

Corolário 12.2.2. *Os operadores auto-adjuntos formam um subespaço de dimensão $n(n+1)/2$ do espaço vetorial $\mathcal{L}(E)$ onde $n = \dim E$.*

Demonstração. Segundo Corolário 6.4.9 isomorfismos, assim aquele em Teorema 12.2.1, preservam dimensão e $\dim \mathcal{S}(n) = n(n+1)/2$ segundo (3.2.2). □

Comentário 12.2.3. Seja $E = F \oplus G$. Pode-se usar Teorema 12.2.1, simetria da matriz, para provar Lema 12.1.1 o que diz que

$$P := P_{F,G} \in \mathcal{L}(E) \text{ auto-adjunto} \qquad \Leftrightarrow \qquad F \perp G$$

Sejam $\{\varepsilon_1, \ldots, \varepsilon_k\}$ e $\{\varepsilon_{k+1}, \ldots, \varepsilon_n\}$ bases ONs de F e G, respectivamente. Então $\mathcal{X} := \{\varepsilon_1, \ldots, \varepsilon_n\}$ é uma base ON de E se e somente se $F \perp G$. Neste caso a matriz de P tem a forma simétrica

$$\mathbf{p} := [P]_{\mathcal{X}} = \begin{bmatrix} \mathbb{1}_k & \mathcal{O} \\ \mathcal{O} & \mathcal{O}_{n-k} \end{bmatrix} = \mathbf{p}^t$$

o que é equivalente a P sendo auto-adjunto (Teorema 12.2.1).

Exemplo 12.2.4. Sejam dois operadores $A, B \in \mathcal{L}(\mathbb{R}^2)$ dados por

$$A(x, y) = (x, 2y), \qquad B(x, y) = (y, x)$$

Determine quais dos quatro A, B, AB, BA são auto-adjuntos.

Uma solução. Obviamente escolhemos a base canónica $\mathcal{E} = \{e_1, e_2\}$ como base ON. Auto-adjunto é equivalente a simetria da matriz.

(i) $\mathbf{a} := [A] = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}^t = \mathbf{a}^t$ e assim $A = A^*$ é auto-adjunto

(ii) $\mathbf{b} := [B] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \mathbf{b}^t$ e assim $B = B^*$ é auto-adjunto

(iii) $[AB] = [A]\,[B] = \mathbf{ab} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix} =: \mathbf{c}$, então como $\mathbf{c} \neq \mathbf{c}^t$ sabemos que AB não é auto-adjunto

(iv) $[BA] = [B]\,[A] = \mathbf{ba} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} =: \mathbf{d}$, então como $\mathbf{d} \neq \mathbf{d}^t$ sabemos que BA não é auto-adjunto

Outra solução (iii-iv). Não-comutatividade $\mathbf{ab} = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} = \mathbf{ba}$ é equivalente a não-comutatividade $AB \neq BA$ o que é equivalente, segundo Lema 12.0.3, a AB e BA ambos não são auto-adjuntos.

Exercício 12.2.5. Considere a base ordenada $\mathcal{B} = (\xi_1, \xi_2) := ((1,-1),(3,1))$ de $\mathbb{R}^2$. Determine se o operador $A \in \mathcal{L}(\mathbb{R}^2)$ cuja matriz $[A]_{\mathcal{B}}$ é dada por

$$\begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix} = [A]_{\mathcal{B}} =: \mathbf{a}$$

é auto-adjunto, ou não.

Uma solução. A matriz $\mathbf{a}$ é simétrica, mas a base $\mathcal{B}$ não é ON. Precisamos calcular a matriz de A em respeito a uma base ON, logicamente vamos escolher a base mais simples, a base canónica $\mathcal{E}$. Para determinar $[A]_{\mathcal{E}}$ começamos assim

$$\begin{aligned} Ae_1 - Ae_2 = A(e_1 - e_2) = A(1,-1) = \xi_1 \cdot 1 + \xi_2 \cdot 0 = \xi_1 = (1,-1) = e_1 - e_2 \\ 3Ae_1 + Ae_2 = A(3e_1 + e_2) = A(3,1) = \xi_1 \cdot 0 + \xi_2 \cdot 5 = 5(3,1) = 15e_1 + 5e_2 \end{aligned}$$

Adicionamos as identidades para obtermos

$$4Ae_1 + \mathcal{O} = 16e_1 + 4e_2 \qquad \Rightarrow \qquad Ae_1 = e_1 \cdot 4 + e_2 \cdot 1$$

e consequentemente a primeira coluna da matriz

$$\mathbf{b} := [A]_{\mathcal{E}} = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$$

A segunda coluna segue da primeira identidade $Ae_2 = Ae_1 - e_1 + e_2 = e_1 \cdot 3 + e_2 \cdot 2$. Como a matriz $\mathbf{b} \neq \mathbf{b}^t$ não é simétrica, o operador A não é auto-adjunto.

12.3 Teorema espectral

Proposição 12.3.1 (Caso $\dim E = 2$)**.** *Na dimensão dois para todo operador auto-adjunto A existe uma base ON $\mathcal{X} = (\hat{\xi}_1, \hat{\xi}_2)$ composto de autovetores.*

Demonstração. Escolha uma base ON $\mathcal{Y}$ de E. A matriz correspondente

$$\mathbf{a} := [A]_{\mathcal{Y}} = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}$$

é simétrica porque $A = A^*$ é auto-adjunto. O polinômio característico é

$$p_A(\lambda) := p_{[A]_{\mathcal{Y}}}(\lambda) := \lambda^2 - (\alpha + \gamma)\lambda + (\alpha\gamma - \beta^2)$$

e suas raízes são segundo a fórmula (9.2.1) dadas por

$$\lambda_{\pm} = \frac{(\alpha + \gamma) \pm \sqrt{\Delta}}{2}, \qquad \Delta = (\alpha + \gamma)^2 - 4(\alpha\gamma - \beta^2)$$

como $\Delta = (\alpha - \gamma)^2 + 4\beta^2 \geq 0$ é realmente não-negativo.
Caso $\Delta = 0$. Então $\alpha = \gamma$ e $\beta = 0$, e daí $\mathbf{a} = \alpha \mathbb{1}$ e $A = \alpha I_E$. Consequentemente α é o único autovalor e qualquer base ON é composto de autovetores.
Caso $\Delta > 0$. Neste caso $\lambda_- < \lambda_+$ são autovalores diferentes de $[A]_{\mathcal{Y}}$, assim de A. Autovetores correspondentes são ortogonais segundo Teorema 12.1.3. □

Proposição 12.3.2 (Existência de um autovetor)**.** *Todo operador auto-adjunto $A\colon E \to E$ admite um autovetor v.*

Demonstração. Segundo Teorema B.5.1 na dimensão finita todo operador linear admite um subespaço invariante $F \subset E$ de dimensão 1 ou 2. (Isso é trabalhoso.)
Caso dim $F = 1$. Segundo Lema 9.0.4 existe um elemento $v \in F$ e um escalar $\lambda \in \mathbb{R}$ tal que $Av = \lambda v$. Não temos usado que A é auto-adjunto, vamos próximo:
Caso dim $F = 2$. Como $A\colon E \to E$ é auto-adjunto a restrição $A|\colon F \to F$ também é auto-adjunta segundo Lema 12.0.4. Segundo Proposição 12.3.1 existe uma base ON de F composto de dois autovetores de $A|$, assim de A. □

Teorema 12.3.3 (Teorema espectral)**.** *Seja $n = dim E$. Um operador $A \in \mathcal{L}(E)$ é auto-adjunto $A^* = A$ se e somente se existe uma base ON $\mathcal{X} = (\hat{\xi}_1, \dots, \hat{\xi}_n)$ de E composto de autovetores ξ_j de A.*

Demonstração. "⇒" 1) Segundo a prova da Proposição 12.3.2 existe um subespaço $F \subset E$ invariante por A e de dimensão 1 ou 2 e além disso F chega com uma base ON composto de autovetores de A.
2) O complemento ortogonal $F^\perp$ é invariante por A segundo Proposição 12.1.2. Assim a restrição $A|\colon F^\perp \to F^\perp$ existe e, segundo Lema 12.0.4, é auto-adjunto.

Agora repetimos 1) e depois 2). Cada vez a dimensão é reduzida por 1 ou 2. Por isso o processo termina em não mais como $n = \dim E$ iterações.

"⇐" Sejam ξ_i e ξ_j membros da base ON $\mathcal{X}$ com autovalores λ_i e λ_j, então

$$\langle \xi_i, A^*\xi_j \rangle = \langle A\xi_i, \xi_j \rangle = \langle \lambda_i \xi_i, \xi_j \rangle = \lambda_i \delta_{ij} = \lambda_j \delta_{ij} = \langle \xi_i, \lambda_j \xi_j \rangle = \langle \xi_i, A\xi_j \rangle$$

Então os dois operadores A^* e A são iguais segundo Lema 10.1.3. □

Exercício 12.3.4. Seja $A = A^* \in \mathcal{L}(E)$ auto-adjunto e $n = \dim E$. Mostre que

a) se $\mathcal{X} = (\xi_1, \dots, \xi_n)$ é uma base de E composto de autovetores de A, então a matriz correspondente $[A]_{\mathcal{X}} = \operatorname{diag}[\lambda_1, \dots, \lambda_n]$ é diagonal onde $\lambda_1, \dots, \lambda_n$ são os autovalores correspondentes (pode ter dobros).

b) autovalores diferentes $\lambda \neq \mu$ tem autosubespaços ortogonais $E_\lambda \perp E_\mu$

b) $E = \oplus_{\lambda \in \operatorname{spec} A} E_\lambda$ $\qquad E_\lambda := \mathrm{N}(\lambda I_E - A)$

c) $\forall \lambda \in \operatorname{spec} A\colon \ \operatorname{alg}_\lambda(A) = g_\lambda(A)$ $\qquad g_\lambda(A) := \dim E_\lambda$

12.4 Operadores não-negativos

Definição 12.4.1. Um operador auto-adjunto $A = A^* \in \mathcal{L}(E)$ é chamado de **operador não-negativo**, símbolo $\boldsymbol{A \geq 0}$, se $\langle Av, v\rangle \geq 0$ para todo vetor $v \in E$. No caso $\langle Av, v\rangle > 0 \; \forall v \neq \mathcal{O}$ chama-se A de **operador positivo**, símbolo $\boldsymbol{A > 0}$.

Lembre que o complemento ortogonal $v^\perp$ de um vetor não-nulo é um hiperplano de E e assim decompõe E em dois semi-espaços cuja interseção é $v^\perp$. Uma interpretação geométrica do que um operador $A = A^*$ é não-negativo seria que cada um vetor imagem Av aponta no mesmo semi-espaço como o vetor v. Vamos ver que $Av \in v^\perp$ só é possível no caso do vetor nulo $Av = \mathcal{O}$.

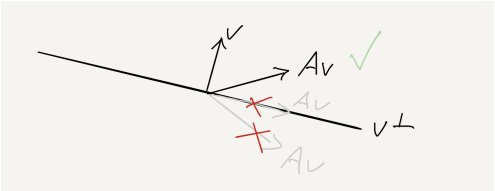


Figura 12.1: Operador não-negativo $A \geq 0$ não inverta direção "ao passado"

Exemplo 12.4.2 (Quadrado de operadores auto-adjuntos)**.** $A = A^* \Rightarrow A^2 \geq 0$

Teorema 12.4.3. *Seja $A = A^* \in \mathcal{L}(E)$ auto-adjunto. Não-negatividade de A é equivalente a não-negatividade de todos os autovalores, em símbolos*

$$A \geq 0 \quad \Leftrightarrow \quad \operatorname{spec} A \subset [0, \infty)$$

Analogamente $A > 0 \;\Leftrightarrow\; \operatorname{spec} A \subset (0, \infty)$.

Demonstração. "⇒" Seja $Av = \lambda v$ onde $v \neq \mathcal{O}$, então $\lambda\langle v, v\rangle = \langle Av, v\rangle \geq 0$. "⇐" Seja $\mathcal{X} = \{\xi_1, \ldots, \xi_n\}$ uma base ON de E composto de autovetores, ou seja $A\xi_i = \lambda_i \xi_i$ onde $\lambda_i \geq 0$. Nesta base escreve $v \in E$ como $v = \sum_{i=1}^n \alpha_i \xi_i$, usando $\langle \xi_i, \xi_j\rangle = \delta_{ij}$ obtemos $\langle Av, v\rangle = \langle \sum_{i=1}^n \alpha_i A\xi_i, \sum_{j=1}^n \alpha_j \xi_j\rangle = \sum_{i=1}^n \alpha_i^2 \lambda_i \geq 0$. □

Corolário 12.4.4. *Seja $A \geq 0$ e $v \in E$, então*

$$\langle Av, v\rangle = 0 \quad \Rightarrow \quad Av = \mathcal{O}$$

(*Escrito em outros símbolos para $A \geq 0$:* $Av \perp v \;\Rightarrow\; v \in \mathrm{N}(A)$.)

Demonstração. "⇒" Como na prova do teorema anterior seja $\mathcal{X} = \{\xi_1, \ldots, \xi_n\}$ uma base ON de E composto de autovetores de A. Suponhamos que os primeiros k formam uma base $\mathcal{X}_0 = \{\xi_1, \ldots, \xi_k\}$ do núcleo $\mathrm{N}(A) = E_0$; possivelmente $\mathcal{X}_0 = \emptyset$. Escrevemos $v = \sum_{i=1}^n \alpha_i \xi_i$. Entao a identidade

$$0 = \langle Av, v\rangle = \sum_{i=k+1}^{n} \alpha_i^2 \underbrace{\lambda_i}_{>0}$$

implica que $\alpha_{k+1} = \cdots = \alpha_n = 0$. Assim

$$Av = \alpha_1 \underbrace{\lambda_1}_{0} \xi_1 + \cdots + \alpha_k \underbrace{\lambda_k}_{0} \xi_k + \underbrace{\alpha_{k+1}}_{0} \lambda_{k+1}\xi_{k+1} + \cdots + \underbrace{\alpha_n}_{0} \lambda_n \xi_n = \mathcal{O}$$

□

Corolário 12.4.5. *$A > 0 \Leftrightarrow A \geq 0$ e A é invertível*

Demonstração. "⇒" Logicamente $A > 0$ implica $A \geq 0$. Segundo Teorema 12.4.3 os autovalores de A são > 0, assim $\mathrm{N}(A) = E_0 = \{\mathcal{O}\}$. Consequentemente A é injetivo, equivalentemente sobrejetivo, então um isomorfismo, assim invertível segundo Proposição 6.4.4.
"⇐" Seja $A \geq 0$ invertível, particularmente $\mathrm{N}(A) = \{\mathcal{O}\}$. Dado $v \in E$ não-nulo, então $Av \neq \mathcal{O}$ e dai $\langle Av, v\rangle \neq 0$ conforme Corolario 12.4.4. De outro lado $\langle Av, v\rangle \geq 0$ segundo à hipótese $A \geq 0$. □

Teorema 12.4.6 (Raíz quadrada não-negativa / positiva)**.** *Todo operador não-negativo admite uma única raíz quadrada não-negativa: Dado $A \geq 0$, então existe um único $B \geq 0$, chamado de **a raíz quadrada não-negativa** de A, tal que $B^2 = A$. Vale $B > 0 \Leftrightarrow A > 0$ e B é chamado de raíz quadrada **positiva**.*

Notações comuns para a raíz quadrada não-negativa são $\sqrt{A}$ ou $A^{\frac{1}{2}}$.

Demonstração. Seja $A \geq 0$. Existência. Sabemos que A tem autovalores $\lambda_i \geq 0$ e os autosubespaços E_{λ_i} são dois-a-dois ortogonais. Com efeito

$$\operatorname{spec} A = \{\lambda_1, \ldots, \lambda_r\} \subset [0, \infty)$$

onde $r \leq n := \dim E$ e

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$$

veja os Teoremas 12.4.3 e 12.3.3 e Exercício 12.3.4. Escrevemos $v \in E$ na forma $v = v_1 + \cdots + v_r$ para autovetores únicos $v_i \in E_{\lambda_i}$, veja Teorema 2.3.4. Assim $Av = \lambda_1 v_1 + \cdots + \lambda_r v_r$. Definimos o candidato para a raíz quadrada assim

$$B\colon E \to E, \quad v \mapsto \sqrt{\lambda_1} v_1 + \cdots + \sqrt{\lambda_r} v_r$$

Obtemos imediatamente que $B^2 = A$, com efeito

$$B^2 v = B(\sqrt{\lambda_1} v_1 + \cdots + \sqrt{\lambda_r} v_r) = \sqrt{\lambda_1}^2 v_1 + \cdots + \sqrt{\lambda_r}^2 v_r = Av$$

Escrevendo $w \in E$ analogamente como $w = w_1 + \cdots + w_r$ seque que $B = B^*$:

$$\langle Bv, w\rangle = \sum_{i=1}^{r} \left\langle \sqrt{\lambda_j} v_j, w_j \right\rangle = \sum_{i=1}^{r} \left\langle v_j, \sqrt{\lambda_j} w_j \right\rangle = \langle v, Bw\rangle$$

onde temos usado ortogonalidade $E_{\lambda_i} \perp E_{\lambda_j}$ na igualdade 1 e 3. Similarmente

$$\langle Bv, v\rangle = \sum_{i=1}^{r} \left\langle v_j, \sqrt{\lambda_j} v_j \right\rangle = \sum_{i=1}^{r} \sqrt{\lambda_j}\, |v_j|^2 \geq 0$$

o que conclui a prova de $B \geq 0$.

UNICIDADE. Suponha que um operador $C \geq 0$ satisfaz $C^2 = A$.
1) Os autosubespaços E_{λ_i} de A são invariante por C: Dado $v \in E_{\lambda_i}$, segue $Cv \in E_{\lambda_i}$ da comutatividade $AC = CA$ a qual vale segundo a hipótese $C^2 = A$.
2) A restrição $C_i := C|_{E_{\lambda_i}}$ iguale $\sqrt{\lambda_i}\, I$: A restrição existe segundo 1) e é auto-adjunto segundo Lema 12.0.4 usando a hipótese $C = C^*$. Conforme o teorema espectral existe uma base ON do domínio E_{λ_i} de C_i composto de autovetores de C_i. Assim resta mostrar que C_i admite só um autovalor e ele é $\sqrt{\lambda_i}$. Existência de um autovalor é garantido por Proposição 12.3.2. Suponha então que $C_i\xi = \mu\xi$ onde $\xi \in E_{\lambda_i}$ e $\mu \in \mathbb{R}$. A hipótese $C \geq 0$ implica que $\mu \geq 0$. Como

$$\lambda_i \xi = A\xi = C(C\xi) = C(\mu\xi) = \mu(C\xi) = \mu^2 \xi$$

e como $\xi \neq \mathcal{O}$ segue que $\mu = \sqrt{\lambda_i}$.
3) Escrevendo $v \in E$ mais uma vez na forma $v = v_1 + \cdots + v_r$ obtemos que

$$Cv = Cv_1 + \cdots + Cv_r \overset{2)}{=} \sqrt{\lambda_1} v_1 + \cdots + \sqrt{\lambda_r} v_r \overset{\text{def.}}{=} Bv$$

Prova-se analogamente o caso $A > 0$. □

Comentário 12.4.7.

a) $(R_{90^\circ})^2 = R_{90^\circ} R_{90^\circ} = R_{180^\circ} = -\mathbb{1}_{\mathbb{R}^2} < 0$ (não todo quadrado B^2 é ≥ 0)
(consequentemente R_{90° não é auto-adjunto)

b) vale $\mathbf{cc} = \mathbb{1}_2 \geq 0$ (existência de outras raízes $\mathbf{c} \not\geq 0$)
- para a matriz $\mathbf{c} = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}$ ($\mathbf{c} \not\geq 0$ como falta auto-adjunto)
- para a matriz $\mathbf{c} = -\mathbb{1}_2$ ($\mathbf{c} \not\geq 0$ como $\langle \mathbf{c}e_1, e_1 \rangle = -1$)
- para a matriz $\mathbf{c} = \mathbb{1}_2$ (a raíz quadrada única)

Lema 12.4.8 (Diagonalização simultânea)**.** *Dado auto-adjuntos $A, B \in \mathcal{L}(E)$:*

$$AB = BA \quad \Leftrightarrow \quad \exists \text{ base ON composto de autovetores comuns a } A \text{ e } B$$

Teorema 12.4.9. *Sejam E e F espaços vetoriais de dimensões finitas e munido de produtos internos. Então $A^*A \in \mathcal{L}(E)$ e $AA^* \in \mathcal{L}(F)$ são operadores não-negativos (positivos $\Leftrightarrow$ A invertível) e* $\text{posto}(A^*A) = \text{posto}(AA^*) = \text{posto}(A)$.

Corolário 12.4.10. *Seja $A \in \mathcal{L}(E, F)$.*

a) *A injetivo $\Leftrightarrow$ A^*A invertível*

b) *A sobrejetivo $\Leftrightarrow$ AA^* invertível*

Exercício 12.4.11. Dado $\mathbf{a} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} : \mathbb{R}^3 \to \mathbb{R}^2$, mostre que $\mathbf{a}\mathbf{a}^t > 0$ e $\mathbf{a}^t\mathbf{a} \geq 0$.

12.5 Teorema dos valores singulares

No seguinte trata-se de uma extensão do teorema espectral a operadores lineares gerais, auto-adjunto ou não.

Teorema 12.5.1 (Teorema dos valores singulares)**.** *Sejam E e F espaços vetoriais de dimensões finitas e munido de produtos internos. Seja $A \in \mathcal{L}(E,F)$ e seja $r = \text{posto}(A)$. Então vale o seguinte:*
*Existe uma base ON $\mathcal{X} = \{\xi_1, \ldots, \xi_n\}$ de E e uma base ON $\mathcal{Y} = \{\eta_1, \ldots, \eta_n\}$ de F e os chamados **valores singulares** $\sigma_1, \ldots, \sigma_r \in (0, \infty)$ de A, tal que*

$$\begin{array}{lll} A\xi_i = \sigma_i \eta_i & A^*\eta_i = \sigma_i \xi_i & i = 1, \ldots, r \\ A\xi_j = \mathcal{O} & A^*\eta_j = \mathcal{O} & i = 1, \ldots, r \end{array}$$

12.6 Exercícios

Para todos os exercícios E é um espaço vetorial de dimensão $n < \infty$, munido de produto interno.

1. Sejam $A, B \in \mathcal{L}(E)$ auto-adjuntos tais que $\langle Av, v\rangle = \langle Bv, v\rangle$, para todo $v \in E$. Prove que $A = B$.

2. Determine $A = A^* \in \mathcal{L}(\mathbb{R}^3)$ tal que

   $$A(2,-1,-2) = (1,1,13) \qquad \text{e} \qquad A(3,-6,-6) = (3,21,33),$$

   sabendo que o traço de A é 5.

3. Seja $A \in \mathcal{L}(\mathbb{R}^3)$ tal que $A(4,4,-2) = (10,-2,-2)$ e

   $$A(4,-2,4) = (-2,10,-2), \quad A(1,-2,-2) = (1,1,-5).$$

   Prove que $A^* = A$.

4. Sejam $A, B \in \mathcal{L}(E)$ auto-adjuntos. Prove que $AB + BA$ é auto-adjunto. Que se pode dizer sobre $AB - BA$?

5. Sejam $A, B \in \mathcal{L}(E)$ auto-adjuntos tais que BA é diagonalizável.[1] Prove que AB também é diagonalizável.

   [Dica: Veja §11.5 Exercício 5 (b).]

6. Seja $P \in \mathcal{L}(E)$ uma projeção ortogonal. Encontre uma raíz quadrada não-negativa de P. É única?

   [Dica: Lembre-se que $2P = I + S$, onde $S \in \mathcal{L}(E)$ é a reflexão ortogonal em torno de $F := \text{Im}(P)$.]

[1] Uma transformação linear $A \in \mathcal{L}(E)$ é chamada **diagonalizável** se, e somente se, existe uma base $\mathcal{U}$ de E tal que a matriz $\mathbf{a} = [A]_{\mathcal{U}}$ de A é da forma diagonal. Neste caso, os elementos de $\mathcal{U}$ são os autovetores de A e a diagonal da matriz $\mathbf{a}$ contém os autovalores de A.

Capítulo 13

Operadores ortogonais

Neste Capítulo 13 consideramos operadores lineares

$$A\colon E \to F$$

exclusivamente entre espaços vetoriais de dimensão finita com produtos internos

$$E = (E, \langle\cdot,\cdot\rangle_E)\,, \qquad F = (F, \langle\cdot,\cdot\rangle_F)$$

e dimensões $n = \dim E$ e $m = \dim F$. Dévido aos produtos internos o corpo sempre será $\mathbb{K} = \mathbb{R}$. Em vez de $\langle\cdot,\cdot\rangle_E$ ou $\langle\cdot,\cdot\rangle_F$ escrevemos geralmente $\langle\cdot,\cdot\rangle$, o contexto vai indicar do qual produto interno trata-se, aquele de E ou F.

Vamos estudar aqueles operadores as quais *preservam o produto interno* no sentido que o produto de dois vetores é igual ao produto das imagens deles

$$\langle v, w\rangle = \langle Av, Aw\rangle$$

para todos os vetores $v, w \in E$. Tal operador A é chamado de **operador ortogonal** (motivado pelo caso das matrizes) ou de **isometria**.

13.1 Matrizes ortogonais

Definição 13.1.1. Uma matriz $\mathbf{u} = (u_{ij}) \in \mathrm{M}(m \times n)$ é chamado de **matriz ortogonal** se suas colunas $\mathbf{u}_{\bullet 1}, \dots, \mathbf{u}_{\bullet n} \in \mathbb{R}^m$ formam um conjunto ON.[1]

Observe que as n colunas de uma matriz $m \times n$ ortogonal formam uma base ON da imagem da matriz no $\mathbb{R}^m$.

Corolário 13.1.2. *Para uma matriz ortogonal* $\mathbf{u} = (u_{ij}) \in \mathrm{M}(m \times n)$ *vale que*

(i) $n \le m$ — *menos ou igual colunas como linhas*

(ii) $\mathrm{posto}(\mathbf{u}) = n$ — *posto = número de colunas*

[1] como devem ter norma 1 todas as colunas de uma matriz ortogonal são não-nulas

(iii) $\mathbf{u}$ *é injetiva* *invertível no caso quadrado* $n = m$

Demonstração. (i) ON e não-nulo implica que $\{\mathbf{u}_{\bullet 1}, \ldots, \mathbf{u}_{\bullet n}\}$ é LI, daí o número n de elementos deve ser menor ou igual à dimensão do espaço ambiente $\mathbb{R}^m$.
(ii) $\text{posto}(\mathbf{u}) := \dim \text{Im}(\mathbf{u}) = \text{pc}(\mathbf{u}) = n$. (iii) Segundo o teorema de núcleo e imagem $\dim \mathbb{R}^n = \dim \text{N}(\mathbf{u}) + \dim \text{Im}(\mathbf{u})$, daí $\dim \text{N}(\mathbf{u}) = 0$. □

Lema 13.1.3. $\mathbf{u}$ *matriz ortogonal* $\Leftrightarrow$ $\mathbf{u}^t\mathbf{u} = \mathbb{1}$

Demonstração. $\mathbf{u} \in \text{M}(m \times n)$ ortogonal $:\Leftrightarrow$ $\langle \mathbf{u}_{\bullet j}, \mathbf{u}_{\bullet k} \rangle = \delta_{jk}$ $\forall j, k$, mas

$$\delta_{jk} = \langle \mathbf{u}_{\bullet j}, \mathbf{u}_{\bullet k} \rangle = \sum_{i=1}^{m} u_{ij} u_{ik} = \sum_{i=1}^{m} (\mathbf{u}^t)_{ji} u_{ik} = (\mathbf{u}^t\mathbf{u})_{jk}$$

□

Lema 13.1.4 (Matrizes ortogonais quadradas)**.** *Para uma matriz quadrada* $\mathbf{u} \in \text{M}(n \times n)$ *são equivalente*

(i) $\mathbf{u}$ *ortogonal*

(ii) $\mathbf{u}^{-1}$ *existe e iguale* $\mathbf{u}^t$

(iii) *as colunas de* $\mathbf{u}$ *formam um conjunto ON*

(iv) *as linhas de* $\mathbf{u}$ *e formam um conjunto ON*

Demonstração. (i) $\Leftrightarrow$ (ii) A identidade $\mathbf{u}^t\mathbf{u} = \mathbb{1}$ entre matrizes *quadradas* significa que $\mathbf{u}^t$ e $\mathbf{u}$ são inversas uma da outra. (i) $\Leftrightarrow$ (iii) Definição 13.1.1
(iii) $\Leftrightarrow$ (iv) Note que como $\mathbf{u}^{-1}$ existe $\mathbf{u}^t\mathbf{u} = \mathbb{1} \Leftrightarrow \mathbf{u}\mathbf{u}^t = \mathbb{1}$, mas a segunda identidade significa (analogamente à prova de Lema 13.1.3) que as linhas de $\mathbf{u}$ são todas não-nulas e formam um conjunto ON. □

Exercício 13.1.5 (O **grupo ortogonal** $\text{O}(n)$)**.** Mostre que o conjunto $\text{O}(n)$ das matrizes ortogonais quadradas $n \times n$ munido do produto matrix é um grupo.

Exercício 13.1.6. Matrizes de passagem $\mathbf{p}$ entre bases ONs são ortogonais.

Exercício 13.1.7 ($\text{O}(1)$)**.** Mostre que $\text{O}(1) = \{-1, +1\} = \mathbb{S}^n$ onde a esfera unitária $\mathbb{S}^n \subset \mathbb{R}^{n-1}$ é composto dos pontos da distancia 1 da origem.

Comentário 13.1.8. O grupo ortogonal $\text{O}(2)$ é composto de dois círculos

$$\text{O}(2) = \{\text{rotações } R_\theta\} \dot\cup \{\text{reflexões ortogonais } S_{L_\theta}\} = \mathbb{S}^1 \dot\cup \mathbb{S}^1$$

onde L_θ é a reta em $\mathbb{R}^2$ passando $\mathcal{O}$ e formando o ângulo θ com o eixo-x.

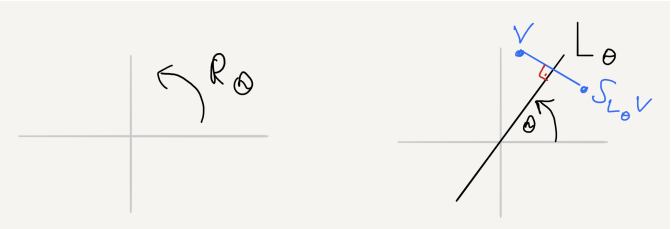


Figura 13.1: Grupo ortogonal O(2) composto de rotações e reflexões ortogonais

13.2 Operadores ortogonais

Teorema 13.2.1. *Para um operador $A \in \mathcal{L}(E, F)$ são equivalente*

(i) $|Av| = |v|\ \forall v \in E$ — *"A preserva norma"*

(ii) $|Au - Av| = |u - v|\ \forall u, v \in E$ — *"A preserva distância"*

(iii) $\langle Au, Av\rangle = \langle u, v\rangle\ \forall u, v \in E$ — *"A preserva produto interno"*

(iv) $A^*A = I_E$ — *"A^* é inversa à esquerda de A"*

(v) *a matriz $[A]_{\mathcal{X},\mathcal{Y}}$ é ortogonal se $\mathcal{X}$ e $\mathcal{Y}$ são bases ONs*

(vi) *$A\mathcal{X}$ é um subconjunto ON de F se $\mathcal{X}$ é uma base ON de E*

Demonstração. Teorema B.6.1 □

Operadores em E

Definição 13.2.2. Chama-se $A \in \mathcal{L}(E, F)$ um **operador ortogonal** se A satisfaz uma (portanto todas as) afirmações de Teorema 13.2.1.

Observação 13.2.3. $A \in \mathcal{L}(E)$ ortogonal $\Leftrightarrow\ A^* = A^{-1}$

Lema 13.2.4. *Seja $A \in \mathcal{L}(E)$ ortogonal, então*

a) $\lambda \in \operatorname{spec} A \Rightarrow \lambda \in \{-1, +1\}$ — *"$\operatorname{spec} A = \emptyset$ é possível: $A = R_{90°}$"*

b) $E_{-1} \perp E_1$ — *"$E_{\pm 1} = \{\mathcal{O}\}$ é possível: $A = R_{90°}$"*

Exercício 13.2.5. Seja $S \in \mathcal{L}(E)$, mostre que como ilustrado na Figura 13.2 duas das três propriedades seguintes implicam a terceira

(i) $S = S^{-1}$ $(\Leftrightarrow S^2 = I)$ — "involução"

(ii) $S^* = S$ — "auto-adjunto"

(iii) $S^* = S^{-1}$ — "ortogonal"

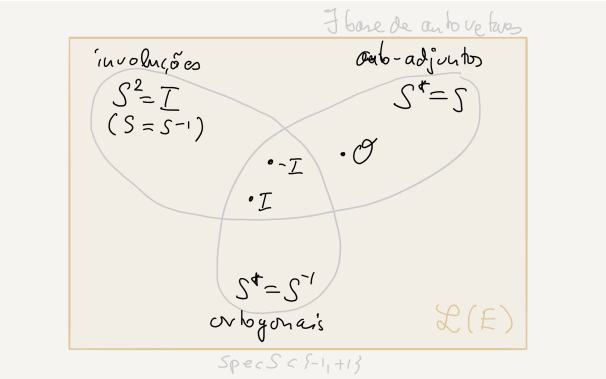


Figura 13.2: Na interseção estão as reflexões ortogonais $S_{E_1,E_{-1}}$ e S_{E_{-1},E_1}

13.2.1 Forma normal

Teorema 13.2.6 (Forma normal de operadores ortogonais)**.** *Para $A \in \mathcal{L}(E)$ ortogonal existe uma base ON $\mathcal{X}$ de E tal que a matriz de A é da forma bloco*

$$[A]_{\mathcal{X}} = \begin{bmatrix} \mathbb{1}_k & & & & \\ & -\mathbb{1}_\ell & & & \\ & & R_{\theta_1} & & \\ & & & \ddots & \\ & & & & R_{\theta_r} \end{bmatrix}, \qquad R_{\theta_j} = \begin{bmatrix} \cos\theta_j & \sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{bmatrix}$$

onde $k + \ell + 2r = n = \dim E$ e todas outras entradas são nulas.

Corolário 13.2.7. *No caso de $\dim E$ ímpar todo operador ortogonal A em E possui um autovalor $\lambda = 1$ ou $\lambda = -1$, em símbolos $\{-1, +1\} \cap \operatorname{spec} A \neq \emptyset$.*

13.3 Decomposição polar

Teorema 13.3.1 (Decomposição polar $A = PU$)**.** *Dado um operador linear A em E, então*

a) *existe um operador não-negativo $P \geq 0$ em E e* *lembre que $P = P^*$*

b) *existe um operador ortogonal U em E*

tal que

$$A = PU$$

Se A é invertível, então P e U são únicos e dados pelas fórmulas

$$P := \sqrt{AA^*}, \qquad U := P^{-1}A$$

Exercício 13.3.2. Seja $A \in \mathcal{L}(E)$ e seja $F \subset E$ um subespaço invariante por A, ou seja $AF \subset F$. Então vale

a) A ortogonal $\Rightarrow AF^\perp \subset F^\perp$

b) A invertível $\Rightarrow A^{-1}F \subset F$

Exercício 13.3.3. Ache a decomposição polar $\mathbf{a} = \mathbf{pu}$ da matriz

$$\mathbf{a} = \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}$$

13.4 Exercícios

1. Dê os seguintes exemplos:

   (a) Uma matriz invertível cujas linhas são duas a duas ortogonais mas as colunas não são.

   (b) Uma matriz (não-quadrada) cujas linhas são ortogonais e têm a mesma norma, mas as colunas não são ortogonais.

   (c) Uma matriz cujas linhas (e colunas) são duas a duas ortogonais mas as normas das linhas são diferentes.

2. Para quaisquer bases ortonormais $\mathcal{X} = (\xi_1, \ldots, \xi_n)$ e $\mathcal{Y} = (\eta_1, \ldots, \eta_n)$ de E, prove que existe um operador *ortogonal* $A \in \mathcal{L}(E)$ tal que

   $$A\xi_1 = \eta_1, \quad \ldots, \quad A\xi_n = \eta_n.$$

   No caso $E = \mathbb{R}^3$ e se as bases dadas são formadas pelos vetores

   $$\xi_1 = \tfrac{1}{3}(1,2,2), \qquad \xi_2 = \tfrac{1}{3}(2,1,-2), \qquad \xi_3 = \tfrac{1}{3}(2,-2,1),$$

   $$\eta_1 = \tfrac{1}{7}(2,3,6), \qquad \eta_2 = \tfrac{1}{7}(6,2,-3), \qquad \eta_3 = \tfrac{1}{7}(3,-6,2),$$

   determine a matriz de A na base canônica $\mathcal{E} = (e_1, e_2, e_2)$ de $\mathbb{R}^3$.

3. Se uma matriz triangular é ortogonal, prove que ela é diagonal e seu quadrado é igual à matriz identidade.

4. Seja $\mathbf{a} = \begin{bmatrix} a_1 & \ldots & a_n \end{bmatrix} \in M(1 \times 4)$ tal que $a_1^2 + \cdots + a_n^2 = 1$. Prove que $\mathbf{a}^t\mathbf{a} \in M(n \times n)$ é uma matriz de uma projeção ortogonal. Determine a imagem e o núcleo dessa projeção.

5. Ache uma matriz ortogonal 4×4 cujos elementos são todos da forma $\pm\frac{1}{2}$.

6. Obtenha a *decomposição polar*[2] da matriz

   $$\mathbf{a} = \begin{bmatrix} \sqrt{2} & 1 & 1 \\ -\sqrt{2} & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}.$$

[2] $\mathbf{a} = \mathbf{pu}$ onde $\mathbf{u}$ é ortogonal e $\mathbf{p}$ satisfaz $\mathbf{p}^t = \mathbf{p}$ e $\langle \mathbf{p}v, v\rangle \geq 0$ para todo vetor v.

Capítulo 14

Produto hermitiano

Neste Capítulo 14 consideramos exclusivamente espaços vetoriais complexos

$$Z = (Z, +, \cdot, \mathbb{C}), \qquad W = (W, +, \cdot, \mathbb{C})$$

então o corpo são os números complexos. Suponhamos que as dimensões

$$n = \dim Z < \infty, \qquad m = \dim W < \infty$$

são finitas (exceto quando especificamos diferentemente). Como na dimensão finita produtos hermitianos existem, podemos supor que $Z = (Z, (\cdot, \cdot))$, também $W = (W, (\cdot, \cdot))$, é munido de um produto hermitiano.

14.1 Definições

Números complexos

Um **número complexo** é uma expressão da forma $c = \alpha + i\beta$ onde $\alpha, \beta \in \mathbb{R}$ são números reais e i é um símbolo. Os números reais definidos assim

$$\mathrm{Re}(\alpha + i\beta) := \alpha, \qquad \mathrm{Im}(\alpha + i\beta) := \beta$$

chama-se de **parte real** e **parte imaginário** do número complexo. O número

$$\overline{\alpha + i\beta} := \alpha - i\beta$$

é chamado de **complexo conjugado** de um número complexo. Observe que um número complexo igualando seu próprio conjugado é um número real $c = \bar{c} \in \mathbb{R}$.

Seja $\mathbb{C}$ o conjunto de todos os números complexos. Adição e multiplicação de números complexos são definidas assim

$$\begin{aligned}(\alpha + i\beta) + (\gamma + i\delta) &:= (\alpha + \gamma) + i(\beta + \delta)\\ (\alpha + i\beta) \cdot (\gamma + i\delta) &:= (\alpha\gamma - \beta\delta) + i(\alpha\delta + \beta\gamma)\end{aligned}$$

É útil abreviar uns tipos especiais de números complexos como

$$0 + i\beta =: i\beta, \qquad \alpha + i0 =: \alpha, \qquad \alpha + i1 =: \alpha + i,$$

Assim i abrevia o número complexo $0 + i1$. E consequentemente

i^2 abrevia $\qquad (0 + i1)^2 = (0 + i1)(0 + i1) = (0^2 - 1^2) + i(0 \cdot 1 + 1 \cdot 0) = -1$

Matrizes complexas

A **complexa conjugada** de uma matriz $\mathbf{a} = (a_{ij}) \in \mathrm{M}(m \times n; \mathbb{C})$ é a matriz $\bar{\mathbf{a}} = (\overline{a_{ij}})$ cujas entradas são os complexos conjugados das entradas de $\mathbf{a}$.

Uma **matriz hermitiana** é uma matriz quadrada $\mathbf{a} \in \mathrm{M}(n \times n; \mathbb{C})$ tal que $\mathbf{a} = \bar{\mathbf{a}}^t$. A diagonal de uma matriz hermitiana tem entradas reais $a_{ii} = \bar{a}_{ii}$.

Produto hermitiano

Definição 14.1.1 (Produto hermitiano)**.** Um **produto hermitiano**, ou **produto $\mathbb{C}$-interno**, num espaço vetorial complexo Z (a dimensão pode ser infinita) é uma função de duas variáveis complexas

$$(\cdot, \cdot) \colon Z \times Z \to \mathbb{C}$$

a qual satisfaz três axiomas para todos os $z, w, \tilde{z}, \tilde{w} \in Z$ e $c \in \mathbb{C}$, isto é

$\overline{\texttt{(SIM)}}$ $\quad \langle z, w\rangle = \overline{\langle w, z\rangle}$ (simetria-cc)

(SL) $\quad \langle z + \tilde{z}, w\rangle = \langle z, w\rangle + \langle \tilde{z}, w\rangle$
$\langle z, w + \tilde{w}\rangle = \langle z, w\rangle + \langle z, \tilde{w}\rangle$ (sesquilinearidade)
$\langle cz, w\rangle = c\langle z, w\rangle$
$\langle z, cw\rangle = \bar{c}\langle z, w\rangle$

(POS) $\quad z \neq \mathcal{O} \Rightarrow \langle z, z\rangle > 0$ (positividade)

As duas identidades em cinza seguem da identidade anterior junto com $\overline{\texttt{(SIM)}}$. Neste caso Z é chamado de **espaço vetorial com produto hermitiano**.

Exemplo 14.1.2 (Produtos hermitianos)**.**

(a) $\mathbb{C}^n$ munido de $(z, w) := z_1\bar{w}_1 + \cdots + z_n\bar{w}_n$

(b) $C^0([a, b], \mathbb{C})$ munido de $(f, g) := \int_a^b f(x)\overline{g(x)}\, dx$

Exercício 14.1.3. Todo espaço vetorial complexo de dimensão finita possui um produto hermitiano.

[Dica: Pegue base $\mathcal{B}$ de E, exprime $z \in Z$ nesta base, Exemplo 14.1.2 (a).]

O seguinte teorema é o fundamento para o teorema espectral complexo, o Teorema 14.2.6 embaixo.

Teorema 14.1.4 (Operadores complexos são triangularizáveis)**.** *Todo operador linear num espaço vetorial complexo de dimensão finita é triangularizável, ou seja, existe uma base na qual a matriz do operador é triangular. A base pode ser obtida ON se pegamos um produto hermitiano.*

Exercício 14.1.5. Dado $A \in \mathcal{L}(Z)$, mostre que se $A^k = \mathcal{O}$ para algum $k > n = \dim Z$, entao $A^n = \mathcal{O}$. [Dica: Teorema 14.1.4.]

Exercício 14.1.6. Mostre que um operador $A \in \mathcal{L}(Z)$ é **nilpotente** (a saber $A^k = \mathcal{O}$ para algum $k \in \mathbb{N}$) se, e somente se, todos os autovalores são nulos.

14.2 Adjunta complexa $A^\dagger$

Definição 14.2.1. A **adjunta complexa** de um operador linear $A\colon Z \to W$ entre espaços vetoriais complexos é o único operador linear $A^\dagger\colon W \to Z$ tal que

$$(Az, w) = (z, A^\dagger w)$$

para todos os $z \in Z$ e $w \in W$.

Teorema 14.2.2 (Regras básicas para a adjunta complexa)**.**

(i) $I = I^\dagger$

(ii) $(A + B)^\dagger = A^\dagger + B^\dagger$

(iii) $(cA)^\dagger = \bar{c}\, A^\dagger$

(iv) $(BA)^\dagger = A^\dagger B^\dagger$

(v) $(A^\dagger)^\dagger = A$

Demonstração. Analogamente como Teorema 11.1.5. □

Exercício 14.2.3 (A matriz da adjunta complexa $A^\dagger$)**.** Dado $A \in \mathcal{L}(Z, W)$ e bases ONs $\mathcal{X}$ de Z e $\mathcal{Y}$ de W, mostre que $\mathbf{a}$ é a matriz de A se, e somente se a matriz da adjunta complexa é a matriz transposta complexo conjugada, em símbolos

$$\mathbf{a} = [A]_{\mathcal{X},\mathcal{Y}} \quad \Leftrightarrow \quad \bar{\mathbf{a}}^t = \left[A^\dagger\right]_{\mathcal{Y},\mathcal{X}}$$

Exercício 14.2.4 (A adjunta complexa $\mathbf{a}^\dagger$ de uma matriz)**.** Considere uma matriz $\mathbf{a} \in \mathrm{M}(m \times n; \mathbb{C})$ como transformação linear $\mathbf{a}\colon \mathbb{C}^n \to \mathbb{C}^m$. Mostre que a adjunta complexa é a matriz transposta complexo conjugada, em símbolos

$$\mathbf{a}^\dagger = \bar{\mathbf{a}}^t$$

14.2.1 Operadores normais

Definição 14.2.5. Um operador $A \in \mathcal{L}(Z)$ que comuta $A^\dagger A = AA^\dagger$ com sua adjunta complexa é chamado de **operador normal**.

Teorema 14.2.6 (Teorema espectral complexo)**.** *Um operador $A \in \mathcal{L}(Z)$ é normal $A^\dagger A = AA^\dagger$ se e somente se existe uma base ON $\mathcal{X} = (\hat{\xi}_1, \ldots, \hat{\xi}_n)$ de E composto de autovetores ξ_j de A.*

Entao operadores normais são diagonalizáveis. No seguinte estudamos dois classes de operadores, hermitianos (generalizando auto-adjunto) e unitários (generalizando ortogonal), ambos as quais são normais e assim herdam o teorema espectral complexo, então diagonalizabilidade.

14.2.2 Operadores hermitianos (complexo auto–adjuntos)

Definição 14.2.7. Um operador $A \in \mathcal{L}(Z)$ que iguale $A = A^\dagger$ a sua adjunta complexa é chamado de **operador hermitiano** ou **complexo auto-adjunto**.

Como hermitiano implica normal $AA^\dagger = AA = A^\dagger A$ o teorema espectral complexo vale.

Corolário 14.2.8. *Para operadores hermitianos vale o teorema espectral complexo, Teorema 14.2.6, e assim eles são diagonalizáveis.*

Teorema 14.2.9. *Seja $A \in \mathcal{L}(Z)$, entao*

$$A^\dagger = A \qquad \Leftrightarrow \qquad \mathbf{a} := [A]_{\mathcal{X}} = \bar{\mathbf{a}}^t \text{ para todas bases ONs } \mathcal{X} \text{ de } Z$$

14.2.3 Operadores unitários (complexo ortogonais)

Definição 14.2.10. Um operador $U \in \mathcal{L}(Z)$ tal que $U^\dagger = U^{-1}$ é chamado de **operador unitário** ou **complexo ortogonal**. Como caso especial, uma **matriz unitária** é uma matriz quadrada complexa tal que $\mathbf{u}^{-1} = \bar{\mathbf{u}}^t$.

Como unitário implica normal $UU^\dagger = UU^{-1} = I = U^{-1}U = U^\dagger U$ o teorema espectral complexo vale.

Corolário 14.2.11. *Para operadores unitários vale o teorema espectral complexo, Teorema 14.2.6, e assim eles são diagonalizáveis.*

Exercício 14.2.12. $U^\dagger = U^{-1} \Leftrightarrow (Uz, U\tilde{z}) = (z, \tilde{z})\ \forall z, \tilde{z} \in Z$.

Parte IV

Apêndices

Apêndice A

Revisão de matrizes e sistemas lineares

Seja $\mathbb{R}$ o conjunto dos números reais.

Comentário A.0.13 (Corpos gerais $\mathbb{K}$)**.** As construções neste Apêndice A para matrizes e listas cujas entradas são números reais funcionam do mesmo jeito para entradas num corpo geral $\mathbb{K}$. Veja por exemplo [Lan93, Chap. VIII].

A.1 Matrizes

Para alunos da matemática recomendo o tratamento excelente [Art91, Chapter 1], para os outros alunos recomendo [San12, Capítulos 1 e 2].

Definição A.1.1 (Matrizes $m \times n$)**.** Uma **matriz**, mais detalhado uma **matriz $m \times n$**, é um quadro numérico de m linhas e n colunas

$$\mathbf{a} = (a_{ij}) := \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

onde os a_{ij} são números reais chamado de **entradas da matriz**. Por definição a entrada a_{ij} está localizada na i-ésima linha e j-ésima coluna, para memorizar

$$a_{\text{linha coluna}}$$

Uma **matriz quadrada** é uma matriz do tipo $n \times n$. Denotamos de

$$\mathrm{M}(m \times n) := \left\{ \mathbf{a} = (a_{ij}) \;\middle|\; a_{ij} \in \mathbb{R},\, i = 1, \dots m,\, j = 1, \dots, n \right\}$$

o conjunto de todas as matrizes reais $\mathbf{a} = (a_{ij})$ do tipo $m \times n$. Definimos a **adição** de duas matrizes entrada por entrada, ou seja

$$\mathbf{a} + \mathbf{b} = (a_{ij}) + (b_{ij}) := (c_{ij}), \qquad c_{ij} := a_{ij} + b_{ij}$$

Definimos a **multiplicação escalar** de uma matriz por um número real γ, chamado de **escalar** por tradição, também entrada por entrada, ou seja

$$\gamma\mathbf{a} = \gamma\,(a_{ij}) := (c_{ij})\,, \qquad c_{ij} := \gamma a_{ij}$$

A matriz cujas entradas todas são o número nulo 0 é chamada de **matriz nula**, símbolo $\mathbf{0}$. Se na matriz nula $n \times n$ colocamos o número 1 ao longo da diagonal obtemos a **matriz identidade** $\mathbb{1} = \mathbb{1}_n$. Denotamos de $-\mathbf{a} := (-1)\mathbf{a}$ a matriz cujas entradas são os negativos $-a_{ij}$ das entradas a_{ij} da matriz $\mathbf{a}$. Como $-\mathbf{a} + \mathbf{a} = \mathbf{0}$ dizemos que $-\mathbf{a}$ e $\mathbf{a}$ são **inversos aditivos** um do outro.

Definição A.1.2 (Matriz transposta)**.** A **transposta** de uma matriz $\mathbf{a} \in \mathrm{M}(m \times n)$ é a matriz $\mathbf{a}^t$ com entradas $(\mathbf{a}^t)_{ij} = a_{ji}$. Ou seja, a transposta de uma matriz tem como linhas as colunas da matriz original.

Definição A.1.3 (Linhas e colunas de matrizes)**.** Seja $\mathbf{a} = (a_{ij}) \in \mathrm{M}(m \times n)$ uma matriz $m \times n$. Note-se que o primeiro índice i de uma entrada a_{ij} indica a linha e o segundo j a coluna dela. Tendo isso na vista vamos denotar a ***k*-ésima coluna**, respectivamente a ***ℓ*-ésima linha**, de uma matriz $\mathbf{a}$ com os símbolos

$$\mathbf{a}_{\bullet k} := \begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix}, \qquad \mathbf{a}_{\ell\bullet} := \begin{bmatrix} a_{\ell 1} & \dots & a_{\ell n} \end{bmatrix} \tag{A.1.1}$$

Temos escolhido o símbolo $\bullet$ para sugerir "este índice é aberto" – ele corre e assim gera uma lista, ou vertical ou horizontal dependente se $\bullet$ fica no primeiro ou no segundo lugar. Assim podemos escrever a matriz $\mathbf{a}$ nas formas seguintes

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}_{1\bullet} \\ \vdots \\ \mathbf{a}_{m\bullet} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_{\bullet 1} & \dots & \mathbf{a}_{\bullet n} \end{bmatrix}$$

Definição A.1.4 (Espaço-coluna e espaço-linha)**.** Seja $\mathbf{a} = (a_{ij}) \in \mathrm{M}(m \times n)$ uma matriz $m \times n$. O **espaço-coluna** é o conjunto de todas as somas das colunas $\mathbf{a}_{\bullet k}$ da matriz decorado com fatores escalares $\gamma_k \in \mathbb{R}$, em símbolos

$$\text{Esp-col}(\mathbf{a}) := \{\gamma_1\mathbf{a}_{\bullet 1} + \cdots + \gamma_n\mathbf{a}_{\bullet n} \mid \gamma_1, \dots, \gamma_n \in \mathbb{R}\} \subset \mathrm{M}(m \times 1)$$

Analogamente no **espaço-linha** usa-se as linhas da matriz $\mathbf{a}$, ou seja

$$\text{Esp-lin}(\mathbf{a}) := \{\gamma_1\mathbf{a}_{1\bullet} + \cdots + \gamma_n\mathbf{a}_{n\bullet} \mid \gamma_1, \dots, \gamma_n \in \mathbb{R}\} \subset \mathrm{M}(1 \times n)$$

Produto matriz

Para duas matrizes $\mathbf{a}$ de tipo $m \times n$ e $\mathbf{b}$ de tipo $k \times p$ pode-se definir o chamado **produto matriz** no caso que os índices $n = k$ coincidem:

$$\mathrm{M}(m \times n) \times \mathrm{M}(n \times p) \to \mathrm{M}(m \times p), \quad (\mathbf{a}, \mathbf{b}) \mapsto \mathbf{ab} := (c_{ij}) \tag{A.1.2}$$

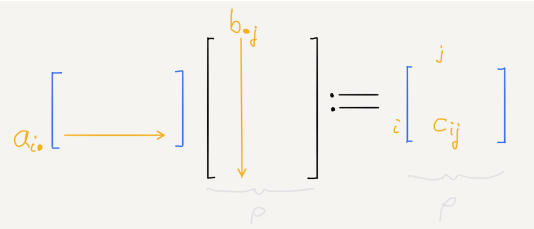


Figura A.1: Produto matriz – o número de colunas de **a** iguale o de linhas de **b**

onde

$$c_{ij} := \underbrace{\mathbf{a}_{i\bullet}}_{1\times n}\underbrace{\mathbf{b}_{\bullet j}}_{n\times 1} := a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

Lema A.1.5 (Umas propriedades do produto matriz)**.** *Vale o seguinte*

(i) $(\mathbf{cb})\mathbf{a} = \mathbf{c}(\mathbf{ba})$

(ii) $\mathbf{c}(\mathbf{a}+\mathbf{b}) = \mathbf{ca} + \mathbf{cb}$ *e* $(\mathbf{a}+\mathbf{b})\mathbf{c} = \mathbf{ac} + \mathbf{bc}$

(iii) $\mathbf{a}\mathbb{1}_n = \mathbf{a}$ *e* $\mathbb{1}_m\mathbf{a} = \mathbf{a}$ $\mathbf{a} \in \mathrm{M}(m \times n)$

(iv) $\mathbf{b}(\gamma\mathbf{a}) = \gamma(\mathbf{ba})$

para todos $\gamma \in \mathbb{R}$ *e matrizes* $\mathbf{a}, \mathbf{b}, \mathbf{c}$ *tal que as operacoes fazem sentido.*

Definição A.1.6 (Matriz inversa)**.** Uma matriz *quadrada* $\mathbf{a} \in \mathrm{M}(n \times n)$ admite uma inversa se existe uma matriz quadrada $\mathbf{b}$ tal que $\mathbf{ab} = \mathbb{1}_n$.[1] Neste caso tal $\mathbf{b}$ é único e usa-se o símbolo $\mathbf{a}^{-1}$ para denotar $\mathbf{b}$; veja Seção 6.4.1.

Definição A.1.7 (Matrizes quadradas comutando)**.** Dizemos que duas matrizes quadradas $\mathbf{a}, \mathbf{b} \in \mathrm{M}(n \times n)$ **comutam** se $\mathbf{ab} = \mathbf{ba}$.

A.2 Escalonamento de matrizes segundo Gauss

Definição A.2.1 (Operações elementares (oe))**.** Pode-se aplicar para as linhas de uma matriz três tipos de operações, as chamadas **operações elementares**:

$(\mathrm{oe1})_{\updownarrow}$ trocar duas linhas

$(\mathrm{oe2})_{\cdot}$ multiplicar uma linha com um escalar α

$(\mathrm{oe3})_{+}$ adicionar uma linha para uma outra

Teorema A.2.2. *O espaço linha não muda quando aplicar* (oe) *'s a uma matriz.*

Demonstração. Óbvio da Definição A.1.4 de Esp-lin. □

[1] No caso se matrizes *quadrada* a condição $\mathbf{ab} = \mathbb{1}_n$ é equivalente à condição $\mathbf{ba} = \mathbb{1}_n$.

Processo de escalonamento – método de Gauss (*1777 †1855)

Chama-se uma matriz **escalonada** se em cada linha o primeiro elemento não-nulo está à esquerda do primeiro elemento não-nulo da próxima linha. Exemplos

escalonadas: $\begin{bmatrix} 1 & 4 & 0 & 0 \\ 0 & 3 & 9 & 5 \\ 0 & 0 & 0 & 3 \end{bmatrix}$, $\begin{bmatrix} 1 & 4 & 0 & 0 \\ 0 & 0 & 3 & 9 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ não é: $\begin{bmatrix} 1 & 4 & 0 & 0 \\ 0 & 3 & 9 & 5 \\ 2 & 0 & 0 & 3 \end{bmatrix}$

Numa matriz *escalonada* os primeiros elementos não-nulos das linhas são chamados de **pivôs** da matriz escalonada.

Definição A.2.3. Uma matriz pode ser transformada numa matriz escalonada aplicando operações elementares. O processo é repetir os três passos seguintes:

1. Localiza a primeira coluna não-nula e nela o primeiro elemento não-nulo, dizemos a. Troca a linha de a e a primeira linha.
2. Embaixo de a anulamos todo elemento não-nulo, dizemos b: Multiplique a linha de b com $-a/b$, depois adiciona a linha de a. Continue até todos os elementos embaixo de a são nulos.
3. Esqueça a linha e a coluna de a e trata a matriz reduzida começando de novo com passo 1.

O processo de escalonar uma matriz $\mathbf{a}$ termina com *uma* matriz escalonada – não é única – denotado $\mathbf{a}_{\text{esc}}$ e ilustrada na Figura 3.1 onde $*$ simboliza entradas qualquais, nulos ou não-nulos.

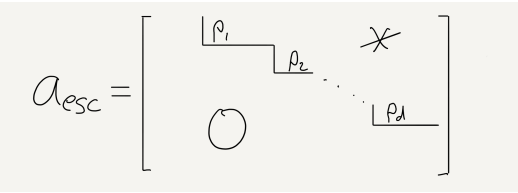


Figura A.2: Uma matriz escalonada $\mathbf{a}_{\text{esc}}$ com pivôs $p_1, \ldots, p_d \neq 0$

Exemplo A.2.4. Ilustramos o escalonamento. Seja Li a i-ésima linha.

$$\mathbf{a} = \begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 1 \\ 4 & 0 & -2 \end{bmatrix} \xrightarrow{1.\ L1 \leftrightarrow L2} \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 4 & 0 & -2 \end{bmatrix} \xrightarrow{2.\ -\frac{2}{4} L3} \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ -2 & 0 & 1 \end{bmatrix} \xrightarrow{2.na\,L3\ adic.\,L1} \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$

$$\xrightarrow{3.esq.linha\ e\ col.de\ 2} \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$

e agora começamos de novo com passo 1 tratando a matriz reduzida

$$\xrightarrow{1.\ L2 \leftrightarrow L2} \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \xrightarrow{2.\ -\frac{1}{1} L3} \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & -1 & -2 \end{bmatrix} \xrightarrow{2.na\,L3\ adic.\,L2} \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

A matriz escalonada reduzida

O processo de escalonar uma matriz $\mathbf{a}$ não lida a um resultado único: por exemplo, se $\mathbf{a}_{\text{esc}}$ é uma matriz escalonada, então $7\mathbf{a}_{\text{esc}}$ também é.

Definição A.2.5 (A matriz escalonada reduzida)**.** Chama-se uma matriz escalonada **reduzida**, símbolo $\mathbf{a}_{\text{esc}-\text{red}}$, se cada um pivô é 1 e as entradas acima dele são nulos. A matriz nula já é escalonada reduzida.

A matriz escalonada reduzida é única: Se uma matriz $\mathbf{a}$ não possui pivô, então trata-se da matriz nula e

$$\mathbf{0}_{\text{esc}} = \mathbf{0}_{\text{esc}-\text{red}} = \mathbf{0}$$

Dado então uma matriz não-nula $\mathbf{a}$. Calcule uma forma escalonada $\mathbf{a}_{\text{esc}}$. Na matriz $\mathbf{a}_{\text{esc}}$ multiplique a primeira linha com $1/p_1$ onde p_1 é o pivô da primeira linha. Se a linha 2 não é nula, então faça o mesmo na linha 2. Depois aplique operações elementares à linha 2 para anular cada uma entrada acima do pivô dois. Continue até não tem mais linhas não-nulas.

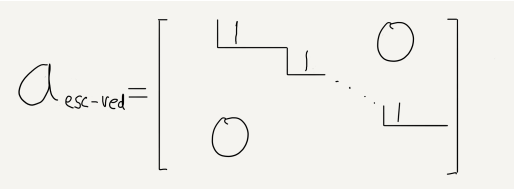


Figura A.3: A matriz escalonada reduzida $\mathbf{a}_{\text{esc}-\text{red}}$

Exercício A.2.6. Suponha que $\mathbf{a}$ é uma matriz quadrada $n \times n$ e uma matriz escalonada $\mathbf{a}_{\text{esc}}$ possui n pivôs. Neste caso a matriz escalonada reduzida $\mathbf{a}_{\text{esc}-\text{red}} = \mathbb{1}_n$ é a matriz identidade.

A.2.1 Aplicações de escalonamento

- Cálculo da dimensão do subespaço gerado por m vetores — Seção 3.1.3.
- Cálculo do posto de uma transformação linear — Seção 6.1.
- Resolução de sistemas lineares — Seções A.3 e 8.1.
- Cálculo da matriz inversa (Gauss-Jordan) — Seção A.4.

A.3 Sistemas Lineares

Seja $\mathbf{a}$ uma matriz $m \times n$ e $b = (b_1, \ldots, b_m)$ uma lista ordenada de m números reais. Agora mesclamos ás n colunas de $\mathbf{a}$ a lista b como a $(n+1)$-ésima coluna

para obter a chamada **matriz aumentada**, notação

$$[\mathbf{a} : b] := \begin{bmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \dots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{bmatrix}$$

Definição A.3.1. Suponha uma matriz $\mathbf{a}$ tipo $m \times n$ e uma lista $b = (b_1, \dots, b_m)$ são dadas. Queremos saber se existe uma solução $x = (x_1, \dots, x_n)$ do **sistema linear (SL) de m equações a n incógnitas $x_1, \dots, x_n$**

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \quad\vdots \qquad\qquad\quad \vdots \qquad\quad \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases} \tag{A.3.1}$$

ou, equivalentemente, da equação correspondente $\mathbf{a}x = b$ entre matrizes.[2]

É útil chamar a matriz aumentada $[\mathbf{a} : b]$ o **sistema linear** definido por (A.3.1). A lista $b = (b_1, \dots, b_m)$ é chamada de **inomogeneidade** do sistema linear. O caso $b = \mathcal{O} = (0, \dots, 0)$ chama-se de **sistema linear homogêneo (SLH)**.

O sistema linear (A.3.1) pode ser escrito equivalentemente na forma

$$x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{A.3.2}$$

O lado esquerdo é um exemplo de uma chamada "combinação linear" das colunas $\mathbf{a}_{\bullet 1}, \dots, \mathbf{a}_{\bullet n}$ da matriz $\mathbf{a}$, veja (A.1.1), representando o vetor b – um conceito fundamental o qual vamos tratar no próximo parágrafo.

Comentário A.3.2. Note-se que o lado esquerdo de (A.3.2) corre sobre toda a imagem da matriz $\mathbf{a}$ se variamos x sobre todas as listas. Então um SL $[\mathbf{a} : b]$ tem uma solução se e somente se a lista b é elemento da imagem da matriz $\mathbf{a}$.

Lembramos do curso MA141 "Geometria Analítica" o seguinte

Lema A.3.3. *Uma lista x é solução do sistema linear $[\mathbf{a} : b]$ se e somente se x é solução do sistema linear associado à matriz escalonada $[\mathbf{a} : b]_{\text{esc}}$.*

Ideia de demonstração. (Veja por exemplo Artin "Algebra" (1991), p. 13.)
As operações elementares podem ser escrito como matrizes invertíveis $\mathbf{e}$. O resultado de uma operação elementar numa matriz $\mathbf{a}$ então é a matriz $\mathbf{ea}$. Assim $\mathbf{a}_{\text{esc}} = \mathbf{pa}$ onde $\mathbf{p}$ é da forma $\mathbf{p} := \mathbf{e}_1 \dots \mathbf{e}_k$. Além disso $\mathbf{e}[\mathbf{a} : b] = [\mathbf{ea} : \mathbf{e}b]$, então $[\mathbf{a} : b]_{\text{esc}} = [\mathbf{pa} : \mathbf{p}b]$. '$\Rightarrow$' Se $\mathbf{a}x = b$, então $\mathbf{pa}x = \mathbf{p}b$. '$\Leftarrow$' Use $\mathbf{p}^{-1}$. $\square$

[2] Sempre se escrevemos $\mathbf{a}x = b$ consideramos x e b como matrizes colunas $n \times 1$ e $m \times 1$, respectivamente, e $\mathbf{a}x$ é o produto matriz.

Resolução de um sistema linear usando escalonamento

Comentário A.3.4. Para resolver o SL $\mathbf{a}x = b$

- escalone a matriz aumentada $[\mathbf{a} : b]$
- obtendo uma matriz escalonada $[\mathbf{a} : b]_{\text{esc}} =: [\tilde{\mathbf{a}} : \tilde{b}]$.
- Resolve o SL $\tilde{\mathbf{a}}x = \tilde{b}$ "de baixo para cima", veja Exemplo A.3.5 que segue.
- Uma lista x é solução de $\tilde{\mathbf{a}}x = \tilde{b}$ se e somente se x é solução de $\mathbf{a}x = b$.

Exemplo A.3.5. Para encontrar as soluções $(x, y, z) \in \mathbb{R}^3$ do sistema linear

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

primeiro, formamos a matriz aumentada $[\mathbf{a} : b]$ onde $b = (0, 0, 0) =: \mathcal{O}$, segundo, escalonamos ela, e terceiro, **resolvemos "de baixo para cima"**. Segundo Lema A.3.3 uma solução de $[\mathbf{a} : b]_{\text{esc}}$ também resolve $[\mathbf{a} : b]$, e vice versa. Exemplo A.2.4 mostra as matrizes $\mathbf{a}$ e $\mathbf{a}_{\text{esc}}$. Note-se que no caso especial quando um sistema linear é *homogêneo*, ou seja $b = \mathcal{O}$, vale a fórmula seguinte

$$[\mathbf{a} : \mathcal{O}]_{\text{esc}} = [\mathbf{a}_{\text{esc}} : \mathcal{O}]$$

No nosso caso o lado direito desta fórmula representa o SLH

$$\begin{cases} 2x + y + z = 0 \\ y + 2z = 0 \\ 0 = 0 \end{cases}$$

Resolução "de baixo para cima":

Linha 3. Começamos embaixo com a ultima linha $0x + 0y + 0z = 0$ a qual não representa nenhuma restrição para x, y, z.

Linha 2. Progredimos para cima, ou seja para a linha dois $y + 2z = 0$. Escolha uma variável para ser a variável dependente da(s) outra(s) variáveis, as quais variam livremente no corpo. No nosso caso só tem uma outra e o corpo é $\mathbb{R}$. Escolhemos por exemplo como variável dependente $y = y(z) = -2z$ como função da variável z a qual varia livremente sobre os números reais, ou seja $z \in \mathbb{R}$.

Linha 1. Progredimos para cima, ou seja para a primeira linha

$$0 = 2x + y(z) + z = 2x - 2z + z = 2x - z$$

lembrando que $z \in \mathbb{R}$ é livre. Então $x = x(z) = \frac{1}{2}z$ para qualquer $z \in \mathbb{R}$.

Conclusão. Toda solução do SL é da forma

$$\begin{bmatrix} x(z) \\ y(z) \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{2}z \\ -2z \\ z \end{bmatrix} = z \begin{bmatrix} \frac{1}{2} \\ -2 \\ 1 \end{bmatrix}$$

onde $z \in \mathbb{R}$ é um número real arbitrário. Então o SL não tem só uma solução – tem uma para cada um número real z. Isso conclui o Exemplo A.3.5.

A.4 Cálculo da matriz inversa – Gauss-Jordan

Esta Seção A.4 aplica só para matrizes *quadradas*.

Proposição A.4.1. *Caso uma matriz $n \times n$ (quadrada) $\mathbf{a}$ admite uma inversa, encontra-se a inversa assim: Considere a matriz $[\mathbf{a} : \mathbb{1}]$ obtida por escrever a matriz identidade $\mathbb{1} = \mathbb{1}_n$ à direita de $\mathbf{a}$. Aplique as três operações elementares* (oe1 – oe3)*, veja Definição A.2.1, até a matriz modificada tem a forma $[\mathbb{1} : \mathbf{c}]$ para uma matriz $\mathbf{c}$. Neste caso $\mathbf{c} = \mathbf{a}^{-1}$ é a inversa buscada.*

$$\left[\mathbf{a} : \mathbb{1}\right] \xrightarrow{\text{(oe)}} \cdots \xrightarrow{\text{(oe)}} \left[\mathbb{1} : \underbrace{\mathbf{c}}_{\mathbf{a}^{-1}}\right]$$

Dica: *É aconselhável escalonar a matriz $\mathbf{a}$ como passo intermediário e chegar numa matriz da forma $[\mathbf{a}_{\text{esc}} : \mathbf{d}]$ onde $\mathbf{a}_{\text{esc}}$ é um escalonamento de $\mathbf{a}$. Depois elimine todas as entradas acima da diagonal para ao fim chegar em $[\mathbb{1} : \mathbf{c}]$.*

Ideia de demonstração. (Veja por exemplo Artin "Álgebra" (1991), p. 17.)
As operações elementares podem ser escrito como matrizes invertíveis $\mathbf{e}$. O resultado de uma operação elementar numa matriz $\mathbf{a}$ então é a matriz $\mathbf{ea}$. Assim reduzir $\mathbf{a}$ para a matriz identidade $\mathbb{1}$ traduz num produto matriz $\mathbf{e}_1 \dots \mathbf{e}_k \mathbf{a} = \mathbb{1}$. Aplicando $\mathbf{a}^{-1}$ da direita obtemos $\mathbf{e}_1 \dots \mathbf{e}_k \mathbb{1} = \mathbf{a}^{-1}$. Esta identidade diz que aplicando as mesmas operações elementares na mesma ordem à matriz identidade $\mathbb{1}$ obtém-se a matriz inversa a^{-1}. □

A.4.1 O determinante de matrizes quadradas

Antes de começar o processo descrito na Proposição A.4.1 deve saber que a matriz é invertível. Nas dimensões 2 e 3 a ferramenta mais útil para checar é o determinante.

Definição A.4.2 (Matrizes)**.** Nas dimensões 1, 2, 3 define-se $\det[a_{11}] = a_{11}$ e

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \overbrace{a_{11}a_{22}}^{\text{diagonal}} - \overbrace{a_{21}a_{12}}^{\substack{\text{anti} \\ \text{-diagonal}}}$$

e

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{array}{c} \overbrace{a_{11}a_{22}a_{33}}^{\text{diagonal}} - \overbrace{a_{31}a_{22}a_{13}}^{\text{anti-diagonal}} \\ +a_{12}a_{23}a_{31} - a_{32}a_{23}a_{11} \\ +a_{13}a_{21}a_{32} - a_{33}a_{21}a_{12} \end{array}$$

Note como o primeiro índice dos a_{ij}'s embaixo do produto da diagonal / anti-diagonal fica constante e o segundo índice muda ciclicamente.

Exercício A.4.3. Seja $\mathbf{a} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ e $\mathbf{m} = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$. Prove que

1. $\det(\mathbf{am}) = \det \mathbf{a} \cdot \det \mathbf{m}$ [cálculo direto];

2. $\det \mathbf{a} \neq 0 \iff \mathbf{a}$ é invertível;

3. $\det(\mathbf{m}^{-1}\mathbf{am}) = \det \mathbf{a}$, para todo $\mathbf{m}$ invertível.

Pode-se definir o determinante de matrizes quadradas do qualquer tamanho $n \times n$; veja por exemplo [San12, § 2.2].

Teorema A.4.4. *Seja* $\mathbf{a}$ *uma matriz* $n \times n$ *(quadrada), então são equivalente*

$$\det \mathbf{a} \neq 0 \qquad \Leftrightarrow \qquad \mathbf{a} \text{ invertível } (\mathbf{a}^{-1} \text{ existe})$$

O determinante respeita transposta e produtos, ou seja

$$\det \mathbf{a}^t = \det \mathbf{a}, \qquad \det \mathbf{ba} = (\det \mathbf{b})\,(\det \mathbf{a}) \tag{A.4.1}$$

Em particular, a ordem no produto matriz não importa $\det \mathbf{ba} = \det \mathbf{ab}$.

Demonstração. Veja por exemplo [San12, Teoremas 2.14 e 2.15]. Para matrizes com entradas num corpo veja [Lan93, VIII Prop. 4.16]. □

A.5 Exercícios

1. Use escalonamento para resolver o sistema linear

$$\begin{aligned} x + 3y + z &= 1 \\ 2x + 6y + 9z &= 7 \\ 2x + 8y + 8z &= 6 \end{aligned}$$

   nas incógnitas $x, y, z \in \mathbb{R}$.

2. Determine a inversa da matriz $\mathbf{a} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ caso existisse.[3]

3. Decida quais das matrizes possuem inversa e calcule quando existir:

$$\mathbf{a} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 4 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 8 \end{bmatrix}.$$

[3] $\det \mathbf{a} = 1 \neq 0$ então $\mathbf{a}^{-1}$ existe, resultado $\mathbf{a}^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$

Apêndice B

Demonstrações restantes

B.1 Espaços vetoriais

Lema B.1.1 (Lema 1.1.5)**.** *Seja $(G, *)$ um grupo. Então vale o seguinte.*
1) *O elemento neutro é único.*
2) *Os elementos inversos são únicos.*
3) *Para todos os elementos $f, g, h \in G$ vale:*
   a) $f * g = f * h \Rightarrow g = h$ *(lei da corte)*
   b) $f * g = f \Rightarrow g = e$
   c) $f * g = e \Rightarrow g = \bar{f}$

Demonstração. 1) Se $e, \tilde{e} \in G$ satisfazem o axioma (elemento neutro), então usando o axioma para e e depois para $\tilde{e}$ obtemos que $e = e * \tilde{e} = \tilde{e}$.
2) Seja $g \in G$. Se $\bar{g}, \tilde{g} \in G$ satisfazem o axioma (inverso) para g, então obtemos

$$\bar{g} = e * \bar{g} = \underbrace{(\tilde{g} * g)}_{=e} * \bar{g} = \tilde{g} * \underbrace{(g * \bar{g})}_{=e} = \tilde{g} * e = \tilde{g}$$

usando (elem. neutro) no início e fim, $(\text{inverso})_{\tilde{g}}$, (associatividade), $(\text{inverso})_{\bar{g}}$.

3) a) $g = e * g = (\bar{f} * f) * g = \bar{f} * (f * g) \overset{\text{hip.}}{=} \bar{f} * (f * h) = (\bar{f} * f) * h = e * h = h$.
b) Use a) com $h = e$. c) Use a) com $h = \bar{f}$. □

Lema B.1.2 (Lema 1.1.10)**.** *Seja $\mathbb{K}$ um corpo e $0 \in K$ é o elemento neutro da adição. Então $0\beta = 0$ e $\beta 0 = 0$ para todos os elementos $\beta \in \mathbb{K}$.*

Demonstração. Seja $\beta \in \mathbb{K}$, denotamos o inverso aditivo de $-\beta$. Então

$$\beta \overset{(\text{el.n.})_{\cdot}}{=} 1\beta \overset{(\text{el.n.})_{+}}{=} (1+0)\beta \overset{(\text{distr.})}{=} 1\beta + 0\beta \overset{(\text{el.n.})_{\cdot}}{=} \beta + 0\beta$$

Usamos esta identidade para obter a segunda igualdade no seguinte

$$0 \overset{(\text{inv.})}{=} (-\beta) + \beta = -\beta + (\beta + 0\beta) \overset{(\text{ass.})_{+}}{=} (-\beta + \beta) + 0\beta \overset{(\text{inv.})}{=} 0 + 0\beta \overset{(\text{el.n.})}{=} 0\beta$$

□

Lema B.1.3 (Lema 1.1.20)**.** *Para o vetor nulo $\mathcal{O} \in E$ de um espaço vetorial e o elemento neutro aditivo $0 \in \mathbb{K}$ do corpo vale o seguinte.*

(i) $\alpha\mathcal{O} = \mathcal{O}$ *para todos os escalares* $\alpha \in \mathbb{K}$.

(ii) $0v = \mathcal{O}$ *para todos os vetores* $v \in E$.

(iii) *Para todo o escalar* $\alpha \in \mathbb{K}$ *e todo o vetor* $w \in E$ *são equivalentes:*

$$\alpha w = \mathcal{O} \quad \Leftrightarrow \quad \alpha = 0 \text{ ou } w = \mathcal{O}$$

Demonstração. (i) Caso $\alpha = 0$. Como $\alpha\mathcal{O} + 0\mathcal{O} = (\alpha + 0)\mathcal{O} = \alpha\mathcal{O}$, então $0\mathcal{O} = \mathcal{O}$ pela lei da corte (Lema B.1.1 3b) para $(G, *) = (E, +)$).
Caso $\alpha \neq 0$. Tal α tem um inverso aditivo, notação α^{-1}. Seja $v \in E$, então

$$v \overset{\text{(comp.)}}{=} 1v \overset{(\text{inv.})_{\mathbb{K},\cdot}}{=} (\alpha\alpha^{-1})v \overset{\text{(comp.)}}{=} \alpha(\alpha^{-1}v)$$

Usando este resultado no início e no fim do seguinte obtemos que

$$v + \alpha\mathcal{O} = \alpha(\alpha^{-1}v) + \alpha\mathcal{O} \overset{(\text{distr.})_E}{=} \alpha((\alpha^{-1}v) + \mathcal{O}) \overset{(\text{el.n.})_{E,+}}{=} \alpha(\alpha^{-1}v) = v$$

Então $\alpha\mathcal{O} = \mathcal{O}$ pela lei da corte (Lema B.1.1 3b) para $(G, *) = (E, +)$).

(ii) Como $v + 0v = 1v + 0v = (1 + 0)v = 1v = v$ a lei da corte diz que $0v = \mathcal{O}$.

(iii) '$\Rightarrow$' Suponha $\alpha w = \mathcal{O}$. Caso $\alpha = 0$, pronto. Caso $\alpha \neq 0$ concluímos que

$$w \overset{\text{(comp.)}}{=} 1w \overset{(\text{el.n.})_{\mathbb{K},\cdot}}{=} (\alpha^{-1}\alpha)w \overset{\text{(comp.)}}{=} \alpha^{-1}(\alpha w) \overset{\text{hip.}}{=} \alpha^{-1}\mathcal{O} \overset{\text{(i)}}{=} \mathcal{O}$$

'$\Leftarrow$' Se $w = \mathcal{O}$, então $\alpha\mathcal{O} \overset{\text{(i)}}{=} \mathcal{O}$, pronto. Se $\alpha = 0$, então $0w \overset{\text{(ii)}}{=} \mathcal{O}$, pronto. □

Corolário B.1.4 (Corolário 1.1.21)**.** *Para todos os $\alpha \in \mathbb{K}$ e $w \in E$ vale:*

a) $\alpha(-w) = -(\alpha w)$

b) $(-\alpha)w = -(\alpha w)$

Demonstração. a) Temos que mostrar que a soma de αw e o candidato para ser seu inverso aditivo iguale o vetor nulo. Com efeito

$$\alpha w + \alpha(-w) \overset{(\text{distr.})_E}{=} \alpha(w + (-w)) \overset{(\text{el.n.})_{E,+}}{=} \alpha\mathcal{O} = \mathcal{O}$$

onde o último passo é parte (i) de Lema B.1.3.
b) Temos o objetivo análogo de chegar ao vetor nulo, com efeito

$$\alpha w + (-\alpha)w \overset{(\text{distr.})_E}{=} (\alpha + (-\alpha))w \overset{(\text{el.n.})_{\mathbb{K},+}}{=} 0w = \mathcal{O}$$

onde o último passo é parte (ii) de Lema B.1.3. □

B.2 Subespaços

Lema B.2.1 (Lema 2.2.4)**.** *Todo subconjunto LI* $\{u, v\} \subset \mathbb{R}^2$ *gera* $\mathbb{R}^2$*.*

Demonstração. Vai ter 4 passos. I. Os vetores u, v não são múltiplos um do outro: Suponha por absurdo que $u = \alpha v$ para um $\alpha \in \mathbb{R}$. Então $1u + (-\alpha)v = 1\alpha v - (\alpha v) = \mathcal{O}$ contradizendo LI. II. $u \neq \mathcal{O}$: Caso contrario $u = \mathcal{O} = 0v$ contradizendo I. III. $v \neq \mathcal{O}$: Análogo. IV. Seja $v \in \mathbb{R}^2$. Caso $w = \mathcal{O}$ escrevemos $w = 0u$, pronto. Caso $v \neq \mathcal{O}$: Agora identificamos $\mathbb{R}^2$ com o plano usando dois eixos OXY, veja Figura 2. Segundo II. e III. temos duas retas $\mathbb{R}u$ e $\mathbb{R}v$ passando ambas a origem O, mas não são iguais segundo I. Recebemos um paralelogramo com dois lados parte das retas e dois vértices sendo $\mathcal{O}$ e v; pensa Figura 2 com OX e OY substituto para Ou e Ov. Então a flecha v é a soma de duas flechas do paralelogramo, uma flecha sendo um múltiplo de u e a outra de v. Pronto. □

Teorema B.2.2 (Teorema 2.3.4)**.** *Sejam* $F_1, F_2 \subset F$ *três subespaços de um espaço vetorial* E*, então são equivalentes*

$$F = F_1 \oplus F_2 \quad \Leftrightarrow \quad \forall f \in F,\ \exists!\ f_1 \in F_1,\ f_2 \in F_2 \ \textit{tal que}\ \ f = f_1 + f_2$$

Demonstração. '⇒' Seja $f \in F$. Como hipótese temos duas informações, a saber (i) $F = F_1 + F_2$ e (ii) $F_1 \cap F_2 = \{\mathcal{O}\}$, dando existência e unicidade.
Existência: De (i) sabemos que $f = f_1 + f_2$ para um $f_1 \in F_1$ e um $f_2 \in F_2$.
Unicidade. Suponha que $f = \tilde{f}_1 + \tilde{f}_2$ também para um $\tilde{f}_1 \in F_1$ e um $\tilde{f}_2 \in F_2$. Então $F_1 \ni f_1 - \tilde{f}_1 = \tilde{f}_2 - f_2 \in F_2$. Assim cada um lado pertence a ambos espaços, então a $F_1 \cap F_2$ o qual segundo (ii) iguale $\{\mathcal{O}\}$. Como não tem outro elemento, cada um lado deve ser o vetor nulo.
'⇐' $F_1 + F_2 = F$: A hipótese *existência* disponibiliza a primeira inclusão $F \subset F_1 + F_2 \subset F$ e a segunda vale como $F_1, F_2 \subset F$.
$F_1 \cap F_2 = \{\mathcal{O}\}$: Seja $f \in F_1 \cap F_2$, a mostrar $f = \mathcal{O}$. Note que $f \in F$ como $F_1, F_2 \subset F$. Então segundo a propriedade do vetor nulo

$$\underbrace{f}_{\in F_1} + \underbrace{\mathcal{O}}_{\in F_2} = f = \underbrace{\mathcal{O}}_{\in F_1} + \underbrace{f}_{\in F_2} \tag{B.2.1}$$

Mas pela hipótese *unicidade* escrever f como soma de um elemento de F_1 e um elemento de F_2 é único, então $f = \mathcal{O}$ e $\mathcal{O} = f$. □

B.3 Bases – SLH

Teorema B.3.1 (Teorema 3.1.11)**.** *Dado uma matriz* $\mathbf{a} \in \mathrm{M}(m \times n; \mathbb{K})$*. Se tem menos linhas (equações) como colunas (incógnitas), em símbolos* $m < n$*, então o sistema linear homogêneo (SLH)*

$$(*) \qquad \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \quad\vdots \qquad\qquad \vdots \qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = 0 \end{cases}$$

admite soluções $x = (x_1, \ldots, x_n)$ não triviais (não todos x_j nulos).

Demonstração. Se todos os coeficientes a_{ij} são nulos, então todos os elementos $x \in \mathbb{K}^n$ são soluções. Sejam então não todos coeficientes nulos: A prova usa indução sobre o número m de equações.
$m = 1$: Em $a_{11}x_1 + \cdots + a_{1n}x_n = 0$ temos pelo menos dois incógnitas segundo nossa hipótese $n > m = 1$. Além disso, pelo menos um dos coeficientes é não-nulo, dizemos $a_{1n} \neq 0$ (caso fosse um outro renomeamos eles). Então

$$\left(x_1, \ldots, x_{n-1}, -\frac{a_{11}}{a_{1n}}x_1 - \cdots - \frac{a_{1,n-1}}{a_{1n}}x_{n-1}\right)$$

é uma solução para cada um $(x_1, \ldots, x_{m-1}) \in \mathbb{K}^{m-1}$.
$m - 1 \Rightarrow m$: Caso todos os coeficientes da última equação em $(*)$ são nulos, então as primeiras $m-1$ equações tem uma solução não-trivial x pela hipótese da indução (x também resolve a ultima equação: os coeficientes dela são nulos).

Suponha então que pelo menos um coeficiente da última equação em $(*)$ não é nulo, dizemos $a_{mn} \neq 0$. Nas primeiras $n-1$ equações de $(*)$ substitua x_n por

$$x_* := -\frac{a_{m1}}{a_{mn}}x_1 - \cdots - \frac{a_{m,n-1}}{a_{mn}}x_{n-1}$$

para obter um SLH de $m-1$ equações a $\tilde{n} := n - 1 > m - 1$ incógnitas. O qual tem uma solução $(x_1, \ldots, x_{n-1}) \neq (0, \ldots, 0)$ pela hipótese $m-1$ da indução. Verifica-se que $(x_1, \ldots, x_{n-1}, x_*)$ é uma solucao nao-trivial de $(*)$. □

B.4 Transformações lineares

Lema B.4.1 (Lema 4.3.10)**.** *Trabalhamos no plano Π identificado com $\mathbb{R}^2$ mediante um sistema ortogonal de coordenadas. A projeção ortogonal sobre a reta L_a é denotada de $P = P_{L_a} \colon \mathbb{R}^2 \to \mathbb{R}^2$ e dada por (4.3.2). Sua matriz é*

$$\mathbf{p}_a := [P_{L_a}] = \frac{1}{1+a^2}\begin{bmatrix} 1 & a \\ a & a^2 \end{bmatrix} \tag{B.4.1}$$

onde $[P_{L_a}] := [P_{L_a}]_{\mathcal{E},\mathcal{E}}$ denota a matriz em respeito à base canônica.

Demonstração. Seja $a \in \mathbb{R}$. Dado um elemento $v = (x, y) \in \mathbb{R}^2$, denota sua imagem sob P de $(X, Y) := Pv \in L_a = \{(x, ax) \mid x \in \mathbb{R}\}$. Assim X e Y são funções de (x, y) e $Y = aX$. Resta determinar a função $X(x, y)$. Vamos provar

$$X(x, y) = \frac{1}{1+a^2}\,x + \frac{a}{1+a^2}\,y, \qquad (x, y) \in \mathbb{R}^2 \tag{B.4.2}$$

Segundo o Teorema de Pitágoras a distância $\operatorname{dist}(\mathcal{O}, v)$ entre a origem $\mathcal{O} = (0, 0)$ e o vetor $v = (x, y)$ é dada por $\sqrt{x^2 + y^2}$. Assim, usando Pitágoras de novo na

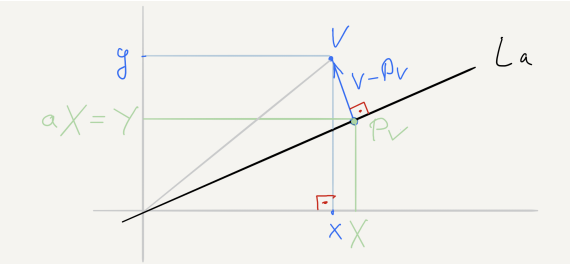


Figura B.1: Dois ângulos retângulos - usando Pitágoras duas vezes

igualdade dois, veja Figura B.1, obtemos

$$\begin{aligned} x^2+y^2 &= \operatorname{dist}(\mathcal{O},v)^2 \\ &= \operatorname{dist}(\mathcal{O},Pv)^2+\operatorname{dist}(v,Pv)^2 \\ &= X^2+(aX)^2+\big(\text{comprimento}^2 \text{ do vetor } v-Pv=(x-X,y-aX)\big) \\ &= X^2+(aX)^2+(x-X)^2+(y-aX)^2 \\ &= X^2+(aX)^2+x^2-2xX+X^2+y^2-2aYX+a^2X^2 \end{aligned}$$

o que é equivalente a

$$X(x,y)^2(1+a^2)=X(x,y)(x+ay)$$

Caso $X(x,y)\neq 0$**.** Divida por $X(x,y)$ e $1+a^2$ para obter (B.4.2).
Caso $X(x,y)=0$**.** Então $Y(x,y)=aX(x,y)=0$ e assim $Pv=(X,Y)=\mathcal{O}$. Como a projeção é ortogonal o ponto v deve ser localizado na reta $(L_a)^{\perp}_{\mathcal{O}}$ ortogonal a L_a e passando a origem. Mas esta reta resulta de L_a mediante uma rotação por $\pi/2$ $(90°)$, em símbolos

$$(L_a)^{\perp}_{\mathcal{O}}=R_{\pi/2}L_a=\left\{\begin{bmatrix}0&-1\\1&0\end{bmatrix}\begin{bmatrix}t\\at\end{bmatrix}:t\in\mathbb{R}\right\}=\{(-at,t)\mid t\in\mathbb{R}\}$$

Então v é da forma $(-at,t)$ para um $t\in\mathbb{R}$ e tal par satisfaz (B.4.2) também.

Para concluir note-se que os coeficientes na última identidade de

$$Pe_1=P(1,0)=(X(1,0),Y(1,0))=X(1,0)e_1+Y(1,0)e_2$$

disponibiliza a primeira coluna da matriz (B.4.1) e analogamente

$$Pe_2=X(0,1)e_1+Y(0,1)e_2$$

disponibiliza a segunda coluna. Para obter os valores de X e Y nos pontos $(1,0)$ e $(0,1)$ usa-se a fórmula (B.4.2). □

B.5 Existência de subespaço invariante ($\mathbb{K} = \mathbb{R}$)

Nesta seção suponhamos que A é um operador linear num espaço vetorial real E de dimensão finita $n \geq 1$.

Teorema B.5.1 (Teorema 9.3.1)**.** *Um operador linear A em E admite um subespaço invariante F de dimensão* 1 *ou* 2*.*

A prova do teorema precisa de preparação.

Definição B.5.2. Seja $\mathbb{K}$ um corpo. Um polinômio da forma

$$p = p(x) = a_0 + a_1 x + \cdots + 1 \cdot x^m \in \mathcal{P}(\mathbb{K})$$

é chamado de **mónico** (o coeficiente a_m da potencia máxima deve ser $1 \in \mathbb{K}$). Um polinômio não-constante p é chamado de **irredutível** se não é um produto

$$p \neq p_1 \dots p_k$$

de polinômios p_i de grau $< m$ e é chamado de **redutível** no caso contrário.

Observações:
1) Todo polinômio de grau 1 – chamado de **fator linear** – é irredutível.
2) Ser irredutível depende do corpo. Por exemplo, os polinômios de grau dois

$$x^2 - 1 = (x+1)(x-1), \qquad x^2 + 1 = (x+i)(x-i)$$

são redutível como polinômios complexos (elementos de $\mathcal{P}(\mathbb{C})$), mas só o primeiro é redutível como polinômio real (elemento de $\mathcal{P}(\mathbb{R})$).

Teorema B.5.3 (Teorema fundamental da álgebra)**.** *Todo polinômio p não-constante com coeficientes complexos tem uma raíz complexa, equivalentemente p pode ser escrito na forma*

$$p(x) = c(x - \alpha_1) \cdots (x - \alpha_m) \tag{B.5.1}$$

onde $c, \alpha_1, \dots, \alpha_m$ sao números complexos.

Demonstração. Para existência de uma raíz complexa veja [Art91, Thm. 9.1, p. 527] e para a representação equivalente como produto de fatores lineares veja [Art91, Ex. 1.6, p. 391] cuja prova usa [Art91, Cor. 3.20, p. 358]. □

Corolário B.5.4. *Um polinômio mónico real p de grau $m \geq 1$ fatoriza como produto $p = p_1 \dots p_k$ de polinômios mónicos reais irredutíveis p_i de grau* 1 *ou* 2*.*

Demonstração. Um polinômio real é complexo e assim p tem a forma (B.5.1) no teorema fundamental acima com $c = 1$ como p é mónico. Os α_i são complexos o que não exclui real. Se um α_i não é real o complexo conjugado $\bar{\alpha}_i$ também é uma raíz, assim igual a um dos α_j, porque p é um polinômio real. Mas o produto $(x - \alpha_i)(x - \bar{\alpha}_i) = x^2 - (\alpha_i + \bar{\alpha}_i)x + \alpha_i \bar{\alpha}_i$ é um polinômio real. Assim os fatores em (B.5.1) ou são linear e real ou são complexos mas combinam em pares cujo produto é um polinômio quadrático real (irredutível). □

Definição B.5.5. Dado um operador linear $A \in \mathcal{L}(E)$ e um polinômio real

$$p = p(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathcal{P}(\mathbb{R})$$

definimos um operador linear em E assim

$$p(A) := a_0 I_E + a_1 A + \cdots + a_n A^n \in \mathcal{L}(E)$$

Lema B.5.6. *Dado $A \in \mathcal{L}(E)$, então existe um polinômio mónico irredutível $q \in \mathcal{P}(\mathbb{R})$ de grau 1 ou 2 e existe um vetor não-nulo $v \in E$ tal que $q(A)v = \mathcal{O}$.*

Demonstração. Seja $n = \dim E$, então como o espaço vetorial $\mathcal{L}(E)$ tem dimensão n^2, veja Corolário 4.1.15, o conjunto de $n^2 + 1$ elementos

$$\{I_E, A, A^2, \ldots, A^{n^2}\}$$

é LD. Por isso existem coeficientes reais α_i, não todos nulos, tal que

$$\mathcal{O} = \alpha_0 I_E + \alpha_1 A + \cdots + \alpha_{n^2} A^{n^2}$$

Seja α_m o coeficiente não nulo do maior índice. O caso $m = 0$ é impossível como $\alpha_0 I_E = \mathcal{O}$ implicaria o absurdo $I_E = {\alpha_0}^{-1}\mathcal{O} = \mathcal{O}$. Então $m \geq 1$ e definindo $\beta_j := \alpha_j/\alpha_m$ obtemos que

$$\mathcal{O} = \beta_0 I_E + \beta_1 A + \cdots + \beta_{m-1} A^{m-1} + A^m =: p(A) \in \mathcal{L}(E)$$

O correspondente polinômio real

$$p(\lambda) := \beta_0 + \beta_1 x + \cdots + \beta_{m-1} x^{m-1} + x^m$$

é mónico e de grau $m \geq 1$. Segundo Corolário B.5.4 obtemos que

$$p = p_1 \ldots p_k$$

onde os p_i são polinômios mónicos reais irredutíveis de grau 1 ou 2. Como

$$\mathcal{O} = p(A) = p_1(A) \ldots p_k(A)$$

pelo menos um dos operadores na direita não é invertível, dizemos $p_i(A)$. (Caso contrario o operador nulo $\mathcal{O}$ é invertível – absurdo.) Daí $q(A) := p_i(A) \colon E \to E$ não é bijetivo, assim não injetivo, ou seja $\mathrm{N}(q(A)) \neq \{\mathcal{O}\}$. □

Demonstração de Teorema B.5.1 (Existe subespaço invariante de dim. 1 ou 2). Dado $A \in \mathcal{L}(E)$, segundo Lema B.5.6 existe um polinômio mónico irredutível $q \in \mathcal{P}(\mathbb{R})$ de grau 1 ou 2 e existe um vetor não-nulo $v \in E$ tal que

$$\mathcal{O} = q(A)v \qquad (*)$$

Caso q tem grau 1. Então q é da forma $q(x) = x - \lambda$ para um $\lambda \in \mathbb{R}$. Daí segue de $(*)$ que $\mathcal{O} = q(A)v = (A - \lambda I_E)v = Av - \lambda v$. Por linearidade de A a

reta $F := \mathbb{R}v$ é um subespaço invariante por A.

Caso q tem grau 2. O polinômio é da forma $q(x) = x^2+\alpha x+\beta$ para constantes $\alpha, \beta \in \mathbb{R}$ onde $\beta \neq 0$ como q é irredutível. (No caso contrario $q(x) = x(x + a)$ é redutível.) Então

$$\mathcal{O} \overset{(*)}{=} q(A)v = AAv + \alpha Av + \beta v \qquad (**)$$

Se $Av = \mathcal{O}$ obtemos a contradição $\mathcal{O} = \beta v$.

• $\{v, Av\}$ é LI: Suponha por absurdo que $Av = \mu v$ para um $\mu \in \mathbb{R} \setminus \{0\}$. Então obtemos que

$$\mathcal{O} \overset{(**)}{=} \mu^2 v + \alpha\mu v + \beta v = (\underbrace{\mu^2 + \alpha\mu + \beta}_{=q(\mu)}) \underbrace{v}_{\neq \mathcal{O}} \qquad \Rightarrow \qquad q(\mu) = 0$$

mas um polinômio irredutível de grau 2 não pode ter uma raíz real.

• O subespaço F gerado por $\{v, Av\}$ tem dimensão 2 e é invariante por A: Como A é linear é suficiente mostrar $Av \in F$ e $A(Av) \in F$. Com efeito

$$Av \in F, \qquad A(Av) \overset{(**)}{=} -\alpha(Av) - \beta v \in F$$

Isso finaliza a prova de Teorema B.5.1. □

B.6 Operadores ortogonais

Teorema B.6.1. *Para um operador $A \in \mathcal{L}(E, F)$ são equivalente*

(i) $|Av| = |v|$ $\forall v \in E$ — *"A preserva norma"*

(ii) $|Au - Av| = |u - v|$ $\forall u, v \in E$ — *"A preserva distância"*

(iii) $\langle Au, Av\rangle = \langle u, v\rangle$ $\forall u, v \in E$ — *"A preserva produto interno"*

(iv) $A^*A = I_E$ — *"A^* é inversa à esquerda de A"*

(v) *a matriz $[A]_{\mathcal{X},\mathcal{Y}}$ é ortogonal se $\mathcal{X}$ e $\mathcal{Y}$ são bases ONs*

(vi) *$A\mathcal{X}$ é um subconjunto ON de F se $\mathcal{X}$ é uma base ON de E*

Demonstração. □

Referências Bibliográficas

[Art91] Michael Artin. *Algebra*. Prentice Hall, Inc., Englewood Cliffs, NJ, 1991.

[Hir21] Oliver Hirsch. Die Psychologie der Gedankenkontrolle, des Mentizids und der Gehirnwäsche. Zugang pdf, January 2021.

[Koe85] Max Koecher. *Lineare Algebra und analytische Geometrie*. Grundwissen Mathematik 2. Springer-Verlag, Berlin Heidelberg New York Tokyo, Zweite Auflage,

[Lan93] Serge Lang. *Algebra. 3rd ed.* Reading, MA: Addison Wesley, 1993.

[Lim05] Elon Lages Lima. *Geometria Analítica e Álgebra Linear*. Coleção Matemática Universitária. Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, Segunda Edição, 2005.

[Lim11] Elon Lages Lima. *Álgebra Linear*. Coleção Matemática Universitária. Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, Oitava Edição, 2011.

[Mee56] Joost A.M. Meerloo. *The Rape of the Mind – the Psychology of Thought Control*. 1956. access pdf.

[Pul12] Petronio Pulino. *Álgebra Linear e suas Aplicações*. Notas da Aula, UNICAMP, 2012. Acessível no site www.ime.unicamp.br/∼pulino/ALESA.

[San12] Reginaldo J. Santos. Matrizes, Vetores e Geometria Analítica. Manuscrito, UFMG, 03 2012.

Índice Remissivo

$A > 0$ operador positivo, 138
$A \geq 0$ operador não-negativo, 138
$A\,|_F$ restrição de $A \in (X)$, 98
A^* adjunta, 126
$A^\dagger$ adjunta complexa, 151
A^{-1} inversa, 46, 79
$Av := A(v)$ operador linear, 43
$\mathcal{A} = \mathcal{A}(n)$ matrizes anti-simétricas, 28, 40, 129
$\sqrt{A}$ raíz quadrada positiva de $A > 0$, 139
$A_f^{\mathcal{B}} = A_f \in \mathcal{L}(E, F)$, 47
$AX := \{Ax \mid x \in X\}$ imagem do conjunto X sob A, 73
$\mathbb{C} \ni c = \alpha + i\beta$ números complexos, 149
$C^0(\mathbb{R})$ funções contínuas $\mathbb{R} \to \mathbb{R}$, 24
$C^k(\mathbb{R})$ funções k vezes continuamente diferenciáveis, 24
$C^\infty(\mathbb{R})$ funções suaves $\mathbb{R} \to \mathbb{R}$, 24
CL combinação linear, 18
$(E, +, \cdot, \mathbb{K})$ espaço vetorial, 13
E_λ autosubespaço, 99
$\mathcal{E}^n := \{e_1, \ldots, e_n\}$ base canônica, 15, 26, 30
$\mathcal{E}^\infty := \{e_1, e_2, \ldots\}$ base canônica, 15, 26, 30
$\mathcal{E}^{m\times n} := \{\mathbf{e^{ij}}\}_{i,j}$ base canônica, 30
Esp-col($\mathbf{a}$), 16, 156
$\mathcal{F}(X, \mathbb{K}) := \{f \mid f : X \to \mathbb{K}\}$, 17
Fix(r) conjunto dos pontos fixos, 86
aFix(r) conjunto dos pontos anti-fixos, 86
$(G, *)$ grupo, 8
H_α hiperplano no $\mathbb{R}^n$, 24, 31, 83
Im(A) imagem, 73
$\mathrm{Im}(\alpha + i\beta) := \beta$ parte imaginário, 149
$I = I_E$ operador identidade, 45
$(\mathbb{K}, +, \cdot)$ corpo, 9
$\mathbb{K}v$ reta passando v e $\mathcal{O}$, 24
LI/LD linearmente in/dep., 19
$\mathcal{L}(E, F)$ operadores lineares, 45
$\mathcal{L}(E)$ operadores lineares em E, 45
$\mathrm{M}(m \times n)$ matrizes $m \times n$, 16, 28, 30, 155
$\mathcal{A}, \mathcal{S}$ matrizes anti-/simétricas, 28, 40, 129
N(A) núcleo, 73
O(n) grupo ortogonal, 144
$\mathcal{O}$ vetor nulo, 13
(oe) operações elementares, 157
$\mathcal{P}(\mathbb{K})$ polinômios, 17
$\mathcal{P}(\mathbb{R}), \mathcal{P}_n(\mathbb{R})$ polinômios reais e aqueles do grau $\leq n$, 24, 31
$P = P_{L_a} \in \mathcal{L}(\mathbb{R}^2)$ projeção ortogonal sobre a reta L_a, 57
$\mathbf{p}_a \in \mathcal{L}(\mathbb{R}^2)$ matriz da projeção ortogonal, 58
$P_{F,G} : E \to E$ projeção sobre F, 87
$\mathrm{Re}(\alpha + i\beta) := \alpha$ parte real, 149
$\mathbb{R}^n$ listas ordenadas de n reais, 15, 30
$\mathbb{R}^\infty$ sequências reais, 15
$\mathbb{R}_0^\infty$ quase todos membros nulos, 15, 24, 30
$R_\theta \in \mathcal{L}(\Pi_O)$ rotação no plano, 55
$\mathbf{r}_\theta \in \mathcal{L}(\mathbb{R}^2)$ matriz da rotação, 57
$\mathbb{S}^n \subset \mathbb{R}^{n+1}$ esfera unitária, 144
$\mathcal{S} = \mathcal{S}(n)$ matrizes simétricas, 28, 40, 129, 134
$S = S_{L_a} \in \mathcal{L}(\mathbb{R}^2)$ reflexão em torno da reta L_a, 58
$\mathbf{s}_a \in \mathcal{L}(\mathbb{R}^2)$ matriz da reflexão, 58

$S^2 = I_E : E \to E$ involução, 85
$S_{F,G} := P_{F,G} - P_{G,F}$ involução (ou reflexão), 89
$\mathcal{SC}(E) = \{(F, G) \mid F \oplus G = E\}$, 85
SL sistema linear, 75, 160
SLH sistema linear homogêneo, 34, 160, 167
TL transformação (ou operador) linear, 43
$\mathbb{1} = \mathbb{1}_n = \mathrm{diag}(1, \ldots, 1) \in \mathrm{M}(n \times n; \mathbb{K})$ matriz identidade, 16, 156
$\forall$, $\exists$, $\exists!$ "para todos", "existe", "existe unicamente", 3
$\mathrm{alg}_\lambda(A)$ multiplicidade algébrica, 102
$\rightarrowtail$ injetivo, 85
$\twoheadrightarrow$ sobrejetivo, 85
$:=$ "definido por", 3
$\simeq$ isomorfismo, 79
$\langle X \rangle$ subespaço gerado por X, 25
$\langle v_1, \ldots, v_\ell \rangle := \langle \{v_1\} \cup \cdots \cup \{v_\ell\} \rangle$, 25
$|X|$ número de elementos de um conjunto X, 7
$|\alpha|$ absoluto de um número α, 3
$\mathbf{a} = (a_{ij})$ matriz, 16, 155
$\mathbf{a}^t = ((\mathbf{a}^t)_{ij} = a_{ji})$ matriz transposta, 156
$\mathbf{a}_{\bullet k}, \mathbf{a}_{k\bullet}$ k-ésima coluna, linha, 16, 156
$\mathbf{a}_{\mathrm{esc-red}}$ matriz escalonada reduzida, 159
$\mathbf{a}_{\mathrm{esc}}$ matriz escalonada, 158
a_{ij} i-ésima linha, j-ésima coluna, 16, 155
$\det A$ onde $A \in \mathcal{L}(E)$, 68
$\det \mathbf{a}$ determinante da matriz quadrada $\mathbf{a}$, 162
δ_{ij} símbolo de Kronecker, 48
$\mathrm{diag}\,[\lambda_1, \ldots, \lambda_n]$ matriz diagonal, 97
$e_i \in \mathbb{K}^n$ i-ésimo vetor canônico, 15
$\mathbf{e^{ij}}_{\mp} = \frac{1}{2}(\mathbf{e^{ij}} \mp (\mathbf{e^{ij}})^t) \in \mathcal{A}/\mathcal{S}$, $i \leq j$, 40
E^* espaço dual de E, 48
$\mathbf{g}_\mathcal{B} = [g]_\mathcal{B}$ matriz do produto interno, 114
$g_\lambda(A) := \dim E_\lambda$ multiplicidade geométrica, 99
$[a, b]$, (a, b) intervalo fechado, aberto, 3
$[\mathbf{a} : b]$ matriz aumentada, 160
$[v]_\mathcal{B}$ vetor coordenada do vetor v na base $\mathcal{B}$, 33
$[v]_\mathcal{B}$ vetor coordenada do vetor v na base $\mathcal{B}$, 62
$[v] := [v]_{\mathcal{E}^m}$ na base canônica, 33
$[v] := [v]_{\mathcal{E}^m}$ na base canônica, 62
$[A]_{\mathcal{U},\mathcal{V}}$ matriz de $A \in \mathcal{L}(E, F)$, 63
$[A] = [A]_{\mathcal{E}^n,\mathcal{E}^m}$ nas bases canôn., 51
$[A]_\mathcal{U} := [A]_{\mathcal{U},\mathcal{U}}$, 63
$[A] := [A]_\mathcal{E}$ na base canônica, 63
$p_A(\lambda)$ caso $\dim E = 2$, 103
$p_A(\lambda)$ polinômio característico de A, 102
$\langle \cdot, \cdot \rangle_0$ produto euclidiano em $\mathbb{R}^n$, 110
$\langle \cdot, \cdot \rangle_*$ produto interno em E^*, 112
$\mathrm{posto}(A) := \dim \mathrm{Im}(A)$, 74
pr_F projeção ortogonal sobre subespaço F, 121
pr_u projeção ortogonal sobre reta $\mathbb{R}u$, 117
$\mathrm{pc}(\mathbf{a})$ posto-coluna, 50
$\mathrm{pl}(\mathbf{a})$ posto-linha, 50
$p(A)$ polinômio do operador A, 105, 171
$\times$, $v \times w$ produto vetorial, 70
$\times$, $X \times Y$ produto cartesiano, 8
$Y^{\times k} := Y \times \cdots \times Y$, 8, 31
$\mathrm{spec}\, A$ spectro de $A \in \mathcal{L}$, 98
$F \oplus G$ soma direta, 27
$X + Y$ soma de subconjuntos, 27
$\mathrm{tr}\, \mathbf{a} := \sum_{i=1}^n a_{ii}$ traço, 40, 71, 129
$\dot\cup$ união de conjuntos disjuntos, 7
$\hat{v}$ vetor unitário, 111
$|z| = \sqrt{a^2 + b^2}$ absoluto de $z = a + ib$, 10
$\bar{z}$ complexo conjugado, 149
$\bar{z} = a - ib$ complexo conjugado de $z = a + ib$, 10
(Δ) desigualdade triangular, 111
(POS) positividade, 111
(SIM) simetria, 111
(Δ) desigualdade triangular, 111
(HOM) homogeneidade, 111
(POS) positividade, 111
(SIM) simetria-cc, 150
(BL) bi-linearidade, 109

`(POS)` positividade, 109, 150
`(SIM)` simetria, 109
`(SL)` sesquilinearidade, 150

absoluto
  de um número
    complexo, 10
adição
  de funções, 17
adjunta, 126
  complexa $A^\dagger$, 151
auto-adjunto, 133
autovalor, 98
  multiplicidade algébrica, 102
  multiplicidade geométrica, 99
autovetor, 98

base, 29
  canônica, 15, 26, 30
    das matrizes $m \times n$, 30
  extensão, 120
  ordenada, 29, 46
  ortonormal (ON), 116
bijetivo, 46, 79

canônico, 112
caraterística de um corpo, 10
combinação linear, 18
  trivial, 18
complemento ortogonal, 121
complexo conjugado, 10, 149
composição
  de funções, 13
comutar
  matrizes, 157
comutativo
  diagrama –, 61
conjunto, 7
  composto de elementos $x_1, \ldots, x_\ell$, 3, 7
  finito, 7
  linearmente independente LI, 19
  ordenado, 7
  que gera, 25
conjuntos
  interseção de –, 7
  união de –, 7
convolução, 45
coordenadas
  de um vetor, 33
  no plano, 2
corpo, 9
  caraterística do –, 10

decomposição
  de vetores, 27
Descartes, 3
determinante
  de uma matriz quadrada, 162
  de uma transformação linear, 68
diagonalizável, 97, 101, 141
diagrama
  comutativo, 61
dimensão, 30
distância, 111
  entre dois pontos, 111
  no $\mathbb{R}^2$, 117
dualidade, 112

eixo, 2
elemento neutro
  aditivo, 9
  multiplicativo, 9
escalar, 156
escalares, 13
espaço dual E^*, 48
espaço métrico, 111
espaço vetorial, 13
  base, 29
  com produto hermitiano, 150
  com produto interno, 109
  dimensão, 30
  normado, 111
  real ($\mathbb{K} = \mathbb{R}$), 49
  subespaço, 23
  trivial, 14
espaço-coluna, 16, 156
espaço-linha, 17, 156
extensão
  base ON, 120

fechado sob uma operação, 23

flechas
  equipolentes, 1
flechas equipolentes, 1
funções
  adição de –, 17
  composição de –, 13
  multiplicação de –, 13
funcional
  $\mathbb{K}$-linear, 48
  linear, 48
    real, 48

gráfico, 91
Gram-Schmidt (GS), 118
grau
  de um polinômio, 17
grupo, 8
  abeliano, 9
  ortogonal, 144

hermitiano, 152
hiper, 31
hiperplano, 24, 31, 83
homomorfismo, 43
homotetia, 64

imagem, 73
independência linear, 18
injetivo, 46, 85
interseção de conjuntos, 7
invariante
  subespaço –, 98
inversa, 79
  à direita, 76
  à esquerda, 78
  de um operador linear, 46
invertível, 79
  matriz –, 17
  transformação linear, 46
involução, 85
  $S_{F,G}$, 89
  $S_{F,G}$ em torno de F, 89
  linear, 88
isometria, 143
isomorfismo, 46, 79
  inversa, 46

Kronecker
  símbolo de –, 48

lei
  da corte, 8
linearidade, 43
linearmente in/dependente, 19

métrica, 111
  induzida, 111
mónico
  polinômio –, 17
matriz, 155
  adição, 155
  anti-/simétrica, 28, 40
  aumentada, 160
  complexa conjugada, 150
  de passagem, 65
  de passagem entre bases, 61
  de uma transformação linear, 51, 63
  entradas da –, 16, 155
  escalonada, 158
    pivôs, 158
  escalonada reduzida, 159
  hermitiana, 150
  identidade $\mathbb{1}$, 16, 156
  invertível, 17
  linhas e colunas, 16, 156
  multiplicação escalar, 156
  operações elementares numa –, 157
  positiva, 114
  produto –, 156
  projeção ortogonal, 58
  quadrada, 155
  reflexão, 58
  rotação, 57
  traço de uma – quadrada, 40, 71, 129
  transposta $\mathbf{a}^t$, 156
  unitária, 152
matriz ortogonal, 143
matrizes
  anti-simétricas, 28, 40, 129
  comutam, 157

semelhante, 67
simétricas, 28, 40, 129, 134
triangulares
inferiores, 40
monômios, 26, 31
multiplicação
de funções, 13
multiplicação escalar, 13
multiplicidade
algébrica, 102
geométrica, 99

núcleo, 73
número
complexo, 10, 149
não-negativo
operador auto-adjunto –, 138
nilpotente
operador –, 151
norma, 111
induzida, 111
normal
operador –, 151

operações elementares numa matriz, 157
operador
auto-adjunto
não-negativo, 138
positivo, 138
hermitiano, 152
identidade, 45
normal, 130, 151
ortogonal, 143
unitário, 152
operador (linear)
= transformação linear, 4
operador linear, 43
auto-adjunto, 133
em E, 45
inversa, 46
operador linear – veja **transformação linear**, 97
operador ortogonal, 145
origem, 15
ortogonal
complemento –, 121
matriz –, 143
operador –, 143, 145
subconjunto –, 116
vetores, 116
ortonormal (ON)
base –, 116
subconjunto –, 116

par
de subespaços complementares, 85, 87
parte
imaginária, 10
real, 10
perpendicular
vetores, 116
pivôs, 158
polinômio, 17, 24
de grau $\leq n$, 24
grau de um –, 17
irredutível, 170
mónico, 17, 170
raíz, 102
determinar, 104
polinômio característico, 102, 103
ponto
anti-fixo, 86
fixo, 86
positiva
matriz –, 114
positivo
operador auto-adjunto –, 138
posto
coluna, 50
linha, 50
transformação linear, 74
produto
cartesiano, 8, 31
escalar (= interno), 109
euclidiano, 110
hermitiano ($\mathbb{C}$-interno), 150
interno, 109
matriz, 156
vetorial $\times$, 70
produto interno

induzido em E^*, 112
projeção, 85, 87
$P_{F,G}$ sobre F, 87, 134
ortogonal, 57, 134
sobre reta, 117
sobre subespaço, 121
projeção ortogonal
matriz da –, 58

raíz de um polinômio, 102
determinar, 104
raíz quadrada positiva de $A \geq 0$, 139
reflexão
$S_{F,G}$, 89
em torno de uma reta, 58
matriz da –, 58
relação
de equivalência, 80
restrição, 98
rotação, 55
matriz de –, 57
no plano, 1

semelhante
matrizes –, 67
sistema de
coordenadas
no plano, 2
sistema de coordenadas, 62
Cartesianas, 3
sistema de geradores, 25
sistema linear (SL), 75, 160
inomogeneidade, 160
resolver "de baixo para cima", 161
sistema linear homogêneo (SLH), 34, 160
sobrejetivo, 46, 85
soma
de subconjuntos, 27
direta, 27
spectro, 98
subconjunto
ortogonal, 116
ortonormal (ON), 116
translação de –, 27
subconjuntos
herdam LI, 32
soma de –, 27
subespaço
invariante, 98
sistema de geradores, 25
vetorial, 23
subespaços
complementares, 85, 87
soma direta de –, 27

teorema
decomposição única de vetores, 27
traço, 40, 71, 129
transformação linear
= operador (linear), 4
adjunta, 126
auto-adjunto, 133
diagonalizável, 97
gráfico, 91
inversa, 79
invertível, 79
matriz de uma –, 63
restrição, 98
transformação linear (TL), 43
translação, 27
transposta, 156

união de conjuntos, 7
unitária
matriz –, 152
unitário
operador –, 152

verdade vazia, 19
vetor
como colecao de flechas, 1
decomposição, 27
unitário, 111
vetor coordenada, 33, 62
vetor nulo, 13
vetores, 13
ortogonais/perpendicular, 116