Álgebra Linear

Notas de aula

Versão de 10 de outubro de 2026

Pedro Lauridsen Ribeiro[1]

Centro de Matemática,
Computação e Cognição (CMCC)
Universidade Federal do ABC (UFABC)

[1]Email: pedro.ribeiro@ufabc.edu.br

Conteúdo

Notação, Convenções e Roteiro **iv**

I Espaços vetoriais **1**
1 Motivação 1
2 Definição e exemplos 8
2.1 Axiomas de um espaço vetorial 8
2.2 Exemplos 12
2.3 Subespaços vetoriais 15
3 (In)dependência linear 20
3.1 Prelúdio: somatórias num espaço vetorial 20
3.2 Combinações lineares e (in)dependência linear 23
3.3 Bases e dimensão 28

II Produtos escalares **37**
1 Definição e exemplos 37
1.1 O produto escalar canônico de $\mathbb{R}^n$ 37
1.2 Axiomas do produto escalar 37
2 Ortogonalidade e ortonormalidade 39
3 Geometria do produto escalar 41
3.1 Normas e ângulos 41
3.2 Projeções ortogonais 44
3.3 Ortonormalização de Gram-Schmidt 46

III Transformações lineares **49**
1 Definição e exemplos 49
2 Espaços vetoriais e produto de transformações lineares 51
3 A matriz de uma transformação linear 53
3.1 Valores de uma transformação linear numa base 53
3.2 Multiplicação de matrizes 54
3.3 Mudança de base 58
4 O Teorema do Núcleo e da Imagem 59
5 Pseudoinversas 63

IV Sistemas lineares **65**

V Determinantes, autovalores e autovetores **67**
1 Determinantes 67
2 Autovalores e autovetores 67

3 Decomposição em valores singulares 67

A Números complexos 69

B Permutações 71

Bibliografia 73

Prefácio

Estas são notas de aula para a disciplina MCTB001 – Álgebra Linear, ministrada no terceiro quadrimestre letivo de 2026.

Comentários, correções, sugestões e críticas dos estudantes e/ou leitores são bem-vindas e até encorajadas. Quaisquer destas podem ser enviadas com prazer ao autor pelo email

pedro.ribeiro@ufabc.edu.br.

As presentes Notas serão mantidas em estado de movimento – atualizações serão feitas com frequência, e a última versão poderá ser encontrada na minha página Web

https://pedroribeiro.prof.ufabc.edu.br/

São Bernardo do Campo, 10 de outubro de 2026
Pedro Lauridsen Ribeiro

Notação, Convenções e Roteiro

I

Espaços vetoriais

1. Motivação

O conceito de *vetor* origina-se em Geometria e em Física – a saber, quantidades como deslocamento, velocidade, aceleração, força, etc.. são descritas pelo ente geométrico que conhecemos como vetor. Dessas quantidades, talvez a noção de *deslocamento* seja a que expressa de maneira mais elementar o conceito de vetor, portanto usar-la-emos como motivação para tal conceito.

Veremos ao elaborar a noção de deslocamento que vetores são melhor caracterizados não tanto pela sua interpretação geométrica, mas pelas *operações* que podemos aplicar a tais objetos.

Consideremos dois conjuntos não-vazios E, V, que serão nossa arena para uma forma rudimentar de geometria que iremos refinar de maneira gradual:

- E = espaço de *pontos* $A, B, C, \ldots$;
- V = espaço de *vetores de deslocamento* $\vec{x}, \vec{y}, \vec{z}, \ldots$.

A noção de *deslocamento* de um ponto $A \in E$ a outro ponto $B \in E$ é fornecida abstratamente por uma função

$$\vec{d} : E \times E \ni (A, B) \mapsto \vec{d}(A, B) = \overrightarrow{AB} \in V$$

– a chamada *função deslocamento* ou *função translação* de E – que assumimos satisfazer as seguintes propriedades:

(i) (unicidade do deslocamento a partir de um ponto inicial dado) Dados $A \in E$, $\vec{x} \in V$, existe um *único* ponto $B \in E$ tal que $\overrightarrow{AB} = \vec{x}$. Ou seja, para cada $A \in E$ a função

$$\vec{d}(A, \cdot) : E \ni B \mapsto \vec{d}(A, B) = \overrightarrow{AB}$$

é *bijetora*. Denotamos esse ponto B por

$$B = A + \vec{x} = \text{ } \textit{translação} \text{ ou } \textit{deslocamento} \text{ de } A \text{ por } \vec{x} \text{ ,}$$

de modo que $B = A + \overrightarrow{AB}$ para $A, B \in E$ quaisquer. Nesse caso, dizemos também que A é o *ponto inicial* do deslocamento e B, o *ponto final* do deslocamento. Em particular, se $\overrightarrow{AB} = \overrightarrow{AC}$ então $B = C$.

(ii) (Lei do Paralelogramo) Dados $A, B, C, D \in A$ quaisquer, temos que

$$\overrightarrow{AB} = \overrightarrow{CD} \text{ se e somente se } \overrightarrow{AC} = \overrightarrow{BD} \text{ .}$$

Podemos entender tal noção de deslocamento como sendo "em linha reta" – a discussão a seguir dará mais substância a essa interpretação. Notar que a operação de translação de um ponto por um vetor só é bem definida graças a (i).

A propriedade (i) nos permite identificar os conjuntos E e V mediante a escolha de um ponto arbitrário $O \in E$, denominado *origem*, através da correspondência biunívoca

$$E \ni A \mapsto \overrightarrow{OA} \in V \; ; \quad V \ni \vec{x} \mapsto O + \vec{x} \in E \; .$$

De fato, como vimos acima,

$$A = O + \overrightarrow{OA} \; , \quad \vec{x} = \overrightarrow{O(O+\vec{x})} \; .$$

A propriedade (ii), por sua vez, pode ser entendida como uma noção de *paralelismo*: deslocamentos a partir de dois pontos iniciais dados são iguais se e somente se o deslocamento entre esses dois pontos iniciais for igual ao deslocamento entre os dois pontos finais, tal como lados opostos de um mesmo paralelogramo (daí o nome).

A noção de deslocamento nos permite definir uma operação binária em V

$$V \times V \ni (\vec{x}, \vec{y}) \mapsto \vec{x} + \vec{y} \in V \; ,$$

denominada *soma vetorial*. A saber, dados $\vec{x}, \vec{y} \in V$ e $A \in E$, definimos

$$A + (\vec{x} + \vec{y}) = (A + \vec{x}) + \vec{y} \; .$$

Equivalentemente, se $B = A + \vec{x}$ e $C = B + \vec{y}$, de modo que $\vec{x} = \overrightarrow{AB}$ e $\vec{y} = \overrightarrow{BC}$, temos que $\vec{x} + \vec{y} = \overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$. Ou seja, a soma vetorial é o vetor de deslocamento obtido após dois deslocamentos sucessivos.

A pergunta que precisamos responder agora é se a definição da soma vetorial dada acima depende da escolha do ponto inicial A. Mostraremos agora que *não* é o caso, justificando com isso a nossa escolha de notação. De fato, dados $A, A' \in E$ e $\vec{x} \in V$, definamos $B = A + \vec{x}$, $C = B + \vec{y}$, $B' = A' + \vec{x}$ e $C' = B' + \vec{y}$ – precisamos mostrar que $\overrightarrow{AC} = \overrightarrow{A'C'}$. Com efeito,

$$\vec{x} = \overrightarrow{AB} = \overrightarrow{A'B'} \; , \quad \vec{y} = \overrightarrow{BC} = \overrightarrow{B'C'} \; ,$$

de modo que (ii) implica

$$\overrightarrow{AA'} = \overrightarrow{BB'} = \overrightarrow{CC'}$$

e portanto

$$\overrightarrow{AC} = \overrightarrow{A'C'} \; ,$$

como desejado.

A operação de soma vetorial em V satisfaz as seguintes propriedades:

(a) (*comutatividade*) Dados $\vec{x}, \vec{y} \in V$ quaisquer, temos que $\vec{x} + \vec{y} = \vec{y} + \vec{x}$;

(b) (*associatividade*) Dados $\vec{x}, \vec{y}, \vec{z} \in V$ quaisquer, temos que $(\vec{x}+\vec{y})+\vec{z} = \vec{x}+(\vec{y}+\vec{z})$;

(c) (*existência de elemento neutro*) Existe um vetor $\vec{0} \in V$ tal que $\vec{x} + \vec{0} = \vec{x}$ para todo $\vec{x} \in V$. Notar que se $\vec{0}, \vec{0}' \in V$ são tais que $\vec{x} + \vec{0} = \vec{x} + \vec{0}' = \vec{x}$ para todo $\vec{x} \in V$, tomando $\vec{x} = \vec{0}$ resulta em $\vec{0} = \vec{0} + \vec{0}' \overset{(a)}{=} \vec{0}$. Logo, um vetor $\vec{0} \in V$ que satisfaça (c) é *único*, devido a (a) – denominamos esse vetor o *vetor nulo* ou *(vetor) zero* de V;

(d) (*existência de oposto*) Dado $\vec{x} \in V$, existe um vetor $-\vec{x} \in V$ tal que $\vec{x} + (-\vec{x}) = \vec{0}$. Notar que se existem $\vec{y}, \vec{y}' \in V$ tais que $\vec{x} + \vec{y} = \vec{x} + \vec{y}' = \vec{0}$, podemos escrever

$$\vec{y}+(\vec{x}+\vec{y}) = \vec{y}+\vec{0} \overset{(c)}{=} \vec{y} \overset{(c)}{=} \vec{y}+(\vec{x}+\vec{y}') \overset{(b)}{=} (\vec{y}+\vec{x})+\vec{y}' \overset{(a)}{=} (\vec{x}+\vec{y})+\vec{y}' = \vec{0}+\vec{y}' \overset{(a)}{=} \vec{y}'+\vec{0} \overset{(c)}{=} \vec{y}' .$$

Logo, para cada $\vec{x} \in V$ um vetor $-\vec{x}$ satisfazendo (d) é *único* – denominamos esse vetor o *oposto* de $\vec{x}$ ou simplesmente *menos* $\vec{x}$. A propriedade (d) nos permite definir a *subtração* $\vec{x} - \vec{y} = \vec{x} + (-\vec{y})$ para $\vec{x}, \vec{y} \in V$ quaisquer.

Vejamos como (i)–(ii) implicam (a)–(d).

(a) Dados $A \in E$, $\vec{x}, \vec{y} \in V$, definimos $B = A + \vec{x}$, $C = B + \vec{y} = A + (\vec{x} + \vec{y})$ e $B' = A + \vec{y}$, $C' = B' + \vec{x} = A + (\vec{y} + \vec{x})$, de modo que

$$\vec{x} = \overrightarrow{AB} = \overrightarrow{B'C'} , \quad \vec{y} = \overrightarrow{BC} = \overrightarrow{AB'}$$

e portanto, devido a (ii),

$$\overrightarrow{AB} = \overrightarrow{B'C} = \overrightarrow{B'C'} .$$

A propriedade (i), por sua vez, implica daí que $C = C'$ e portanto $\overrightarrow{AC} = \vec{x} + \vec{y} = \vec{y} + \vec{x} = \overrightarrow{AC'}$, como desejado.

(b) Dados $A \in E$, $\vec{x}, \vec{y}, \vec{z} \in V$, temos pela definição de soma vetorial que

$$A + ((\vec{x} + \vec{y}) + \vec{z}) = (A + (\vec{x} + \vec{y})) + \vec{z} = ((A + \vec{x}) + \vec{y}) + \vec{z}$$

e

$$A + (\vec{x} + (\vec{y} + \vec{z})) = (A + \vec{x}) + (\vec{y} + \vec{z}) = ((A + \vec{x}) + \vec{y}) + \vec{z} ,$$

de modo que de fato $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$.

(c) O zero consiste num deslocamento *trivial*, ou seja, não saímos do ponto inicial dado. Mais precisamente, dado $A \in E$, definimos $\vec{0} = \overrightarrow{AA}$, de modo que (ii) implica trivialmente que $\vec{0} = \overrightarrow{AA} = \overrightarrow{A'A'}$ para $A, A' \in E$ quaisquer. Veremos agora que tal $\vec{0}$ satisfaz (c) – de fato, dado $\vec{x} \in V$, escrevendo $B = A + \vec{x}$ temos que $\vec{x} + \vec{0} = \overrightarrow{AB} + \overrightarrow{BB} = \overrightarrow{AB} = \vec{x}$, como desejado.

(d) Dados $\vec{x} \in V$, $A \in E$, o oposto de $\vec{x}$ corresponde ao deslocamento *no sentido oposto*. Mais precisamente, escrevendo $B = A + \vec{x}$, definimos $-\vec{x} = \overrightarrow{BA}$. Tal definição é imediatamente sugerida pela prova de (c) apresentada acima, e prontamente implica (d): $\vec{x} + (-\vec{x}) = \overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \vec{0}$.

I.1 Observação. Conversamente, a fórmula $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$ e a comutatividade (a) da soma vetorial juntamente com a propriedade (i) implicam a Lei do Paralelogramo (ii). De fato, sejam $A, B, C, D \in E$ tais que $\overrightarrow{AB} = \overrightarrow{CD} = \vec{x}$. Definindo $\vec{y} = \overrightarrow{AC}$, seja (por (i)) o ponto $D' = B + \vec{y}$, de modo que $\overrightarrow{AC} = \overrightarrow{BD'} = \vec{y}$. Temos então que comutatividade (a) e as fórmulas $\overrightarrow{AD'} = \overrightarrow{AB} + \overrightarrow{BD'}$, $\overrightarrow{AD} = \overrightarrow{AC} + \overrightarrow{CD}$ implicam

$$\vec{x} + \vec{y} = \overrightarrow{AB} + \overrightarrow{BD'} = \overrightarrow{AD'} \overset{(a)}{=} \vec{y} + \vec{x} = \overrightarrow{AC} + \overrightarrow{CD} = \overrightarrow{AD} \ ,$$

de modo que, devido a (i), temos $D' = D$ e portanto $\overrightarrow{AC} = \overrightarrow{BD}$. Trocando os papeis de B e C, obtemos por um argumento análogo que $\overrightarrow{AC} = \overrightarrow{BD}$ implica por (i) e (a) que $\overrightarrow{AB} = \overrightarrow{CD}$. Notar, contudo, que a validade da fórmula $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$ é ligada ao fato de que a operação de soma vetorial não depende do ponto inicial A, algo que provamos acima usando (i) e (ii). Isso mostra que a Lei do Paralelogramo é intimamente atrelada à comutatividade da soma vetorial em V tal como a definimos. Notamos, nesse contexto, que é comum na literatura denominar a propriedade de comutatividade (a) da soma vetorial também por "lei do paralelogramo".

I.2 Observação. Em vista da identificação de E com V obtida mediante uma escolha de origem $O \in E$ e da propriedade (c), vemos que tal ponto é identificado justamente com $\vec{0} = \overrightarrow{OO}$ nesse caso. Isso nos mostra que a única diferença entre E e V é que V possui uma escolha *canônica* de origem, determinada pela propriedade (c).

As propriedades (a)–(d) implicam que a soma vetorial tem exatamente o mesmo comportamento que a soma de números reais, e portanto lidaremos com a primeira da mesma maneira que fazemos com a segunda – somas de mais de dois vetores podem ser escritas em qualquer ordem e com qualquer inserção de parênteses sem que o resultado mude.

Em particular, notar que para cada $n \in \mathbb{N}$, $\vec{x} \in V$ existe um vetor $n\vec{x} \in V$ definido aplicando-se o deslocamento por $\vec{x}$ n vezes sucessivas. Mais precisamente, $n\vec{x}$ é definido recursivamente pela fórmula

$$1\vec{x} = \vec{x} \ , \quad (n+1)\vec{x} = n\vec{x} + \vec{x} \ .$$

Em particular, obtemos daí que $n\vec{0} = \vec{0}$ para $n \in \mathbb{N}$ qualquer (de fato, $1\vec{0} = \vec{0}$. Supondo que a hipótese de indução em n (HI) $n\vec{0} = \vec{0}$ é válida, provemos a sua validade com n substituído por $n+1$: $(n+1)\vec{0} = n\vec{0} + \vec{0} = n\vec{0} = \vec{0}$). Comutatividade

(a) e associatividade (b) mostram que os n deslocamentos por $\vec{x}$ podem ser somados em qualquer ordem sem que o resultado mude, de modo que obtemos as seguintes propriedades: dados $\vec{x}, \vec{y} \in V$, $n, m \in \mathbb{N}$ quaisquer, temos que

(e) $m\vec{x}+n\vec{x} = (m+n)\vec{x}$; (de fato, temos por definição que $(m+1)\vec{x} = m\vec{x}+\vec{x} = m\vec{x}+1\vec{x}$. Supondo que a hipótese de indução em n (HI) $m\vec{x} + n\vec{x} = (m+n)\vec{x}$ é válida, provemos a sua validade com n substituído por $n+1$: $m\vec{x} + (n+1)\vec{x} = m\vec{x} + (n\vec{x} + \vec{x}) \overset{(b)}{=} (m\vec{x} + n\vec{x}) + \vec{x} \overset{(HI)}{=} (m+n)\vec{x} + \vec{x} = (m+n+1)\vec{x})$

(f) $n\vec{x}+n\vec{y} = n(\vec{x}+\vec{y})$; (de fato, temos por definição que $\vec{x}+\vec{y} = 1\vec{x}+1\vec{y} = 1(\vec{x}+\vec{y})$. Supondo que a hipótese de indução em n (HI) $n\vec{x} + n\vec{y} = n(\vec{x}+\vec{y})$ é válida, provemos a sua validade com n substituído por $n+1$: $(n+1)\vec{x} + (n+1)\vec{y} = (n\vec{x}+\vec{x})+(n\vec{y}+\vec{y}) \overset{(b)}{=} ((n\vec{x}+\vec{x})+n\vec{y})+\vec{y} \overset{(b)}{=} (n\vec{x}+(\vec{x}+n\vec{y}))+\vec{y} \overset{(a)}{=} (n\vec{x}+(n\vec{y}+\vec{x}))+\vec{y} \overset{(b)}{=} ((n\vec{x}+n\vec{y})+\vec{x})+\vec{y} \overset{(b)}{=} (n\vec{x}+n\vec{y})+(\vec{x}+\vec{y}) \overset{(HI)}{=} n(\vec{x}+\vec{y})+(\vec{x}+\vec{y}) = (n+1)(\vec{x}+\vec{y}))$

(g) $m(n\vec{x}) = (mn)\vec{x}$; (de fato, temos por definição que $n\vec{x} = 1(n\vec{x}) = (1n)\vec{x}$. Supondo que a hipótese de indução em m (HI) $m(n\vec{x}) = (mn)\vec{x}$ é válida, provemos a sua validade com m substituído por $m+1$: $(m+1)(n\vec{x}) = m(n\vec{x}) + n\vec{x} \overset{(HI)}{=} (mn)\vec{x} + n\vec{x} \overset{(e)}{=} (mn+n)\vec{x} = ((m+1)n)\vec{x})$

(h) $1\vec{x} = \vec{x}$. (parte da definição, mas isolamos essa identidade para conveniência futura).

Dado que a multiplicação de números naturais é definida essencialmente da mesma maneira, podemos entender o vetor $n\vec{x}$ como a *multiplicação* de $\vec{x}$ pelo *fator de escala* – ou, mais concisamente, *escalar* – $n \in \mathbb{N}$, ou seja, podemos chamar $n\vec{x}$ de um *múltiplo escalar* de $\vec{x}$. Em suma, definimos uma nova operação em V, que chamamos de *multiplicação escalar*. Notar que as propriedades (e)–(h) correspondem às seguintes propriedades da multiplicação de números naturais:

(e)+(f) = *distributividade* (i.e. possibilidade de colocar fatores comuns em evidência);

(g) = *associatividade* (i.e. multiplicação de mais de dois fatores pode ser feita com qualquer inserção de parênteses sem que o resultado mude);

(h) = *existência de unidade* (i.e. 1 é o elemento neutro para a multiplicação).

Por outro lado, notar que, no caso da multiplicação escalar, o primeiro fator é um *número* e o segundo fator é um *vetor*, ou seja, os dois fatores em princípio vivem em conjuntos *diferentes*, de modo que não faz sentido falar em comutatividade para essa operação. Portanto, nesse caso devemos definir distributividade *em cada fator separadamente*, tal como feito em (e) e (f).

É possível estender a multiplicação escalar em V para escalares *inteiros* (i.e. possivelmente negativos ou zero) da seguinte maneira: dados $\vec{x} \in V$, $n \in \mathbb{N}$ quaisquer, definimos

$$(-n)\vec{x} = n(-\vec{x}) \ , \quad 0\vec{x} = \vec{0} \ ,$$

de modo que $(-n)\vec{x} = -(n\vec{x})$ pois $n\vec{x} + (-n)\vec{x} = n\vec{x} + n(-\vec{x}) \overset{(f)}{=} n(\vec{x} - \vec{x}) \overset{(d)}{=} n\vec{0} \overset{(c)}{=} \vec{0}$. As propriedades (e)–(g) se estendem então prontamente para escalares inteiros: dados $\vec{x}, \vec{y} \in V$, $p, q \in \mathbb{Z}$ quaisquer,

(e) $p\vec{x} + q\vec{x} = (p+q)\vec{x}$; (de fato, seja $\vec{x} \in V$ qualquer. Se $p = 0$ ou $q = 0$, o resultado segue imediatamente de (c). Notar ainda que se $q = n \in \mathbb{N}$, o resultado pode ser provado por indução em n para $p \in \mathbb{Z}$ qualquer, usando o mesmo argumento do caso de escalares naturais. Finalmente, se $q = -n$ com $n \in \mathbb{N}$, notar que o caso que acabamos de provar implica $(p-n)\vec{x} + n\vec{x} = (p-n+n)\vec{x} = p\vec{x}$ e portanto $(p-n)\vec{x} = p\vec{x} - n\vec{x} = p\vec{x} + (-n)\vec{x}$ para $p \in \mathbb{Z}$ qualquer)

(f) $p\vec{x} + p\vec{y} = p(\vec{x}+\vec{y})$; (se $p = 0$, o resultado segue imediatamente de (c). Se $p = -n$ com $n \in \mathbb{N}$, observar que $(-n)\vec{x} + (-n)\vec{y} = n(-\vec{x}) + n(-\vec{y})$)

(g) $p(q\vec{x}) = (pq)\vec{x}$. (o caso $p = 0$ é trivial, e o caso $q = 0$ segue do fato que $p\vec{0} = \vec{0}$. Notar ainda que se $p = m \in \mathbb{N}$, o resultado pode ser provado por indução em m para $q \in \mathbb{Z}$ qualquer, usando o mesmo argumento do caso de escalares naturais. Finalmente, se $p = -m$ com $m \in \mathbb{N}$, notar que $(-m)(q\vec{x}) = m(-(q\vec{x})) = m((-q)\vec{x}) = (m(-q))\vec{x} = (-mq)\vec{x}$)

O caminho traçado acima sugere a possibilidade de definirmos a multiplicação de vetores pos escalares mais gerais que inteiros. Contudo, isso não pode ser feito sem hipóteses adicionais sobre V e a operação de soma vetorial. Por exemplo, dados $\vec{x} \in V$ diferente de $\vec{0}$ e $n \in \mathbb{N}$, não temos como garantir somente com as hipóteses (i) e (ii) que exista $\vec{y} \in V$ tal que $n\vec{y} = \vec{x}$, tampouco que tal $\vec{y}$ seja único – por exemplo, se existirem $n \in \mathbb{N}$ e $\vec{z} \neq \vec{0}$ tais que $n\vec{z} = \vec{0}$, então $n(\vec{y} + \vec{z}) \overset{(f)}{=} n\vec{y} + n\vec{z} \overset{(c)}{=} n\vec{y} = \vec{x}$. Conversamente, se existirem $\vec{y}, \vec{y}' \in V$ tais que $\vec{y}' \neq \vec{y}$ e $n\vec{y} = n\vec{y}' = \vec{x}$, então $n(\vec{y}' - \vec{y}) \overset{(f)}{=} n\vec{y}' - n\vec{y} = \vec{0}$. Se, contudo, tal $\vec{y}$ existir e for único, podemos definir $\frac{1}{n}\vec{x} = \vec{y}$, de modo que nesse caso devemos ter $\vec{y} \neq 0$. Ou seja, nesse caso podemos dividir o deslocamento por $\vec{x}$ em n deslocamentos iguais e sucessivos. Se a soma vetorial de V satisfizer a propriedade

Tal $\vec{z}$ é denominado um *elemento de torção* de V. Obviamente, devemos ter $n > 1$ para que tal $\vec{z}$ exista.

(iii) Para todo $\vec{x} \in V$, $n \in \mathbb{N}$ existe um *único* $\vec{y} \in V$ tal que $n\vec{y} = \vec{x}$,

então podemos escrever $\vec{y} = \frac{1}{n}\vec{x}$ para qualquer $\vec{x} \in V$, $n \in \mathbb{N}$ e portanto definir

$$\frac{m}{n}\vec{x} = m\left(\frac{1}{n}\vec{x}\right) = \frac{1}{n}(m\vec{x}) \ , \quad \vec{x} \in V \ , \ m \in \mathbb{Z} \ , \ n \in \mathbb{N}$$

(a segunda identidade vem da observação que a propriedade de associatividade (g) para a multiplicação por escalares inteiros implica $n(\frac{m}{n}\vec{x}) = (nm)(\frac{1}{n}\vec{x}) = m\vec{x}$).

1.1 Exercício. *Assumindo as propriedades (i)–(iii), mostre que a multiplicação de vetores em V por escalares racionais tal como definida acima satisfaz as propriedades (e), (f) e (g).*

1.2 Exercício. *Mostre que se as propriedades (i)–(iii) são válidas, então $p\vec{x} = \vec{0}$ com $\vec{x} \in V$, $p \in \mathbb{Q}$ implica $p = 0$ ou $\vec{x} = \vec{0}$. (*Dica: *notar que (i)–(ii) implicam que $m\vec{0} = \vec{0}$ para todo $m \in \mathbb{Z}$)*

Todavia, resultados da geometria euclideana (e.g. o teorema de Pitágoras) apontam que mesmo escalares racionais não são suficientes – é necessário admitir escalares irracionais também. Ao chegarmos nesse ponto, vemos que passa a ser mais produtivo simplesmente assumir que V é munido de uma operação de multiplicação por escalares *reais* satisfazendo as propriedades (e)–(h), portanto abster-nos-emos de explorar as consequências dessas propriedades aqui, dado que isso será feito de maneira sistemática na próxima Seção.

Esperamos neste ponto ter motivado as propriedades algébricas das operações de soma vetorial e multiplicação escalar o suficiente para deixar claro a a importância de seu estudo. Tal ponto de vista nos permite irmos além da geometria: veremos que espaços de funções podem frequentemente ser imbuídos de operações gozando das mesma propriedades, e portanto podem também ser tratados como espaços de vetores. Isso abre inúmeras possibilidades de aplicações: processamento de sinais, aprendizado de máquina, etc.

Finalmente, cabe apontar que também é possível definir a multiplicação de vetores por escalares *complexos*. Essa possibilidade é importante, por exemplo, em aplicações ao processamento de sinais e à Mecânica Quântica. O conteúdo destas Notas se concentrará primariamente no caso de escalares reais, mas a maior parte da nossa discussão pode ser (e será) desenvolvida para escalares complexos em paralelo sem alterações ou com alterações mínimas que serão apontadas quando necessário. A única exceção a esse paralelismo está na discussão sobre o Teorema Espectral no Capítulo V, pois o caso complexo do Teorema Espectral traz uma simplificação crucial devido ao *Teorema Fundamental da Álgebra*, que diz que todo polinômio com coeficientes complexos possui uma raiz complexa. É tecnicamente mais fácil extrair o caso real do Teorema Espectral a partir do caso complexo. A despeito do nome, não existe prova puramente algébrica do Teorema Fundamental da Álgebra – para a conveniência do leitor familiarizado com cálculo em várias (duas) variáveis reais, incluímos uma demonstração desse Teorema no Apêndice, mas pode-se numa primeira leitura considerar esse resultado como uma "caixa preta".

2. Definição e exemplos

Como visto na Seção 1, o aspecto essencial comum aos diferentes exemplos de vetores não é a sua interpretação concreta como entes geométricos, funções, etc., mas as *operações algébricas* que realizamos com tais objetos. Mais precisamente, essas operações satisfazem certas propriedades comuns – chamadas *axiomas* –, de modo que em *qualquer* conjunto munido de operações satisfazendo os axiomas poderemos aplicar qualquer resultado que seja consequência apenas destes. O estudo dos conceitos e resultados decorrentes de tais axiomas é o que chamamos de *Álgebra Linear*.

2.1. Axiomas de um espaço vetorial. No que se segue, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ denota o que chamamos de um *corpo* (de escalares) e seus elementos serão ocasionalmente chamados de *escalares* tal como justificado pela Seção 1. Podemos abstrair os axiomas (a)–(h) para qualquer conjunto não-vazio munido de operações similares, o que nos leva ao conceito central da Álgebra Linear:

Mais em geral, um *corpo* é um conjunto $\mathbb{K}$ munido de operações de soma e produto comutativas e associativas com respectivos elementos neutros $0 \neq 1$, tais que o produto é distributivo com respeito à soma e todo elemento de $\mathbb{K}$ possui oposto com respeito à soma e, se diferente de 0, com respeito ao produto. É possível desenvolver a teoria básica de espaços vetoriais desenvolvida neste Capítulo sobre qualquer corpo, mas não faremos isso aqui.

I.3 Definição. Um *espaço vetorial (sobre $\mathbb{K}$)* é um conjunto $V \neq \varnothing$ munido de duas operações:

- A *soma (vetorial)* $\vec{x} + \vec{y}$ ($\vec{x}, \vec{y} \in V$);
- A *multiplicação escalar* $\alpha\vec{x}$ ($\alpha \in \mathbb{K}$, $\vec{x} \in V$),

satisfazendo as propriedades (a)–(h) listadas a seguir, denominadas *axiomas de espaço vetorial:*

(a) $\vec{x} + \vec{y} = \vec{y} + \vec{x}$ para $\vec{x}, \vec{y} \in V$ *quaisquer* (*comutatividade* da soma vetorial);

(b) $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$ para $\vec{x}, \vec{y}, \vec{z} \in V$ *quaisquer* (*associatividade* da soma vetorial);

(c) Existe um elemento $\vec{0} \in V$ (chamado de (vetor) *zero* ou *vetor nulo*) tal que $\vec{x} + \vec{0} = \vec{x}$ para todo $\vec{x} \in V$ (existência de *elemento neutro* da soma vetorial);

(d) Dado $\vec{x} \in V$, existe $-\vec{x} \in V$ (chamado de *oposto* de $\vec{x}$ ou simplesmente *menos* $\vec{x}$) tal que $\vec{x} + (-\vec{x}) = \vec{0}$ (existência de *oposto* da soma vetorial);

(e) $(\alpha + \beta)\vec{x} = \alpha\vec{x} + \beta\vec{x}$ para $\alpha, \beta \in \mathbb{K}$, $\vec{x} \in V$ *quaisquer* (*distributividade* da multiplicação escalar no *primeiro* fator = com respeito à soma *escalar*). Ou seja, um fator *vetorial* comum pode ser colocado em evidência;

(f) $\alpha(\vec{x}+\vec{y}) = \alpha\vec{x} + \alpha\vec{y}$ para $\alpha \in \mathbb{K}$, $\vec{x}, \vec{y} \in V$ *quaisquer* (*distributividade* da multiplicação escalar no *segundo* fator = com respeito à soma *vetorial*). Ou seja, um fator *escalar* comum pode ser colocado em evidência;

(g) $(\alpha\beta)\vec{x} = \alpha(\beta\vec{x})$ para $\alpha, \beta \in \mathbb{K}$, $\vec{x} \in V$ *quaisquer* (*associatividade* da multiplicação escalar);

(h) $1\vec{x} = \vec{x}$ (existência de *elemento neutro* para a multiplicação escalar).

Neste caso, os elementos de V são chamados de *vetores* em V, os elementos de $\mathbb{K}$ são chamados de *escalares* em V, e as duas operações acima são chamadas de *operações vetoriais* em V. Se $\mathbb{K} = \mathbb{R}$ dizemos que V é um espaço vetorial *real* e se$\mathbb{K} = \mathbb{C}$ dizemos que V é um espaço vetorial *complexo*.

Obviamente, todo espaço vetorial complexo é também um espaço vetorial real. Os axiomas (a)–(h) são análogos às propriedades da soma e produto de números reais ou complexos. Uma propriedade dos últimos que *não* vale para operações vetoriais é a *comutatividade* da multiplicação escalar, pois cada um dos fatores vive num conjunto *diferente* e portanto não faz sentido mudar a ordem dos fatores na multiplicação escalar. Por isso, as propriedades (e), (f) de distributividade são apresentadas separadamente para cada fator. Retornaremos a esse ponto em breve.

Antes de prosseguirmos, cabe um esclarecimento sobre o axioma (d). O último, tal como está escrito, não deixa claro a *qual* zero de V estamos nos referindo, tampouco se cada vetor em V tem um *único* oposto. Felizmente, tais ambiguidades são resolvidas pelas duas consequências de (a)–(c) abaixo:

- Só existe *um* elemento neutro (portanto, justificadamente denotado por $\vec{0}$) para a soma vetorial de V. (de fato, dados vetores $\vec{o}_1, \vec{o}_2 \in V$ tais que

  $$\vec{x} + \vec{o}_1 = \vec{x} + \vec{o}_2 = \vec{x}$$

  para todo vetor $\vec{x} \in V$, tomando $\vec{x} = \vec{o}_1$ obtemos

  $$\vec{o}_1 + \vec{o}_2 = \vec{o}_1$$

  e tomando $\vec{x} = \vec{o}_2$ obtemos

  $$\vec{o}_2 + \vec{o}_1 = \vec{o}_2 \ .$$

  Concluímos daí e do axioma (a) que $\vec{o}_1 = \vec{o}_2$)

- Dado $\vec{x} \in V$, só existe *um* oposto de $\vec{x}$ (portanto, justificadamente denotado por $-\vec{x}$). Em particular, $-(-\vec{x}) = \vec{x}$. (de fato, dado $\vec{x} \in V$ sejam vetores $\vec{y}, \vec{z} \in V$

Por conveniência, denotamos a *subtração* de vetores por $\vec{x} + (-\vec{y}) = \vec{x} - \vec{y}$.

tais que $\vec{x}+\vec{y}=\vec{x}+\vec{z}=\vec{0}$. Somando-se $\vec{z}$ à primeira e última fórmulas, segue que

$$\vec{z}+(\vec{x}+\vec{y})=\vec{z}+\vec{0}\overset{(c)}{=}\vec{z}$$

e

$$\vec{z}+(\vec{x}+\vec{y})\overset{(b)}{=}(\vec{z}+\vec{x})+\vec{y}\overset{(a)}{=}(\vec{x}+\vec{z})+\vec{y}=\vec{0}+\vec{y}\overset{(c)}{=}\vec{y}\,,$$

de onde concluímos que $\vec{y}=\vec{z}$. A última identidade segue disso e de (a); uma prova alternativa desse fato será dada mais adiante)

Como dito no início desta Seção, propriedades das operações vetoriais que dependam *apenas* dos axiomas de espaço vetorial valem para *todos* os espaços vetoriais. Por exemplo, seguem abaixo outras consequências simples dos axiomas (a)–(h):

Identidades que seguem de uma propriedade específica que tenha sido enumerada no texto, tal como os axiomas (a)–(h), terão a propriedade invocada indicada sobre o sinal de igualdade quando necessário.

1.) $0\vec{x}=\vec{0}$. (de fato,

$$0\vec{x}+0\vec{x}\overset{(g)}{=}(0+0)\vec{x}=0\vec{x}\,.$$

Somando $-0\vec{x}$ à primeira e última fórmulas, concluímos que

$$(0\vec{x}+0\vec{x})-0\vec{x}\overset{(b)}{=}0\vec{x}+(0\vec{x}-0\vec{x})\overset{(d)}{=}0\vec{x}+\vec{0}\overset{(c)}{=}0\vec{x}$$

e $0\vec{x}-0\vec{x}\overset{(d)}{=}\vec{0}$, logo $0\vec{x}=\vec{0}$)

2.) $\alpha\vec{0}=\vec{0}$ para $\alpha\in\mathbb{K}$ qualquer. (de fato,

$$\alpha\vec{0}+\alpha\vec{0}\overset{(f)}{=}\alpha(\vec{0}+\vec{0})\overset{(c)}{=}\alpha\vec{0}\,.$$

De maneira análoga à prova da propriedade anterior, somando-se $-\alpha\vec{0}$ à primeira e última fórmulas, concluímos que

$$(\alpha\vec{0}+\alpha\vec{0})-\alpha\vec{0}\overset{(b)}{=}\alpha\vec{0}+(\alpha\vec{0}-\alpha\vec{0})\overset{(d)}{=}\alpha\vec{0}+\vec{0}\overset{(c)}{=}\alpha\vec{0}$$

e $\alpha\vec{0}-\alpha\vec{0}\overset{(d)}{=}\vec{0}$, logo $\alpha\vec{0}=\vec{0}$)

3.) $(-\alpha)\vec{x}=-(\alpha\vec{x})=\alpha(-\vec{x})$ para $\alpha\in\mathbb{K}$ qualquer. Em particular, $(-1)\vec{x}\overset{(h)}{=}-\vec{x}$ e portanto $-(-\vec{x})=\vec{x}$. (de fato,

$$(-\alpha)\vec{x}+\alpha\vec{x}\overset{(e)}{=}(-\alpha+\alpha)\vec{x}=0\vec{x}=\vec{0}$$

e

$$\alpha(-\vec{x})+\alpha\vec{x}\overset{(f)}{=}\alpha(\vec{x}-\vec{x})=\alpha\vec{0}=\vec{0}\,.$$

O resultado desejado segue da unicidade do oposto da soma vetorial de V, provada acima. A última fórmula acima, já provada na nossa discussão sobre a unicidade do oposto, segue então de (h))

4.) Se $\alpha\vec{x} = \vec{0}$, então $\alpha = 0$ ou $\vec{x} = \vec{0}$. Em particular, se $V \neq \{\vec{0}\}$, 1 é o único escalar que satisfaz o axioma (h). (de fato, vimos acima que $0\vec{x} = \vec{0}$. Se, por outro lado, $\alpha \neq 0$, temos que

$$\frac{1}{\alpha}(\alpha\vec{x}) \overset{(e)}{=} \left(\frac{\alpha}{\alpha}\right)\vec{x} = 1\vec{x} \overset{(h)}{=} \vec{x} = \frac{1}{\alpha}\vec{0} = \vec{0}\ ,$$

logo $\vec{x} = \vec{0}$. Para a última afirmação, notar que se $\alpha\vec{x} = \vec{x}$ para todo $\vec{x} \in V$, então

$$\alpha\vec{x} - \vec{x} = \alpha\vec{x} + (-1)\vec{x} \overset{(e)}{=} (\alpha - 1)\vec{x} = \vec{x} - \vec{x} \overset{(d)}{=} \vec{0}\ .$$

Tomando $\vec{x} \neq \vec{0}$ qualquer resulta em $\alpha = 1$)

5.) Se $\alpha\vec{x} = \beta\vec{x}$ e $\vec{x} \neq \vec{0}$, então $\alpha = \beta$. (de fato, se $\alpha\vec{x} = \beta\vec{x}$, então

$$\alpha\vec{x} + (-\beta)\vec{x} \overset{(e)}{=} (\alpha - \beta)\vec{x} = \beta\vec{x} + (-\beta)\vec{x} \overset{(e)}{=} (\beta - \beta)\vec{x} = 0\vec{x} \overset{1.)}{=} \vec{0}\ .$$

Segue de 4.) que $\alpha = \beta$)

6.) Se $\alpha\vec{x} = \alpha\vec{y}$ e $\alpha \neq 0$, então $\vec{x} = \vec{y}$. (de fato, se $\alpha\vec{x} = \alpha\vec{y}$, então

$$\alpha\vec{x} + \alpha(-\vec{y}) \overset{(f)}{=} \alpha(\vec{x} + (-\vec{y})) = \alpha\vec{y} + \alpha(-\vec{y}) \overset{(f)}{=} \alpha(\vec{y} + (-\vec{y})) = \alpha\vec{0} \overset{2.)}{=} \vec{0}\ .$$

Segue de 4.) que $\vec{x} + (-\vec{y}) = 0$ e portanto $\vec{x} = \vec{y}$ pela unicidade do oposto da soma vetorial)

7.) $-(\vec{x} + \vec{y}) = (-\vec{x}) + (-\vec{y})$. (de fato, temos que

$$(\vec{x} + \vec{y}) + ((-\vec{x}) + (-\vec{y})) \overset{(a)}{=} (\vec{x} + \vec{y}) + ((-\vec{y}) + (-\vec{x})) \overset{(b)}{=} \vec{x} + (\vec{y} + ((-\vec{y}) + (-\vec{x})))$$
$$\overset{(b)}{=} \vec{x} + ((\vec{y} + (-\vec{y})) + (-\vec{x})) \overset{(d)}{=} \vec{x} + (\vec{0} + (-\vec{x})) \overset{(a)+(c)}{=} \vec{x} + (-\vec{x}) \overset{(d)}{=} \vec{0}\ ,$$

logo a afirmação segue da unicidade do oposto da soma vetorial)

8.) $\vec{x} + \vec{x} = 2\vec{x}$. Mais em geral, $(n + 1)\vec{x} = n\vec{x} + \vec{x}$ para todo $n \in \mathbb{N}$. (de fato, $2\vec{x} = (1 + 1)\vec{x} \overset{e}{=} 1\vec{x} + 1\vec{x} \overset{(h)}{=} \vec{x} + \vec{x}$ e $(n + 1)\vec{x} \overset{(e)}{=} n\vec{x} + 1\vec{x} \overset{(h)}{=} n\vec{x} + \vec{x}$)

As consequências (4.) e (5.) são chamadas de *leis de cancelamento* (à direita e à esquerda, respectivamente). As consequências (1.)–(8.) implicam juntamente com os axiomas (a)–(h) que podemos manipular expressões envolvendo operações vetoriais da mesma maneira que expressões análogas envolvendo respectivamente *somas e produtos de números reais*, com as seguintes ressalvas:

- Como dito acima, não faz sentido em geral falar de "comutatividade" para a *multiplicação escalar* pois cada fator vive num conjunto diferente se $V \neq \mathbb{K}$. Por consequência, os axiomas de distributividade (e) e (f) são *distintos* e *independentes*, ao contrário do que ocorre com *o* axioma de distributividade (no singular) para o produto de números reais.

- Da mesma forma, não faz sentido falar de "divisão por vetores" se $V \neq \mathbb{K}$, apenas por escalares diferentes de zero. Não obstante, é possível sim cancelar um fator vetorial comum diferente de zero *nos dois lados de uma igualdade em* V, por meio da lei de cancelamento à direita (5.) acima.

Em particular, podemos inserir ou retirar parênteses, mudar a ordem de parcelas numa soma vetorial de dois ou mais vetores, manipular sinais, colocar fatores comuns em evidência, etc.. Faremos isso de maneira tácita de agora em diante.

I.4 Observação. Se E é um conjunto não-vazio, V é um espaço vetorial e $\vec{d} : V \times V \to E$ é uma função satisfazendo as propriedades (i)–(ii) da Seção 1 acima (i.e. $\vec{d}$ é uma *função deslocamento* em E), tal que a soma vetorial derivada a partir de $\vec{d}$ *coincide* com a soma vetorial de V, dizemos que E é um *espaço afim* modelado em V.

2.2. Exemplos de espaços vetoriais. Dentre os exemplos de espaços vetoriais *reais*, podemos citar:

1.) $V = \mathbb{R}^n$ = conjunto das "listas ordenadas" de n números reais $x_1, \ldots, x_n$ (o exemplo seguinte dará um significado mais preciso a essa ideia)

$$\begin{gathered} \vec{x} = (x_1, \ldots, x_n) \\ \Rightarrow x_j = j\text{-ésima componente (canônica) de } \vec{x}\ , \\ j = 1, \ldots, n\ . \end{gathered}$$

Por "ordenadas" entende-se que trocando duas componentes de $\vec{x}$ de lugar mudamos a lista, a menos que tais componentes sejam iguais. Os elementos de $\mathbb{R}^n$ são denominados por vezes *vetores n-dimensionais*. As (candidatas a) operações vetoriais (por vezes ditas *canônicas*) de $\mathbb{R}^n$ são definidas "componente a componente":

- A soma vetorial de $\vec{x} = (x_1, \ldots, x_n)$ e $\vec{y} = (y_1, \ldots, y_n) \in V$ é dada por

$$\begin{gathered} \vec{x} + \vec{y} = (x_1 + y_1, \ldots, x_n + y_n) \\ \Rightarrow x_j + y_j = j\text{-ésima componente de } \vec{x} + \vec{y}\ , \\ j = 1, \ldots, n\ ; \end{gathered}$$

- A multiplicação (pelo) escalar ($\alpha \in \mathbb{R}$) do vetor $\vec{x} = (x_1, \ldots, x_n) \in V$ é dada por

$$\begin{gathered} \alpha\vec{x} = (\alpha x_1, \ldots, \alpha x_n) \\ \Rightarrow \alpha x_j = j\text{-ésima componente de } \alpha\vec{x}\ , \\ j = 1, \ldots, n\ . \end{gathered}$$

Provaremos a validade dos axiomas (a)–(h) observando que este exemplo é caso particular do exemplo seguinte.

2.) Dado $A \neq \emptyset$, seja V o *espaço das funções de* A em $\mathbb{R}$:

$$V = \mathbb{R}^A = \{f : A \to \mathbb{R}\} .$$

Definimos em tal V as (candidatas a) *operações vetoriais pontuais* (a seguir, $f, g : A \to \mathbb{R}$, $p \in A$, $\alpha \in \mathbb{R}$ são quaisquer)

- Soma vetorial:
$$(f + g)(p) = f(p) + g(p) ;$$
- Multiplicação escalar:
$$(\alpha f)(p) = \alpha f(p) .$$

A validade dos axiomas (a)–(h) para tais operações segue imediatamente da validade dos axiomas correspondentes no contradomínio comum dos elementos de $\mathbb{R}^A$ (i.e. $\mathbb{R}$). De fato, dados $f, g, h \in V$, $p \in A$ e $\alpha, \beta \in \mathbb{R}$ quaisquer, então:

(a) $(f + g)(p) = f(p) + g(p) = g(p) + f(p) = (g + f)(p)$;

(b) $(f + (g + h))(p) = f(p) + (g + h)(p) = f(p) + (g(p) + h(p)) = (f(p) + g(p)) + h(p) = (f + g)(p) + h(p) = ((f + g) + h)(p)$;

(c) $0(p) \equiv 0$ (i.e. $0(p) = 0$ para todo $p \in A$) satisfaz $(f + 0)(p) = f(p) + 0 = f(p)$ para todo $p \in A$;

(d) Dado $f \in V$, $(-f)(p) = -f(p)$ $(p \in A)$ satisfaz $(f + (-f))(p) = f(p) - f(p) = 0 = 0(p)$ para todo $p \in A$;

(e) $((\alpha + \beta)f)(p) = (\alpha + \beta)f(p) = \alpha f(p) + \beta f(p) = (\alpha f)(p) + (\beta f)(p) = ((\alpha f) + (\beta f))(p)$;

(f) $(\alpha(f + g))(p) = \alpha(f + g)(p) = \alpha(f(p) + g(p)) = \alpha f(p) + \alpha g(p) = (\alpha f)(p) + (\alpha g)(p)$;

(g) $((\alpha\beta)f)(p) = (\alpha\beta)f(p) = \alpha(\beta f(p)) = \alpha(\beta f)(p) = (\alpha(\beta f))(p)$;

(h) $(1f)(p) = 1f(p) = f(p)$.

Em particular, se $A = \{1, \ldots, n\}$, $n \in \mathbb{N}$ então $V = \mathbb{R}^n$ e as operações vetoriais pontuais de $\mathbb{R}^n$ coincidem com as operações vetoriais de $\mathbb{R}^n$ indicadas no exemplo (i) acima, logo as últimas satisfazem os axiomas (a)–(h). No caso $n = 1$ temos que cada $\vec{x} \in \mathbb{R}^1$ é unicamente determinado por $\vec{x}(1) = x$, portando podemos identificar $\mathbb{R}^1$ com $\mathbb{R}$. Mediante essa identificação, as operações vetoriais de $\mathbb{R}$ coincidem com a soma e produto usuais respectivamente. Em particular, $\mathbb{R}$ munido da soma e produto usuais é também um espaço vetorial real.

3.) Mais em geral, se $A \neq \varnothing$ e W é um espaço vetorial, seja V o *espaço das funções em A a valores em W*:

$$V = W^A = \{\vec{f} : A \to W\} .$$

Definimos em tal V as (candidatas a) *operações vetoriais pontuais* como no exemplo (ii) (a seguir, $\vec{f}, \vec{g} : A \to \mathbb{R}$, $p \in A$, $\alpha \in \mathbb{R}$ são quaisquer):

- Soma vetorial:
$$(\vec{f} + \vec{g})(p) = \vec{f}(p) + \vec{g}(p) ;$$
- Multiplicação escalar:
$$(\alpha\vec{f})(p) = \alpha\vec{f}(p) .$$

Tal como no exemplo (2.) acima, a validade dos axiomas (a)–(h) para tais operações segue imediatamente da validade dos axiomas correspondentes no contradomínio comum W. Os cálculos são exatamente os mesmos.

4.) Este exemplo pode ser visto como uma variante de (1.) ou mesmo de (3.). Seja $V = \mathscr{M}_{m\times n}(\mathbb{R})$ o *espaço das matrizes* (reais) com m linhas e n colunas

$$A = [A_{ij}] = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix} .$$

Dizemos que $A_{ij} = [A]_{ij}$ é a *entrada* de A na i-ésima linha e j-ésima coluna. Notemos que V nada mais é do que $\mathbb{R}^{mn}$ com seus elementos representados graficamente numa tabela retangular ao invés de uma lista ordenada linear. Com efeito, podemos identificar $A = [A_{ij}] \in V$ com $\vec{x} = (x_1, \ldots, x_{mn}) \in \mathbb{R}^{mn}$ através da fórmula

$$A_{ij} = x_{n(i-1)+j} , \quad i = 1, \ldots, m , \ n = 1, \ldots, n ,$$

de modo que visualmente

$$\vec{x} = \big(x_1 = A_{11}, \ldots, x_n = A_{1n}, x_{n+1} = A_{21}, \ldots, x_{2n} = A_{2n}, \ldots, x_{n(m-1)+1} = A_{m1}, \ldots, x_{nm} = A_{mn}\big) .$$

Assim sendo, as operações vetoriais de V (que chamaremos de *operações vetoriais matriciais*) são as mesmas de $\mathbb{R}^{mn}$ mediante tal identificação: se $A, B \in V$ e $\alpha \in \mathbb{R}$, então

$$(A + B)_{ij} = A_{ij} + B_{ij} ,$$
$$(\alpha A)_{ij} = \alpha A_{ij} .$$

Em suma, as operações vetoriais de V são definidas "entrada a entrada". Segue imediatamente dos exemplos anteriores que elas satisfazem os axiomas (a)–(h). Alternativamente, podemos identificar V com $(\mathbb{R}^m)^n$ e nesse caso as operações vetoriais pontuais de $(\mathbb{R}^m)^n$ se reduzem às operações vetoriais matriciais definidas acima pois se $\vec{A} \in (\mathbb{R}^m)^n$ então podemos identificar $\vec{A}(j)$ com a j-ésima coluna de $A \in V$ para cada $j = 1, \ldots, n$. Mais precisamente, dados $A = [A_{ij}], B = [B_{ij}] \in V$, $\alpha \in \mathbb{R}$ quaisquer, reiterando que $\vec{A}(j)(i) = A_{ij}$, $\vec{B}(j)(i) = B_{ij}$ temos que

$$(\vec{A}(j)+\vec{B}(j))(i) = \vec{A}(j)(i)+vB(j)(i) = A_{ij}+B_{ij}\ ,\ (\alpha\vec{A}(j))(i) = \alpha\vec{A}(j)(i) = \alpha A_{ij}$$

para $i = 1, \ldots, m$, $j = -1, \ldots, n$ quaisquer, portanto

$$(\overrightarrow{A+B})(j) = \vec{A}+\vec{B}\ ,\ \overrightarrow{\alpha A} = \alpha\vec{A}\ .$$

Em particular, essa identificação deixa claro a dependência de $\vec{A} \in (\mathbb{R}^m)^n$ em duas variáveis (linha e coluna). Essa segunda maneira de definir matrizes será útil no Capítulo III.

Substituindo $\mathbb{R}$ por $\mathbb{C}$ em cada um dos exemplos acima, obtemos exemplos correspondentes de espaços vetoriais complexos.

2.3. Subespaços vetoriais. Mais exemplos (e contraexemplos) podem ser obtidos a partir dos dados na Subseção 2.2 a partir da seguinte

I.5 Definição. Seja V um espaço vetorial, e $W \subset V$. Dizemos que W é *subespaço vetorial* de V se satisfizer as seguintes condições, denominadas *axiomas de subespaço vetorial*:

(i') $W \neq \emptyset$;

(ii) Se $\vec{x}, \vec{y} \in W$ então $\vec{x}+\vec{y} \in W$;

(iii) Se $\vec{x} \in W$ e $\alpha \in \mathbb{K}$ então $\alpha\vec{x} \in W$.

Se $W \neq V$, dizemos que W é subespaço vetorial *próprio* de V.

Mostraremos que todo subespaço vetorial $W \subset V$ é um espaço vetorial se munido das operações vetoriais herdadas do espaço vetorial *ambiente* V (o que sempre assumiremos ser o caso). De fato, os axiomas (a), (b) e (e)–(h) são claramente satisfeitos em W. O que nos resta fazer para obter a validade dos axiomas (c) e (d) em W é provar, respectivamente, que:

- $\vec{0} \in W$. (se $\vec{x} \in W$, segue de (iii) que $0\vec{x} = \vec{0} \in W$)

- Se $\vec{x} \in W$, então $-\vec{x} \in W$. (se $\vec{x} \in W$, segue de (iii) que $-\vec{x} = (-1)\vec{x} \in W$)

Em particular, W tem o mesmo zero e os mesmos opostos para seus elementos que em V. Notar também que não há perda de generalidade de assumir no lugar de (i') o axioma

(i) $\vec{0} \in W$

na Definição I.5. Como (i) é mais fácil de verificar na prática, faremos tacitamente essa substituição daqui em diante. Uma vantagem de lidar com subespaços vetoriais é que verificar se $W \subset V$ é subespaço vetorial de V é bem mais simples do que verificar todos os axiomas (a)–(h) em W.

I.6 Observação. Lembrando que todo espaço vetorial *complexo* é também um espaço vetorial *real* (basta restrigirmo-nos a escalares reais na multiplicação escalar), notar que podemos ter um subespaço vetorial real $W \subset V$ de um espaço vetorial complexo V que *não* é um subespaço vetorial (complexo) de V. Um exemplo típico é $V = \mathbb{C}^n$, $W = \mathbb{R}^n$ (veremos outros mais adiante) – com efeito, nesse caso os axiomas (i) e (ii) são satisfeitos mas (iii) é *violado* para escalares não-reais, pois se $\vec{0} \neq \vec{x} \in W$ então $i\vec{x} \notin W$. Doravante, quando dizemos que $W \subset V$ é um subespaço vetorial de um espaço vetorial V (sem qualificação), subentender-se-á que os corpos de escalares de V e de W são os *mesmos*.

Exibiremos agora alguns (contra)exemplos de subespaços vetoriais. Os exemplos (1.)–(7.) abaixo são todos reais.

0.) V qualquer, $W = \{\vec{0}\}$: claramente $\vec{0} + \vec{0} = \vec{0}$ pelo axioma (c) e vimos que $\alpha\vec{0} = \vec{0}$ para todo $\alpha \in \mathbb{K}$, logo W é subespaço vetorial de V – o chamado *subespaço trivial* ou *subespaço zero* de V, que denotamos simplesmente por 0.

1.) $V = \mathbb{R}^3$, $W = \{(x, y, z) \in V \mid z = 0\}$: claramente $\vec{0} = (0, 0, 0) \in W$. Além disso, temos que se $(x, y, 0), (x', y', 0) \in W$ e $\alpha \in \mathbb{R}$ então

$$(x, y, 0) + (x', y', 0) = (x + x', y + y', 0) \in W$$

e

$$\alpha(x, y, 0) = (\alpha x, \alpha y, 0) \in W \ .$$

2.) $V = \mathbb{R}^2$, $W = \{(x, y) \in V \mid y = x^2\}$: neste caso, $\vec{0} = (0, 0) \in W$ e $(1, 1) \in W$ mas $(2, 2) = 2(1, 1) \notin W$, ou seja, W não é fechado por multiplicação escalar e portanto *não* é subespaço vetorial de V.

3.) (generaliza (2.)) Se $V = \mathbb{R}^n$ e $W = \{\vec{x} = (x_1, \ldots, x_n) \in V \mid \alpha_1 x_1 + \cdots + \alpha_n x_n = 0\}$, então W é subespaço vetorial de V. Isso será provado no exemplo seguinte.

4.) (generaliza (3.)) Se $V = \mathbb{R}^n$ e

$$W = \{\vec{x} = (x_1, \ldots, x_n) \in V \mid \vec{x} \text{ é solução de (I.1)}\} ,$$

onde (I.1) é o sistema linear *homogêneo*

$$\text{(I.1)} \qquad \begin{cases} A_{11}x_1 & + & \cdots & + & A_{1n}x_n & = 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{m1}x_1 & + & \cdots & + & A_{mn}x_n & = 0 \end{cases} .$$

Claramente $\vec{0} = (0, \ldots, 0) \in W$. Além disso, se $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n)$ pertencem a W e $\alpha \in \mathbb{R}$, então

$$\begin{cases} A_{11}(x_1 + y_1) & + & \cdots & + & A_{1n}(x_n + y_n) & = 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{m1}(x_1 + y_1) & + & \cdots & + & A_{mn}(x_n + y_n) & = 0 \end{cases}$$

e

$$\begin{cases} A_{11}(\alpha x_1) & + & \cdots & + & A_{1n}(\alpha x_n) & = 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{m1}(\alpha x_1) & + & \cdots & + & A_{mn}(\alpha x_n) & = 0 \end{cases} ,$$

logo $\vec{x} + \vec{y}$ e $\alpha\vec{x}$ também pertencem a W. Como (3.) é um caso particular de (4.) (mais precisamente, quando $m = 1$), segue daí que o exemplo (3.) é de fato um subespaço vetorial. Outra maneira de ver que W é subespaço vetorial de V, como discutiremos mais adiante, é notar que

$$W = \bigcap_{i=1}^{m} W_i ,$$

onde W_i é o conjunto dos vetores $\vec{x} = (x_1, \ldots, x_n)$ em V que satisfazem a i-ésima equação $A_{i1}x_1 + \cdots + A_{in}x_n = 0$ do sistema linear homogêneo acima e portanto uma instância do exemplo (3.). Como veremos abaixo que a intersecção de uma família arbitrária de subespaços vetoriais de V é também subespaço vetorial de V, segue que W é subespaço vetorial de V. Veremos mais adiante no Capítulo II que este é um exemplo genérico: *todo* subespaço vetorial de $\mathbb{R}^n$ é o espaço de soluções de *algum* sistema linear homogêneo. Por outro lado, se o lado direito de alguma das equações de (I.1) *não* fosse zero (i.e. o sistema passasse a ser *não-homogêneo*), então $\vec{0}$ *não* pertenceria a W pois nesse caso $\vec{0}$ não pode ser solução dessa equação. Logo, nesse caso W *não* pode ser subespaço vetorial de V.

5.) Se $V = \mathbb{R}^{\mathbb{R}}$ e

$$W = \mathcal{P}_{\leq k}(\mathbb{R}) = \{f \in V \mid f(t) = a_0 + a_1 t + \cdots + a_k t^k \ , \ a_0, a_1, \ldots, a_k, t \in \mathbb{R}\}$$

(conjunto dos polinômios de ordem $\leq k \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ com coeficientes reais), então W é subespaço vetorial de V. Claramente $\vec{0} = (f \equiv 0) \in W$; ademais, se $f(t) = a_0 + a_1 t + \cdots + a_k t^k$, $g(t) = b_0 + b_1 t + \cdots + b_k t^k$, onde $a_0, \ldots, a_k, b_0, \ldots, b_k, t \in \mathbb{R}$, e $\alpha \in \mathbb{R}$ quaisquer, então

$$\begin{aligned}(f+g)(t) &= a_0 + a_1 t + \cdots + a_k t^k + a_0 + a_1 t + \cdots + a_k t^k \\ &= (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_k + b_k)t^k\end{aligned}$$

e

$$(\alpha f)(t) = \alpha(a_0 + a_1 t + \cdots + a_k t^k) = \alpha a_0 + (\alpha a_1)t + \cdots + (\alpha a_k)t^k \ ,$$

portanto $f + g, \alpha f \in W$ se $f, g \in W$, $\alpha \in \mathbb{R}$ quaisquer. Concluímos destarte que W é subespaço vetorial de V. Obviamente, $W = \mathbb{R}$ se $k = 0$. Por outro lado, notar que o conjunto $\mathcal{P}_k(\mathbb{R}) = \{f \in V \mid f(t) = a_0 + a_1 t + \cdots + a_k t^k \ , \ a_0, a_1, \ldots, a_k, t \in \mathbb{R} \ , \ a_k \neq 0\}$ dos polinômios de ordem $k \in \mathbb{N}$ com coeficientes reais não é subespaço vetorial de V pois $\vec{0} \notin \mathcal{P}_k(\mathbb{R})$. Mais em geral, o conjunto

$$\mathcal{P}(\mathbb{R}) = \bigcup_{k=0}^{\infty} \mathcal{P}_{\leq k}(\mathbb{R})$$

dos polinômios com coeficientes reais é subespaço vetorial de V pois $\mathcal{P}_{\leq k}(\mathbb{R}) \subset \mathcal{P}_{\leq m}(\mathbb{R})$ se $k \leq m$ e portanto para $f \in \mathcal{P}_{\leq k}(\mathbb{R})$, $g \in \mathcal{P}_{\leq m}(\mathbb{R})$, $\alpha \in \mathbb{R}$ quaisquer temos que $f + g \in \mathcal{P}_{\leq \max(k,m)}(\mathbb{R})$, $\alpha f \in \mathcal{P}_{\leq k}(\mathbb{R})$.

6.) Se $V = \mathbb{R}^{\mathbb{R}}$ e

$$W = \mathcal{C}^0(\mathbb{R}) = \{f \in V \mid f \text{ contínua}\} \ ,$$

então W é subespaço vetorial de V. Obviamente $f \equiv 0$ é contínua; ademais, se $f, g \in V$ são contínuas, então $f + g$ e αf também são. Concluímos destarte que W é de fato subespaço vetorial de V.

7.) Se $V = \mathbb{R}^{\mathbb{R}}$ e

$$W = \{f \in V \mid f'(t) \text{ existe para todo } t \in \mathbb{R}\}$$

(conjunto das funções deriváveis em $\mathbb{R}$), então W é subespaço vetorial de V. Obviamente $f \equiv 0$ é derivável; ademais, se $f, g \in V$ são deriváveis, então $(f+g)'(t) = f'(t) + g'(t)$ e $(\alpha f)'(t) = \alpha f'(t)$ para $t, \alpha \in \mathbb{R}$ quaisquer, portanto $f + g, \alpha f \in W$ nesse caso. Concluímos destarte que W é de fato subespaço vetorial de V.

Exemplos similares a (1.)–(6.) no caso complexo são obtidos substituindo $\mathbb{R}$ por $\mathbb{C}$ onde quer que o primeiro apareça. No exemplo (6.), notar que identificamos $\mathbb{C}$ com $\mathbb{R}^2$ como espaço vetorial real, de modo que a noção de continuidade para funções $f : \mathbb{C} \to \mathbb{C}$ conincide com a noção de continuidade para aplicações bidimensionais $f : \mathbb{R}^2 \to \mathbb{R}^2$. Detalharemos a discussão desse exemplo no Capítulo II.

A situação descrita no exemplo (5.) acima admite a seguinte abstração. Sejam agora V um espaço vetorial sobre $\mathbb{K}$ e $W_1, W_2 \subset V$ subespaços vetoriais de V. Claramente $W_1 \cap W_2$ é também, nesse caso, subespaço vetorial de V: de fato, se $\vec{x}, \vec{y} \in W_1 \cap W_2$ e $\alpha \in \mathbb{K}$, então $\vec{x}, \vec{y} \in W_1$ e $\vec{x}, \vec{y} \in W_2$, logo $\vec{x} + \vec{y}, \alpha\vec{x} \in W_1$ e $\vec{x} + \vec{y}, \alpha\vec{x} \in W_2$, provando a asserção acima.

Mais em geral, se $\{W_j \mid j \in J\}$ é uma família arbitrária de subespaços vetoriais de V, temos que a *intersecção* de todos os membros dessa família

$$W = \bigcap_{j \in J} W_j = \{\vec{x} \in V \mid \vec{x} \in W_j \text{ para todo } j \in J\}$$

também é subespaço vetorial de V (*exercício:* verifique usando o mesmo argumento do caso $J = \{1, 2\}$ acima).

Notar, contudo, que a *união* $W_1 \cup W_2$ de dois subespaços vetoriais W_1, W_2 de V satisfaz os axiomas (i) e (iii) mas *não* é necessariamente um subespaço vetorial de V pois a soma de um vetor em W_1 e um vetor em W_2 pode não pertencer a $W_1 \cup W_2$ – por exemplo, se $V = \mathbb{R}^2$, considere os subespaços vetoriais de V dados por

$$W_1 = \{(x, 0) \mid x \in \mathbb{R}\}\ ,\ W_2 = \{(0, y) \mid y \in \mathbb{R}\}$$

(*exercício:* prove que W_1 e W_2 são de fato subespaços vetoriais de V). Temos nesse caso que $(1, 0) \in W_1$ e $(0, 1) \in W_2$ mas $(1, 1) = (1, 0) + (0, 1) \notin W_1 \cup W_2$, logo $W_1 \cup W_2$ *não* é subespaço vetorial de V. Podemos corrigir esse problema definindo a *soma* $W_1 + W_2$ de W_1 e W_2 como

$$W_1 + W_2 = \{\vec{x} \in V \mid \vec{x} = \vec{y} + \vec{z}\ ,\ \vec{y} \in W_1\ ,\ \vec{z} \in W_2\}\ .$$

Notar que $\vec{y}, \vec{z}$ na definição de $W_1 + W_2$ podem *não* ser únicos!

Segue que, ao contrário de $W_1 \cup W_2$, $W_1 + W_2$ é subespaço vetorial de V: se $\vec{x}, \vec{x}' \in W_1 + W_2$ e $\alpha \in \mathbb{K}$, então $\vec{x} = \vec{y} + \vec{z}$ e $\vec{x}' = \vec{y}' + \vec{z}'$ para alguma escolha de $\vec{y}, \vec{y}' \in W_1$ e $\vec{z}, \vec{z}' \in W_2$. Desta forma, temos que

$$\vec{x}+\vec{x}' = (\vec{y}+\vec{z})+(\vec{y}'+\vec{z}') = (\vec{y}+\vec{y}')+(\vec{z}+\vec{z}') \in W_1+W_2\ ,\ \alpha\vec{x} = \alpha(\vec{y}+\vec{z}) = \alpha\vec{y}+\alpha\vec{z} \in W_1+W_2\ ,$$

logo a afirmação segue. W_1+W_2 também pode ser entendido como o *menor* subespaço vetorial de V que contém W_1 e W_2, pois qualquer subespaço vetorial $W \supset W_1 \cup W_2$ de V contém $W_1 + W_2$: dado $\vec{x} = \vec{x}_1 + \vec{x}_2 \in W_1 + W_2$ qualquer, $\vec{x}_1 \in W_1 \subset W$, $\vec{x}_2 \in W_2 \subset W$, concluímos de (ii) que $\vec{x} \in W$ (alternativamente, poderíamos definir

Notar que $\vec{y} = \vec{y} + \vec{0}$, $\vec{z} = \vec{0}+\vec{z} \in W_1+W_2$ para $\vec{y} \in W_1$, $\vec{z} \in W_2$ quaisquer.

$W_1 + W_2$ como a intersecção de todos os subespaços vetoriais de V que contém $W_1 \cup W_2$).

Assim como no caso de intersecções de subespaços vetoriais, é possível definir a *soma* de uma família *arbitrária* $\{W_j \mid j \in J\}$ de subespaços vetoriais W_j de V:

Tal como no caso $J = \{1, 2\}$, notar que os vetores $\vec{x}_i$ na definição de $\sum_{j\in J} W_j$ podem *não* ser únicos!

$$\sum_{j\in J} W_j = \{\vec{x} \in V \mid \vec{x} = \vec{x}_1 + \cdots + \vec{x}_k \ , \vec{x}_i \in W_{j_i} \ , \ j_1, \ldots, j_k \in J \ , \ k \in \mathbb{N}\}$$

(*exercício:* verifique usando o mesmo argumento do caso $J = \{1, 2\}$ acima que $\sum_{j\in J} W_j$ é subespaço vetorial de V). Tal como no caso particular $J = \{1, 2\}$ acima, temos que $\sum_{j\in J} W_j$ é o menor subespaço vetorial de V que contém $\cup_{j\in J} W_j$, ou (equivalentemente) a intersecção de todos os subespaços vetoriais de V que contém $\cup_{j\in J} W_j$.

3. (In)dependência linear

3.1. Prelúdio: somatórias num espaço vetorial. É conveniente relembrarmos agora a definição de *somatória*, aqui devidamente estendida para vetores num espaço vetorial V. Essa operação é definida recursivamente, ou seja, por *indução* no número de parcelas em V. Dados $i_0, k \in \mathbb{Z}$, escrevemos

$$\sum_{i=i_0}^{k} \vec{y}_i = \begin{cases} \vec{0} & (i_0 > k) \\ \sum_{i=i_0}^{k-1} \vec{y}_i + \vec{y}_k & (i_0 \leq k) \end{cases} \quad , \quad \vec{y}_{i_0}, \ldots, \vec{y}_k \in V \ .$$

(notar que a coleção de vetores $\vec{y}_{i_0}, \ldots, \vec{y}_k$ é *vazia* se $i_0 > k$!) Em particular, $\sum_{i=i_0}^{i_0} \vec{y}_i = \vec{y}_{i_0}$. Mais em geral, se B é um conjunto finito, denotamos seu número de elementos por $|B|$. Se $B \subset A$, $A \neq \varnothing$ e $\vec{x} : A \to V$, escrevemos $\vec{x}_a = \vec{x}(a)$ para cada $a \in A$. Assim, podemos escrever a *somatória* dos vetores $\vec{x}_a$ *ao longo de* B como

$$\sum_{a\in B} \vec{x}_a = \begin{cases} \vec{0} & (B = \varnothing \Leftrightarrow |B| = 0) \\ \sum_{i=1}^{k} \vec{x}_{a_i} & (B = \{a_1, \ldots, a_k\} \Rightarrow |B| = k > 0) \end{cases} .$$

Como a soma vetorial é comutativa e associativa, $\sum_{a\in B} \vec{x}_a$ não depende de como enumeramos os elementos de B, portanto podemos omitir a escolha de enumeração da notação. A prova detalhada desse fato a partir dos axiomas (a.) e (b.) é elementar mas tediosa, portanto ela será relegada ao Apêndice B, onde as técnicas necessárias são desenvolvidas. Para uso futuro e como prática com manipulação de somatórias num espaço vetorial, provaremos agora algumas propriedades destas.

I.7 Lema. *Seja V um espaço vetorial, $\vec{y}, \vec{y}_i, \vec{z}_i, \vec{x}_{i,j} \in V$, $\alpha, \alpha_i \in \mathbb{K}$, $i = i_0, \ldots, k$, $j = j_0, \ldots, l$, $i_0, j_0, k, l \in \mathbb{Z}$ quaisquer. Então valem as seguintes fórmulas:*

(i) $\sum_{i=i_0}^{k}(\vec{y}_i+\vec{z}_i)=\sum_{i=i_0}^{k}\vec{y}_i+\sum_{i=i_0}^{k}\vec{z}_i$;

(ii) $\sum_{i=i_0}^{k}\vec{0}=\vec{0}$;

(iii) $\sum_{i=i_0}^{k}\alpha_i\vec{y}=\left(\sum_{i=i_0}^{k}\alpha_i\right)\vec{y}$;

(iv) $\sum_{i=i_0}^{k}\alpha\vec{y}_i=\alpha\sum_{i=i_0}^{k}\vec{y}_i$;

(v) $\sum_{i=i_0}^{k}\left(\sum_{j=j_0}^{l}\vec{x}_{i,j}\right)=\sum_{j=j_0}^{l}\left(\sum_{i=i_0}^{k}\vec{x}_{i,j}\right)$.

Demonstração. Todos os argumentos a seguir serão por indução no número k de parcelas. No caso de (v), essa fórmula será demonstrada por indução em k para l qualquer.

(i) Se $k<i_0$ segue da observação de que, nesse caso, $\sum_{i=i_0}^{k}(\vec{y}_i+\vec{z}_i)=\vec{0}$ e $\sum_{i=i_0}^{k}\vec{y}_i+\sum_{i=i_0}^{k}\vec{z}_i=\vec{0}+\vec{0}\overset{(c)}{=}\vec{0}$. Para o passo de indução se $k\geq i_0$, suponhamos válida a hipótese de indução (HI) $\sum_{i=i_0}^{k-1}(\vec{y}_i+\vec{z}_i)=\sum_{i=i_0}^{k-1}\vec{y}_i+\sum_{i=i_0}^{k-1}\vec{z}_i$. Nesse caso, temos que

$$\begin{aligned}\sum_{i=i_0}^{k}(\vec{y}_i+\vec{z}_i)&=\sum_{i=i_0}^{k-1}(\vec{y}_i+\vec{z}_i)+(\vec{y}_k+\vec{z}_k)\overset{(a)}{=}\sum_{i=i_0}^{k-1}(\vec{y}_i+\vec{z}_i)+(\vec{z}_k+\vec{y}_k)\\&\overset{(b)}{=}\left[\sum_{i=i_0}^{k-1}(\vec{y}_i+\vec{z}_i)+\vec{z}_k\right]+\vec{y}_k\overset{(HI)}{=}\left[\left(\sum_{i=i_0}^{k-1}\vec{y}_i+\sum_{i=i_0}^{k-1}\vec{z}_i\right)+\vec{z}_k\right]+\vec{y}_k\\&\overset{(b)}{=}\left[\sum_{i=i_0}^{k-1}\vec{y}_i+\left(\sum_{i=i_0}^{k-1}\vec{z}_i+\vec{z}_k\right)\right]+\vec{y}_k=\left(\sum_{i=i_0}^{k-1}\vec{y}_i+\sum_{i=i_0}^{k}\vec{z}_i\right)+\vec{y}_k\\&\overset{(a)}{=}\left(\sum_{i=i_0}^{k}\vec{z}_i+\sum_{i=i_0}^{k-1}\vec{y}_i\right)+\vec{y}_k\overset{(b)}{=}\sum_{i=i_0}^{k}\vec{z}_i+\left(\sum_{i=i_0}^{k-1}\vec{y}_i+\vec{y}_k\right)\\&=\sum_{i=i_0}^{k}\vec{z}_i+\sum_{i=i_0}^{k}\vec{y}_i\overset{(a)}{=}\sum_{i=i_0}^{k}\vec{y}_i+\sum_{i=i_0}^{k}\vec{z}_i\ ,\end{aligned}$$

como desejado.

(ii) Se $k<i_0$ a fórmula segue diretamente da definição de somatória. Para o passo de indução se $k\geq i_0$, suponhamos válida a hipótese de indução (HI) $\sum_{i=i_0}^{k-1}\vec{0}=\vec{0}$. Nesse caso, segue que

$$\sum_{i=i_0}^{k}\vec{0}=\sum_{i=i_0}^{k-1}\vec{0}+\vec{0}\overset{(HI)}{=}\vec{0}+\vec{0}\overset{(c)}{=}\vec{0}\ ,$$

como desejado.

(iii) Se $k < i_0$ segue da observação de que, nesse caso, $\sum_{i=i_0}^{k} \alpha_i \vec{y} = \vec{0}$ e $\sum_{i=i_0}^{k} \alpha_i = 0$, portanto $\left(\sum_{i=i_0}^{k} \alpha_i\right) \vec{y} = 0\vec{y} = \vec{0}$. Para o passo de indução se $k \geq i_0$, suponhamos válida a hipótese de indução (HI) $\sum_{i=i_0}^{k-1} \alpha_i \vec{y} = \left(\sum_{i=i_0}^{k-1} \alpha_i\right) \vec{y}$. Nesse caso, temos que

$$\begin{aligned}\sum_{i=i_0}^{k} \alpha_i \vec{y} &= \sum_{i=i_0}^{k-1} \alpha_i \vec{y} + \alpha_k \vec{y} \overset{(HI)}{=} \left(\sum_{i=i_0}^{k-1} \alpha_i\right) \vec{y} + \alpha_k \vec{y} \overset{(e)}{=} \left(\sum_{i=i_0}^{k-1} \alpha_i + \alpha_k\right) \vec{y} \\ &= \left(\sum_{i=i_0}^{k} \alpha_i\right) \vec{y} \,,\end{aligned}$$

como desejado.

(iv) Se $k < i_0$ segue da observação de que, nesse caso, $\sum_{i=i_0}^{k} \alpha \vec{y}_i = \vec{0}$ e $\sum_{i=i_0}^{k} \vec{y}_i = \vec{0}$, portanto $\alpha \sum_{i=i_0}^{k} \vec{y}_i = \alpha\vec{0} = \vec{0}$. Para o passo de indução se $k \geq i_0$, suponhamos válida a hipótese de indução (HI) $\sum_{i=i_0}^{k-1} \alpha \vec{y}_i = \alpha \sum_{i=i_0}^{k-1} \vec{y}_i$. Nesse caso, temos que

$$\begin{aligned}\sum_{i=i_0}^{k} \alpha \vec{y}_i &= \sum_{i=i_0}^{k-1} \alpha \vec{y}_i + \alpha \vec{y}_k \overset{(HI)}{=} \alpha \sum_{i=i_0}^{k-1} \vec{y}_i + \alpha \vec{y}_k \overset{(f)}{=} \alpha \left(\sum_{i=i_0}^{k-1} \vec{y}_i + \vec{y}_k\right) \\ &= \alpha \sum_{i=i_0}^{k} \vec{y}_i \,,\end{aligned}$$

como desejado.

(v) Se $k < i_0$ notar que, nesse caso, $\sum_{i=i_0}^{k} \left(\sum_{j=j_0}^{l} \vec{x}_{i,j}\right) = \vec{0}$ e $\sum_{j=j_0}^{l} \left(\sum_{i=i_0}^{k} \vec{x}_{i,j}\right) = \sum_{j=j_0}^{l} \vec{0} = \vec{0}$, onde a última igualdade segue da fórmula (ii) provada acima. Para o passo de indução se $k \geq i_0$, suponhamos válida a hipótese de indução (HI) $\sum_{i=i_0}^{k-1} \left(\sum_{j=j_0}^{l} \vec{x}_{i,j}\right) = \sum_{j=j_0}^{l} \left(\sum_{i=i_0}^{k-1} \vec{x}_{i,j}\right)$. Nesse caso, temos que

$$\begin{aligned}\sum_{i=i_0}^{k} \left(\sum_{j=j_0}^{l} \vec{x}_{i,j}\right) &= \sum_{i=i_0}^{k-1} \left(\sum_{j=j_0}^{l} \vec{x}_{i,j}\right) + \sum_{j=j_0}^{l} \vec{x}_{k,j} \overset{(HI)}{=} \sum_{j=j_0}^{l} \left(\sum_{i=i_0}^{k-1} \vec{x}_{i,j}\right) + \sum_{j=j_0}^{l} \vec{x}_{k,j} \\ &= \sum_{j=j_0}^{l} \left(\sum_{i=i_0}^{k-1} \vec{x}_{i,j} + \vec{x}_{k,j}\right) = \sum_{j=j_0}^{l} \left(\sum_{i=i_0}^{k} \vec{x}_{i,j}\right) \,,\end{aligned}$$

onde na última linha usamos a fórmula (i) obtida acima, provando assim a fórmula desejada.

□

As fórmulas (i) e (v) do Lema I.7 podem ser entendidas como casos particulares da invariância de somatórias por permutações de suas parcelas que mencionamos

logo antes desse Lema, com (i) sendo caso particular de (v) quando $j_0 = 1$ e $l = 2$. Mais precisamente, se arranjarmos os vetores $\vec{x}_{i,j}$ numa tabela retangular tal como as entradas de uma matriz, com $\vec{x}_{i,j}$ sendo a entrada na i-ésima linha e j-ésima coluna dessa tabela, o lado esquerdo de (v) representa a somatória ordenando as parcelas $\vec{x}_{i,j}$ *linha após linha* e o lado direito de (v), por sua vez, representa a somatória ordenando as parcelas $\vec{x}_{i,j}$ *coluna após coluna*.

3.2. Combinações lineares e (in)dependência linear.

I.8 Definição. Seja V um espaço vetorial, $\vec{x}_1, \ldots, \vec{x}_k \in V$ e $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$. A *combinação linear* (c.l.) de $\vec{x}_1, \ldots, \vec{x}_k$ *com coeficientes* $\vec{x}_1, \ldots, \vec{x}_k$ (nessa ordem) é o vetor

$$\vec{x} = \alpha_1\vec{x}_1 + \cdots + \alpha_k\vec{x}_k = \sum_{i=1}^{k} \alpha_i\vec{x}_i \ .$$

Se $\alpha_1 = \cdots = \alpha_k = 0$ (nesse caso, segue do Lema I.7 (ii) que $\vec{x} = \vec{0}$), dizemos que a c.l. acima é *trivial*. Do contrário, i.e. se $\alpha_i \neq 0$ para algum $i = 1, \ldots, k$, dizemos que a c.l. acima é *não-trivial*. Dizemos que $\vec{x} \in V$ é (ou podemos expressar $\vec{x}$ como) uma *c.l. de* $\vec{x}_1, \ldots, \vec{x}_k$ se $\vec{x}$ for da forma acima para *alguma* escolha de coeficientes $\alpha_1, \ldots, \alpha_k$.

Notar que uma c.l. de um único vetor $\vec{x}$ é simplesmente um múltiplo escalar de $\vec{x}$.

I.9 Definição. Seja V um espaço vetorial, e $\vec{x}_1, \ldots, \vec{x}_k \in V$. Dizemos que $\vec{x}_1, \ldots, \vec{x}_k$ são *linearmente dependentes* (l.d.) se existe uma c.l. *não-trivial* de $\vec{x}_1, \ldots, \vec{x}_k$ que é igual a $\vec{0}$, i.e. existem $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ *não todos zero* (ou seja, $\alpha_j \neq 0$ para *algum* $j = 1, \ldots, k$) tais que

$$\sum_{i=1}^{k} \alpha_i\vec{x}_i = \vec{0} \ .$$

Do contrário, i.e. se

$$\sum_{i=1}^{k} \alpha_i\vec{x}_i = \vec{0} \ \Rightarrow \ \alpha_1 = \cdots = \alpha_k = 0 \ ,$$

dizemos que $\vec{x}_1, \ldots, \vec{x}_k$ são *linearmente independentes* (l.i.).

Por exemplo, dois vetores $\vec{x}_1, \vec{x}_2$ são l.i. se e somente se eles são *não-colineares*, ou seja, nenhum deles é múltiplo escalar do outro. Essa interpretação de (in)dependência linear pode ser generalizada para mais de dois vetores da seguinte maneira:

I.10 Proposição. *Seja V um espaço vetorial e $\vec{x}_1, \ldots, \vec{x}_k \in V$. São equivalentes as seguintes afirmações:*

(a) $\vec{x}_1, \ldots, \vec{x}_k$ *são l.d.;*

(b) Existem $\alpha_i, \beta_i \in \mathbb{K}$, $i = 1, \ldots, k$ *tais que* $\alpha_j \neq \beta_j$ *para algum* $j \in \{1, \ldots, k\}$ *e*

$$\sum_{i=1}^{k} \alpha_i \vec{x}_i = \sum_{i=1}^{k} \beta_i \vec{x}_i \ ;$$

(c) Algum $\vec{x}_j$ *é c.l. dos outros* $\vec{x}_i$*'s , ou seja, existe um* $j \in \{1, \ldots, k\}$ *e* $\alpha_i \in \mathbb{K}$, $i = 1, \ldots, k$, $i \neq j$ *tais que*

$$\vec{x}_j = \sum_{\substack{i=1 \\ i \neq j}}^{k} \alpha_i \vec{x}_i = \sum_{i \in \{i' \in \{1,\ldots,k\} \ | \ i' \neq j\}} \alpha_i \vec{x}_i \ .$$

Demonstração. Provaremos as seguintes equivalências:

(a)⇔(b) Suponhamos que $\vec{x}_1, \ldots, \vec{x}_k$ são l.d. e sejam $\delta_1, \ldots, \delta_k \in \mathbb{K}$ tais que $\delta_j \neq 0$ para algum j e $\sum_{i=1}^{k} \delta_i \vec{x}_i = \vec{0}$. Dados $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$, defina $\beta_i = \alpha_i + \delta_i$, $i = 1, \ldots, k$, de modo que $\beta_j \neq \alpha_j$. Nesse caso, temos que

$$\begin{aligned} \sum_{i=1}^{k} \beta_i \vec{x}_i &= \sum_{i=1}^{k} (\alpha_i + \delta_i)\vec{x}_i = \sum_{i=1}^{k} (\alpha_i \vec{x}_i + \delta_i \vec{x}_i) \\ &= \sum_{i=1}^{k} \alpha_i \vec{x}_i + \sum_{i=1}^{k} \delta_i \vec{x}_i = \sum_{i=1}^{k} \alpha_i \vec{x}_i + \vec{0} \\ &= \sum_{i=1}^{k} \alpha_i \vec{x}_i \ . \end{aligned}$$

Conversamente, se $\alpha_i, \beta_i \in \mathbb{K}$, $i = 1, \ldots, k$ são tais que $\beta_j \neq \alpha_j$ para algum j e $\vec{x} = \sum_{i=1}^{k} \beta_i \vec{x}_i = \sum_{i=1}^{k} \alpha_i \vec{x}_i$, defina $\delta_i = \beta_i - \alpha_i$, de modo que $\delta_j \neq 0$ e

$$\begin{aligned} \sum_{i=1}^{k} \beta_i \vec{x}_i - \sum_{i=1}^{k} \alpha_i \vec{x}_i &= \sum_{i=1}^{k} \beta_i \vec{x}_i + \sum_{i=1}^{k} (-\alpha_i)\vec{x}_i = \sum_{i=1}^{k} (\beta_i \vec{x}_i - \alpha_i \vec{x}_i) \\ &= \sum_{i=1}^{k} (\beta_i - \alpha_i)\vec{x}_i = \sum_{i=1}^{k} \delta_i \vec{x}_i \\ &= \vec{x} - \vec{x} = \vec{0} \ , \end{aligned}$$

logo $\vec{x}_1, \ldots, \vec{x}_k$ são l.d..

(a)⇔(c) Suponhamos que $\vec{x}_1, \ldots, \vec{x}_k$ são l.d. e sejam $\delta_1, \ldots, \delta_k \in \mathbb{K}$ tais que $\delta_j \neq 0$ para algum j e $\sum_{i=1}^{k} \delta_i \vec{x}_i = \vec{0}$. Podemos então escrever

$$\vec{0} = \sum_{i=1}^{k} \delta_i \vec{x}_i = \delta_j \vec{x}_j + \sum_{\substack{i=1 \\ i \neq j}}^{k} \delta_i \vec{x}_i$$

$$\Rightarrow \vec{x}_j = -\frac{1}{\delta_j} \sum_{\substack{i=1 \\ i \neq j}}^{k} \delta_i \vec{x}_i = \sum_{\substack{i=1 \\ i \neq j}}^{k} (-\frac{\delta_i}{\delta_j}) \vec{x}_i \ .$$

Conversamente, se existem $\alpha_i \in \mathbb{K}$, $i = 1, \ldots, k$, $i \neq j$ tais que $\vec{x}_j = \sum_{\substack{i=1 \\ i \neq j}}^{k} \alpha_i \vec{x}_i$ para algum j, então definindo $\alpha_j = -1 \neq 0$ temos que $\sum_{i=1}^{k} \alpha_j \vec{x}_j = \vec{0}$, logo $\vec{x}_1, \ldots, \vec{x}_k$ são l.d..

□

Notemos que ao provar a Proposição I.10 obtivemos uma assertiva (aparentemente) mais forte: se V um espaço vetorial e $\vec{x}_1, \ldots, \vec{x}_k \in V$, então são equivalentes as seguintes afirmações:

(a) $\vec{x}_1, \ldots, \vec{x}_k$ são l.d.;

(b') Para *quaisquer* $\alpha_i \in \mathbb{K}$, $i = 1, \ldots, k$ existem $\beta_i \in \mathbb{K}$, $i = 1, \ldots, k$ tais que $\beta_j \neq \alpha_j$ para algum $j \in \{1, \ldots, k\}$ e

$$\sum_{i=1}^{k} \alpha_i \vec{x}_i = \sum_{i=1}^{k} \beta_i \vec{x}_i \ .$$

Para ver isso, notar que (b') ⇒ (b) ⇒ (a) ⇒ (b'), onde a primeira implicação é óbvia e as outras duas foram obtidas na prova da Proposição I.10. Essa Proposição pode ser também lida como sua contrapositiva, substituindo (b) por (b') para conveniência futura: se V é um espaço vetorial e $\vec{x}_1, \ldots, \vec{x}_k \in V$, então são equivalentes as seguintes afirmações:

(¬a) $\vec{x}_1, \ldots, \vec{x}_k$ são l.i.;

(¬b') Se $\alpha_i, \beta_i \in \mathbb{K}$, $i = 1, \ldots, k$ são tais que $\sum_{i=1}^{k} \alpha_i \vec{x}_i = \sum_{i=1}^{k} \beta_i \vec{x}_i$, então $\beta_i = \alpha_i$ para todo $i = 1, \ldots, k$;

(¬c) Nenhum $\vec{x}_j$ pode ser escrito como c.l. dos outros $\vec{x}_i$'s.

Em particular, $\vec{x}_1, \ldots, \vec{x}_k$ serem l.i. é o mesmo que dizer que existe *no máximo uma* maneira de escrever um vetor qualquer em V como c.l. de $\vec{x}_1, \ldots, \vec{x}_k$. A afirmação (¬c) no caso $k = 2$ se reduz à equivalência dada acima entre a independência linear e a não-colinearidade de $\vec{x}_1, \vec{x}_2 \in V$.

I.11 Definição. Se $S \subset V$, dizemos que S é *linearmente dependente* (l.d.) se existem $\vec{x}_1, \ldots, \vec{x}_k \in S$ *l.d.*. Do contrário, i.e. se quaisquer $\vec{x}_1, \ldots, \vec{x}_k \in S$ são *l.i.*, dizemos que S é *linearmente independente* (l.i.).

Podemos listar as seguintes propriedades de subconjuntos l.d. e l.i.:

(i) $S = \varnothing$ é l.i.; (não existe uma lista de vetores l.d. em $\varnothing$)

(ii) $S = \{\vec{0}\}$ é l.d.; (pois $\vec{0} = 1\vec{0}$ é uma c.l. não-trivial de $\vec{0}$)

(iii) Se $S = \{\vec{x}\}$ com $\vec{x} \neq \vec{0}$, então S é l.i.; (se $\alpha\vec{x} = \vec{0}$, nesse caso necessariamente $\alpha = 0$)

(iv) Se S é l.i., então qualquer subconjunto $\tilde{S}$ de S é l.i.;

(v) Se S é l.d. e $S \subset \tilde{S} \subset V$, então $\tilde{S}$ é l.d.. Em particular, qualquer $S \ni \vec{0}$ (e.g. subespaços vetoriais de V) é l.d..

(vi) Se S é *finito*, i.e. $S = \{\vec{x}_1, \ldots, \vec{x}_n\}$, então S é l.d. (resp. l.i.) se e somente se $\vec{x}_1, \ldots, \vec{x}_n$ são l.d. (resp. l.i.). (obviamente S l.i. $\Rightarrow \vec{x}_1, \ldots, \vec{x}_n$ l.i. e $\vec{x}_1, \ldots, \vec{x}_n$ l.d. $\Rightarrow S$ l.d.. Conversamente, se $\vec{x}_1, \ldots, \vec{x}_n$ são l.i., sejam $\vec{y}_1, \ldots, \vec{y}_k \in S$ e $\beta_1, \ldots, \beta_k \in \mathbb{K}$ tais que

$$\sum_{j=1}^{k} \beta_j \vec{y}_j = \vec{0} \, .$$

Defina

$$\alpha_i = \begin{cases} \beta_j & (\vec{x}_i = \vec{y}_j \text{ para algum } j = 1, \ldots, k) \\ 0 & \text{de outra forma} \end{cases} , \quad i = 1, \ldots, n \, .$$

Então

$$\sum_{j=1}^{k} \beta_j \vec{y}_j = \sum_{i=1}^{n} \alpha_i \vec{x}_i = \vec{0} \, ,$$

logo $\beta_1 = \cdots = \beta_k = 0$ e portanto $\vec{y}_1, \ldots, \vec{y}_k$ são l.i.. Finalmente, se S é l.d., sejam $\vec{y}_1, \ldots, \vec{y}_k \in S$ e $\beta_1, \ldots, \beta_k \in \mathbb{K}$ não todos zero tais que

$$\sum_{j=1}^{k} \beta_j \vec{y}_j = \vec{0} \, .$$

Definindo α_i como acima, $i = 1, \ldots, n$, concluímos que $\alpha_1, \ldots, \alpha_n$ não são todos zero mas

$$\sum_{i=1}^{n} \alpha_i \vec{x}_i = \sum_{j=1}^{k} \beta_j \vec{y}_j = \vec{0} \, ,$$

logo $\vec{x}_1, \ldots, \vec{x}_n$ são l.d.)

Citemos dois exemplos importantes de conjuntos l.i.:

1.) Sejam $A \neq \varnothing$, $V = \mathbb{R}^A$ e $S = \{f_p \mid p \in A\}$, onde

$$f_p(q) = \begin{cases} 0 & (q \neq p) \\ 1 & (q = p) \end{cases} , \quad q \in A .$$

Sejam então $p_1, \ldots, p_k \in A$, $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tais que $\sum_{i=1}^k \alpha_i f_{p_i} = \vec{0}$, i.e.

$$\left(\sum_{i=1}^k \alpha_i f_{p_i} \right)(q) = \sum_{i=1}^k \alpha_i f_{p_i}(q) = 0 \text{ para todo } q \in A .$$

Tomando em particular $q = p_j$, $j = 1, \ldots, k$, temos que

$$\sum_{i=1}^k \alpha_i f_{p_i}(p_j) = \alpha_j = 0 , \quad j = 1, \ldots, k ,$$

logo S é l.i.. Em particular, se $A = \{1, \ldots, n\}$ então $V = \mathbb{R}^n$ e S é a chamada *base canônica* de $\mathbb{R}^n$:

$$S = \{\vec{e}_1, \ldots, \vec{e}_n\} , \quad \vec{e}_j = (0, \ldots, 0, \underset{\substack{\uparrow \\ j\text{-ésima componente}}}{1}, 0, \ldots, 0) , \; j = 1, \ldots, n .$$

Segue do argumento geral apresentado acima que a base canônica de $\mathbb{R}^n$ é l.i..

2.) Sejam $V = \mathscr{C}^0(\mathbb{R})$ = espaço vetorial das funções contínuas em $\mathbb{R}$ (exemplo (7.) de subespaço vetorial de $\mathbb{R}^{\mathbb{R}}$ apresentado na Subseção 2.3 acima) e $S = \{f_k(t) = t^k \mid k \in \mathbb{N}_0\}$, de modo que uma c.l. $f = \sum_{j=0}^k \alpha_j f_j$ de elementos de S (lembrando que alguns α_j's podem ser zero!) é um polinômio

$$f(t) = \sum_{j=0}^k \alpha_j t^j .$$

Se $f = 0$, lembrando que um polinômio de grau k (com coeficientes reais) tem no máximo k raízes reais, concluímos daí que $\alpha_j = 0$ para todo $j = 0, \ldots, k$ e portanto S é l.i.. Notar que o mesmo argumento permanece válido se substituirmos o domínio $\mathbb{R}$ por qualquer intervalo da reta real com interior não-vazio.

Podemos provar essa afirmação e a anterior simultaneamente restringindo f a um intervalo fechado e aplicando o teorema de Rolle k vezes sucessivas (*exercício!*).

Mais uma vez, substituindo $\mathbb{R}$ por $\mathbb{C}$ em (1.) e (2.) acima fornece exemplos de conjuntos l.i. em espaços vetoriais complexos. No caso (2.), podemos tratar as partes real e imaginária de um polinômio com coeficientes complexos *restrito a* $\mathbb{R}$ como polinômios (com coeficientes) reais e aplicar o mesmo raciocínio do caso

real (*exercício:* preencha os detalhes). Um caminho alternativo para provar que um polinômio de grau $k \in \mathbb{N}$ tem no máximo k raízes, válido tanto para $\mathbb{K} = \mathbb{R}$ como para $\mathbb{K} = \mathbb{C}$, é aplicar k vezes o algoritmo de divisão de Euclides a um polinômio $f(t)$ de grau k por polinômios de primeiro grau da forma $t - \alpha$, onde α é uma raiz de $f(t)$. Para detalhes, ver o Apêndice A.

Cabe aqui apontar que, tal como no caso de (in)dependência linear de listas de vetores, o conceito de subconjuntos l.d. e l.i. depende da escolha de corpo de escalares no caso $\mathbb{K} = \mathbb{C}$. Um exemplo genérico é o seguinte: suponha que V é um espaço vetorial *complexo* e $\varnothing \neq S \subset V$ um subconjunto l.i. (com respeito a $\mathbb{C}$). Defina

$$iS = \{i\vec{x} \,|\vec{x} \in S\} \ .$$

Como S é l.i., segue imediatamente que $S \cap iS = \varnothing$. Além disso, $S \cup iS$ é claramente l.d. (com respeito a $\mathbb{C}$), mas mostraremos agora que $S \cup iS$ é *l.i. com respeito a* $\mathbb{R}$. Com efeito, sejam $\vec{x}_1, \ldots, \vec{x}_k \in S \cup iS$ (de modo que $\vec{x}_j \in S$ ou $\vec{x}_j \in iS$ mas $\vec{x}_j \notin S \cap iS (= \varnothing)$ para todo $j = 1, \ldots, k$) e $\beta_1, \ldots, \beta_k \in \mathbb{R}$ tais que $\sum_{j=1}^{k} \beta_j \vec{x}_j = \vec{0}$. Defina

$$\alpha_j = \begin{cases} \beta_j & (\vec{x}_j \in S) \\ i\beta_j & (\vec{x}_j \in iS) \end{cases} \ , \ \vec{y}_j = \begin{cases} \vec{x}_j & (\vec{x}_j \in S) \\ -i\vec{x}_j & (\vec{x}_j \in iS) \end{cases} \ , \ j = \ldots, k \ ,$$

de modo que $\vec{y}_j \in S$ para todo $j = 1, \ldots, k$ e

$$\sum_{j=1}^{k} \beta_j \vec{x}_j = \vec{0} = \sum_{j=1}^{k} \alpha_j \vec{y}_j \ .$$

Como S é l.i., temos que $\alpha_1 = \cdots = \alpha_k = 0$ e portanto $\beta_1 = \cdots = \beta_k = 0$, logo $\vec{x}_1, \ldots, \vec{x}_k$ são l.i. com respeito a $\mathbb{R}$. Como essa lista de vetores foi escolhida arbitrariamente em $S \cup iS$, concluímos que $S \cup iS$ é de fato l.i. com respeito a $\mathbb{R}$.

3.3. Subespaços vetoriais gerados por um subconjunto, bases de um espaço vetorial, dimensão.

I.12 Definição. Seja V espaço vetorial, $S \subset V$. O *subespaço vetorial gerado* por S (também conhecido como *varredura linear* de S) é dado por

$$L(S) = \begin{cases} \{\vec{0}\} & (S = \varnothing) \\ \left\{ \vec{x} = \sum_{i=1}^{k} \alpha_i \vec{x}_i \;\middle|\; \alpha_1, \ldots, \alpha_k \in \mathbb{K} \ , \ \vec{x}_1, \ldots, \vec{x}_k \in V \right\} & (S \neq \varnothing) \end{cases} .$$

Usa-se também a notação alternativa $L(S) = \mathrm{span}\,(S)$, onde "*span*" = "varredura" em inglês.

Mostraremos que $L(S)$ é subespaço vetorial de V. Obviamente isso é verdade se $S = \varnothing$, então podemos assumir que $S \neq \varnothing$. Se $\vec{x} \in S$, então claramente $0\vec{x} = \vec{0} \in L(S)$. Resta então apenas mostrar que se $\vec{x}, \vec{y} \in L(S)$ e $\alpha \in \mathbb{R}$, então $\vec{x} + \vec{y}$, $\alpha\vec{x} \in L(S)$. De fato, nesse caso podemos escrever

$$\vec{x} = \sum_{i=1}^{m} \alpha_i \vec{x}_i \ , \ \vec{y} = \sum_{j=1}^{n} \beta_j \vec{y}_j$$

com $\vec{x}_1, \ldots, \vec{x}_m, \vec{y}_1, \ldots, \vec{y}_n \in S$, $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n \in \mathbb{K}$. Podemos, contudo, escrever $\vec{x}, \vec{y}$ como c.l.'s da *mesma* lista de vetores de S, similarmente à prova da propriedade (vi) obtida logo após a Definição I.11. A saber, defina

$$\tilde{S} = \{\vec{x}_1, \ldots, \vec{x}_m\} \cup \{\vec{y}_1, \ldots, \vec{y}_n\} = \{\vec{z}_1, \ldots, \vec{z}_p\} \ ,$$

$$\alpha'_k = \begin{cases} \alpha_i & (\vec{z}_k = \vec{x}_i \text{ para algum } i = 1, \ldots, m) \\ 0 & \text{(de outra forma)} \end{cases} \ , \quad k = 1, \ldots, p \ ,$$

$$\beta'_k = \begin{cases} \beta_j & (\vec{z}_k = \vec{y}_j \text{ para algum } j = 1, \ldots, n) \\ 0 & \text{(de outra forma)} \end{cases} \ , \quad k = 1, \ldots, p \ ,$$

de modo que

$$\vec{x} = \sum_{k=1}^{p} \alpha'_k \vec{z}_k \ , \ \vec{y} = \sum_{k=1}^{p} \beta'_k \vec{z}_k \ .$$

Uma vez feito isso, segue que

$$\begin{aligned} \vec{x} + \vec{y} &= \sum_{k=1}^{p} \alpha'_k \vec{z}_k + \sum_{k=1}^{p} \beta'_k \vec{z}_k = \sum_{k=1}^{p} (\alpha'_k + \beta'_k)\vec{z}_k \in L(S) \ , \\ \alpha\vec{x} &= \alpha \sum_{k=1}^{p} \alpha'_k \vec{z}_k = \sum_{k=1}^{p} (\alpha\alpha'_k)\vec{z}_k \in L(S) \ , \end{aligned} \tag{I.2}$$

logo $L(S)$ é subespaço vetorial de V. Outra maneira de definir $L(S)$ é como o *menor* subespaço vetorial de V contendo S. De fato, se W é um subespaço vetorial de V contendo S, então claramente $W \supset L(S)$. Alternativamente, graças a isso podemos também definir $L(S)$ como a interseção de todos os subespaços vetoriais de V contendo S.

(I.2) mostra que, em termos dos coeficientes das c.l.'s de vetores em $S \neq \varnothing$, as operações vetoriais em $L(S)$ são exatamente as operações vetoriais *pontuais* de $F(S, \mathbb{R})$. Além disso, a Proposição I.10 nos diz que se S é *l.i.*, então a representação dos vetores em $L(S)$ em termos de c.l.'s de vetores em S é *única*. Ou seja, se S é l.i. então S fornece um "sistema de coordenadas lineares" para $L(S)$ no sentido de que não há duas escolhas diferentes de coeficientes para o mesmo vetor em $L(S)$ e as operações lineares de V agem "coeficiente a coeficiente", tal como em $\mathbb{R}^n$. Isso motiva a seguinte

I.13 Definição. Seja V espaço vetorial, $S \subset V$. Se S é l.i. e $L(S) = V$, dizemos que S é *base* de V.

Obviamente, todo $S \subset V$ l.i. é base de $L(S)$. Mais em geral, como $L(S) \subset V$ para todo $S \subset V$, para verficar se um subconjunto l.i. $S \subset V$ é base de V basta checar se todo $\vec{x} \in V$ é c.l. de vetores em S.

Dentre (contra)exemplos de bases, podemos citar:

1.) $V = \mathbb{R}^n$, $S = \{\vec{e}_1, \ldots, \vec{e}_n\}$ = base canônica de $\mathbb{R}^n$. Já mostramos acima que tal S é l.i., logo resta apenas mostrar que todo $\vec{x} = (x_1, \ldots, x_n) \in V$ é c.l. dos vetores em S. De fato,

$$\begin{aligned}\vec{x} = (x_1, \ldots, x_n) &= (x_1, 0, \ldots, 0) + \cdots + (0, \ldots, 0, x_n) \\ &= x_1(1, 0, \ldots, 0) + \cdots + x_n(0, \ldots, 0, 1) \\ &= \sum_{i=1}^{n} x_i \vec{e}_i \in L(S) \ .\end{aligned}$$

2.) Dado $A \neq \varnothing$, sejam $V = \mathbb{R}^A$, $S = \{f_p \mid p \in A\}$, onde

$$f_p(q) = \begin{cases} 0 & (q \neq p) \\ 1 & (q = p) \end{cases} , \quad q \in A \ .$$

Vimos acima que tal S é l.i.. Contudo,

$$\begin{aligned} L(S) &= F_0(A, \mathbb{R}) \\ &= \{f : A \to \mathbb{R} \mid \text{existe } J \subset A \textit{ finito} \\ &\qquad\qquad \text{tal que } f(p) = 0 \text{ se } p \notin J\} \ . \end{aligned}$$

Primeiramente, é fácil notar que $F_0(A, \mathbb{R})$ é subespaço vetorial de $\mathbb{R}^A$. Além disso, obviamente $S \subset F_0(A, \mathbb{R})$ pois $f_p(q) = 0$ se $q \notin \{p\}$ para todo $p \in A$, logo $L(S) \subset F_0(A, \mathbb{R})$. Conversamente, se $f \in F_0(A, \mathbb{R})$ seja $J = \{p_1, \ldots, p_k\} \subset A$ tal que $f(q) = 0$ se $q \notin J$, de modo que

$$\sum_{i=1}^{k} f(p_i) f_{p_i}(q) = \begin{cases} 0 & (q \notin J) \\ f(p_i) & (q = p_i \text{ para algum } i = 1, \ldots, k) \end{cases} ,$$

logo

$$f = \sum_{i=1}^{k} f(p_i) f_{p_i} \in L(S) \ ,$$

como desejado. Contudo, temos $F_0(A, \mathbb{R}) = \mathbb{R}^A$ se e somente se A for *finito*. Neste caso, sempre podemos tomar $J = A$ na definição de $F_0(A, \mathbb{R})$. Por

outro lado, se A é infinito então claramente a função $f(p) \equiv 1$ (i.e. $f(p) = 1$ para todo $p \in A$) *não* pertence a $F_0(A, \mathbb{R})$. Portanto, neste último caso S não é base de V.

3.) Seja $S \subset V = \mathbb{R}^{\mathbb{R}}$ o subconjunto l.i. dado no exemplo (2.) da Subseção 3.2. Nota-se imediatamente que $L(S) = \mathcal{P}(\mathbb{R})$ = espaço dos polinômios com coeficientes reais (ver exemplo (5.) da Subseção 2.3).

O último exemplo acima mostra que se S é base de um espaço vetorial V, então V expresso pelos coeficientes das c.l.'s de elementos de S pode ser *identificado* como espaço vetorial com $F_0(S, \mathbb{R})$. Mais precisamente, se $\vec{x} = \sum_{i=1}^{k} \alpha_i \vec{z}_i$, onde $\vec{z}_i \in S$, $\alpha_i \in \mathbb{R}$, $i = 1, \ldots, k$, definimos $f_{\vec{x}} \in F_0(S, \mathbb{R})$ como

$$f_{\vec{x}}(\vec{z}) = \begin{cases} \alpha_i & (\vec{z} = \vec{z}_i \text{ para algum } i = 1, \ldots, k) \\ 0 & (\text{de outra forma}) \end{cases} .$$

Denominamos $f_{\vec{x}}(\vec{z})$ a *componente de $\vec{x}$ ao longo de $\vec{z}$ em S*. Conversamente, se $f \in F_0(S, \mathbb{R})$ e $S_0 = \{\vec{z}_1, \ldots, \vec{z}_k\} \subset S$ for tal que $f(\vec{z}) = 0$ se $\vec{z} \notin S_0$, definimos $\vec{x}_f \in V$ como

$$\vec{x}_f = \sum_{\vec{z} \in S_0} f(\vec{z})\vec{z} = \sum_{i=1}^{k} f(\vec{z}_i)\vec{z}_i \ .$$

Desta forma, temos que $f_{\vec{x}_f} = f$ e $\vec{x}_{f_{\vec{x}}} = \vec{x}$ (*exercício:* verifique!). Mais ainda, temos que $f_{\vec{y}+\vec{z}} = f_{\vec{y}} + f_{\vec{z}}$, $f_{\alpha\vec{y}} = \alpha f_{\vec{y}}$ para quaisquer $\vec{y}, \vec{z} \in V$, $\alpha \in \mathbb{R}$ e $\vec{x}_{g+h} = \vec{x}_g + \vec{x}_h$, $\vec{x}_{\alpha g} = \alpha \vec{x}_g$ para quaisquer $g, h \in F_0(S, \mathbb{R})$, $\alpha \in \mathbb{R}$.

Checar se um conjunto l.i. $S \subset V$ é base de V é uma tarefa grandemente facilitada se S é *finito*, graças a dois resultados cruciais:

I.14 Teorema. *Seja V um espaço vetorial e $S \subset V$ l.i. Se $S = \{\vec{x}_1, \ldots, \vec{x}_n\}$, então quaisquer $n+1$ vetores $\vec{y}_1, \ldots, \vec{y}_{n+1} \in L(S)$ são* l.d..

Demonstração. O resultado será provado por indução em n. No caso $n = 1$, i.e. $S = \{\vec{x}_1\}$ (em particular, $\vec{x}_1 \neq \vec{0}$) e $L(S) = \{\alpha_1\vec{x}_1 \mid \alpha_1 \in \mathbb{R}\}$. Sejam então $\vec{y}_1 = \alpha_1\vec{x}_1, \vec{y}_2 = \beta_1\vec{x}_1 \in L(S)$ – se $\alpha_1 = 0$ ou $\beta_1 = 0$ então $\vec{y}_1, \vec{y}_2$ são obviamente l.d.; se, por outro lado, $\alpha_1, \beta_1 \neq 0$, podemos escrever

$$\vec{y}_2 = \frac{\beta_1}{\alpha_1}\vec{y}_1 \ \Rightarrow \ \vec{y}_2 - \frac{\beta_1}{\alpha_1}\vec{y}_1 = \vec{0} \ ,$$

logo $\vec{y}_1, \vec{y}_2$ também são l.d. nesse caso. Assumamos agora que o Teorema vale no caso $n = k - 1$ – provaremos que o mesmo vale no caso $n = k$. Escrevamos

$$\vec{y}_j = \sum_{i=1}^{k} A_{ij}\vec{x}_i \ , \quad j = 1, \ldots, k+1 \ .$$

Há duas possibilidades a considerar:

(i) $A_{1j} = 0$ para todo $j = 1, \ldots, k+1$: neste caso, temos que

$$\vec{y}_j = \sum_{i=2}^{n} A_{ij}\vec{x}_i \ , \quad j = 1, \ldots, k+1$$

e portanto $\vec{y}_1, \ldots \vec{y}_{k+1}$ são c.l.'s dos $k-1$ vetores l.i. $\vec{x}_2, \ldots, \vec{x}_k$. Pela hipótese de indução, $\vec{y}_1, \ldots, \vec{y}_{k+1}$ são l.d..

(ii) $A_{11} \neq 0$ (isso sempre pode ser obtido reordenando-se os $\vec{y}_j$'s): defina

$$c_j = \frac{A_{1j}}{A_{11}} \ , \quad j = 2, \ldots, k+1 \ ,$$

de modo que

$$c_j\vec{y}_1 = \sum_{i=1}^{k} c_j A_{i1}\vec{x}_i = A_1^j\vec{x}_1 + \sum_{i=2}^{k} c_j A_{i1}\vec{x}_i$$

e portanto

$$c_j\vec{y}_1 - \vec{y}_j = \sum_{i=2}^{k} (c_j A_{i1} - A_{ij})\vec{x}_i \ , \quad j = 2, \ldots, k+1 \ .$$

Logo, os k vetores $c_j\vec{y}_1 - \vec{y}_j$, $j = 2, \ldots, k+1$ são c.l.'s dos $k-1$ vetores l.i. $\vec{x}_2, \ldots, \vec{x}_k$. Pela hipótese de indução, existem $\alpha_2, \ldots, \alpha_{k+1} \in \mathbb{R}$ não todos zero tais que

$$\sum_{j=2}^{k+1} \alpha_j(c_j\vec{y}_1 - \vec{y}_j) = \left(\sum_{j=2}^{k+1} \alpha_j c_j\right)\vec{y}_1 - \sum_{j=2}^{k+1} \alpha_j\vec{y}_j = \vec{0} \ .$$

A segunda expressão claramente é uma c.l. não-trivial de $\vec{y}_1, \ldots, \vec{y}_{k+1}$, logo esses vetores são l.d..

□

I.15 Lema. *Seja $S \subset V$ l.i. e $\vec{y} \notin L(S)$. Então $S \cup \{\vec{y}\}$ também é l.i..*

Demonstração. Sejam $\vec{x}_1, \ldots, \vec{x}_k \in S$, e $\alpha_1, \ldots, \alpha_{k+1} \in \mathbb{R}$ tais que

$$\sum_{i=1}^{k} \alpha_i\vec{x}_i + \alpha_{k+1}\vec{y} = \vec{0} \ .$$

Há duas possibilidades:

(i) $\alpha_{k+1} = 0$: então $\sum_{i=1}^{k} \alpha_i\vec{x}_i = \vec{0}$ e portanto $\alpha_1 = \cdots = \alpha_k = 0$, logo $\vec{x}_1, \ldots, \vec{x}_k, \vec{y}$ são l.i.;

(ii) $\alpha_{k+1} \neq 0$: podemos escrever

$$\vec{y} = -\frac{1}{\alpha_{k+1}} \sum_{i=1}^{k} \alpha_i \vec{x}_i = \sum_{i=1}^{k} \left(-\frac{\alpha_i}{\alpha_{k+1}} \right) \vec{x}_i \in L(S)$$

(absurdo).

□

Como consequência desses dois resultados, temos o seguinte critério para bases finitas:

I.16 Corolário (critério dimensional para bases). *Seja V um espaço vetorial e $S \subset V$ l.i. Se $S = \{\vec{x}_1, \ldots, \vec{x}_n\}$, então $\tilde{S} \subset L(S)$ l.i. é base de $L(S)$ (i.e. $L(\tilde{S}) = L(S)$) se e somente se $\tilde{S}$ tiver o* mesmo número *de vetores que S.*

Demonstração. Considere $\tilde{S} \subset L(S)$ l.i. tal que $L(\tilde{S}) = L(S)$. Pelo Teorema I.14, $\tilde{S}$ não pode ter mais vetores do que S, pois do contrário teríamos $n + 1$ vetores de $\tilde{S}$ em $L(S)$ que portanto são l.d., o que é absurdo pois $\tilde{S}$ é l.i.. Logo, podemos escrever $\tilde{S} = \{\vec{y}_1, \ldots, \vec{y}_k\}$, $1 \leq k \leq n$. Pelo mesmo motivo, S não pode ter mais vetores do que $\tilde{S}$ pois do contrário teríamos $k + 1$ vetores de S em $L(\tilde{S})$ que portanto são l.d., o que é absurdo pois S é l.i..

Conversamente, seja $\tilde{S} \subset L(S)$ l.i. tal que $\tilde{S} = \{\vec{y}_1, \ldots, \vec{y}_n\}$, e suponha que existe $\vec{y}_{n+1} \in L(S) \setminus L(\tilde{S})$. Pelo Lema I.15, $\vec{y}_1, \ldots, \vec{y}_{n+1}$ são l.i., o que é absurdo pelo Teorema I.14. □

Em particular, o Corolário I.16 nos diz que o número de vetores de uma base *finita* de V *não depende de S*, apenas de V. Em suma, o número de "coordenadas lineares" de V é o mesmo para *todas* as bases de V se uma delas (logo, todas) for(em) finita(s). Isso motiva a seguinte

I.17 Definição. Seja V espaço vetorial. Se V possui uma base finita S, dizemos que V *tem dimensão finita* (notação: $\dim(V) < \infty$). Caso contrário, dizemos que V *tem dimensão infinita* (notação: $\dim(V) = \infty$). A *dimensão de V* é o número $\dim(V) \in \mathbb{N} \cup \{\infty\}$ dado por

$$\dim(V) = \begin{cases} 0 & V = \{\vec{0}\} \\ n = |S| & S = \{\vec{x}_1, \ldots, \vec{x}_n\} \text{ base de } V \\ \infty & \dim(V) = \infty \end{cases}.$$

Exemplos:

1.) $\dim(\mathbb{R}^n) = n$;

2.) Mais em geral, dado $A \neq \emptyset$, então

$$\dim(F_0(A, \mathbb{R})) = \begin{cases} |A| & A \text{ finito} \\ \infty & A \text{ infinito} \end{cases}.$$

De fato, a base $S = \{f_p \mid p \in A\}$ de $F_0(A, \mathbb{R})$ tem claramente o mesmo número de elementos que A. Em particular, se A é infinito, então $F_0(A, \mathbb{R})$ e portanto $\mathbb{R}^A$ contém um subconjunto l.i. infinito. Pelo Teorema I.10, isso implica que tanto $F_0(A, \mathbb{R})$ quanto $\mathbb{R}^A$ tem dimensão infinita se A é infinito.

3.) Se $V = \mathscr{C}^0(\mathbb{R})$, lembrando que $S = \{f_k(t) = t^k \mid k \in \mathbb{N}_0\} \subset V$ é l.i. (exemplo (2.) apresentado na Subseção 3.2 acima) e infinito, concluímos tal como em (2.) logo acima que $\dim(V) = \infty$. O mesmo resultado é válido se substituirmos o domínio $\mathbb{R}$ por qualquer intervalo.

Notar que se W é subespaço vetorial de V e $\dim(V) < \infty$, então a dimensão de W não pode ser maior que a de V, pois uma base de W não pode ter mais vetores do que uma base de V. Isso é evidente em vista do Lema I.15 e do Corolário I.16. Esses resultados, por sua vez, fornecem um algoritmo para a obtenção de uma base de V se V tem dimensão finita. O método funciona por indução no número de vetores.

0.) Tome $S_0 = \emptyset$. Se $W = \{\vec{0}\}$, não há mais nada a fazer – tome $S = S_0$ base de V. Se, por outro lado, existe $\vec{0} \neq \vec{x}_1 \in V$, defina $S_1 = \{\vec{x}_1\}$. Esse conjunto é l.i.. Tome $k = 1$.

1.) O passo de indução é o seguinte: seja $S_k = \{\vec{x}_1, \ldots, \vec{x}_k\} \subset V$ l.i., $k \geq 1$. Se $L(S_k) = V$ (ou seja, $k = \dim(V)$), podemos parar e tomar $S = S_k$ como base de V. Do contrário, existe $\vec{x}_{k+1} \in V$ tal que $\vec{x}_{k+1} \notin L(S_k)$. Pelo Lema I.15, $S_{k+1} = S_k \cup \{\vec{x}_{k+1}\}$ é l.i..

2.) Repita o passo 1.) trocando k por $k + 1$ até encontrarmos uma base S de V da maneira descrita nesse passo.

Pelo Corolário I.16, e pela observação acima, podemos repetir o passo de indução (1.) no máximo $\dim(V)$ vezes, então uma base de V é de fato encontrada após um número finito de passos. Notar que podemos iniciar o algoritmo acima começando com S_k para qualquer $0 \leq k \leq \dim(V)$. Nesse caso, se $L(S_k) = W$, a base S obtida possui a propriedade de que existe um subconjunto de S (a saber, S_k) que gera W. Tal base é dita *adaptada* a(o subespaço vetorial) W.

Por exemplo, seja $V = \mathbb{R}^3$ e $W = \{\vec{x} = (x, y, z) \in V \mid y = z\}$ (*exercício:* verifique que W é subespaço vetorial de V). Notar que $(1, 0, 0), (0, 1, 1) \in W$ são l.i. pois são não-colineares, portanto pelo Teorema I.14 $\dim(W) \geq 2$. Como

$(0, 1, 0) \notin W$, concluímos que $W \neq V$ e portanto $\dim(W) = 2$, de modo que $S_2 = \{(1, 0, 0), (0, 1, 1)\}$ é base de W pelo Corolário I.16 e $(0, 1, 0) \notin L(S_2)$. Assim, $S = S_3 = S_2 \cup \{(0, 1, 0)\}$ é base de V.

No próximo Capítulo, usaremos informação geométrica adicional para aprimorar nossa escolha de bases.

II

Produtos escalares

1. Definição e exemplos

1.1. O produto escalar canônico de $\mathbb{R}^n$. Dados vetores $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$, o *produto escalar* ou *produto interno* (canônico) de $\vec{x}$ e $\vec{y}$ é dado por

$$\vec{x} \cdot \vec{y} = \langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^{n} x_i y_i \ .$$

O produto escalar satisfaz as seguintes propriedades, denominadas *axiomas do produto escalar*: se $\vec{x}, \vec{y}, \vec{z} = (z_1, \ldots, z_n) \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ são quaisquer, então

(a) $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i y_i = \sum_{i=1}^n y_i x_i = \langle \vec{y}, \vec{x} \rangle$; (*simetria*)

(b) $\langle \vec{x}, \vec{x} \rangle = \sum_{i=1}^n x_i^2 > 0$ se $\vec{x} \neq \vec{0}$; (*positividade definida*)

(c_2) $\langle \vec{x}, \vec{y} + \vec{z} \rangle = \sum_{i=1}^n x_i (y_i + z_i) = \sum_{i=1}^n (x_i y_i + x_i z_i) = \sum_{i=1}^n x_i y_i + \sum_{i=1}^n x_i z_i = \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{z} \rangle$, $\langle \vec{x}, \alpha \vec{y} \rangle = \sum_{i=1}^n x_i (\alpha y_i) = \alpha \sum_{i=1}^n x_i y_i = \alpha \langle \vec{x}, \vec{y} \rangle$. (*linearidade na segunda variável*)

Os axiomas (a) e (c_2) juntos implicam

(c_1) $\langle \vec{x} + \vec{y}, \vec{z} \rangle \overset{(a)}{=} \langle \vec{z}, \vec{z} + \vec{y} \rangle \overset{(c_2)}{=} \langle \vec{z}, \vec{x} \rangle + \langle \vec{z}, \vec{y} \rangle \overset{(a)}{=} \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$, $\langle \alpha \vec{x}, \vec{y} \rangle \overset{(a)}{=} \langle \vec{y}, \alpha \vec{x} \rangle \overset{(c_2)}{=} \alpha \langle \vec{y}, \vec{x} \rangle \overset{(a)}{=} \alpha \langle \vec{x}, \vec{y} \rangle$. (*linearidade na primeira variável*)

As propriedades (c_1) e (c_2) juntas são chamadas de (c) *bilinearidade*, que pode ser lida sucintamente como: *podemos colocar as operações vetoriais de $\mathbb{R}^n$ "para fora" de um dos argumentos do produto escalar canônico, mantendo o outro argumento fixo.*

1.2. Axiomas do produto escalar. Mais em geral, se V é um espaço vetorial e $\omega : V \times V \to \mathbb{R}$ é uma função de duas variáveis em V a valores em $\mathbb{R}$ satisfazendo os axiomas

(a) $\omega(\vec{x}, \vec{y}) = \omega(\vec{y}, \vec{x})$; (*simetria*)

(b) $\omega(\vec{x}, \vec{x}) > 0$ se $\vec{x} \neq \vec{0}$; (*positividade definida*)

(c_2) $\omega(\vec{x}, \vec{y} + \vec{z}) = \omega(\vec{x}, \vec{y}) + \omega(\vec{x}, \vec{z})$, $\omega(\vec{x}, \alpha \vec{y}) = \alpha \omega(\vec{x}, \vec{y})$; (*linearidade na segunda variável*)

dizemos que ω é um produto escalar ou produto interno em V.

Se ω satisfaz apenas (c), dizemos que ω é uma *forma bilinear* em V.

Tal como no caso do produto escalar canônico em $\mathbb{R}^n$, os axiomas (a) e (c_2) juntos implicam

(c_1) $\omega(\vec{x}+\vec{y},\vec{z}) = \omega(\vec{x},\vec{z}) + \omega(\vec{y},\vec{z})$; $\omega(\alpha\vec{x},\vec{y}) = \alpha\omega(\vec{x},\vec{y})$, (*linearidade na primeira variável*)

e chamamos os axiomas (c_1) e (c_2) juntos de (c) *bilinearidade*. Outras consequências notáveis de (a) e (c_2), que usaremos com frequência, são:

(d) $\omega(\vec{0},\vec{x}) \overset{(a)}{=} \omega(\vec{x},\vec{0}) = \omega(\vec{x},0\vec{0}) \overset{(c_2)}{=} 0\omega(\vec{x},\vec{0}) = 0$ para todo $\vec{x} \in V$. Em particular, segue de (b) que $\omega(\vec{x},\vec{x}) \geq 0$ para todo $\vec{x} \in V$, com igualdade se e somente se $\vec{x} = \vec{0}$.

(e_2) Dados $\vec{x},\vec{y}_1,\ldots,\vec{y}_k \in V$, $\alpha_1,\ldots,\alpha_k \in \mathbb{R}$ quaisquer, temos que

$$\omega\left(\vec{x}, \sum_{i=1}^{k} \alpha_i\vec{y}_i\right) = \sum_{i=1}^{k} \alpha_i\omega(\vec{x},\vec{y}_i) \ .$$

(de fato, o caso inicial $k = 0$ segue de (d) acima. Para obter o passo de indução $k-1 \to k$, $k > 0$ suponhamos a validade da hipótese de indução (HI) $\omega\left(\vec{x}, \sum_{i=1}^{k-1} \alpha_i\vec{y}_i\right) = \sum_{i=1}^{k-1} \alpha_i\omega(\vec{x},\vec{y}_i)$. Nesse caso, temos que

$$\begin{aligned}
\omega\left(\vec{x}, \sum_{i=1}^{k} \alpha_i\vec{y}_i\right) = \omega\left(\vec{x}, \sum_{i=1}^{k-1} \alpha_i\vec{y}_i + \alpha_k\vec{y}_k\right) &\overset{(c_2)}{=} \omega\left(\vec{x}, \sum_{i=1}^{k-1} \alpha_i\vec{y}_i\right) + \omega(\vec{x},\alpha_k\vec{y}_k) \\
&\overset{(c_2)}{=} \omega\left(\vec{x}, \sum_{i=1}^{k-1} \alpha_i\vec{y}_i\right) + \alpha_k\omega(\vec{x},\vec{y}_k) \overset{(HI)}{=} \sum_{i=1}^{k-1} \alpha_i\omega(\vec{x},\vec{y}_i) + \alpha_k\omega(\vec{x},\vec{y}_k) \\
&= \sum_{i=1}^{k} \alpha_i\omega(\vec{x},\vec{y}_i) \ ,
\end{aligned}$$

como desejado)

(e_1) Dados $\vec{x}_1,\ldots,\vec{x}_k,\vec{y} \in V$, $\alpha_1,\ldots,\alpha_k \in \mathbb{R}$ quaisquer, temos que

$$\omega\left(\sum_{i=1}^{k} \alpha_i\vec{x}_i, \vec{y}\right) = \sum_{i=1}^{k} \alpha_i\omega(\vec{x}_i,\vec{y}) \ .$$

(de fato, segue que

$$\begin{aligned}
\omega\left(\sum_{i=1}^{k} \alpha_i\vec{x}_i, \vec{y}\right) &\overset{(a)}{=} \omega\left(\vec{y}, \sum_{i=1}^{k} \alpha_i\vec{x}_i\right) \\
&\overset{(e_2)}{=} \sum_{i=1}^{k} \alpha_i\omega(\vec{y},\vec{x}_i) \\
&\overset{(a)}{=} \sum_{i=1}^{k} \alpha_i\omega(\vec{x}_i,\vec{y}) \ ,
\end{aligned}$$

como desejado)

De maneira análoga à bilinearidade (c), as propriedades (e_1) e (e_2) podem ser lidas da seguinte forma: podemos colocar c.l.'s "para fora" de um argumento de ω mantendo o outro argumento fixo.

Não é difícil produzir exemplos de produtos escalares em V se V tem dimensão finita. Se $S = \{\vec{e}_1, \ldots, \vec{e}_n\}$ é base de V, dados $\vec{x} = \sum_{i=1}^n x_i\vec{e}_i, \vec{y} = \sum_{i=1}^n y_i\vec{e}_i \in V$ podemos escrever

$$\langle \vec{x}, \vec{y}\rangle_S = \sum_{i=1}^n x_i y_i \ .$$

Pelos mesmo cálculos empregados acima para o produto escalar canônico em $\mathbb{R}^n$, verificamos que $\langle \vec{x}, \vec{y}\rangle_S$ é um produto escalar em V, denominado *produto escalar associado a S*. Se $V = \mathbb{R}^n$ e S é a base canônica de $\mathbb{R}^n$, então claramente o produto escalar canônico é o produto escalar associado à base canônica de $\mathbb{R}^n$.

Vejamos agora um exemplo de produto escalar num espaço vetorial de dimensão infinita. Seja $V = \mathscr{C}^0([a, b])$ = espaço vetorial das funções contínuas no intervalo $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$. Definamos

$$\omega(f, g) = \int_a^b f(t)g(t)\, dt \ .$$

Tal ω é claramente (a) simétrica. Como a integral (de Riemann) em $[a, b]$ satisfaz para quaisquer $f, g \in V$, $\alpha \in \mathbb{R}$

$$\int_a^b (f(t) + g(t))\, dt = \int_a^b f(t)\, dt + \int_a^b g(t)\, dt \ , \ \int_a^b \alpha f(t)\, dt = \alpha \int_a^b f(t)\, dt$$

e para qualquer $f \in V$

$$\int_a^b f(t)^2\, dt > 0 \text{ se } f \neq 0 \ ,$$

segue daí que ω é (b) positiva definida e (c) bilinear (*exercício:* verifique a última afirmação usando as fórmulas dadas acima), portanto ω é de fato um produto escalar em V.

2. Ortogonalidade e ortonormalidade

A noção de produto escalar associado a uma base de um espaço vetorial pode ser vista de outra maneira ao indagarmo-nos se um produto escalar *dado* é associado a alguma base:

II.1 Definição. Seja V um espaço vetorial, $S \subset V$ e ω um produto escalar em V. Dizemos que S é:

- *Ortogonal* (com respeito a ω) se $\omega(\vec{x}, \vec{y}) = 0$ para todo $\vec{x}, \vec{y} \in S, \vec{x} \neq \vec{y}$.

- *Ortonormal* (o.n.) (com respeito a ω) se, além disso, $\omega(\vec{x}, \vec{x}) = 1$ para todo $\vec{x} \in S$.

Dois vetores $\vec{x}, \vec{y} \in V$ são ditos *ortogonais* ou *perpendiculares* (com respeito a ω) se $\omega(\vec{x}, \vec{y}) = 0$ – também diz-se nesse caso que $\vec{x}$ (resp. $\vec{y}$) é *ortogonal* ou *perpendicular* a $\vec{y}$ (resp. $\vec{x}$) (com repseito a ω). Um vetor $\vec{x} \in V$ é dito *ortogonal* ou *perpendicular* a $\varnothing \neq S \subset V$ (com respeito a ω) se $\vec{x}$ é ortogonal a *todo* $\vec{y} \in S$ (com respeito a ω).

Segue da definição II.1 que $S = \varnothing$ é (trivialmente) o.n.. No restante do texto, assumiremos sem prejuízo que todo subconjunto ortogonal dado é *não-vazio*, a menos que seja dito o contrário.

Pode-se facilmente mostrar que um subconjunto ortogonal $S \subset V$ que *não contém* $\vec{0}$ é l.i.. Em particular, devido a (b), todo subconjunto o.n. de V é l.i.. De fato, se $\vec{x}_1, \ldots, \vec{x}_k \in S$, $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ são tais que

$$\sum_{i=0}^{k} \alpha_i \vec{x}_i = \vec{0} \ ,$$

então

$$=\overset{d}{=} \omega(\vec{x}_j, \vec{0}) \overset{e_2}{=} \sum_{i=1}^{n} \alpha_i \omega(\vec{x}_j, \vec{x}_i) = \alpha_j \omega(\vec{x}_j, \vec{x}_j) \ .$$

Os pontos denotam as posições dos dois argumentos do produto escalar.

Como $\vec{x}_j \neq \vec{0}$, concluímos a partir de (b) que $\alpha_j = 0$ para cada $j = 1, \ldots, k$.

Notar agora que, dada uma base (finita) $S = \{\vec{e}_1, \ldots, \vec{e}_n\}$ de V, temos que $\langle \cdot, \cdot \rangle_S$ é o *único* produto escalar em V com respeito ao qual S é *o.n.*. De fato, é fácil ver que

$$\langle \vec{e}_i, \vec{e}_j \rangle_S = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases} \ .$$

Conversamente, se ω é um produto escalar em V tal que

$$\omega(\vec{e}_i, \vec{e}_j) = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases} \ ,$$

$\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i, \vec{y} = \sum_{j=1}^{n} y_j \vec{e}_j \in V$ quaisquer, concluímos que

$$\omega(\vec{x}, \vec{y}) = \omega\left(\sum_{i=1}^{n} x_i \vec{e}_i, \sum_{j=1}^{n} y_j \vec{e}_j \right) \overset{(e_1)}{=} \sum_{i=1}^{n} x_i \omega\left(\vec{e}_i, \sum_{j=1}^{n} y_j \vec{e}_j \right)$$
$$\overset{(e_2)}{=} \sum_{i=1}^{n} \left(\sum_{j=1}^{n} x_i y_j \omega(\vec{e}_i, \vec{e}_j) \right) = \sum_{i=1}^{n} x_i y_i = \langle \vec{x}, \vec{y} \rangle_S \ .$$

Isso implica que, dado um produto escalar ω em V e uma base o.n. (finita) $S \subset V$ com respeito a ω, temos que $\omega(\vec{x}, \vec{y}) = \langle \vec{x}, \vec{y} \rangle_S$ para todo $\vec{x}, \vec{y} \in V$.

Uma das vantagens de bases ortogonais é que é simples obter as componentes de um vetor em tal base. Para simplificar a discussão, assumiremos de agora em diante que $\dim(V) = n < \infty$. Com efeito, se $S = \{\vec{e}_1, \ldots, \vec{e}_n\} \subset V$ é uma base ortogonal de V com respeito a um produto escalar ω em V (ou seja, S é ortogonal com respeito a ω, $\vec{0} \notin S$ e $L(S) = V$) e $\vec{x} \in V$ é representado na base S) pela c.l.

$$\vec{x} = \sum_{j=1}^{n} x_j \vec{e}_j \ , \quad x_j \in \mathbb{R} \ , \ j = 1, \ldots, n \ ,$$

segue que

$$\omega(\vec{e}_i, \vec{x}) \overset{(e_2)}{=} \sum_{j=1}^{n} x_j \omega(\vec{e}_i, \vec{e}_j) = x_i \omega(\vec{e}_i, \vec{e}_i) \ ,$$

para todo $i = 1, \ldots, n$, de modo que

$$x_i = \frac{\omega(\vec{e}_i, \vec{x})}{\omega(\vec{e}_i, \vec{e}_i)} \ , \quad i = 1, \ldots, n$$

e portanto

$$\vec{x} = \sum_{j=1}^{n} \frac{\omega(\vec{e}_j, \vec{x})}{\omega(\vec{e}_j, \vec{e}_j)} \vec{e}_j$$

para todo $\vec{x} \in V$. Logo, se $\vec{x}, \vec{y} \in V$, concluímos a partir da discussão no parágrafo anterior que

$$\omega(\vec{x}, \vec{y}) = \sum_{j=1}^{n} \frac{\omega(\vec{e}_j, \vec{x})}{\omega(\vec{e}_j, \vec{e}_j)} \frac{\omega(\vec{e}_j, \vec{y})}{\omega(\vec{e}_j, \vec{e}_j)} \ .$$

Em particular, se S é o.n., temos para quaisquer $\vec{x}, \vec{y} \in V$ que

$$\vec{x} = \sum_{j=1}^{n} \frac{\omega(\vec{e}_j, \vec{x})}{\omega(\vec{e}_j, \vec{e}_j)} \vec{e}_j \ , \ \omega(\vec{x}, \vec{y}) = \sum_{j=1}^{n} \omega(\vec{e}_j, \vec{x}) \omega(\vec{e}_j, \vec{y}) \ .$$

3. Geometria do produto escalar

Como mencionado no início deste Capítulo, o produto escalar pode ser entendido como uma "régua e transferidor" num espaço vetorial – mais precisamente, ele nos permite medir *comprimentos* e *ângulos*. De agora em diante, passaremos a denotar um produto escalar arbitrário mas *fixo* num espaço vetorial V pela mesma notação $\langle \vec{x}, \vec{y} \rangle$ usada para o produto escalar canônico de $\mathbb{R}^n$, e neste último caso tal produto escalar será sempre entendido como o canônico para evitar confusão (se for necessário usar outro produto escalar em $\mathbb{R}^n$, usaremos uma notação diferente para este).

3.1. Norma euclidiana de um vetor e ângulo entre dois vetores. A desigualdade de Cauchy-Schwarz.

- O "comprimento" de um vetor $\vec{x}$ é dado pela sua *norma euclidiana*

$$\|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle} \ .$$

  Em particular, se $V = \mathbb{R}^n$,

$$\|\vec{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2} \ .$$

  Levando em consideração a ortogonalidade da base canônica, podemos pensar nessa fórmula como uma extensão n-dimensional do teorema de Pitágoras.

- O "ângulo" θ entre dois vetores $\vec{x}, \vec{y} \neq \vec{0}$ em V pode ser obtido pela "lei dos cossenos"

$$\langle \vec{x}, \vec{y} \rangle = \|\vec{x}\| \cdot \|\vec{y}\| \cos(\theta) \ .$$

O que garante que a norma euclidiana satisfaz as propriedades básicas de uma noção de distância e que a "lei dos cossenos" pode ser usada para *definir* θ é a seguinte desigualdade fundamental, conhecida como *desigualdade de Cauchy-Schwarz*: dados $\vec{x}, \vec{y} \in V$ quaisquer, temos que

$$\langle \vec{x}, \vec{y} \rangle^2 \leq \langle \vec{x}, \vec{x} \rangle \langle \vec{y}, \vec{y} \rangle \ , \tag{II.1}$$

com igualdade se e somente se $\vec{x}, \vec{y}$ são l.d., ou seja, $\vec{x}$ é múltiplo escalar de $\vec{y}$ ou vice-versa.

Prova de (II.1). A prova usa apenas os axiomas de produto escalar (a), (b) e (c_2) (logo, (c_1)). Para quaisquer $\vec{x}, \vec{y} \in V$, $t \in \mathbb{R}$, temos

$$\begin{aligned} P(t) = \langle \vec{x} + t\vec{y}, \vec{x} + t\vec{y} \rangle &\overset{(c_1)}{=} \langle \vec{x}, \vec{x} + t\vec{y} \rangle + t\langle \vec{y}, \vec{x} + t\vec{y} \rangle \\ &\overset{(c_2)}{=} \langle \vec{x}, \vec{x} \rangle + t\langle \vec{x}, \vec{y} \rangle + t\langle \vec{y}, \vec{x} \rangle + t^2 \langle \vec{y}, \vec{y} \rangle \\ &\overset{(a)}{=} \langle \vec{x}, \vec{x} \rangle + 2t\langle \vec{x}, \vec{y} \rangle + t^2 \langle \vec{y}, \vec{y} \rangle \overset{(b)}{\geq} 0 \ . \end{aligned}$$

Se $\vec{y} = \vec{0}$, (II.1) é trivialmente satisfeita, pois nesse caso ambos os lados de (II.1) são zero. Se $\vec{y} \neq \vec{0}$ (equivalentemente, por (b), $\langle \vec{y}, \vec{y} \rangle > 0$), então $P(t)$ é um polinômio de segundo grau em t que satisfaz $P(t) \geq 0$ para todo $t \in \mathbb{R}$. Isso significa que $P(t)$ tem *no máximo uma* raiz real. Em termos do *discriminante* Δ de $P(t)$

$$P(t) = at^2 + bt + c \implies \Delta = b^2 - 4ac = 4\langle \vec{x}, \vec{y} \rangle^2 - 4\langle \vec{y}, \vec{y} \rangle \langle \vec{x}, \vec{x} \rangle$$

isso é o mesmo que $\Delta \leq 0$, e nesse caso

$$\langle \vec{x}, \vec{y} \rangle^2 - \langle \vec{x}, \vec{x} \rangle \langle \vec{y}, \vec{y} \rangle \leq 0$$

como desejado. Finalmente, o caso de igualdade de (II.1) é precisamente a situação na qual $\vec{y} = \vec{0}$ ou $P(t)$ tem *exatamente uma* raiz real t_0 – nesse último caso,

$$P(t_0) = 0 = \langle \vec{x} + t_0\vec{y}, \vec{x} + t_0\vec{y} \rangle \overset{(b)}{\Rightarrow} \vec{x} + t_0\vec{y} = \vec{0} \, .$$

Em ambos os casos, concluímos que $\vec{x}, \vec{y}$ são l.d.. □

Exploremos agora algumas das consequências da desigualdade de Cauchy-Schwarz (II.1). Tirando a raiz quadrada de (II.1) dos dois lados, concluímos que

$$|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \cdot \|\vec{y}\| \, .$$

Essa forma de (II.1) garante a validade da *desigualdade triangular* para a norma euclidiana: dados $\vec{x}, \vec{y} \in V$ quaisquer,

$$\begin{aligned} \|\vec{x} + \vec{y}\|^2 &= \langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle \\ &\overset{(c_1)}{=} \langle \vec{x}, \vec{x} + \vec{y} \rangle + \langle \vec{y}, \vec{x} + \vec{y} \rangle \\ &\overset{(c_2)}{=} \langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle \\ &\overset{(a)}{=} \langle \vec{x}, \vec{x} \rangle + 2\langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{y} \rangle \\ &\overset{\text{(II.1)}}{\leq} \|\vec{x}\|^2 + 2\|\vec{x}\| \cdot \|\vec{y}\| + \|\vec{y}\|^2 = (\|\vec{x}\| + \|\vec{y}\|)^2 \end{aligned}$$

$$\Rightarrow \|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| \, . \tag{II.2}$$

Essa desigualdade, juntamente com as propriedades de *homogeneidade*

$$\|\alpha\vec{x}\| = \sqrt{\langle \alpha\vec{x}, \alpha\vec{x} \rangle} \overset{(c_1)+(c_2)}{=} \sqrt{\alpha^2 \langle \vec{x}, \vec{x} \rangle} = |\alpha| \cdot \|\vec{x}\| \tag{II.3}$$

e *não-degenerescência*

$$\|\vec{x}\| = 0 \overset{(b)}{\Rightarrow} \vec{x} = \vec{0} \, , \tag{II.4}$$

válidas para $\vec{x} \in V$, $\alpha \in \mathbb{R}$ quaisquer, são análogas às propriedades do valor absoluto (módulo) de números reais, e garantem que a norma euclidiana satisfaz as condições mínimas que se espera de uma noção de *comprimento*. Assim, podemos definir a *distância (euclidiana)* entre dois vetores $\vec{x}, \vec{y} \in V$ como o comprimento $\|\vec{x} - \vec{y}\|$ da diferença entre $\vec{x}$ e $\vec{y}$.

3.1 Exercício. *Use as propriedades* (II.2)–(II.4) *da norma euclidiana para provar a chamada* desigualdade triangular reversa*: dados $\vec{x}, \vec{y} \in V$ quaisquer,*

$$|\|\vec{x}\| - \|\vec{y}\|| \leq \|\vec{x} - \vec{y}\| \, . \tag{II.5}$$

*(*Dica: *o argumento é o mesmo usado para provar que $||x| - |y|| \leq |x - y|$ usando as propriedades correspondentes do módulo)*

Lembrar que $\theta = \arccos(t) = (\cos|_{[0,\pi]})^{-1}(t)$, $t \in [-1,1]$ = imagem de $\cos(\theta)$, e portanto $\theta \in [0,\pi]$.

A desigualdade de Cauchy-Schwarz (II.1) também implica que *se* $\vec{x}, \vec{y} \neq \vec{0}$, então

$$-1 \leq \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \cdot \|\vec{y}\|} \leq 1 \ .$$

Em outras palavras, se $\vec{x}, \vec{y} \neq \vec{0}$, podemos definir o ângulo θ entre $\vec{x}$ e $\vec{y}$ como

$$\theta = \arccos\left(\frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \cdot \|\vec{y}\|} \right) \ ,$$

pois a desigualdade de Cauchy-Schwarz nesse caso implica que o argumento de arccos na expressão acima sempre assume valores no domínio dessa função. A escolha dessa função trigonométrica inversa é consistente com as propriedades geométricas do produto escalar:

- O caso $\theta = \pi/2$ ocorre precisamente quando $\vec{x}$ e $\vec{y}$ são *ortogonais*;
- Os casos $\theta = 0, \pi$ ocorrem precisamente no caso de igualdade de (II.1), ou seja, quando $\vec{x}, \vec{y}$ são l.d.. Nessa situação, concluímos que $\theta = 0$ (resp. $\theta = \pi$) quando $\vec{y} = \lambda \vec{x}$ com $\lambda > 0$ (resp. $\lambda < 0$).

3.2. Projeções ortogonais. A definição do ângulo entre dois vetores $\vec{0} \neq \vec{x}, \vec{y} \in V$ é consistente com a interpretação geométrica do produto escalar: se $S = \{\vec{e}_1, \ldots, \vec{e}_n\}$ é uma base o.n. de V, então

$$\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i \ \Rightarrow \ \langle \vec{e}_j, \vec{x} \rangle = \sum_{i=1}^{n} x_i \langle \vec{e}_j, \vec{e}_i \rangle = x_j$$

para todo $j = 1, \ldots, n$, de modo que

$$\vec{x} = \sum_{i=1}^{n} \langle \vec{e}_i, \vec{x} \rangle \vec{e}_i \ .$$

Ou seja, $\langle \vec{e}_i, \vec{x} \rangle$ é a *componente de* $\vec{x}$ *ao longo de* $\vec{e}_i$ em S, tal como esperado da interpretação geométrica do cosseno. Além disso, há uma conexão íntima entre as componentes de um vetor numa base o.n. e a noção de distância euclidiana: dados $\vec{x}, \vec{y} \in V$ com $\|\vec{x}\| = 1$, temos que $\langle \vec{x}, \vec{y} \rangle \vec{x}$ é o múltiplo escalar de $\vec{x}$ *mais próximo* de $\vec{y}$. Para ver isso, escrevamos

$$\vec{y} = \underbrace{\vec{y} - \langle \vec{x}, \vec{y} \rangle \vec{x}}_{=\vec{y}_1^{\perp}} + \underbrace{\langle \vec{x}, \vec{y} \rangle \vec{x}}_{=\vec{y}_1} \ .$$

Notar agora que

$$\langle \vec{x}, \vec{y}_1^{\perp} \rangle = \langle \vec{x}, \vec{y} - \langle \vec{x}, \vec{y} \rangle \vec{x} \rangle = \langle \vec{x}, \vec{y} \rangle - \langle \vec{x}, \vec{y} \rangle \langle \vec{x}, \vec{x} \rangle = 0 \ .$$

Dado $\vec{z} = \alpha\vec{x}$, temos que

$$\begin{aligned}\|\vec{y} - \vec{z}\|^2 &= \langle \vec{y} - \vec{z}, \vec{y} - v z\rangle \\ &= \langle \vec{y}_1^{\perp} + \vec{y}_1 - \alpha\vec{x}, \vec{y}_1^{\perp} + \vec{y}_1 - \alpha\vec{x}\rangle \\ &= \langle \vec{y}_1^{\perp}, \vec{y}_1^{\perp}\rangle + (\langle \vec{x}, \vec{y}\rangle - \alpha)\langle \vec{y}_1^{\perp}, \vec{x}\rangle \\ &\quad + (\langle \vec{x}, \vec{y}\rangle - \alpha)\langle \vec{x}, \vec{y}_1^{\perp}\rangle + (\langle \vec{x}, \vec{y}\rangle - \alpha)^2\langle \vec{x}, \vec{x}\rangle \\ &= \|\vec{y}_1^{\perp}\|^2 + (\langle \vec{x}, \vec{y}\rangle - \alpha)^2 \ .\end{aligned}$$

Segue da identidade acima (que é nada mais, nada menos do que o teorema de Pitágoras) que o menor valor possível para $\|\vec{y} - \vec{z}\|$ é atingido precisamente quando $\alpha = \langle \vec{x}, \vec{y}\rangle$, e é precisamente nesse caso que $\vec{y} - \vec{z}$ é *ortogonal a* $\vec{x}$. Podemos generalizar esse raciocínio para mais de um vetor: dado um conjunto o.n. $S = \{\vec{e}_1, \ldots, \vec{e}_k\} \subset V$, definamos

$$P_S(\vec{x}) = \sum_{i=1}^{k} \langle \vec{e}_i, \vec{x}\rangle \vec{e}_i \ .$$

A aplicação $P_S : V \to V$ possui as seguintes propriedades:

- P_S é linear: dados $\vec{x}, \vec{y} \in V$, $\alpha \in \mathbb{R}$ quaisquer, temos que

  Uma discussão geral de aplicações lineares pode ser encontrada no Capítulo III.

  $$P_S(\vec{x} + \vec{y}) = \sum_{i=1}^{k} \langle \vec{e}_i, \vec{x} + \vec{y}\rangle \vec{e}_i = P_S(\vec{x}) + P_S(\vec{y})$$

  e

  $$P_S(\alpha\vec{x}) = \sum_{i=1}^{k} \langle \vec{e}_i, \alpha\vec{x}\rangle \vec{e}_i = \alpha P_S(\vec{x}) \ .$$

- $\vec{x} \in L(S)$ se e somente se $P_S(\vec{x}) = \vec{x}$: se $\vec{x} = \sum_{i=1}^{k} x_i \vec{e}_i$, já vimos acima que necessariamente $x_i = \langle \vec{e}_i, \vec{x}\rangle$ para todo $i = 1, \ldots, k$. Conversamente, obviamente $\sum_{i=1}^{k} \langle \vec{e}_i, \vec{x}\rangle \vec{e}_i \in L(S)$, logo $\vec{x} = P_S(\vec{x})$ implica $\vec{x} \in L(S)$. Em particular, $P_S = P_{\tilde{S}}$ se $S, \tilde{S}$ são conjuntos o.n. tais que $L(S) = L(\tilde{S})$, ou seja, P_S depende *apenas* do subespaço vetorial $L(S)$ e do produto escalar.

- $\langle \vec{x} - P_S(\vec{x}), P_S(\vec{y})\rangle = 0$ para todo $\vec{x}, \vec{y} \in V$: de fato,

  $$\begin{aligned}\langle \vec{x} - P_S(\vec{x}), P_S(\vec{y})\rangle &= \left\langle \vec{x} - \sum_{i=1}^{k} \langle \vec{e}_i, \vec{x}\rangle \vec{e}_i, \sum_{j=1}^{k} \langle \vec{e}_j, \vec{y}\rangle \vec{e}_j \right\rangle \\ &= \sum_{j=1}^{k} \left(\langle \vec{e}_j, \vec{y}\rangle \langle \vec{x}, \vec{e}_j\rangle - \langle \vec{e}_j, \vec{y}\rangle \sum_{i=1}^{k} \langle \vec{e}_i, \vec{x}\rangle \langle \vec{e}_i, \vec{e}_j\rangle \right) \\ &= \sum_{j=1}^{k} \left(\langle \vec{e}_j, \vec{y}\rangle \langle \vec{x}, \vec{e}_j\rangle - \langle \vec{e}_j, \vec{y}\rangle \langle \vec{x}, \vec{e}_j\rangle \right) = 0 \ .\end{aligned}$$

Em particular, obtemos daí mais uma versão do *teorema de Pitágoras*:

$$\begin{aligned}\|\vec{x}\|^2 &= \langle(\vec{x} - P_S(\vec{x})) + P_S(\vec{x}),\, (\vec{x} - P_S(\vec{x})) + P_S(\vec{x})\rangle \\ &= \langle \vec{x} - P_W(\vec{x}), \vec{x} - P_S(\vec{x})\rangle + \langle \vec{x} - P_S(\vec{x}), P_S(\vec{x})\rangle \\ &\quad + \langle P_S(\vec{x}), \vec{x} - P_S(\vec{x})\rangle + \langle P_S(\vec{x}), P_S(\vec{x})\rangle \\ &= \|\vec{x} - P_S(\vec{x})\|^2 + \|P_S(\vec{x})\|^2 \,,\end{aligned}$$

Em vista das propriedades acima, chamamos $P_S = P_W$ de *projeção ortogonal ao longo do subespaço vetorial* $W = L(S)$. Podemos também caracterizar P_W em termos da distância euclidiana: dado um vetor $\vec{x} \in V$ qualquer, $P_W(\vec{x})$ é o vetor em W *mais próximo de* $\vec{x}$. De fato, se $\vec{y}(= P_W(\vec{y})) \in W$, então por um cálculo similar ao feito acima,

$$\begin{aligned}\|\vec{x} - \vec{y}\|^2 = \|\vec{x} - P_W(\vec{y})\|^2 &= \langle \vec{x} - P_W(\vec{y}), \vec{x} - P_W(\vec{y})\rangle \\ &= \langle(\vec{x} - P_W(\vec{x})) + P_W(\vec{x} - \vec{y}), (\vec{x} - P_W(\vec{x})) + P_W(\vec{x} - \vec{y})\rangle \\ &= \langle \vec{x} - P_W(\vec{x}), \vec{x} - P_W(\vec{x})\rangle + \langle \vec{x} - P_W(\vec{x}), P_W(\vec{x} - \vec{y})\rangle \\ &\quad + \langle P_W(\vec{x} - \vec{y}), \vec{x} - P_W(\vec{x})\rangle + \langle P_W(\vec{x} - \vec{y}), P_W(\vec{x} - \vec{y})\rangle \\ &= \|\vec{x} - P_W(\vec{x})\|^2 + \|P_W(\vec{x} - \vec{y})\|^2 \,,\end{aligned}$$

Portanto, o menor valor possível para $\|\vec{x} - \vec{y}\|$ se $\vec{y} \in W$ é atingido precisamente quando $P_W(\vec{x} - \vec{y}) = \vec{0}$ e portanto $P_W(\vec{x}) = P_W(\vec{y}) = \vec{y}$.

3.3. Ortonormalização de bases e construção de bases ortonormais. O método de Gram-Schmidt. Projeções ortogonais fornecem um método para transformar um conjunto l.i. $\tilde{S} = \{\vec{x}_1, \ldots, \vec{x}_k\} \subset V$ num conjunto *o.n.* $S = \{\vec{e}_1, \ldots, \vec{e}_k\}$ tal que se

$$\tilde{S}_l = \{\vec{x}_1, \ldots, \vec{x}_l\} \,,\; S_l = \{\vec{e}_1, \ldots, \vec{e}_l\} \,, \quad l = 1, \ldots, k \,,$$

então $L(\tilde{S}_l) = L(S_l)$ para *todo* $l = 1, \ldots, k$. Isso garante que as c.l.'s dos vetores em $\tilde{S}$ que expressam os vetores em S são as mais simples possíveis computacionalmente e podem ser escritas de maneira *recursiva*, i.e., por indução no número de vetores k em $\tilde{S}$. Tal procedimento é chamado de *ortonormalização de Gram-Schmidt*.

A construção de S procede da seguinte maneira.

- Defina $\vec{e}_1 = \frac{1}{\|\vec{x}_1\|}\vec{x}_1 =$ *normalização* de $\vec{x}_1$, e $S_1 = \{\vec{e}_1\}$. Notar que $\|\vec{e}_1\| = \frac{\|\vec{x}_1\|}{\|\vec{x}_1\|} = 1$ (logo, S_1 é o.n. e portanto l.i.) e $\alpha_1\vec{x}_1 = \alpha_1\|\vec{x}_1\|\vec{e}_1$ (logo, $L(\tilde{S}_1) = L(S_1)$). Se $k = 1$, não há mais nada a fazer.

- Se $k > 1$, façamos a seguinte hipótese de indução: suponhamos que existe um conjunto o.n. $S_l = \{\vec{e}_1, \ldots, \vec{e}_l\}$, $1 \leq l < k$, tal que $L(S_l) = L(\tilde{S}_l)$. Defina

$$\begin{aligned}\vec{y}_{l+1} &= \vec{x}_{l+1} - P_{S_l}(\vec{x}_{l+1}) \\ &= \vec{x}_{l+1} - \sum_{i=1}^{k} \langle \vec{e}_i, \vec{x}_{l+1} \rangle \vec{e}_i \ , \\ \vec{e}_{l+1} &= \frac{1}{\|\vec{y}_{l+1}\|} \vec{y}_{l+1} \ .\end{aligned}$$

  Segue que $\|\vec{e}_{l+1}\| = 1$ e $\langle \vec{e}_j, \vec{e}_{l+1} \rangle = \langle \vec{e}_j, \vec{y}_{l+1} \rangle = 0$ para todo $j = 1, \ldots, l$, portanto $S_{l+1} = S_l \cup \{\vec{e}_{l+1}\} = \{\vec{e}_1, \ldots, \vec{e}_{l+1}\}$ é o.n. (logo, l.i.). Obviamente $\vec{e}_{l+1} \in L(\tilde{S}_{l+1})$, logo $S_{l+1} \subset L(\tilde{S}_{l+1})$. Como S_{l+1} é l.i. e tem o mesmo número de vetores que a base $\tilde{S}_{l+1}$ de $L(\tilde{S}_{l+1})$, concluímos que $L(S_{l+1}) = L(\tilde{S}_{l+1})$, como desejado. Se $l + 1 = k$, acabamos, do contrário repetir o procedimento acima com $l + 1$ no lugar de l.

Em síntese, a ortonormalização de Gram-Schmidt consiste no seguinte algoritmo:

(i) Seja $\tilde{S} = \{\vec{x}_1, \ldots, \vec{x}_k\}$ um subconjunto *l.i.* de V.

(ii) Defina recursivamente (i.e. por indução em $l = 1, \ldots, k$)

$$\vec{e}_1 = \frac{1}{\|\vec{x}_1\|} \vec{x}_1 \ , \ \vec{e}_{l+1} = \frac{1}{\|\vec{y}_{l+1}\|} \vec{y}_{l+1} \ ,$$

$$\vec{y}_{l+1} = \vec{x}_{l+1} - \sum_{i=1}^{l} \langle \vec{e}_l, \vec{x}_{l+1} \rangle \vec{e}_l = \vec{x}_{l+1} - P_{S_l}(\vec{x}_{l+1}) \ ,$$

onde $S_l = \{\vec{e}_1, \ldots, \vec{e}_l\}$, $\tilde{S}_l = \{\vec{x}_1, \ldots, \vec{x}_l\}$, $l = 1, \ldots, k$.

(iii) O resultado, como mostramos acima, é um conjunto *o.n.* $S = S_k = \{\vec{e}_1, \ldots, \vec{e}_k\}$ tal que $L(S_l) = L(\tilde{S}_l)$ para todo $l = 1, \ldots, k$.

Podemos inclusive integrar a ortonormalização de Gram-Schmidt à própria construção de um conjunto l.i. $\tilde{S}$ tal como feito no final da Seção 3, dado que ali obtemos $\tilde{S}_{l+1}$ recursivamente a partir de $\tilde{S}_l$. Tudo que precisamos fazer é substituir $\vec{x}_{l+1}$ por $\vec{e}_{l+1}$ no passo de indução, pois $\vec{e}_{l+1}$ depende apenas de S_l e $\vec{x}_{l+1}$. Isso permite uma construção *direta* de um conjunto o.n. e, em particular, de bases o.n.'s de subespaços vetoriais (de dimensão finita) de V se este é munido de um produto escalar.

Em vista de sua importância, de agora em diante sempre assumiremos que nossos espaços vetoriais V são munidos de um produto escalar *fixo*, denotado por $\langle \vec{x}, \vec{y} \rangle$,

$\vec{x}, \vec{y} \in V$, e no caso $V = \mathbb{R}^n$ tal produto escalar será sempre o canônico. Se for necessário mudar o produto escalar ou se tal notação causar confusão, passaremos a usar uma notação diferente de maneira explícita.

III

Transformações lineares

1. Definição e exemplos

Sejam V, W espaços vetoriais (reais). Uma aplicação $T : V \to W$ é dita uma *aplicação linear* ou uma *transformação linear* (t.l.) de V em W se, dados $\vec{x}, \vec{x}' \in V$, $\alpha \in \mathbb{R}$ quaisquer, temos:

(a) $T(\vec{x} + \vec{x}') = T(\vec{x}) + T(\vec{x}')$;

(b) $T(\alpha\vec{x}) = \alpha T(\vec{x})$.

Ou seja, podemos "por para fora de T" as operações vetoriais de V (notando que, no lado direito de (a), (b) passamos a empregar as operações vetoriais de W). Isso implica que podemos fazer o mesmo para qualquer c.l. de vetores em V: por indução no número k de vetores na c.l., temos que

$$\begin{aligned} T\left(\sum_{j=1}^{k} \alpha_j \vec{x}_j\right) &= T\left(\sum_{j=1}^{k-1} \alpha_j \vec{x}_j + \alpha_k \vec{x}_k\right) \\ &= T\left(\sum_{j=1}^{k-1} \alpha_j \vec{x}_j\right) + \alpha_k T(\vec{x}_k) \\ &= \sum_{j=1}^{k} \alpha_j T(\vec{x}_j) \end{aligned}$$

para $\vec{x}_1, \ldots, \vec{x}_k \in V$, $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ quaisquer.

Se o contradomínio W de T coincide com o seu domínio V, dizemos apenas que T é uma t.l. *em* V. Se $W = \mathbb{R}$, dizemos que T é um *funcional linear* ou uma *1-forma* em V.

De agora em diante, se T é uma t.l., empregaremos a notação simplificada $T(\vec{x}) = T\vec{x}$ (i.e. sem os parênteses ao redor do argumento $\vec{x}$ de T) sempre que esta não causar confusão. Como exemplos de t.l.'s, podemos citar:

(i) $T(\vec{x}) = \vec{0}$. De fato, $T(\vec{x} + \vec{x}') = \vec{0} = \vec{0} + \vec{0} = T(\vec{x}) + T(\vec{x}')$ e $T(\alpha\vec{x}) = \vec{0} = \alpha\vec{0} = \alpha T(\vec{x})$ para $\vec{x}, \vec{x}' \in V$, $\alpha \in \mathbb{R}$ quaisquer. Essa t.l. é conhecida como a *aplicação zero* de V em W, denotada simplesmente por $\vec{0}$. Se $W = \mathbb{R}$, escrevemos $T = 0$.

(ii) ($W = V$) $T(\vec{x}) = \vec{x}$. De fato, $T(\vec{x} + \vec{x}') = \vec{x} + vx' = T(\vec{x}) + T(\vec{x}')$ e $T(\alpha\vec{x}) = \alpha\vec{x} = \alpha T(\vec{x})$ para $\vec{x}, \vec{x}' \in V$, $\alpha \in \mathbb{R}$ quaisquer. Essa t.l. é conhecida como a (aplicação) identidade $T = \mathbb{1}_V = \mathbb{1}$ de V.

Mais em geral, dado $A \neq \varnothing$, podemos definir a (aplicação) identidade $\mathbb{1}_A : A \to A$ de A como $\mathbb{1}_A(p) = p$, $p \in A$ qualquer. Quando não houver confusão, podemos escrever $\mathbb{1}_A = \mathbb{1}$.

(iii) ($W = \mathbb{R}$) Dado $\vec{e} \in V$, seja

$$T\vec{x} = \langle \vec{e}, \vec{x} \rangle \ .$$

Devido à linearidade do produto escalar de V na segunda variável, claramente

$$T(\vec{x} + \vec{x}') = \langle \vec{e}, \vec{x} + \vec{x}' \rangle = \langle \vec{e}, \vec{x} \rangle + \langle \vec{e}, \vec{x}' \rangle = T\vec{x} + T\vec{x}'$$

e

$$T(\alpha\vec{x}) = \langle \vec{e}, \alpha\vec{x} \rangle = \alpha \langle \vec{e}, \vec{x} \rangle = \alpha T\vec{x}$$

para $\vec{x}, \vec{x}' \in V$, $\alpha \in \mathbb{R}$ quaisquer, logo T é uma 1-forma em V. Veremos adiante que essa situação é genérica: se $\dim(V) < \infty$, dada qualquer 1-forma T em V temos um *unico* vetor $\vec{e}_T \in V$ tal que $T\vec{x} = \langle \vec{e}_T, \vec{x} \rangle$.

(iv) Mais em geral, se $\vec{f}_1, \ldots, \vec{f}_m \in W$, $\vec{g}_1, \ldots, \vec{g}_m \in V$, defina

$$T\vec{x} = \sum_{i=1}^{m} \langle \vec{g}_i, \vec{x} \rangle \vec{f}_i \ .$$

Por um cálculo análogo ao empregado no exemplo (iii), é fácil concluir que T é uma t.l. de V em W: dados $\vec{x}, \vec{x}' \in V$, $\alpha \in \mathbb{R}$ quaisquer, temos que

$$\begin{aligned} T(\vec{x} + \vec{x}') &= \sum_{i=1}^{m} \langle \vec{g}_i, \vec{x} + \vec{x}' \rangle \vec{f}_i \sum_{i=1}^{m} \\ &= \sum_{i=1}^{m} (\langle \vec{g}_i, \vec{x} \rangle + \langle \vec{g}_i, \vec{x}' \rangle) \vec{f}_i \\ &= \langle \vec{g}_i, \vec{x} \rangle \vec{f}_i + \sum_{i=1}^{m} \langle \vec{g}_i, \vec{x}' \rangle \vec{f}_i = T\vec{x} + T\vec{x}' \end{aligned}$$

e

$$T(\alpha\vec{x}) = \sum_{i=1}^{m} \langle \vec{g}_i, \alpha\vec{x} \rangle \vec{f}_i = \sum_{i=1}^{m} \alpha \langle \vec{g}_i, \vec{x} \rangle \vec{f}_i = \alpha \langle \vec{g}_i, \vec{x} \rangle \vec{f}_i = \alpha T\vec{x} \ .$$

Um exemplo desse tipo no caso $W = V$ é a projeção ortogonal $T = P_X$ sobre um subespaço vetorial $X \subset V$ – nesse caso, $\tilde{S} = \{\vec{f}_i = \vec{g}_i \mid i = 1, \ldots, m = \dim(X)\}$ é uma base o.n. de X. Na verdade, assim como no exemplo (iii), veremos adiante que tal situação é genérica: se T é uma t.l. de V em W e $\tilde{S} = \{\vec{f}_1, \ldots, \vec{f}_m\}$ é uma base de W, então existe uma *única* escolha de vetores $\vec{g}_1, \ldots, \vec{g}_m \in V$ tal que T tem a forma acima.

(v) ($W = \mathbb{R}^n$, $n = \dim(V)$) Dada uma base $S = \{\vec{e}_1, \ldots, \vec{e}_n\}$ de V, para cada $\vec{x} = \sum_{j=1}^{n} x_j \vec{e}_j$ definimos

$$T\vec{x} = (x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i \vec{f}_i \ ,$$

onde $\tilde{S} = \{\vec{f}_1, \ldots, \vec{f}_n\}$ é a base *canônica* de $\mathbb{R}^n$. Segue de (I.2) (ou, alternativamente, notando que T é da forma descrita no exemplo (iv) acima se escolhermos o produto escalar associado a S em V, ou seja, se S for o.n.) que T é uma t.l. de V em $\mathbb{R}^n$. Esse exemplo mostra que o sistema de coordenadas para V obtido por uma escolha de base S é *linear* justamente no sentido de ser uma t.l. de V em $\mathbb{R}^n$.

2. Espaços vetoriais e produto de transformações lineares

Denotamos respectivamente por

$$L(V,W) = \{T : V \to W \mid T \text{ t.l.}\}\ ,$$
$$L(V) = L(V,V)\ , \quad V' = L(V,\mathbb{R})$$

o espaço das t.l.'s de V em W, o espaço das t.l.'s em V e o espaço das 1-formas em V. V' é também conhecido como o *(espaço) dual* de V. Notar agora que:

- $L(V,W)$ é *subespaço vetorial* de $F(V,W) = \{\vec{f} : V \to W\}$ com respeito às operações vetoriais *pontuais*

$$(T_1 + T_2)(\vec{x}) = T_1\vec{x} + T_2\vec{x}\ ,\ (\beta T_1)(\vec{x}) = \beta T_1\vec{x}\ ,$$

$T_1, T_2 \in L(V,W)$, $\vec{x} \in V$, $\beta \in \mathbb{R}$ quaisquer. Ou seja, a soma pontual $T_1 + T_2$ e a multiplicação escalar pontual βT_1 são t.l.'s se T_1, T_2 o são. De fato, dados $\vec{x}, \vec{x}' \in V$, $\alpha \in \mathbb{R}$ quaisquer, temos que

$$\begin{aligned}
(T_1 + T_2)(\vec{x} + \vec{x}') &= T_1(\vec{x} + \vec{x}') + T_2(\vec{x} + \vec{x}') \\
&\overset{(a)}{=} (T_1\vec{x} + T_1\vec{x}') + (T_2\vec{x} + T_2\vec{x}') \\
&= (T_1\vec{x} + T_2\vec{x}) + (T_1\vec{x}' + T_2\vec{x}') \\
&= (T_1 + T_2)(\vec{x}) + (T_1 + T_2)(\vec{x}')\ , \\
(\beta T_1)(\vec{x} + \vec{x}') &= \beta T_1(\vec{x} + \vec{x}') \\
&\overset{(a)}{=} \beta(T_1\vec{x} + T_1\vec{x}') \\
&= \beta T_1\vec{x} + \beta T_1\vec{x}' \\
&= (\beta T_1)(\vec{x}) + (\beta T_1)(\vec{x}')
\end{aligned}$$

e

$$\begin{aligned}
(T_1 + T_2)(\alpha\vec{x}) &= T_1(\alpha\vec{x}) + T_2(\alpha\vec{x}) \\
&\overset{(b)}{=} \alpha T_1\vec{x} + \alpha T_2\vec{x} \\
&= \alpha(T_1\vec{x} + T_2\vec{x}) = \alpha(T_1 + T_2)(\vec{x})\ , \\
(\beta T_1)(\alpha\vec{x}) &= \beta T_1(\alpha\vec{x}) \\
&\overset{(b)}{=} \beta\alpha T_1\vec{x} \\
&= \alpha(\beta T_1)(\vec{x})\ .
\end{aligned}$$

Obviamente, o zero de $L(V,W)$ é a aplicação zero de V em W, apresentada no exemplo (i) acima.

- A *composta* de duas t.l.'s é também uma t.l.. Mais precisamente, sejam V, W, X epaços vetoriais (reais), e $T_1 : V \to W$, $T_2 : W \to X$ t.l.'s. Escrevendo $T_2T_1 = T_2 \circ T_1 : V \to X$, i.e. $T_2T_1(\vec{x}) = T_2(T_1\vec{x})$, $\vec{x} \in V$, temos que

$$\begin{aligned} T_2T_1(\vec{x}+\vec{x}') = T_2(T_1(\vec{x}+\vec{x}')) &\overset{(a)}{=} T_2(T_1\vec{x}+T_1\vec{x}') \\ &\overset{(a)}{=} T_2(T_1\vec{x}) + T_2(T_1\vec{x}') = T_2T_1\vec{x} + T_2T_1\vec{x}' \end{aligned}$$

e

$$\begin{aligned} T_2T_1(\alpha\vec{x}) = T_2(T_1(\alpha\vec{x})) &\overset{(a)}{=} T_2(\alpha T_1\vec{x}) \\ &\overset{(a)}{=} \alpha T_2(T_1\vec{x}) = \alpha T_2T_1\vec{x}\ . \end{aligned}$$

para $\vec{x}, \vec{x}' \in V$, $\alpha \in \mathbb{R}$ quaisquer. Como a composição de aplicações é *associativa* e temos que

$$(T_2+T_2')T_1\vec{x} = (T_2+T_2')(T_1\vec{x}) = T_2T_1\vec{x} + T_2'T_1\vec{x}$$

e

$$\begin{aligned} T_2(T_1+T_1')\vec{x} = T_2((T_1+T_1')\vec{x}) = T_2(T_1\vec{x}+T_1'\vec{x}) \\ \overset{(a)}{=} T_2T_1\vec{x} + T_2T_1'\vec{x} \end{aligned}$$

para $T_1, T_1' \in L(V,W)$, $T_2, T_2' \in L(W,X)$, $\vec{x} \in V$ quaisquer, vemos que podemos considerar T_2T_1 como um produto (associativo e distributivo), e portanto passaremos a denominá-lo como tal. Tal produto, contudo, *não é comutativo*: se o domínio de T_1 não contiver a imagem de T_2, sequer faz sentido falar de T_1T_2. Mais ainda, mesmo se $V = W = X$, *não* é necessariamente verdade que $T_1T_2 = T_2T_1$.

Notar que a distributividade com respeito ao *segundo* fator depende crucialmente do fato que T_2 é uma t.l. e não vale para a composição de aplicações mais gerais, ao contrário da distributividade com respeito ao *primeiro* fator, que depende apenas da definição de soma pontual de aplicações.

- Seja $T : V \to W$ *bijetora*, de modo que existe uma única aplicação *inversa* $T^{-1} : W \to V$ satisfazendo $T^{-1} \circ T = \mathbb{1}_V$ e $T \circ T^{-1} = \mathbb{1}_W$. Se V é um espaço vetorial, então $\mathbb{1}_V$ é claramente uma t.l. em V. Além disso, se T é uma t.l., então T^{-1} *também* é: dados $\vec{x}, \vec{x}' \in V$, $\alpha \in \mathbb{R}$ quaisquer, temos que

$$\begin{aligned} T^{-1}(\vec{x}+\vec{x}') &= T^{-1}(T(T^{-1}\vec{x}) + T(T^{-1}\vec{x}')) \\ &\overset{(a)}{=} T^{-1}(T(T^{-1}(\vec{x}) + T^{-1}(\vec{y}))) \\ &= T^{-1}(\vec{x}) + T^{-1}(\vec{x}') \end{aligned}$$

e

$$\begin{aligned} T^{-1}(\alpha\vec{x}) &= T^{-1}(\alpha T(T^{-1}(\vec{x}))) \\ &\overset{(b)}{=} T^{-1}(T(\alpha T^{-1}(\vec{x}))) \\ &= \alpha T^{-1}(\vec{x})\ . \end{aligned}$$

3. A matriz de uma transformação linear

3.1. Determinação de uma transformação linear por seus valores numa base. Mostraremos agora que uma t.l. $T : V \to W$ é *unicamente* determinada pelos seus valores numa base de V. Seja $S = \{\vec{e}_1, \ldots, \vec{e}_n\}$ uma base de V; dado

$$\vec{x} = \sum_{j=1}^{n} x_j \vec{e}_j \in V \ ,$$

temos que

$$T\vec{x} = T\left(\sum_{j=1}^{n} x_j \vec{e}_j\right) = \sum_{j=1}^{n} x_j T(\vec{e}_j) \ . \tag{III.1}$$

Assim, para definir uma t.l. $T : V \to W$, é suficiente fornecer uma lista de n vetores $T(\vec{e}_1), \ldots, T(\vec{e}_n)$, que serão as respectivos valores de T em $\vec{e}_1, \ldots, \vec{e}_n$.

Seja agora $\tilde{S} = \{\vec{f}_1, \ldots, \vec{f}_m\}$ uma base de W, de modo que

$$T\vec{e}_j = \sum_{i=1}^{m} A_{ij} \vec{f}_i \ , \quad j = 1, \ldots, . \tag{III.2}$$

Em outras palavras, $A_{ij} \in \mathbb{R}$ é a componente de $T\vec{e}_j$ ao longo de $\vec{f}_i$ em $\tilde{S}$. Segue daí que

$$T\vec{x} = \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} A_{ij} \vec{f}_i\right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} A_{ij} x_j \vec{f}_i\right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} A_{ij} x_j\right) \vec{f}_i \ . \tag{III.3}$$

Suponhamos que S é *o.n.*; definindo

$$\vec{g}_i = \sum_{j=1}^{n} A_{ij} \vec{e}_j \ , \quad i = 1, \ldots, m \ , \tag{III.4}$$

temos que

$$T\vec{x} = \sum_{i=1}^{m} \langle \vec{g}_i, \vec{x} \rangle \vec{f}_i \tag{III.5}$$

e portanto mostramos que *toda* t.l. $T : V \to W$ é da forma descrita no exemplo (iv) acima. Em particular, se $W = \mathbb{R}$ e portanto $m = 1$, temos que

$$T\vec{e}_j = A_{1j} \in \mathbb{R} \ , \quad j = 1, \ldots, n$$

e portanto

$$\vec{e}_T = \vec{g}_1 = \sum_{j=1}^{n} (T\vec{e}_j) \vec{e}_j$$

é o *único* vetor em V tal que $T\vec{x} = \langle \vec{e}_T, \vec{x} \rangle$ para todo $\vec{x} \in V$, tal como afirmado no exemplo (iii) acima. A unicidade segue do seguinte argumento: se $\vec{e} \in V$ satisfaz $T\vec{x} = \langle \vec{e}, \vec{x} \rangle$ para todo $\vec{x} \in V$, então

$$\langle \vec{e} - \vec{e}_T, \vec{x} \rangle = \langle \vec{e}, \vec{x} \rangle - \langle \vec{e}_T, \vec{x} \rangle = T\vec{x} - T\vec{x} = 0$$

para todo $\vec{x} \in V$. Em particular, tomando $\vec{x} = \vec{e} - \vec{e}_T$ concluímos que $\|\vec{e} - \vec{e}_T\|^2 = 0$, ou seja, $\vec{e} = \vec{e}_T$. Esse resultado é conhecido como *Lema de Riesz*. Obviamente, $T\vec{x} = \vec{0}$ se e somente se $\vec{x}$ for *ortogonal* a $\vec{e}_T$.

3.2. Vetores-linha, vetores-coluna e multiplicação de matrizes. Retornando ao caso geral, notemos agora que (III.3) fornece ainda uma expressão das componentes de $T\vec{x}$ em $\tilde{S}$ em termos das componentes de $\vec{x}$ em S como uma *multiplicação de matrizes* (revisaremos esse conceito em breve): dado $\vec{x} = \sum_{j=1}^{n} x_j \vec{e}_j \in V$, definamos

$$\vec{x}_S = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

I.e. uma matriz com uma única coluna.

o *vetor-coluna* de $\vec{x}$ em S. A última identidade em (III.3) nos diz que

$$(T\vec{x})_{\tilde{S}} = A\vec{x}_S \ ,$$

cujo lado direito denota a multiplicação matricial de $\vec{x}_S$ pela matriz $A = [A_{ij}]$, denominada a *matriz de T nas bases $S, \tilde{S}$*.

Em particular, se $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ e $S, \tilde{S}$ são as respectivas bases canônicas, então $\vec{x}_S$ é simplesmente a matriz de uma coluna cuja entrada na i-ésima linha é a i-ésima componente de $\vec{x}$, e a ação de qualquer t.l. T de $\mathbb{R}^n$ em $\mathbb{R}^m$ assume a forma

$$\begin{aligned} \vec{x} = \sum_{j=1}^{n} x_j \vec{e}_j = (x_1, \ldots, x_n) &= ((\vec{x}_S)_{11}, \ldots, (\vec{x}_S)_{n1}) \\ \Rightarrow \ T\vec{x} = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} A_{ij} x_j \right) \vec{f}_i &= \big(((T\vec{x})_{\tilde{S}})_{11}, \ldots, ((T\vec{x})_{\tilde{S}})_{m1} \big) \\ &= \big((A\vec{x}_S)_{11}, \ldots, (A\vec{x}_S)_{m1} \big) \ . \end{aligned}$$

Se $W = V$ e $\tilde{S} = S$, dizemos apenas que A é a *matriz de T em S*. Se $W = \mathbb{R}$ e $\tilde{S} = \{1\}$ é a base canônica de $\mathbb{R}$, então

$$A = [T\vec{e}_1 \ \cdots \ T\vec{e}_n]$$

é denominada nesse caso de *vetor-linha* de T em S, cujas entradas são as componentes do vetor $\vec{e}_T$ dado pelo Lema de Riesz se S for o.n..

I.e. uma matriz com uma só linha.

Lembremos que o conjunto $M_{m\times n}(\mathbb{R})$ das matrizes (reais) com m linhas e n colunas é um espaço vetorial (ver o exemplo (iv) da Subseção 2.2). Além disso, podemos definir a *transposta* de $A \in M_{m\times n}(\mathbb{R})$ como a matriz A^T com "linhas e colunas de A trocadas", i.e. com entradas

$$(A^T)_{ij} = A_{ji} \ , \quad i = 1, \dots, m \ , \ j = 1, \dots, n \ ,$$

e o *produto* de $B \in M_{p\times m}(\mathbb{R})$ por $A \in M_{m\times n}(\mathbb{R})$ como a matriz $BA \in M_{p\times n}(\mathbb{R})$ com entradas

$$(BA)_{kj} = \sum_{i=1}^{m} B_{ki}A_{ij} \ , \quad i = 1, \dots, m \ , \ j = 1, \dots, n \ .$$

Notar que BA só faz sentido se o número de linhas de A for igual ao número de colunas de B, e que tal produto *não* é comutativo.

3.1 Exercício. *(a) Mostre que o produto de matrizes é associativo e distributivo, no mesmo sentido que o produto de t.l.'s. (*Dica: *mostre isso para cada entrada)*

(b) Seja a matriz identidade $\mathbb{1} = \mathbb{1}_n \in M_{n\times n}(\mathbb{R})$ *dada por*

$$\mathbb{1}_{ij} = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases} \ .$$

*Mostre que $\mathbb{1}A = A$, $B\mathbb{1} = B$ para $A \in M_{n\times p}(\mathbb{R})$, $B \in M_{m\times n}(\mathbb{R})$ quaisquer. (*Dica: *mostre isso para cada entrada)*

(c) Sejam as matrizes 2×2 $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ *e* $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. *Mostre que $BA = -AB$.*

Observamos que cada uma das operações algébricas de matrizes descritas acima expressa uma operação correspondente de t.l.'s. Mais precisamente:

- Se $\tilde{S}$ é *o.n.*, temos que $A_{ij} = \langle \vec{f}_i, T\vec{e}_j \rangle$ para todo $i = 1, \dots, m$, $j = 1, \dots, n$. Nesse caso, se a t.l. $T^* : W \to V$ é dada por $T^*\vec{f}_i = \vec{g}_i$, $i = 1, \dots, m$, onde $\vec{g}_1, \dots, \vec{g}_m$ são os vetores de V aparecendo na representação (III.5) de T acima, temos que

(III.6) $$\vec{y} = \sum_{i=1}^{m} y_i\vec{f}_i \in W \ \Rightarrow \ T^*\vec{y} = \sum_{i=1}^{m} y_i\vec{g}_i = \sum_{i=1}^{n} \langle \vec{f}_i, \vec{y} \rangle \vec{g}_i$$

(note o paralelo com a fórmula (III.5)!) e portanto

(III.7) $$\langle T^*\vec{y}, \vec{x} \rangle = \left\langle \sum_{i=1}^{m} y_i\vec{g}_i, \vec{x} \right\rangle = \sum_{i=1}^{m} y_i \langle \vec{g}_i, \vec{x} \rangle = \langle \vec{y}, T\vec{x} \rangle$$

para todo $\vec{x} \in V, \vec{y} \in W$. Notar que a primeira e última expressões em (III.7) *não* dependem de $S, \tilde{S}$! Em particular, se $\tilde{T} : W \to V$ satisfaz $\langle \tilde{T}\vec{y}, \vec{x}\rangle = \langle \vec{y}, T\vec{x}\rangle$ para todo $\vec{x} \in V, \vec{y} \in W$, então

$$\langle \tilde{T}\vec{y} - T^*\vec{y}, \vec{x}\rangle = \langle \tilde{T}\vec{y}, \vec{x}\rangle - \langle T^*\vec{y}, \vec{x}\rangle = \langle \vec{y}, T\vec{x}\rangle - \langle \vec{y}, T\vec{x}\rangle = 0$$

para todo $\vec{x} \in V, \vec{y} \in W$. Em particular, tomando $\vec{x} = \tilde{T}\vec{y} - T^*\vec{y}$, concluímos que $\|\tilde{T}\vec{y} - T^*\vec{y}\|^2 = 0$ e portanto $\tilde{T}\vec{y} = T^*\vec{y}$ para todo $\vec{y} \in W$. A t.l. T^* de W em V é denominada a *adjunta* de T, e a matriz B de T^* em $\tilde{S}, S$ é igual à *transposta* A^T da matriz A de T em $S, \tilde{S}$:

$$B_{ji} = (A^T)_{ji} = A_{ij} \ , \quad i = 1, \ldots, m \ , \ j = 1, \ldots, n \ .$$

Em particular, vemos que a identidade $A^{TT} = (A^T)^T = A$ expressa o fato que $T^{**} = (T^*)^* = T$, que por sua vez segue da unicidade da adjunta que obtivemos a partir de (III.7) acima no caso em que substituímos T por T^*.

- Sejam $T_1, T_2 : V \to W$ t.l.'s de V em W, $\beta \in \mathbb{R}$ quaisquer, $S = \{\vec{e}_1, \ldots, \vec{e}_n\}$ base de V e $\tilde{S} = \{\vec{f}_1, \ldots, \vec{f}_m\}$ base de W. Se $A = [A_i^j]$ é a matriz de T_1 em $S, \tilde{S}$ e $B = [B_i^j]$ é a matriz de T_2 em $S, \tilde{S}$, então

$$A + B = [A_{ij} + B_{ij}]$$

e

$$\beta A = [\beta A_{ij}]$$

são respectivamente as matrizes de $T_1 + T_2$ e βT_1 em $S, \tilde{S}$. De fato, como A_i^j é a componente de $T_1\vec{e}_j$ ao longo de $\vec{f}_i$ em $\tilde{S}$ e B_i^j é a componente de $T_2\vec{e}_j$ ao longo de $\vec{f}_i$ em $\tilde{S}$, claramente $A_i^j + B_i^j$ é a componente de $T_1\vec{e}_j + T_2\vec{e}_j = (T_1 + T_2)\vec{e}_j$ ao longo de $\vec{f}_i$ em $\tilde{S}$ e βA_i^j é a componente de $\beta T_1\vec{e}_j = (\beta T_1)\vec{e}_j$ ao longo de $\vec{f}_i$ em $\tilde{S}$ para cada $i = 1, \ldots, m$, $j = 1, \ldots, n$.

- Sejam V, W, X espaços vetoriais (reais), $T_1 : V \to W$, $T_2 : W \to X$ t.l.'s, e

$$S = \{\vec{e}_1, \ldots, \vec{e}_n\} \ ,$$
$$\tilde{S} = \{\vec{f}_1, \ldots, \vec{f}_m\} \ ,$$
$$\tilde{\tilde{S}} = \{\vec{h}_1, \ldots, \vec{h}_p\}$$

respectivamente bases de V, W e X. Se $A = [A_{ij}] \in M_{m\times n}(\mathbb{R})$ é a matriz de T_1 em $S, \tilde{S}$ e $B = [B_{ki}] \in M_{p\times m}$ é a matriz de T_2 em $\tilde{S}, \tilde{\tilde{S}}$, então $C = [C_{kj}] = BA =$ produto de B por A é a matriz de T_2T_1 em $S, \tilde{\tilde{S}}$. De fato, como A_{ij} é a componente de $T_1\vec{e}_j$ ao longo de $\vec{f}_i$ em $\tilde{S}$, B_{ki} é a componente de $T_2\vec{f}_i$ ao

longo de $\vec{h}_k$ em $\tilde{S}$ e C_{kj} é a componente de $T_2T_1\vec{e}_j$ ao longo de $\vec{h}_k$ para cada $i = 1, \ldots, m$, $j = 1, \ldots, n$, $k = 1, \ldots, p$, temos que

$$\begin{aligned} T_2T_1\vec{e}_j &= \sum_{k=1}^{p} C_{kj}\vec{h}_k = T_2\left(\sum_{i=1}^{m} A_{ij}\vec{f}_i\right) \\ &= \sum_{i=1}^{m} A_{ij}T_2\vec{f}_i = \sum_{i=1}^{m} A_{ij}\left(\sum_{k=1}^{p} B_{ki}\vec{h}_k\right) \\ &= \sum_{i=1}^{m}\left(\sum_{k=1}^{p} A_{ij}B_{ki}\vec{h}_k\right) \\ &= \sum_{k=1}^{p}\left(\sum_{i=1}^{m} B_{ki}A_{ij}\right)\vec{h}_k\ , \quad j = 1, \ldots, n\ , \end{aligned}$$

logo obtemos a fórmula desejada. No caso em que $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ e $X = \mathbb{R}^p$, a vemos que a multiplicação de matrizes é precisamente a composição de t.l.'s.

- A matriz da *identidade* $\mathbb{1} = \mathbb{1}_V$ de V em S é precisamente a matriz identidade $\mathbb{1} = \mathbb{1}_n$, $n = \dim(V)$, já que $\mathbb{1}\vec{e}_j = \vec{e}_j$ para todo $j = 1, \ldots, n$. Notar que essa matriz é a *mesma* para *todas* as bases de V, por isso usamos a mesma notação para a identidade e sua matriz, posto que isso não causará confusão.

- Se uma t.l. $T : U \to V$ é invertível, vimos acima que sua *inversa* $T^{-1} : W \to V$ é também uma t.l.. Além disso, se $S = \{\vec{e}_1, \ldots, \vec{e}_n\}$ é base de V e $\tilde{S} = \{\vec{f}_1, \ldots, \vec{f}_m\}$ é base de W, A é a matriz de T em $S, \tilde{S}$ e B é a matriz de T^{-1} em $\tilde{S}, S$, então

$$\begin{aligned} T^{-1}T\vec{e}_j = \vec{e}_j &= \sum_{i=1}^{m} A_{ij}T^{-1}\vec{f}_i = \sum_{i=1}^{m} A_{ij}\left(\sum_{l=1}^{n} B_{li}\vec{e}_l\right) \\ &= \sum_{l=1}^{n}\left(\sum_{i=1}^{m} B_{li}A_{ij}\right)\vec{e}_l \ \Rightarrow\ BA = \mathbb{1} \end{aligned}$$

e

$$\begin{aligned} TT^{-1}\vec{f}_i = \vec{f}_i &= \sum_{j=1}^{n} B_{ji}T\vec{e}_j = \sum_{j=1}^{n} B_{ji}\left(\sum_{k=1}^{m} A_{kj}\vec{f}_k\right) \\ &= \sum_{k=1}^{m}\left(\sum_{j=1}^{n} A_{kj}B_{ji}\right)\vec{f}_k \ \Rightarrow\ AB = \mathbb{1}\ , \end{aligned}$$

ou seja, B é a *matriz inversa* A^{-1} de A. Notar que AB e BA só podem ser definidas simultaneamente se $n = m$.

3.3. O efeito de uma mudança de base sobre a matriz de uma transformação linear. Veremos agora como muda a matriz de uma t.l. $T : V \to W$ se mudarmos as bases de V e de W. Para tal, sejam

$$S_1 = \{\vec{e}_1, \ldots, \vec{e}_n\} \ , \ S_2 = \{\vec{e}'_1, \ldots, \vec{e}'_n\}$$

bases de V, e

$$\tilde{S}_1 = \{\vec{f}_1, \ldots, \vec{f}_m\} \ , \ \tilde{S}_2 = \{\vec{f}'_1, \ldots, \vec{f}'_m\}$$

bases de W. Defina as transformações lineares $U : V \to V$, $\tilde{U} : W \to W$ como

$$U\vec{e}_j = \vec{e}'_j \ , \ \tilde{U}\vec{f}_i = \vec{f}'_i \ , \quad i = 1, \ldots, m \ , \ j = 1, \ldots, n \ .$$

Segue que U e $\tilde{U}$ são *invertíveis*, e suas respectivas inversas U^{-1}, $\tilde{U}^{-1}$ são obviamente dadas por

$$U^{-1}\vec{e}'_j = \vec{e}_j \ , \ \tilde{U}^{-1}\vec{f}'_i = \vec{f}_i \ , \quad i = 1, \ldots, m \ , \ j = 1, \ldots, n \ .$$

Escrevemos

$$\vec{e}'_l = U\vec{e}_l = \sum_{j=1}^{n} C_{jl}\vec{e}_j \ , l = 1, \ldots, n \ ,$$

$$\vec{f}'_k = \tilde{U}\vec{f}_k = \sum_{i=1}^{m} D_{ik}\vec{f}_i \ , \quad k = 1, \ldots, m \ ,$$

de modo que $C = [C_{jl}]$ é a matriz de U em S_1 e $D = [D_{ik}]$ é a matriz de $\tilde{U}$ em $\tilde{S}_1$, denominadas respectivamente *matrizes de mudança de base* de S_1 para S_2 e de $\tilde{S}_1$ para $\tilde{S}_2$. Devido à forma particular de U e $\tilde{U}$, temos que

$$U\vec{e}'_l = \sum_{j=1}^{n} C_{jl}U\vec{e}_j = \sum_{j=1}^{n} C_{jl}\vec{e}'_j \ , \quad l = 1, \ldots, n \ ,$$

$$\tilde{U}\vec{f}'_k = \sum_{i=1}^{m} D_{ik}\tilde{U}\vec{f}_i = \sum_{i=1}^{m} D_{ik}\vec{f}'_i \ , \quad k = 1, \ldots, m \ ,$$

ou seja, C é também a matriz de U em S_2 e D é também a matriz de $\tilde{U}$ em $\tilde{S}_2$. Obviamente, U e $\tilde{U}$ (e, portanto, C e D) são invertíveis, e U^{-1} (resp. $\tilde{U}^{-1}$) transforma a base S_2 (resp. $\tilde{S}_2$) de volta na base S_1 (resp. $\tilde{S}_1$), de modo que C^{-1} (resp. D^{-1}) é a matriz de mudança da base S_2 (resp. $\tilde{S}_2$) para a base S_1 (resp. $\tilde{S}_1$). Logo, pelo argumento acima, C^{-1} é a matriz de U^{-1} em S_1 *e* S_2, e D^{-1} é a matriz de $\tilde{U}^{-1}$ em $\tilde{S}_1$ *e* $\tilde{S}_2$.

Segue então que

$$\vec{x} = \sum_{j=1}^{n} x_j\vec{e}_j = \sum_{l=1}^{n} x'_l\vec{e}'_l = \sum_{l=1}^{n} x'_l\left(\sum_{j=1}^{n} C_{jl}\vec{e}_j\right) = \sum_{j=1}^{n}\left(\sum_{l=1}^{n} C_{jl}x'_l\right)\vec{e}_j$$

$$\Rightarrow \vec{x}_{S_2} = C^{-1}\vec{x}_{S_1}$$

para todo $\vec{x} \in V$ e, por um cálculo análogo, $\vec{y}_{\tilde{S}_2} = D^{-1}\vec{y}_{\tilde{S}_1}$ para todo $\vec{y} \in W$. Seja agora $T : V \to W$ uma t.l. cuja matriz em $S_1, \tilde{S}_1$ é dada por A e cuja matriz em $S_2, \tilde{S}_2$ é dada por B:

$$T\vec{e}_j = \sum_{i=1}^{m} A_{ij}\vec{f}_i \ , \ T\vec{e}'_l = \sum_{k=1}^{m} B_{kl}\vec{f}'_k \ .$$

Segue que

$$\begin{aligned} T\vec{e}'_l &= T\left(\sum_{j=1}^{n} C_{jl}\vec{e}_j\right) = \sum_{j=1}^{n} C_{jl}T\vec{e}_j = \sum_{j=1}^{n} C_{jl}\left(\sum_{i=1}^{m} A_{ij}\vec{f}_i\right) \\ &= \sum_{k=1}^{m} B_{kl}\vec{f}'_k = \sum_{k=1}^{m} B_{kl}\left(\sum_{i=1}^{m} D_{ik}\vec{f}_i\right) \\ &= \sum_{i=1}^{m}\left(\sum_{k=1}^{m} D_{ik}B_{kl}\right)\vec{f}_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij}C_{jl}\right)\vec{f}_i \end{aligned}$$

e portanto $DB = AC$, logo

$$B = D^{-1}AC \ .$$

Essa fórmula é particularmente conveniente no caso em que $\tilde{S}_1$ e $\tilde{S}_2$ são *o.n.*, pois nesse caso

$$D_{ik} = \langle\vec{f}_i, \tilde{U}\vec{f}_k\rangle = \langle\vec{f}_i, \vec{f}'_k\rangle = \langle\vec{f}'_k, \vec{f}_i\rangle = \langle\vec{f}'_k, \tilde{U}^{-1}\vec{f}'_i\rangle = (D^{-1})_{ki}$$

para todo $i, k = 1, \ldots, m$ e portanto $D^{-1} = D^T$, de modo que $B = D^T AC$. Da mesma forma, se S_1 e S_2 são o.n., concluímos analogamente que $C^{-1} = C^T$.

4. O Teorema do Núcleo e da Imagem

Seja $T : V \to W$ é uma t.l.. Sendo T uma aplicação de V em W, então podemos falar de imagens e imagens inversas de subconjuntos por T. Mais precisamente, dados $S \subset V$ e $\tilde{S} \subset W$, denotamos a *imagem* de S por T por

$$T(S) = \{\vec{y} = T\vec{x} \in W \mid \vec{x} \in S\}$$

e a *imagem inversa* de $\tilde{S}$ por T por

$$T^{-1}(\tilde{S}) = \{\vec{x} \in V \mid T\vec{x} \in \tilde{S}\} \ .$$

Se $\tilde{S} = \{\vec{y}\}$ é unitário, dizemos simplesmente que $T^{-1}(\{\vec{y}\}) = T^{-1}(\vec{y})$ é a *imagem inversa* de $\vec{y}$ por T.

Cabe aqui uma observação sobre a notação empregada para imagens inversas. Note que T *não* precisa ser invertível! $T^{-1}(\vec{y})$ denota um *subconjunto* de V (que é

vazio se $\vec{y} \notin T(V)$ ou tem *mais de um* elemento se T não for injetora) e pode ser identificado com um vetor em V *somente* no caso em que $T^{-1}(\vec{y})$ é unitário.

Segue dessas definições que

$$\begin{aligned} T(L(S)) &= \left\{ \vec{y} = T\left(\sum_{j=1}^{k} \alpha_j \vec{x}_j \right) \;\middle|\; \alpha_1, \ldots, \alpha_k \in \mathbb{R} \,,\; \vec{x}_1, \ldots, \vec{x}_j \in S \right\} \\ &= \left\{ \vec{y} = \sum_{j=1}^{k} \alpha_j T\vec{x}_j \;\middle|\; \alpha_1, \ldots, \alpha_k \in \mathbb{R} \,,\; \vec{x}_1, \ldots, \vec{x}_j \in S \right\} \\ &= L(T(S)) \end{aligned}$$

– em particular, a *imagem*

$$\mathrm{Im}\,(T) = T(V)$$

de T é subespaço vetorial de W – e que o *núcleo* ("kernel" em inglês)

$$\ker(T) = T^{-1}(\vec{0})$$

de T é subespaço vetorial de V: se $\vec{x}, \vec{x}' \in \ker(T)$ e $\alpha \in \mathbb{R}$ quaisquer, então

$$T(\vec{x} + \vec{x}') = T\vec{x} + T\vec{x}' = \vec{0} + \vec{0} = \vec{0}$$

e

$$T(\alpha\vec{x}) = \alpha T\vec{x} = \alpha\vec{0} = \vec{0} \,,$$

portanto $\vec{x} + \vec{x}'$, $\alpha\vec{x} \in \ker(T)$.

Por definição, T é *sobrejetora* se e somente se

$$\mathrm{Im}\,(T) = W \,.$$

Além disso, T é *injetora* se e somente se

$$\ker(T) = \{\vec{0}\} \,.$$

De fato, se $\ker(T) = \{\vec{0}\}$ e $\vec{x}, \vec{x}' \in V$ satisfazem $T\vec{x} = T\vec{x}'$, temos que $T\vec{x} - T\vec{x}' = T(\vec{x} - \vec{x}') = \vec{0}$ e portanto $\vec{x} - \vec{x}' = \vec{0}$. Conversamente, temos que se $T\vec{x} = \vec{0}$ então $T(\vec{x}+\vec{x}') = T\vec{x}+T\vec{x}' = T\vec{x}'$ para todo $\vec{x}' \in V$. Logo, se T é injetora, nesse caso $\vec{x} = \vec{0}$. A sobrejetividade e a injetividade de uma t.l. $T : V \to W$ podem ser expressas em termos de sua ação sobre (certos) subconjuntos l.i. $S \subset V$:

- T é *sobrejetora* se e somente se, dada uma base S qualquer de V, temos que $L(T(S)) = \mathrm{Im}\,(T) = W$. (a primeira identidade segue do cálculo acima, e a segunda identidade é a definição da sobrejetividade de T)

- T é *injetora* se e somente se, dado um subconjunto l.i. $S \subset V$ qualquer, temos que $T(S)$ *também é l.i.*. (suponha que $T(S)$ é l.i. para todo $S \subset V$ l.i., então $T\vec{x} \neq \vec{0}$ se $\vec{x} \neq \vec{0}$ e portanto T é injetora. Conversamente, se T é injetora e $\vec{x}_1, \ldots, \vec{x}_k$ são l.i., sejam $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tais que $\sum_{j=1}^{k} \alpha_k T(\vec{x}_k) = \vec{0} = T\left(\sum_{j=1}^{k} \alpha_j \vec{x}_j\right)$. Então $\sum_{j=1}^{k} \alpha_j \vec{x}_j = \vec{0}$ e portanto $\alpha_1 = \cdots = \alpha_k = 0$, logo $T(\vec{x}_1), \ldots, T(\vec{x}_k)$ são l.i.)

Em particular, T é *bijetora* se e somente se $T(S)$ é base de W para toda base S de V. Em outras palavras, o exemplo típico de t.l.'s bijetoras ocorre ao efetuarmos uma *mudança de base*. Se V tem dimensão *finita*, ambos os fatos acima são casos particulares de um resultado geral, conhecido como *Teorema do Núcleo e da Imagem*. Dados espaços vetoriais (reais) V, W e $T \in L(V, W)$, o *posto* ("rank" em inglês) de T é dado por

$$R(T) = \dim(\mathrm{Im}\,(T)) \ ,$$

e a *nulidade* de T por

$$N(T) = \dim(\ker(T)) \ .$$

Obviamente $R(T), N(T) \leq \dim(V)$ pois $L(T(S)) = \mathrm{Im}\,(T)$ se S é base de V e $\ker(T)$ é subespaço vetorial de V.

III.1 Teorema (do Núcleo e da Imagem). *Sejam V, W espaços vetoriais (reais), $n = \dim(V) < \infty$, e $T \in L(V, W)$. Então*

$$N(T) + R(T) = \dim(V) = n \ .$$

Demonstração. Definamos $k = N(T)$, $l = R(T)$. Sejam $S_0 = \{\vec{e}_1, \ldots, \vec{e}_k\}$ uma base de $\ker(T)$ e $\tilde{S}_1 = \{\vec{f}_1, \ldots, \vec{f}_l\}$ uma base de $\mathrm{Im}\,(T)$, de modo que $\vec{f}_i = T\vec{e}_{k+i}$ para algum $\vec{e}_{k+i} \in V$, $i = 1, \ldots, l$. Obviamente $\vec{e}_{k+i} \notin \ker(T)$ para todo $i = 1, \ldots, l$; mostraremos agora que definindo $S_1 = \{\vec{e}_{k+1}, \ldots, \vec{e}_{k+l}\}$, temos que $S = S_0 \cup S_1 = \{\vec{e}_1, \ldots, \vec{e}_{k+l}\}$ é *base* de V. Primeiramente, notar que S é l.i.: dados $\alpha_1, \ldots, \alpha_{k+l} \in \mathbb{R}$ tais que $\sum_{j=1}^{k+l} \alpha_j \vec{e}_j = \vec{0}$, temos que

$$T\left(\sum_{j=1}^{k+l} \alpha_j \vec{e}_j\right) = \vec{0} = \sum_{j=1}^{k+l} \alpha_j T\vec{e}_j = \sum_{i=1}^{l} \alpha_{k+i} \vec{f}_i$$

e portanto $\alpha_{k+1} = \cdots = \alpha_{k+l} = 0$ pois $\tilde{S}_1$ é l.i.. Logo, $\sum_{j=1}^{k} \alpha_j \vec{e}_j = \vec{0}$ e portanto $\alpha_1 = \cdots = \alpha_k = 0$ pois S_0 é l.i.. Falta apenas concluir que $L(S) = V$ – para tal, notar que dado $\vec{x} \in V$, podemos escrever

$$T\vec{x} = \sum_{i=1}^{l} x_{k+i} \vec{f}_i = \sum_{i=1}^{l} x_{k+i} T\vec{e}_{k+i} = T\left(\sum_{i=1}^{l} x_{k+i} \vec{e}_{k+i}\right)$$

para uma (única) escolha de $x_{k+1}, \ldots, x_{k+l} \in \mathbb{R}$. Defina

$$\vec{x}_1 = \sum_{i=1}^{l} x_{k+i}\vec{e}_{k+i} \ , \ \vec{x}_0 = \vec{x} - \vec{x}_1 \ .$$

Segue que $T\vec{x}_1 = T\vec{x}$ e portanto $\vec{x}_0 \in \ker(T)$, pois

$$T\vec{x}_1 = \sum_{i=1}^{l} x_{k+i}\vec{f}_i = T\vec{x} \ , \ T\vec{x}_0 = T(\vec{x} - \vec{x}_1) = T\vec{x} - T\vec{x} = \vec{0} \ ,$$

logo podemos escrever $\vec{x}_0 = \sum_{j=1}^{k} x_j\vec{e}_j$ para uma (única) escolha de $x_1, \ldots, x_k \in \mathbb{R}$ e portanto $\vec{x} = \sum_{j=1}^{k+l} x_j\vec{e}_j$, como desejado. □

O Teorema do Núcleo e da Imagem possui várias consequências. No que se segue, V, W são espaços vetoriais (reais) com $\dim(V) < \infty$.

(i) $T \in L(V, W)$ é sobrejetora se e somente se $N(T) = \dim(V) - \dim(W)$. Em particular, se $\dim(W) > \dim(V)$, então T não pode ser sobrejetora.

(ii) $T \in L(V, W)$ é injetora se e somente se $R(T) = \dim(V)$.

(iii) $T \in L(V, W)$ é bijetora se e somente se $R(T) = \dim(V) = \dim(W)$.

(iv) Se $T \in V'$, então $N(T) = \dim(V) - 1$ ou $\dim(V)$, e o último caso ocorre se e somente se $T = 0$.

O fato (iv) acima fornece uma maneira alternativa de obter o Lema de Riesz:

III.2 Lema (de Riesz). *Seja V um espaço vetorial,* $\dim(V) < \infty$. *Dado $T \in V'$, existe um único* vetor $\vec{x}_T \in V$ *tal que $T\vec{x} = \langle \vec{x}_T, \vec{x} \rangle$ para todo $\vec{x} \in V$ − a saber,*

$$\vec{x}_T = \begin{cases} \vec{0} & (T = 0) \\ \frac{T\vec{z}_T}{\|\vec{z}_T\|^2}\vec{z}_T & (T \neq 0) \end{cases} ,$$

onde $0 \neq \vec{z}_T$ é um vetor ortogonal a $\ker(T)$.

Notar que, pelo Teorema do Núcleo e da Imagem, $\vec{z}_T$ é único a menos de um múltiplo escalar.

Demonstração. Comecemos com a questão da *unicidade*. Se $\vec{z}, \vec{z}' \in V$ são tais que $\langle \vec{z}, \vec{x} \rangle = \langle \vec{z}', \vec{x} \rangle = T\vec{x}$ para todo $\vec{x} \in V$, então

$$\langle \vec{z} - \vec{z}', \vec{x} \rangle = \langle \vec{z}, \vec{x} \rangle - \langle \vec{z}', \vec{z} \rangle = T\vec{z} - T\vec{z} = \vec{0}$$

para todo $\vec{x} \in V$. Em particular, tomando $\vec{x} = \vec{z} - \vec{z}'$ obtemos $\|\vec{z} - \vec{z}'\|^2 = 0$ e portanto $\vec{z} = \vec{z}'$. Passando agora à questão da existência, se $T = 0$, obviamente não há outra escolha para $\vec{x}_T$ a não ser $\vec{x}_T = \vec{0}$. Se $T \neq 0$, seja $\vec{z} \in T$ tal que $T\vec{z} \neq \vec{0}$. Podemos então construir uma base o.n. $S_0 = \{\vec{e}_1, \ldots, \vec{e}_{n-1}\}$ de $\ker(T)$ e portanto a projeção

ortogonal $P_{\ker(T)}$ sobre $\ker(T)$. De maneira similar à adotada na ortonormalização de Gram-Schmidt, defina

$$\vec{z}_T = \vec{z} - P_{\ker(T)}\vec{z} \ , \ \vec{e}_n = \frac{1}{\|\vec{z}_T\|}\vec{z}_T \ ,$$

de modo que $S = \{\vec{e}_1, \ldots, \vec{e}_n\}$ é base o.n. de V. Mostraremos agora que se $\vec{x}_T$ tem as propriedades listadas no enunciado do Lema de Riesz, então

$$\vec{x}_T = (T\vec{e}_n)\vec{e}_n = \frac{T\vec{z}_T}{\|\vec{z}_T\|^2}\vec{z}_T \ .$$

De fato, dado $\vec{x} \in V$ qualquer, podemos escrever

$$\vec{x} = P_{\ker(T)}\vec{x} + (\vec{x} - P_{\ker(T)}\vec{x}) \ , \quad \vec{x} - P_{\ker(T)}\vec{x} = \langle \vec{e}_n, \vec{x}\rangle \vec{e}_n$$

e portanto

$$T\vec{x} = TP_{\ker(T)}\vec{x} + \langle \vec{e}_n, \vec{x}\rangle T\vec{e}_n = \langle (T\vec{e}_n)\vec{e}_n, \vec{x}\rangle \ ,$$

como desejado. □

Notar que $\vec{x}_{T_1+T_2} = \vec{x}_{T_1} + \vec{x}_{T_2}$ e $\vec{x}_{\alpha T_1} = \alpha \vec{x}_{T_1}$ para $T_1, T_2 \in V'$, $\alpha \in \mathbb{R}$ quaisquer. Vamos agora estabelecer a relação da fórmula obtida no Lema III.2 com a fórmula que já tínhamos obtido anteriormente: se $T \in V'$ e $S = \{\vec{e}_1, \ldots, \vec{e}_n\}$ é uma base o.n. de V, então

$$V \ni \vec{x} = \sum_{j=1}^{n} x_j \vec{e}_j \ \Rightarrow \ T\vec{x} = \sum_{j=1}^{n} x_j T\vec{e}_j = \langle \vec{x}_T, \vec{x}\rangle$$

$$\Rightarrow \ \vec{x}_T = \sum_{j=1}^{n} (T\vec{e}_j)\vec{e}_j \ .$$

Em particular, $\vec{x}_T$ é ortogonal a $\ker(T)$ e

$$T\vec{x}_T = \sum_{j=1}^{n} (T\vec{e}_j)^2 = \|\vec{x}_T\|^2 \ ,$$

logo $T\vec{x}_T = \vec{0}$ se e somente se $T = 0$.

Aproveitamos o momento para introduzir uma notação alternativa para $\vec{x}_T$, que será a notação padrão daqui em diante. Dados $\vec{x} \in V$, $T \in V'$, definimos $\vec{x}^\flat \in V'$, $T^\sharp \in V$ como

$$\vec{x}^\flat(\vec{x}') = \langle \vec{x}, \vec{x}'\rangle \ (\vec{x}' \in V) \ , \quad T^\sharp = \vec{x}_T \ .$$

Fica claro que $(\vec{x}^\flat)^\sharp = \vec{x}$ e $(T^\sharp)^\flat = T$ para $\vec{x} \in V$, $T \in V'$ quaisquer. Além disso, as aplicações $V \ni \vec{x} \mapsto \vec{x}^\flat \in V'$, $V' \ni T \mapsto T^\sharp \in V$ são t.l.'s (*exercício:* verifique!), portanto denotar-las-emos por *isomorfismos musicais*.

5. Pseudoinversas de transformações lineares. Regressão linear

IV

Sistemas lineares

V

Determinantes, autovalores e autovetores

1. Determinantes

2. Autovalores e autovetores de uma transformação linear

3. Decomposição em valores singulares. Análise de componentes principais

Números complexos e o Teorema Fundamental da Álgebra

A

B

Permutações

Bibliografia

[1] T. M. Apostol, *Calculus, Volume II – Second Edition*. Wiley, 1969.

[2] S. J. Axler, *Linear Algebra Done Right – Fourth Edition*. Springer, 2026. Disponibilizado com acesso aberto em PDF pelo autor em `https://linear.axler.net/LADR4e.pdf` (último acesso: 15 de setembro de 2026).

[3] S. Mac Lane, G. Birkhoff, *Algebra – Third Edition*. AMS Chelsea, 2000.