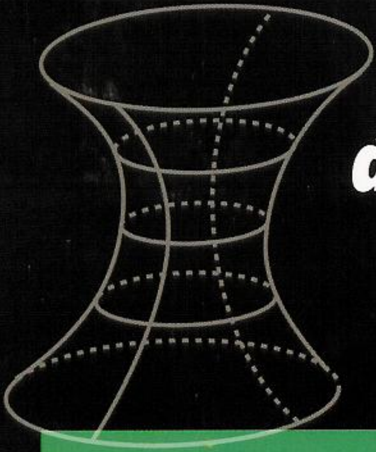
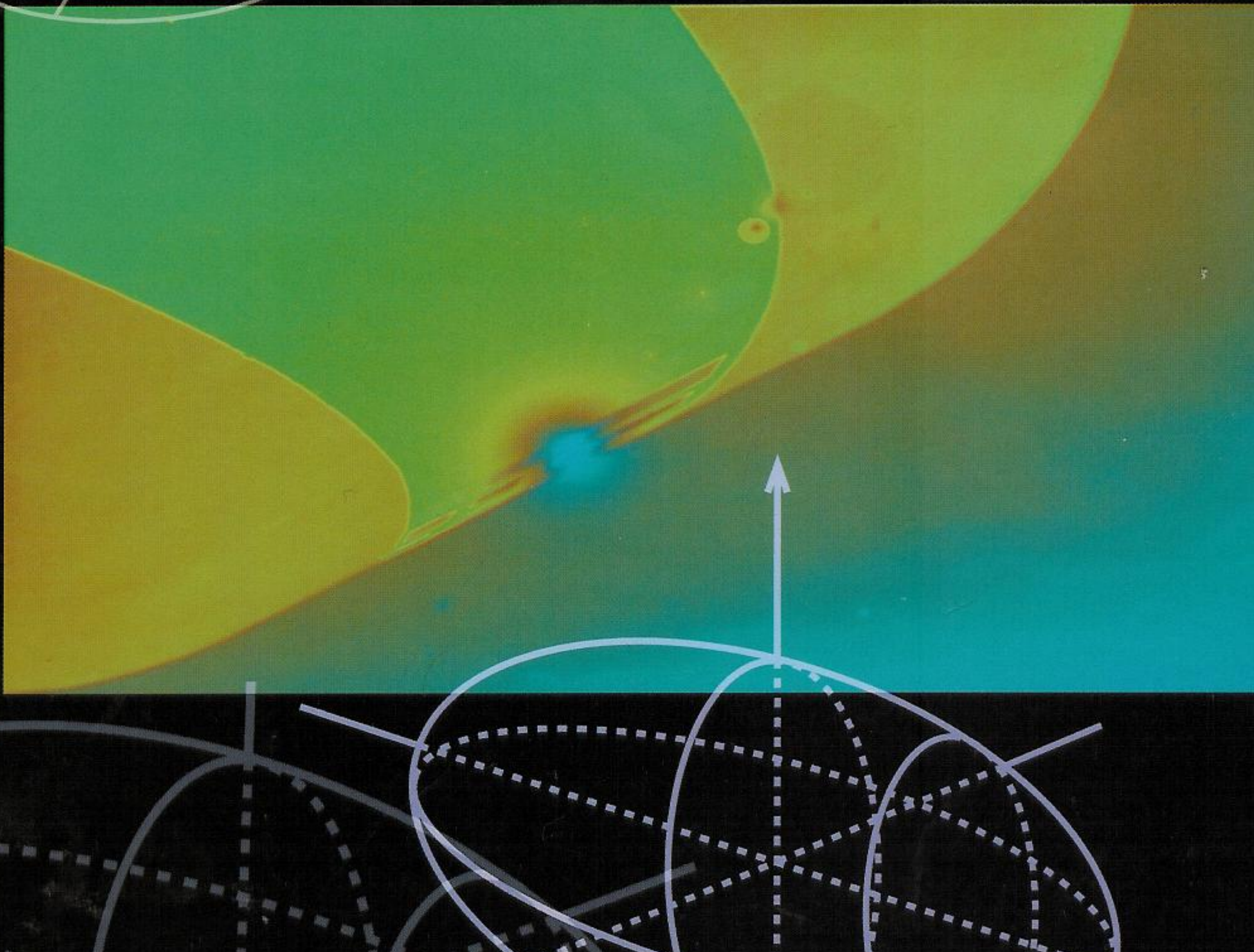
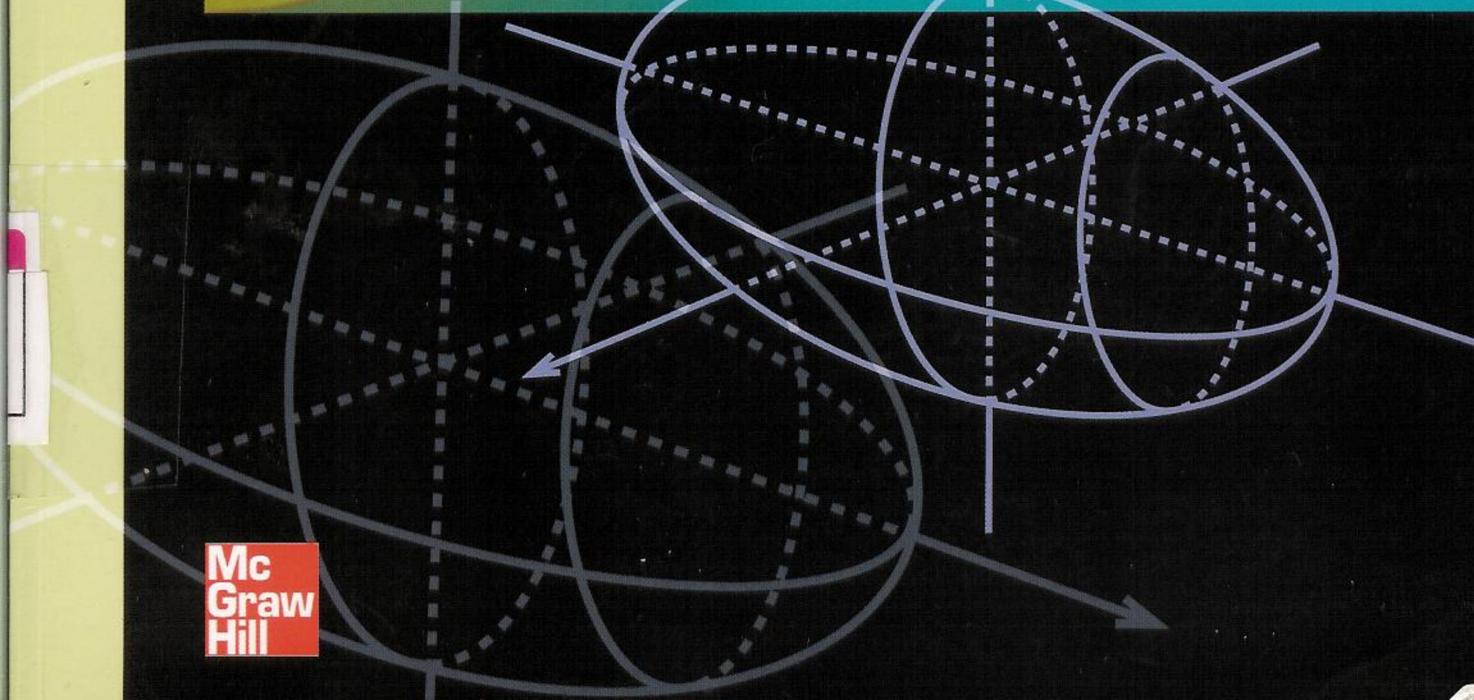


álgebra linear e geometria analítica

ANTÓNIO MONTEIRO





McGraw Hill

Álgebra Linear
e Geometria Analítica

António Monteiro
Professor da Universidade Lusíada,
Lisboa

McGraw-Hill

LISBOA • SÃO PAULO • BOGOTÁ • BUENOS AIRES • GUATEMALA
MADRID • MÉXICO • NOVA IORQUE • PANAMÁ • SAN JUAN • SANTIAGO

AUCKLAND • HAMBURGO • LONDRES • MILÃO • MONTREAL • NOVA DELI
PARIS • SINGAPURA • SYDNEY • TÓQUIO • TORONTO

Álgebra Linear e Geometria Analítica

António Monteiro
Professor da Universidade Lusíada,
Lisboa

McGraw-Hill

LISBOA • SÃO PAULO • BOGOTÁ • BUENOS AIRES • GUATEMALA
MADRID • MÉXICO • NOVA IORQUE • PANAMÁ • SAN JUAN • SANTIAGO
AUCKLAND • HAMBURGO • LONDRES • MILÃO • MONTREAL • NOVA DELI
PARIS • SINGAPURA • SYDNEY • TÓQUIO • TORONTO

McGraw-Hill

A Division of The McGraw-Hill Companies

António J. Monteiro, nascido em Lisboa em 1951, licenciou-se em Matemática na Faculdade de Ciências da Universidade de Lisboa em 1973 e obteve o grau de mestre em Álgebra, na mesma faculdade, em 1989.

Tem exercido uma larga e variada actividade docente, quer no campo da Álgebra quer no da Análise Matemática. Presentemente, é professor na Universidade Lusíada. É autor de diversos artigos e livros no domínio da Álgebra e afins.

Fora do âmbito da Matemática, tem desenvolvido uma intensa actividade no campo da Malacologia, sendo autor e co-autor de numerosos artigos e livros. Tem-se ainda dedicado à literatura fantástica, tendo publicado diversas obras de ficção e outras.

ÁLGEBRA LINEAR E GEOMETRIA ANALÍTICA

Editora McGraw-Hill de Portugal, L.da
Rua Barata Salgueiro, 51-A
1250-043 LISBOA, PORTUGAL
Telef.: (351) 213 553 180
Fax: (351) 213 553 189
e-mail: servico_clientes@mcgraw-hill.com

Depósito legal: 220806/04
ISBN: 972-773-106-6
1E1P05011M01T0
1E2P09021M02T0
1E3P02050M52T5

Coordenação Editorial: Susana Calhau
Coordenação de Produção: Sofia Marques
Composição e Paginação: Nuno de Carvalho
Capa: Claudia Gigliotti
Impressão: Tipografia Lousanense, L.da – Lousã

Impresso em Portugal – Printed in Portugal

Índice geral

Introdução

A disciplina de Álgebra Linear e Geometria Analítica aparece normalmente integrada nos primeiros anos dos currículos de cursos de Matemática, Informática, Engenharia, Física, Química, Economia, Agronomia, etc. Com efeito, a Álgebra Linear fornece um quadro teórico rigoroso adequado ao tratamento de inúmeras questões oriundas das diversas áreas científicas, sendo de salientar as múltiplas ligações com a Análise Matemática. Para além disso, uma tal disciplina constitui, muitas vezes, o primeiro contacto dos estudantes com uma teoria apresentada de modo consideravelmente abstracto, em que a exposição faz um apelo relativamente reduzido à intuição. Nesta perspectiva, a Álgebra Linear desempenha também um papel formativo e propedêutico.

A origem de muitas das técnicas que hoje integram a Álgebra Linear, motivada pela necessidade de resolver eficazmente problemas específicos, perde-se no tempo. Assim, por exemplo, a utilização de matrizes na resolução de sistemas de equações lineares pode encontrar-se já em trabalhos de matemáticos chineses, cerca de 200 anos antes de Cristo (se bem que o termo "matriz" só em 1859 fosse formalmente introduzido por James Sylvester). De resto, já um século antes os babilónios haviam lidado com problemas semelhantes. Por sua vez, em 1683, Leibniz dominava já o conceito de "determinante", que, de resto, era usado, na mesma época, no Japão. A noção fulcral de "espaço vectorial" foi formalizada por Giuseppe Peano em 1888, se bem que a noção de "vector" se encontre, quase um século antes, nos trabalhos de Bolzano.

Ao longo dos séculos, muitos dos mais eminentes matemáticos europeus, entre eles Euler, Lagrange, Gauss, Cayley, Grassmann, Hamilton, Frobenius, Cauchy, Schwarz, Laplace, Cramer, Descartes, etc., contribuíram para o desenvolvimento das ideias fundamentais da Álgebra Linear.

Do que fica dito, infere-se a vantagem de pôr à disposição dos estudantes um texto acessível, onde os tópicos principais, que forçosamente integrarão um qualquer programa de Álgebra Linear, sejam apresentados com rigor e com clareza adequados. É esse o propósito do presente trabalho, que reflecte algumas dezenas de anos de prática pedagógica, no âmbito das matérias tratadas.

A preparação de um texto com estas características apresenta algumas dificuldades, nomeadamente no que respeita à obtenção de um saudável equilíbrio entre a necessária concisão — sem a qual nos arriscaríamos a ver a obra estender-se por um número incomportável de páginas — e a simplicidade da exposição, sem a qual o objectivo fundamental ficaria por atingir. Procurámos que tudo ficasse facilmente compreensível, mesmo quando as óbvias limitações de espaço se podem traduzir em algum nível de abstracção e aridez.

Pretende-se, geralmente, que um livro desta natureza seja, tanto quanto possível, auto-suficiente. Com efeito, a exposição da matéria não requer, regra geral, senão um conhecimento prévio dos conceitos mais elementares da Aritmética. Na verdade, porém, é de pressupor no leitor um certo nível de sofisticação e de prática do raciocínio típico da Matemática. Mais ainda, a regra da auto-suficiência é frequentemente infringida no nível dos exemplos ilustrativos das diversas noções.

Do ponto de vista da utilização deste trabalho como livro de texto para acompanhamento de um curso sobre a matéria focada — que cabe, sem atropelos e com poucas omissões, num curso anual —, impõem-se duas observações: em primeiro lugar, a de que a leitura de um livro, mesmo que de boa qualidade, não dispensa, em regra, a frequência das aulas em que a mesma matéria é exposta; em segundo lugar, a de que o domínio dos assuntos estudados passa pela realização de exercícios de aplicação, tanto de natureza prática como de carácter teórico. Permitimo-nos deixar o presente trabalho despido de uma tal componente prática, pelo facto de ele se encontrar acompanhado, nesse sector, por uma outra obra, *Álgebra Linear e Geometria Analítica — Problemas e Exercícios*, A. Monteiro *et al.*, publicado em 1997 pela Editora McGraw-Hill de Portugal.

Apesar do cuidado com que o texto foi elaborado, é possível que se notem, aqui e ali, imperfeições. Ficaremos, como sempre, muito gratos aos leitores que no-las queiram transmitir.

Finalmente, queremos agradecer à Editora McGraw-Hill de Portugal o interesse com que acolheu o nosso trabalho e felicitá-la pela continuidade da sua acção editorial, verdadeiramente ímpar e inestimável no nosso país.

1 Sistemas de equações lineares

1-1 Generalidades

Muitas situações que encontraremos no estudo da Álgebra Linear, bem como outras que se encontram no estudo das mais diversas matérias, requerem a resolução de sistemas de equações lineares. Por esse motivo, começaremos, precisamente, por analisar este tipo de problema. Ao longo deste capítulo, trabalharemos fundamentalmente no domínio dos números reais, mas os resultados que vamos obter são válidos em muitos outros domínios (nomeadamente, no domínio dos números racionais ou dos números complexos, ou, na sua generalidade, num corpo qualquer).

Como é do conhecimento geral, chama-se "equação" a uma igualdade que envolve uma ou mais quantidades desconhecidas, as "incógnitas"; "resolver" uma equação consiste em determinar os valores que deveríamos atribuir a essas incógnitas, dentro de um domínio preestabelecido, por forma que a equação se transformasse numa igualdade verdadeira. Quando, em vez de uma só equação, temos um certo número (finito) de equações que deverão ser satisfeitas simultaneamente, dizemos que se trata de um "sistema" de equações.

Assim, por exemplo, $7x^2 - x + 8 = 0$, $3xy - 2x + y = 1$, ou $x + 1 = \log x$ são exemplos de equações, enquanto

$$\begin{cases} \operatorname{sen} x = \frac{1}{2} \\ \cos x = -\frac{\sqrt{3}}{2} \end{cases} \quad \text{ou} \quad \begin{cases} 3x + 2y = 7 \\ 5x - 4y = 9 \end{cases}$$

são sistemas de equações.

Chamamos **equação linear**, com as incógnitas $x_1, x_2, \ldots, x_n$ e coeficientes reais (ou complexos, ou racionais, ...) àquela que resulta de igualar a uma constante arbitrária uma expressão polinomial de primeiro grau, com as variáveis $x_1, x_2, \ldots, x_n$ e coeficientes reais (ou complexos, ou racionais, ...), isto é, uma equação da forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

onde $a_1, a_2, \ldots, a_n, b$ são números reais (ou complexos, ou racionais, ...).

Uma **solução** desta equação será, portanto, uma lista $(r_1, r_2, \dots, r_n)$ de números reais (ou complexos, ou racionais, ..., segundo o domínio estabelecido) que satisfaçam a equação, ou seja, tais que se torne verdadeira a igualdade

$$a_1 r_1 + a_2 r_2 + \cdots + a_n r_n = b$$

Numa equação linear, as constantes $a_1, a_2, \dots, a_n$ são os **coeficientes** das incógnitas, enquanto b é o **termo independente**.

No que se segue, consideraremos exclusivamente sistemas de equações lineares. Para representar um sistema de equações lineares, com um número arbitrário de equações e um número arbitrário de incógnitas, usaremos uma notação especialmente concebida para facilitar a distinção entre as várias equações e entre os coeficientes das diversas incógnitas. Essa notação consiste, basicamente, em designar os coeficientes das incógnitas por símbolos da forma a_{ij}, com dois índices: o primeiro desses índices indica a equação em que o coeficiente se encontra, e o segundo índice refere-se à incógnita à qual o coeficiente está associado.

Assim, por exemplo, um sistema com duas equações e três incógnitas escrever-se-á na forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \end{cases}$$

ao passo que um sistema de quatro equações com quatro incógnitas será

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 = b_3 \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4 = b_4 \end{cases}$$

Na prática e a fim de simplificar a escrita, muitas vezes se designam as incógnitas (especialmente se são poucas) por letras distintas, x, y, z, etc., podendo fazer-se o mesmo em relação aos coeficientes daquelas e aos termos independentes.

De modo geral, um sistema com p equações e n incógnitas terá a forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n = b_p \end{cases}$$

É fácil verificar, através de exemplos muito simples, que nem todos os sistemas de equações têm o mesmo comportamento, do ponto de vista da existência e do número de soluções.

Exemplo 1-1

Consideremos o sistema de equações

$$\begin{cases} x - y = 0 \\ x + y = 8 \end{cases}$$

A primeira equação só será satisfeita por um par de números iguais; mas a segunda obriga que a soma desses dois números dê 8. Consequentemente, a única solução do sistema é o par (4,4). ■

Exemplo 1-2

Consideremos o sistema de equações

$$\begin{cases} x + y = 3 \\ 2x + 2y = 7 \end{cases}$$

É fácil verificar que este sistema não tem qualquer solução. Com efeito, se (r_1, r_2) fosse uma solução, haveria de verificar a primeira equação, tendo-se, portanto, $r_1 + r_2 = 3$; mas, então, $2r_1 + 2r_2 = 2(r_1 + r_2) = 2(r_1 + r_2) = 2 \times 3 = 6 \neq 7$, pelo que uma solução da primeira equação nunca verificaria a segunda. ■

Exemplo 1-3

Consideremos o sistema de equações

$$\begin{cases} x + y = 3 \\ 2x + 2y = 6 \end{cases}$$

Vemos imediatamente que os pares (2,1), (0,3), (−1,4), etc., satisfazem ambas as equações. Na verdade, qualquer par de números (r_1, r_2) tal que $r_1 + r_2 = 3$ verifica ambas as equações, já que $2r_1 + 2r_2 = 2(r_1 + r_2) = 2 \times 3 = 6$. ■

Os exemplos anteriores levam-nos a classificar os sistemas de equações lineares da seguinte maneira:

Definição 1-4 Um sistema de equações lineares diz-se:

a) **Possível** quando tem solução.

  a1) **Determinado** quando tem uma única solução (como no Exemplo 1-1).

  a2) **Indeterminado** quando tem mais de uma solução (como no Exemplo 1-3).

b) **Impossível** quando não tem solução (como no Exemplo 1-2).

Dado um sistema de equações lineares, levanta-se, naturalmente, a questão de saber se trata de um sistema possível ou impossível, e, no primeiro caso, se é determinado ou indeterminado. Por sua vez, caso o sistema seja possível, poderemos querer conhecer as respectivas soluções. Veremos, já de seguida, como se pode dar resposta a estas questões.

1-2 O método de Gauss

Naturalmente, a existência e o número de soluções de um dado sistema de equações nem sempre se reconhecem tão facilmente como nos exemplos anteriores.

Assim, por exemplo, dado o sistema de equações

$$\begin{cases} x+3y=1 \\ 2x-y=7 \end{cases}$$

não é fácil saber se ele é possível ou impossível, e, no primeiro caso, se é determinado ou indeterminado. Mas se conseguíssemos transformá-lo num sistema tão simples como os dos exemplos anteriores, já seríamos capazes de chegar a conclusões. Para seguirmos essa ordem de ideias, introduzimos a seguinte definição:

Definição 1-5 Dois sistemas de equações, com o mesmo número de incógnitas (mas não necessariamente com o mesmo número de equações), dizem-se **equivalentes** quando têm precisamente as mesmas soluções.

O problema que se nos apresenta é, portanto, o de passar de um sistema dado para outro, equivalente, mais simples que o primeiro. Ora, para o efeito, há um certo número de transformações admissíveis, isto é, de transformações que levam, de facto, de um sistema de equações a outro, equivalente ao primeiro, a saber:

Teorema 1-6 **Princípios de equivalência de sistemas de equações lineares**

Dado um sistema de equações lineares, obtém-se um sistema equivalente ao dado quando:

a) Se troca a ordem às equações;
b) Se multiplicam ambos os membros de uma dada equação por uma constante diferente de zero;
c) A uma equação se soma, membro a membro, uma outra, eventualmente multiplicada por uma constante qualquer;
d) Se troca a ordem às incógnitas.

Demonstração:

Atendendo às definições, as afirmações a) e d) são evidentes, já que a ordem — quer das equações, quer das incógnitas — não interfere na existência ou não de soluções [tendo em conta, na afirmação d), que a adição goza da propriedade comutativa]. Analisemos, então, as duas restantes afirmações:

b) Suponhamos que, num sistema de equações, trocamos uma equação

$$a_1x_1+a_2x_2+\cdots+a_nx_n=b$$

pela equação

$$ca_1x_1 + ca_2x_2 + \cdots + ca_nx_n = cb$$

que resulta da primeira ao multiplicarmos ambos os membros pela constante $c \neq 0$

Se $(r_1, r_2, ..., r_n)$ é uma solução da primeira equação, teremos

$$a_1r_1 + a_2r_2 + \cdots + a_nr_n = b$$

donde

$$ca_1r_1 + ca_2r_2 + \cdots + ca_nr_n = c(a_1r_1 + a_2r_2 + \cdots + a_nr_n) = cb$$

e, portanto, $(r_1, r_2, ..., r_n)$ é uma solução da segunda equação; reciprocamente, se $(r_1, r_2, ..., r_n)$ é uma solução da segunda equação, teremos

$$ca_1r_1 + ca_2r_2 + \cdots + ca_nr_n = cb$$

pelo que, tendo em conta que $c \neq 0$,

$$a_1r_1 + a_nr_2 + \cdots + a_nr_n = \frac{1}{c}(ca_1r_1 + ca_2r_2 + \cdots + ca_nr_n) = \frac{1}{c}(cb) = b$$

e $(r_1, r_2, ..., r_n)$ é uma solução da primeira equação.

Uma vez que a modificação efectuada não alterou as restantes equações do sistema dado, este conserva-se equivalente ao sistema obtido.

c) Consideremos, agora, num sistema de equações, as equações

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$
$$a_1'x_1 + a_2'x_2 + \cdots + a_n'x_n = b'$$

e vamos substituí-las pelo par

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$
$$(ca_1 + a_1')x_1 + (ca_2 + a_2')x_2 + \cdots + (ca_n + a_n')x_n = cb + b'$$

onde, à segunda equação, somámos, membro a membro, a primeira, multiplicada pela constante c; se for $c = 1$, a transformação efectuada consiste tão-somente em somar, membro a membro, as duas equações, e, obviamente, se for $c = 0$ o segundo par de equações coincide com o primeiro.

Seja $(r_1, r_2, ..., r_n)$ uma solução do primeiro par de equações. Então, de

$$a_1r_1 + a_2r_2 + \cdots + a_nr_n = b$$
$$a_1'r_1 + a_2'r_2 + \cdots + a_n'r_n = b'$$

resulta

$$\begin{aligned}(ca_1 + a_1')r_1 + (ca_2 + a_2')r_2 + \cdots + (ca_n + a_n')r_n &= \\ &= ca_1r_1 + a_1'r_1 + ca_2r_2 + a_2'r_2 + \cdots + ca_nr_n + a_n'r_n = \\ &= c(a_1r_1 + a_2r_2 + \cdots + a_nr_n) + (a_1'r_1 + a_2'r_2 + \cdots + a_n'r_n) = cb + b'\end{aligned}$$

Concluímos acima que qualquer solução do primeiro par de equações satisfaz também o segundo. A recíproca é evidente, já que o primeiro par de equações se obtém a partir do segundo precisamente da mesma maneira que o segundo se obtém a partir do primeiro: bastaria multiplicar ambos os membros da primeira equação por $-c$ e somá-los aos da segunda.

Estes princípios de equivalência podem, então, ser usados para transformar um dado sistema de equações noutro, equivalente, eventualmente mais simples. Por exemplo, pegando no sistema de equações considerado acima, podemos transformá-lo, sucessivamente, da seguinte maneira:

$$\begin{cases} x + 3y = 1 \\ 2x - y = 7 \end{cases} \xrightarrow[\text{c}]{} \begin{cases} x + 3y = 1 \\ -7y = 5 \end{cases} \xrightarrow[\text{b}]{} \begin{cases} x + 3y = 1 \\ y = -\dfrac{5}{7} \end{cases}$$

No primeiro passo, multiplicámos ambos os membros da primeira equação por -2 e somámo-la, membro a membro, à segunda; no segundo passo, multiplicámos ambos os membros da segunda equação por $-1/7$. O sistema obtido é mais fácil de resolver que o primeiro; por exemplo, vemos imediatamente que a segunda variável tem de tomar o valor $-5/7$, e, substituindo este valor na primeira equação, encontramos $x = 22/7$. Concluímos, assim, que o sistema dado é possível e determinado.

Ao aplicarmos as transformações acima descritas, verificamos que as operações [ou trocas de posição, se usarmos os princípios de equivalência a) e d)] efectuadas afectam exclusivamente os coeficientes e os termos independentes das diversas equações. A fim de se simplificar a apresentação dos cálculos, é costume dispor esses coeficientes num quadro; por exemplo, no sistema de equações que temos vindo a considerar, teríamos o quadro

$$\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 2 & -1 & 7 \end{array}\right]$$

Estes quadros são designados por **matrizes**. As transformações efectuadas podem realizar-se a partir destes quadros de coeficientes:

$$\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 2 & -1 & 7 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & -7 & 5 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & 1 & -5/7 \end{array}\right]$$

Definição 1-7 Dado o sistema de equações lineares

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n = b_p \end{cases}$$

o quadro

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}$$

toma o nome de **matriz simples** do sistema dado, ao passo que o quadro

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} & b_p \end{array}\right]$$

é a **matriz completa** ou **matriz ampliada** do sistema. Por construção, os elementos de uma matriz dispõem-se em filas horizontais, a que chamamos as **linhas** da matriz, e, simultaneamente, em filas verticais, que são as **colunas**. O elemento a_{ij} encontra-se na linha número i e na coluna número j.

Deve observar-se que cada linha da matriz completa corresponde a uma das equações do sistema, enquanto cada coluna da matriz simples é constituída pelos coeficientes de uma das incógnitas do sistema, nas sucessivas equações.

Usando esta nova linguagem, os princípios de equivalência de sistemas de equações lineares podem enunciar-se do seguinte modo:

Teorema 1-8 **Princípios de equivalência de sistemas de equações lineares**
Dado um sistema de equações lineares, obtém-se um sistema equivalente ao dado quando:

a) Se troca a ordem às linhas da sua matriz completa;
b) Se multiplica uma linha da matriz completa por uma constante diferente de zero;
c) A uma linha da matriz completa se soma uma outra, eventualmente multiplicada por uma constante qualquer;
d) Se troca a ordem às colunas da matriz simples.

Resta agora provar que a utilização dos quatro princípios de equivalência apresentados permite, de facto, partir de um sistema de equações qualquer e chegar a um outro, equivalente ao primeiro, mas suficientemente simples para que as soluções possam ser encontradas.

Teorema 1-9 Através da utilização dos princípios de equivalência indicados no Teorema 1-8, é sempre possível chegar a um sistema cuja matriz completa tenha a forma

$$\left[\begin{array}{cccccccccc|c}
1 & 0 & 0 & \cdots & 0 & a'_{1,r+1} & a'_{1,r+2} & \cdots & a'_{1n} & & b'_1 \\
0 & 1 & 0 & \cdots & 0 & a'_{2,r+1} & a'_{2,r+2} & \cdots & a'_{2n} & & b'_2 \\
0 & 0 & 1 & \cdots & 0 & a'_{3,r+1} & a'_{3,r+2} & \cdots & a'_{3n} & & b'_3 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & 1 & a'_{r,r+1} & a'_{r,r+2} & \cdots & a'_{rn} & & b'_r \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & & b'_{r+1} \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & & b'_{r+2} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & & b'_p
\end{array}\right]$$

onde $0 \leq r \leq \min(p, n)$.

Demonstração:

O caso de a matriz simples do sistema ser formada toda por zeros é trivial, visto a matriz completa ser já da forma desejada, com $r = 0$. Podemos, portanto, supor que haja elementos não nulos na matriz simples, não havendo perda de generalidade em se admitir que $a_{11} \neq 0$, já que, por meio de eventuais trocas de linhas e de colunas, poderíamos sempre colocar nessa posição um elemento diferente de zero.

Multiplicando então a primeira linha por constantes adequadamente escolhidas e somando com as restantes, podemos chegar a uma matriz da forma

$$\left[\begin{array}{cccc|c}
a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\
0 & \breve{a}_{22} & \cdots & \breve{a}_{2n} & \breve{b}_2 \\
\vdots & \vdots & & \vdots & \vdots \\
0 & \breve{a}_{p2} & \cdots & \breve{a}_{pn} & \breve{b}_p
\end{array}\right]$$

Seguidamente, podemos multiplicar a primeira linha por $1/_{a_{11}}$ e chegamos a

$$\left[\begin{array}{cccc|c} 1 & \breve{a}_{12} & \cdots & \breve{a}_{1n} & \breve{b}_1 \\ 0 & \breve{a}_{22} & \cdots & \breve{a}_{2n} & \breve{b}_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \breve{a}_{p2} & \cdots & \breve{a}_{pn} & \breve{b}_p \end{array}\right]$$

Caso os elementos $\breve{a}_{ij} (i = 2,\ldots,p,\ j = 2,\ldots,n)$ sejam todos nulos, o processo está terminado, tendo-se obtido a matriz desejada. Caso contrário, podemos supor que seja $\breve{a}_{22} \neq 0$ e, com a ajuda desse elemento, anular todos os outros que se encontrem na mesma coluna. Multiplicando depois a linha número 2 por $1/_{\breve{a}_{22}}$, obtemos

$$\left[\begin{array}{ccccc|c} 1 & 0 & \breve{\breve{a}}_{13} & \cdots & \breve{\breve{a}}_{1n} & \breve{\breve{b}}_1 \\ 0 & 1 & \breve{\breve{a}}_{23} & \cdots & \breve{\breve{a}}_{2n} & \breve{\breve{b}}_2 \\ 0 & 0 & \breve{\breve{a}}_{33} & \cdots & \breve{\breve{a}}_{3n} & \breve{\breve{b}}_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \breve{\breve{a}}_{p3} & \cdots & \breve{\breve{a}}_{pn} & \breve{\breve{b}}_p \end{array}\right]$$

O método exposto é então iterado, até se obter a matriz da forma pretendida.

O processo descrito na demonstração é chamado o **método de condensação** (ou **de eliminação**) **de Gauss**. A matriz final diz-se **condensada**.

Corolário 1-10 O sistema de equações representado pela matriz indicada no Teorema 1-9 é:

a) Possível determinado se e só se $b'_{r+1} = b'_{r+2} = \cdots = b'_p = 0$ e $r = n$;
b) Possível indeterminado se e só se $b'_{r+1} = b'_{r+2} = \cdots = b'_p = 0$ e $r < n$;
c) Impossível se e só se $r < n$ e $b'_k \neq 0$, para algum $k > r$.

Demonstração:
Resulta imediatamente do aspecto da matriz condensada. Deve notar-se que, no caso de ser $b'_{r+1} = b'_{r+2} = \cdots = b'_p = 0$ e $r < n$, a matriz condensada tem a forma

$$\left[\begin{array}{ccccccccc|c}
1 & 0 & 0 & \cdots & 0 & a'_{1,r+1} & a'_{1,r+2} & \cdots & a'_{1n} & b'_1 \\
0 & 1 & 0 & \cdots & 0 & a'_{2,r+1} & a'_{2,r+2} & \cdots & a'_{2n} & b'_2 \\
0 & 0 & 1 & \cdots & 0 & a'_{3,r+1} & a'_{3,r+2} & \cdots & a'_{3n} & b'_3 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & a'_{r,r+1} & a'_{r,r+2} & \cdots & a'_{rn} & b'_r \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0
\end{array}\right]$$

e o sistema de equações é equivalente ao que é representado pela matriz

$$\left[\begin{array}{ccccccccc|c}
1 & 0 & 0 & \cdots & 0 & a'_{1,r+1} & a'_{1,r+2} & \cdots & a'_{1n} & b'_1 \\
0 & 1 & 0 & \cdots & 0 & a'_{2,r+1} & a'_{2,r+2} & \cdots & a'_{2n} & b'_2 \\
0 & 0 & 1 & \cdots & 0 & a'_{3,r+1} & a'_{3,r+2} & \cdots & a'_{3n} & b'_3 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & a'_{r,r+1} & a'_{r,r+2} & \cdots & a'_{rn} & b'_r
\end{array}\right]$$

que é o seguinte, supondo-se que, após eventuais trocas de colunas efectuadas, as incógnitas se apresentam pela ordem $y_1, y_2, \ldots, y_n$:

$$\begin{cases}
y_1 + a'_{1,r+1} y_{r+1} + a'_{1,r+2} y_{r+2} + \cdots + a'_{1n} y_n = b'_1 \\
y_2 + a'_{2,r+1} y_{r+1} + a'_{2,r+2} y_{r+2} + \cdots + a'_{2n} y_n = b'_2 \\
\vdots \\
y_r + a'_{r,r+1} y_{r+1} + a'_{r,r+2} y_{r+2} + \cdots + a'_{rn} y_n = b'_r
\end{cases}$$

A solução deste sistema é dada por

$$\begin{cases} y_1 = b_1' - a_{1,r+1}' y_{r+1} - a_{1,r+2}' y_{r+2} - \cdots - a_{1n}' y_n \\ y_2 = b_2' - a_{2,r+1}' y_{r+1} - a_{2,r+2}' y_{r+2} - \cdots - a_{2n}' y_n \\ \vdots \\ y_r = b_r' - a_{r,r+1}' y_{r+1} - a_{r,r+2}' y_{r+2} - \cdots - a_{rn}' y_n \end{cases}$$

sendo $y_{r+1}, y_{r+2}, \ldots, y_n$ arbitrárias.

Por exemplo, consideremos o sistema de equações

$$\begin{cases} x + 2y - 3z + w = 1 \\ -x + z - 5w = 2 \\ 2x + 3y - z + w = 0 \end{cases}$$

A sua matriz completa é

$$\left[\begin{array}{cccc|c} 1 & 2 & -3 & 1 & 1 \\ -1 & 0 & 1 & -5 & 2 \\ 2 & 3 & -1 & 1 & 0 \end{array}\right]$$

Aplicando o método de Gauss, passamos, usando o elemento 1, que se encontra na primeira linha e na primeira coluna, como pivô, para

$$\left[\begin{array}{cccc|c} 1 & 2 & -3 & 1 & 1 \\ -1 & 0 & 1 & -5 & 2 \\ 2 & 3 & -1 & 1 & 0 \end{array}\right] \xrightarrow[L_3 - 2L_1]{L_2 + L_1} \left[\begin{array}{cccc|c} 1 & 2 & -3 & 1 & 1 \\ 0 & 2 & -2 & -4 & 3 \\ 0 & -1 & 5 & -1 & -2 \end{array}\right]$$

A notação usada é sugestiva: "$L_k + cL_t$" significa que à linha L_k somamos a linha L_t, multiplicada pela constante c.

Podemos, depois, trocar a ordem às linhas e multiplicar uma delas por -1 e outra por $1/2$, passando para

$$\left[\begin{array}{cccc|c} 1 & 2 & -3 & 1 & 1 \\ 0 & 2 & -2 & -4 & 3 \\ 0 & -1 & 5 & -1 & -2 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 2 & -3 & 1 & 1 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 1 & -1 & -2 & 3/2 \end{array}\right]$$

Usamos agora o elemento 1, localizado na segunda linha e na segunda coluna, como pivô, obtendo

$$\left[\begin{array}{cccc|c} 1 & 2 & -3 & 1 & 1 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 1 & -1 & -2 & 3/2 \end{array}\right] \xrightarrow[L_3-L_1]{L_1-2L_2} \left[\begin{array}{cccc|c} 1 & 0 & 7 & -1 & -3 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 0 & 4 & -3 & -1/2 \end{array}\right]$$

Finalmente, multiplicamos a terceira linha por $1/4$ e usamos o elemento 1 resultante como pivô, para anular os outros elementos da terceira coluna:

$$\left[\begin{array}{cccc|c} 1 & 0 & 7 & -1 & -3 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 0 & 4 & -3 & -1/2 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 0 & 7 & -1 & -3 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 0 & 1 & -3/4 & -1/8 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 0 & 0 & 17/4 & -17/8 \\ 0 & 1 & 0 & -11/4 & 11/8 \\ 0 & 0 & 1 & -3/4 & -1/8 \end{array}\right]$$

O sistema de equações correspondente a esta última matriz (que é, não esqueçamos, equivalente ao sistema dado) é

$$\begin{cases} x+\dfrac{17}{4}w=-\dfrac{17}{8} \\ y-\dfrac{11}{4}w=\dfrac{11}{8} \\ z-\dfrac{3}{4}w=-\dfrac{1}{8} \end{cases}$$

A solução deste sistema de equações é

$$\begin{cases} x=-\dfrac{17}{8}-\dfrac{17}{4}w \\ y=\dfrac{11}{8}+\dfrac{11}{4}w \\ z=-\dfrac{1}{8}+\dfrac{3}{4}w \end{cases}$$

ficando arbitrário o valor de w. Trata-se, portanto, de um sistema possível, mas indeterminado.

1-3 Determinantes

Consideremos um sistema de equações com o mesmo número de equações e de incógnitas, dito também do "tipo $n \times n$" (dizemos, neste caso, que a respectiva matriz simples é "quadrada"), por exemplo, um sistema com duas equações e duas incógnitas

$$\begin{cases} ax + by = c \\ a'x + b'y = c' \end{cases}$$

Supondo que se tem $a \neq 0$, usemos o método de Gauss para, a partir da matriz completa do sistema dado, obter, sucessivamente, as seguintes:

$$\left[\begin{array}{cc|c} a & b & c \\ a' & b' & c' \end{array}\right] \rightarrow \left[\begin{array}{cc|c} a & b & c \\ 0 & b' - \dfrac{a'b}{a} & c' - \dfrac{a'c}{a} \end{array}\right] =$$

$$= \left[\begin{array}{cc|c} a & b & c \\ 0 & \dfrac{ab' - a'b}{a} & \dfrac{ac' - a'c}{a} \end{array}\right] \rightarrow \left[\begin{array}{cc|c} a & b & c \\ 0 & ab' - a'b & ac' - a'c \end{array}\right]$$

Atendendo a que supusemos $a \neq 0$, a natureza do sistema considerado dependerá do valor das expressões $ab' - a'b$ e $ac' - a'c$. Com efeito, a segunda equação aparece com a forma

$$(ab' - a'b)y = ac' - a'c$$

Observe-se que, se for $ab' - a'b = 0$, $ac' - a'c \neq 0$, o sistema será impossível; no caso de ser $ab' - a'b = 0$, $ac' - a'c = 0$, o valor de y, nesta igualdade, seria arbitrário; finalmente, se for $ab' - a'b \neq 0$, encontraremos o valor fixo

$$y = \frac{ac' - a'c}{ab' - a'b}$$

O número $ab' - a'b$, cujo valor vai, de certo modo, condicionar ou determinar o valor de y, chamar-se-á "determinante".

Analogamente, se partirmos de um sistema de três equações e três incógnitas, digamos

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 \end{cases}$$

podemos, através de um procedimento análogo, chegar facilmente, por exemplo, a uma igualdade da forma

$$\eta z = \zeta$$

com

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

Se tivermos $\eta = 0,\ \zeta \neq 0$, o sistema de equações considerado é impossível; sendo $\eta = \zeta = 0$, o valor de z fica indeterminado, na igualdade obtida; se $\eta \neq 0$, teremos $z = \zeta / \eta$.

Repare-se que a expressão de η é uma soma de 6 parcelas (3 afectadas do sinal $+$, outras 3 do sinal $-$) e que cada parcela é um produto de 3 factores; entre esses factores, há sempre um de cada linha e um de cada coluna da matriz simples do sistema dado.

Para sistematizarmos o assunto, introduzindo uma importante definição, necessitamos previamente de recordar algumas noções do foro do Cálculo Combinatório:

Dados os números naturais 1, 2, ..., n, podemos escrevê-los por uma ordem qualquer, obtendo o que se chama uma "permutação" desses números: $i_1, i_2, ..., i_n$.

De cada vez que se tenha, nessa permutação, $i_k > i_{k+t}$, dizemos que os números i_k e i_{k+t} formam uma "inversão". A permutação $i_1, i_2, ..., i_n$ diz-se "par" ou "ímpar" consoante o número total de inversões que nela figuram é par ou ímpar.

Assim, por exemplo, tomemos os números 1, 2, 3, 4, 5, 6, 7, 8, 9.

Na permutação 7, 2, 8, 4, 5, 9, 1, 6, 3, encontramos as inversões 9-1, 9-6, 9-3, 8-4, 8-5, 8-1, 8-6, 8-3, 7-2, 7-4, 7-5, 7-1, 7-6, 7-3, 6-3, 5-1, 5-3, 4-1, 4-3, 2-1, num total de 20, pelo que a permutação considerada é par.

Podemos agora apresentar a definição seguinte:

Definição 1-11 Dada a matriz quadrada

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

formada por números reais, chama-se **determinante** de A, e representa-se por

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

o número obtido do seguinte modo:

a) Formam-se todos os possíveis produtos de n elementos, usando um factor de cada linha e, simultaneamente, um factor de cada coluna de A;
b) Afecta-se cada produto do sinal + ou do sinal −, consoante as permutações formadas pelos primeiros índices (que correspondem às linhas) e pelos segundos índices (que correspondem às colunas) dos factores usados tenham a mesma paridade ou não;
c) Somam-se as $n!$ parcelas obtidas.

Simbolicamente, podemos escrever:

$$|A| = \sum (-1)^{\sigma} a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_n j_n}$$

onde $i_1, i_2, \ldots, i_n$ e $j_1, j_2, \ldots, j_n$ percorrem todas as possíveis permutações dos números 1, 2, …, n e

$$\sigma = \begin{cases} 0, & \text{se } i_1, i_2, \ldots, i_n \text{ e } j_1, j_2, \ldots, j_n \text{ têm a mesma paridade} \\ 1, & \text{se } i_1, i_2, \ldots, i_n \text{ e } j_1, j_2, \ldots, j_n \text{ têm paridades diferentes} \end{cases}$$

Uma vez que a ordem dos factores, em cada parcela, é arbitrária, podemos ordená-los, por exemplo, por ordem crescente de um dos índices, obtendo

$$|A| = \sum (-1)^{\sigma} a_{1 j_1} a_{2 j_2} \cdots a_{n j_n}$$

onde $j_1, j_2, \ldots, j_n$ percorre todas as possíveis permutações dos números 1, 2, …, n e

$$\sigma = \begin{cases} 0, & \text{se } j_1, j_2, \ldots, j_n \text{ é par} \\ 1, & \text{se } j_1, j_2, \ldots, j_n \text{ é ímpar} \end{cases}$$

ou

$$|A| = \sum (-1)^{\sigma} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}$$

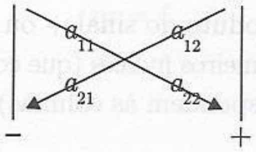
onde $i_1, i_2, \ldots, i_n$ percorre todas as possíveis permutações dos números 1, 2, …, n e

$$\sigma = \begin{cases} 0, & \text{se } i_1, i_2, \ldots, i_n \text{ é par} \\ 1, & \text{se } i_1, i_2, \ldots, i_n \text{ é ímpar} \end{cases}$$

Note-se que, no caso de uma matriz do tipo 2×2, o determinante é simplesmente dado por

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

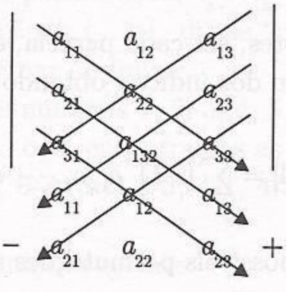
e as duas parcelas desta expressão são obtidas efectuando os dois produtos que se sugerem no esquema seguinte:



Já no caso de uma matriz do tipo 3×3, o determinante é dado por esta expressão:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

Esta expressão pode ser obtida, na prática, por uma regra mnemónica, conhecida por **regra de Sarrus**, que se pode enunciar de diferentes maneiras, mas que apresentaremos como se segue, naquele que é talvez o modo mais fácil de compreender:



O problema de se calcular o determinante de uma matriz maior complica-se rapidamente, já que, para uma matriz do tipo $n\times n$, envolve o uso de uma expressão com $n!$ parcelas. Ora, o factorial aumenta rapidamente com n, o que torna quase impraticável a utilização de tais expressões:

n	$n!$
2	2
3	6
4	24
5	120
6	720
7	5 040
8	40 320
9	368 280
10	3 628 800
…	…

Estudando algumas propriedades dos determinantes, que, naturalmente, resultam da definição dada, será possível, no entanto, encontrar maneiras mais eficientes para calcular os seus valores. Vejamos, então, algumas dessas propriedades:

Teorema 1-12 Seja A uma matriz qualquer, do tipo $n \times n$.

a) Se multiplicarmos uma linha (ou uma coluna) de A por uma constante α, o valor do determinante da matriz resultante obtém-se a partir do determinante de A, multiplicando-o também por α.
b) Se A tem uma linha (ou uma coluna) toda formada por zeros, então $|A| = 0$.
c) Se trocarmos entre si as posições de duas linhas (ou de duas colunas) de A, o valor do determinante da matriz passa ao simétrico.
d) Se A tem duas linhas (ou duas colunas) iguais, então $|A| = 0$.
e) Se A tem duas linhas (ou duas colunas) proporcionais, então $|A| = 0$.
f) Se uma linha (resp.: coluna) de A se pode considerar a soma de duas linhas (resp.: duas colunas) quaisquer, o determinante de A é igual à soma dos determinantes de duas matrizes em que, nessa linha (resp.: coluna) se usa uma parcela de cada vez, mantendo-se as demais.
g) O valor do determinante de A não se altera se a uma linha (resp.: coluna) somarmos uma outra, eventualmente multiplicada por uma constante qualquer.

Demonstração:

a) Seja A' a matriz que resulta de A se multiplicarmos uma das linhas (podemos pensar, concretamente, na linha número 1, já que o raciocínio é análogo para qualquer das linhas ou das colunas) pela constante α. Os elementos da primeira linha de A' terão a forma αa_{1j} $(j = 1, 2, \ldots, n)$. Então:

$$\begin{aligned}|A'| &= \sum (-1)^{\sigma} (\alpha a_{1j_1}) a_{2j_2} \cdots a_{nj_n} = \sum (-1)^{\sigma} \alpha a_{1j_1} a_{2j_2} \cdots a_{nj_n} = \\ &= \alpha \sum (-1)^{\sigma} a_{1j_1} a_{2j_2} \cdots a_{nj_n} = \alpha |A|\end{aligned}$$

b) Resulta imediatamente de a).
c) Dada uma permutação

$$i_1 i_2 i_3 \cdots i_{k-1} i_k i_{k+1} \cdots i_{t-1} i_t i_{t+1} \cdots i_{n-1} i_n$$

dos números $1, 2, 3, \ldots, n$, consideremos a permutação

$$i_1 i_2 i_3 \cdots i_{k-1} i_t i_{k+1} \cdots i_{t-1} i_k i_{t+1} \cdots i_{n-1} i_n$$

que resulta da primeira trocando os índices k e t.

As inversões que os números i_k e i_t formam com os números $i_1 i_2 i_3 \cdots i_{k-1}$ e $i_{t+1} \cdots i_{n-1} i_n$ são as mesmas em ambas as permutações, visto as respectivas posições relativas não se terem alterado.

O mesmo se pode dizer das inversões que os números $i_1 i_2 i_3 \cdots i_{k-1}$, os números $i_{t+1} \cdots i_{n-1} i_n$ e os números $i_{k+1} \cdots i_{t-1}$ formem entre si e uns com os outros.

Se o número i_k fizer, digamos, r inversões com os números $i_{k+1} \cdots i_{t-1}$, na primeira permutação, não fará inversões com os restantes desses números, que são $t-1-k-r$; ao mesmo tempo, o número i_t fará s inversões com os números $i_{k+1} \cdots i_{t-1}$, na primeira permutação, não fazendo inversões com os restantes desses números, que são $t-1-k-s$. Na segunda permutação, cada um dos números i_k e i_t passará a formar inversões com aqueles dos números $i_{k+1} \cdots i_{t-1}$ com os quais não fazia inversão, deixando de formar inversões com os demais. O número total de inversões dos números i_k e i_t com os números $i_{k+1} \cdots i_{t-1}$ na segunda permutação, será, portanto, $(t-1-k-r)+$ $+(t-1-k-s)=2t-2-2k-r-s=2(t-1-k)-(r+s)$, que tem a mesma paridade que o número total de inversões dos números i_k e i_t com os números $i_{k+1} \cdots i_{t-1}$ na primeira permutação, que era $r+s$.

Finalmente, se os números i_k e i_t formam uma inversão na primeira permutação, deixam de a formar na segunda, e vice-versa.

Consequentemente, ao trocarmos entre si as posições de dois índices, numa permutação arbitrária, esta muda de paridade. A tese resulta então imediatamente da definição de determinante.

d) Resulta imediatamente de c).

e) Resulta imediatamente de a) e d).

f) Suponhamos, para fixar ideias, que a primeira linha de A se pode considerar a soma de duas certas linhas (o raciocínio é análogo para qualquer outra linha ou para uma das colunas): teremos então

$$a_{1j} = a'_{1j} + a''_{1j} \quad (j = 1, 2, \ldots, n)$$

Consequentemente,

$$\begin{aligned} |A| = \sum (-1)^{\sigma} a_{1j_1} a_{2j_2} \cdots a_{nj_n} &= \sum (-1)^{\sigma} (a'_{1j_1} + a''_{1j_1}) a_{2j_2} \cdots a_{nj_n} = \\ &= \sum \left((-1)^{\sigma} a'_{1j_1} a_{2j_2} \cdots a_{nj_n} + (-1)^{\sigma} a''_{1j_1} a_{2j_2} \cdots a_{nj_n} \right) = \\ &= \sum (-1)^{\sigma} a''_{1j_1} a_{2j_2} \cdots a_{nj_n} + \sum (-1)^{\sigma} a''_{1j_1} a_{2j_2} \cdots a_{nj_n} \end{aligned}$$

Se compararmos cada parcela desta última soma com a expressão do determinante da matriz A, verificamos que são precisamente os determinantes das matrizes que resultam de A substituindo a primeira linha, respectivamente, pelas linhas $a'_{11} \quad a'_{12} \quad \cdots \quad a'_{1n}$ e $a''_{11} \quad a''_{12} \quad \cdots \quad a''_{1n}$.

g) Representemos, abreviadamente, a matriz A por

$$\begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{bmatrix}$$

onde L_j $(j = 1, 2, ..., n)$ designa as várias linhas da matriz.

Usando as propriedades anteriores, temos, então,

$$\begin{vmatrix} \vdots \\ L_k \\ \vdots \\ L_t + cL_k \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ L_k \\ \vdots \\ L_t \\ \vdots \end{vmatrix} + \begin{vmatrix} \vdots \\ L_k \\ \vdots \\ cL_k \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ L_k \\ \vdots \\ L_t \\ \vdots \end{vmatrix} + 0 = |A| + 0 = |A|$$

As propriedades a), c) e g) do teorema anterior descrevem o efeito, no valor do determinante de uma dada matriz, da realização de certas transformações, a saber: multiplicar linhas (ou colunas) por constantes, trocar entre si linhas (ou colunas) e somar linhas (ou colunas) com outras linhas (ou outras colunas), eventualmente multiplicando-as por constantes.

Ora, estas transformações são análogas às que encontrámos anteriormente, na aplicação do método de Gauss à resolução de um sistema de equações lineares. Mais ainda, nas propriedades anteriores, tudo o que se diz para linhas diz-se também para colunas, e reciprocamente. Este facto resulta de, na definição do determinante de uma matriz, as linhas e as colunas desempenharem precisamente o mesmo papel. Podemos mesmo salientar esta situação mediante o seguinte conceito:

Definição 1-13 Dada a matriz $A = [a_{ij}]$, com $i = 1, 2, ..., p$, $j = 1, 2, ..., n$, chama-se **matriz transposta** de A e representa-se por A^T, à matriz que resulta de escrever as linhas de A como colunas e vice-versa, ou seja, a matriz $A^T [a_{ij}]$, do tipo $n \times p$.

Por exemplo, sendo

$$A = \begin{bmatrix} 3 & 1 & 8 & 6 \\ -2 & 5 & 0 & -3 \\ 8 & 9 & 1 & 7 \end{bmatrix}, \text{ teremos } A^T = \begin{bmatrix} 3 & -2 & 8 \\ 1 & 5 & 9 \\ 8 & 0 & 1 \\ 6 & -3 & 7 \end{bmatrix}$$

Teorema 1-14 Sendo A uma matriz quadrada, tem-se $|A| = |A^T|$.

Demonstração:

Resulta imediatamente da definição de determinante. ■

Voltando, porém, ao problema do cálculo do determinante de uma matriz quadrada qualquer, consideremos um caso particular:

Definição 1-15 Chama-se **matriz triangular (superior)** a uma matriz quadrada em que se tem $a_{ij} = 0$ sempre que $i > j$, ou seja, uma matriz da forma

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Os determinantes de matrizes triangulares são particularmente fáceis de calcular, em virtude do seguinte:

Teorema 1-16 O determinante de uma matriz triangular é igual ao produto dos elementos $a_{ii} (i = 1,2,...,n)$. Estes elementos são chamados os **elementos principais** da matriz (e formam a respectiva **diagonal principal**).

Demonstração:
Consideremos a matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Se aplicarmos a definição de determinante,

$$|A| = \sum (-1)^\sigma a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}$$

vemos imediatamente que são nulas todas as parcelas em que se tenha $i_1 > 1$, pelo que, na soma, basta considerar as parcelas em que $i_1 = 1$:

$$|A| = \sum (-1)^\sigma a_{11} a_{i_2 2} \cdots a_{i_n n} = a_{11} \sum (-1)^\sigma a_{i_2 2} \cdots a_{i_n n}$$

No que se refere aos elementos da coluna número 2, o índice i_2 não pode tomar o valor 1, porque temos já um elemento (que é a_{11}) da primeira linha, e, sempre que se tem $i_2 > 2$, a respectiva parcela anula-se, pelo que basta considerar as parcelas em que $i_2 = 2$:

$$|A| = a_{11}\sum(-1)^{\sigma} a_{22}\cdots a_{i_n n} = a_{11}\sum(-1)^{\sigma} a_{i_2 2}\cdots a_{i_n n}$$

Prosseguindo este raciocínio, verificamos que todas as parcelas da expressão de $|A|$ se anulam, com a possível excepção de $a_{11}a_{22}\cdots a_{nn}$. Esta parcela fica afectada do sinal $+$, visto que nas linhas e nas colunas figura a mesma permutação dos índices, pelo que, como pretendíamos, $|A| = a_{11}a_{22}\cdots a_{nn}$.

Com o auxílio do teorema anterior e do que vimos anteriormente, podemos calcular, com relativa facilidade, o determinante de uma matriz quadrada: basta usar o método de Gauss para transformar a matriz dada numa matriz triangular, tomando a devida nota das modificações que as transformações necessárias acarretem ao valor do seu determinante.

Exemplo 1-17 Vamos calcular o determinante da matriz

$$A = \begin{bmatrix} 1 & -2 & 2 & 1 \\ -1 & 2 & 1 & 3 \\ 1 & -4 & 2 & 3 \\ 1 & 1 & 3 & 2 \end{bmatrix}$$

Usando o processo descrito acima, temos:

$$\begin{vmatrix} 1 & -2 & 2 & 1 \\ -1 & 2 & 1 & 3 \\ 1 & -4 & 2 & 3 \\ 1 & 1 & 3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & -2 & 2 & 1 \\ 0 & 0 & 3 & 4 \\ 0 & -2 & 0 & 2 \\ 0 & 3 & 1 & 2 \end{vmatrix} = -\begin{vmatrix} 1 & -2 & 2 & 1 \\ 0 & -2 & 0 & 2 \\ 0 & 0 & 3 & 4 \\ 0 & 3 & 1 & 2 \end{vmatrix} = -(-2)\begin{vmatrix} 1 & -2 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 3 & 4 \\ 0 & 3 & 1 & 2 \end{vmatrix} =$$

$$= 2\begin{vmatrix} 1 & -2 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 1 & 5 \end{vmatrix} = -2\begin{vmatrix} 1 & -2 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 3 & 4 \end{vmatrix} = -2\begin{vmatrix} 1 & -2 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & -11 \end{vmatrix} =$$

$$= -2\big(1\times1\times1\times(-11)\big) = 22 \ \blacksquare$$

Podemos ainda encontrar um outro método para o cálculo de um determinante. A fim de o apresentarmos, porém, precisamos de, em primeiro lugar, introduzir algumas definições:

Definição 1-18 Dada uma matriz quadrada A, chama-se **determinante complementar** de um dos seus elementos, a_{ij}, ao determinante da matriz que resulta de A suprimindo a linha i e a coluna j. Estando subentendido a que matriz nos referimos, representaremos esse determinante por C_{ij}.

Por exemplo, dada

$$A = \begin{bmatrix} 2 & 5 & 3 \\ 9 & 1 & 6 \\ 4 & 8 & 7 \end{bmatrix}$$

teremos

$$C_{11} = \begin{vmatrix} 1 & 6 \\ 8 & 7 \end{vmatrix} = 7 - 48 = -41, \quad C_{21} = \begin{vmatrix} 5 & 3 \\ 8 & 7 \end{vmatrix} = 35 - 24 = 11$$

$$C_{23} = \begin{vmatrix} 2 & 5 \\ 4 & 8 \end{vmatrix} = 16 - 20 = -4, \quad C_{31} = \begin{vmatrix} 5 & 3 \\ 1 & 6 \end{vmatrix} = 30 - 3 = 27$$

$$C_{33} = \begin{vmatrix} 2 & 5 \\ 9 & 1 \end{vmatrix} = 2 - 45 = -43, \text{ etc.}$$

Definição 1-19 Dada uma matriz quadrada A, chama-se **complemento algébrico** (ou **cofactor**) de um dos seus elementos, a_{ij}, ao número $\Delta_{ij} = (-1)^{i+j} C_{ij}$, ou seja,

$$\Delta_{ij} = \begin{cases} C_{ij}, & \text{se} \quad i + j \text{ é par} \\ -C_{ij}, & \text{se} \quad i + j \text{ é ímpar} \end{cases}$$

No exemplo anterior, será

$$\Delta_{11} = -41, \ \Delta_{21} = -11, \ \Delta_{23} = 4, \ \Delta_{31} = 27, \ \Delta_{33} = -43, \text{ etc.}$$

Usando estes novos conceitos, podemos agora provar o seguinte:

Lema 1-20 Tem-se

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} = a_{11} C_{11} = a_{11} \Delta_{11}$$

Demonstração:
Recorrendo, mais uma vez, à definição de determinante, temos

$$|A| = \sum (-1)^{\sigma} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}$$

Uma vez que se tem $a_{i_1 1} = 0$ sempre que $i_1 > 1$, basta, nesta soma, considerar as parcelas em que $i_1 = 1$, obtendo-se

$$|A| = \sum (-1)^{\sigma} a_{11} a_{i_2 2} \cdots a_{i_n n} = a_{11} \sum (-1)^{\sigma} a_{i_2 2} \cdots a_{i_n n}$$

Note-se que, na permutação $1 i_2 \cdots i_n$, o índice 1 não intervém em qualquer inversão, uma vez que é o menor de todos e se encontra em primeiro lugar. Consequentemente, podemos escrever

$$|A| = a_{11} \sum (-1)^{\sigma} a_{i_2 2} \cdots a_{i_n n}$$

onde $i_2 \cdots i_n$ percorre as permutações dos números $2, \ldots, n$ e se tem

$$\sigma = \begin{cases} 0, & \text{se } i_2, \cdots, i_n \text{ é par} \\ 1, & \text{se } i_2, \cdots, i_n \text{ é ímpar} \end{cases}$$

Mas isto significa que

$$|A| = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} = a_{11} C_{11} = a_{11} \Delta_{11}$$

Lema 1-21

Tem-se, para qualquer $j \in \{1, 2, ..., n\}$,

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = a_{1j} \Delta_{1j}$$

Demonstração:

Através de trocas de colunas, podemos levar a coluna número j para primeiro lugar: basta trocá-la com cada uma das $j-1$ anteriores. Como cada troca de colunas faz trocar o sinal ao determinante, obtemos

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = (-1)^{j-1} \begin{vmatrix} a_{1j} & a_{11} & a_{12} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ 0 & a_{21} & a_{22} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n1} & a_{n2} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} =$$

$$= (-1)^{j-1} a_{1j} \begin{vmatrix} a_{21} & a_{22} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} = (-1)^{1+j} a_{1j} C_{1j} = a_{1j} \Delta_{1j}$$

Lema 1-22 Tem-se, para quaisquer $i, j \in \{1, 2, ..., n\}$,

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = a_{ij} \Delta_{ij}$$

Demonstração:

Através de trocas de linhas, obtemos:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = (-1)^{i-1} \begin{vmatrix} a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ a_{11} & a_{12} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} =$$

$$= (-1)^{i-1} a_{ij} (-1)^{1+j} C_{ij} = (-1)^{i+j} a_{ij} C_{ij} = a_{ij} \Delta_{ij}$$

Lema 1-23 Tem-se, para quaisquer $i,j \in \{1,2,...,n\}$,

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{ij} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} = a_{ij}\Delta_{ij}$$

Demonstração:
A demonstração é análoga à do lema anterior.

Com o auxílio dos lemas anteriores, estamos agora em condições de demonstrar o seguinte teorema, que nos fornece, como anunciámos, um novo método para o cálculo de um determinante:

Teorema 1-24 **Teorema de Laplace**

O determinante de uma matriz é igual à soma dos produtos dos elementos de uma certa linha ou coluna da matriz dada pelos respectivos complementos algébricos, ou seja,

$$|A| = a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in} = a_{1j}\Delta_{1j} + a_{2j}\Delta_{2j} + \cdots + a_{nj}\Delta_{nj}$$

qualquer que seja a linha i ou a coluna j.

Demonstração:
Para fixar ideias, escolhamos a primeira linha de A e notemos que

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} +$$

$$+ \begin{vmatrix} 0 & a_{12} & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} 0 & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11}\Delta_{11} + a_{12}\Delta_{12} + \cdots + a_{1n}\Delta_{1n}$$

Note-se que os complementos algébricos calculados em cada uma das matrizes em que decompusemos a matriz inicial são os mesmos que nesta, já que a primeira linha (a única que sofreu alterações) é, em cada caso, desprezada.

Exemplo 1-25

Vamos usar o teorema de Laplace para calcular um determinante. Por exemplo, efectuando o desenvolvimento ao longo da segunda linha, temos

$$\begin{vmatrix} 1 & 0 & 1 & -1 \\ 5 & -7 & 6 & 8 \\ 1 & 2 & 1 & -1 \\ 0 & 2 & 1 & 1 \end{vmatrix} = 5(-1)^3 \begin{vmatrix} 1 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 1 & 1 \end{vmatrix} + (-7)(-1)^4 \begin{vmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \end{vmatrix} + 6(-1)^5 \begin{vmatrix} 1 & 0 & -1 \\ 1 & 2 & -1 \\ 0 & 2 & 1 \end{vmatrix} +$$

$$+8(-1)^6 \begin{vmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 0 & 2 & 1 \end{vmatrix} = -5(1-2-2+2-2+1) - 7 \times 0 - 6(2-2+2) + 8(2+2-2) =$$

$$= -10 - 12 + 16 = -6$$

Na prática, o método de Laplace, para o cálculo de um determinante, é algo lento (pelo menos mais lento que o método de Gauss, para matrizes com mais de três linhas e três colunas), já que, embora a matriz inicial se desdobre em matrizes mais pequenas, estas aparecem em grande número. Assim, por exemplo, para calcular, pelo método de Laplace, o determinante de uma matriz do tipo 6×6, precisamos de

6 determinantes de matrizes do tipo 5×5

ou

30 determinantes de matrizes do tipo 4×4

ou

120 determinantes de matrizes do tipo 3×3.

Chegando à ordem 3, podemos usar a regra de Sarrus...

É possível calcular um determinante de uma forma mais rápida do que qualquer das anteriores se usarmos um método misto, conjugando os métodos de Gauss e de Laplace. ■

Exemplo 1-26

Vamos calcular um determinante usando o método misto. Por exemplo,

$$\begin{vmatrix} 2 & -1 & 1 & 0 & 2 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & -1 & 0 & 2 \\ 1 & 2 & 1 & -1 & 1 \\ 1 & 2 & 0 & 1 & 1 \end{vmatrix} \underset{G}{=} \begin{vmatrix} 2 & -1 & 1 & 0 & 2 \\ 2 & 3 & 2 & 0 & 2 \\ 1 & 0 & -1 & 0 & 2 \\ 1 & 2 & 1 & -1 & 1 \\ 2 & 4 & 1 & 0 & 2 \end{vmatrix} \underset{L}{=} (-1)(-1)^7 \begin{vmatrix} 2 & -1 & 1 & 2 \\ 2 & 3 & 2 & 2 \\ 1 & 0 & -1 & 2 \\ 2 & 4 & 1 & 2 \end{vmatrix} =$$

$$\underset{G}{=} \begin{vmatrix} 2 & -1 & 1 & 2 \\ 0 & 4 & 1 & 0 \\ -1 & 1 & -2 & 0 \\ 0 & 5 & 0 & 0 \end{vmatrix} \underset{L}{=} 2(-1)^5 \begin{vmatrix} 0 & 4 & 1 \\ -1 & 1 & -2 \\ 0 & 5 & 0 \end{vmatrix} = -2(-5) = 10 \ \blacksquare$$

1-4 A regra de Cramer

A noção de determinante foi sugerida pela consideração de sistemas de equações lineares, com tantas equações como incógnitas. Vamos agora ver de que maneira o cálculo de determinantes intervém, na realidade, na resolução de sistemas de equações. Começaremos por demonstrar um corolário, de carácter técnico, do teorema de Laplace:

Corolário 1-27 Dada uma matriz A e tomando dois índices distintos, i e k, tem-se

$$a_{i1}\Delta_{k1} + a_{i2}\Delta_{k2} + \cdots + a_{in}\Delta_{kn} = 0$$

isto é, a soma dos produtos dos elementos de uma certa linha de A pelos complementos algébricos de uma linha diferente é sempre igual a zero. É válida uma propriedade análoga para as colunas.

Demonstração:

Consideremos a matriz auxiliar A' que resulta de A substituindo os elementos da linha k pelos elementos da linha i; estes elementos aparecem, portanto, em duas linhas da matriz A': uma vez na sua posição inicial, na linha i, e também na linha k.

Pela propriedade d) do Teorema 1-12, sabemos que se tem $|A'| = 0$.

Mas a expressão $a_{i1}\Delta_{k1} + a_{i2}\Delta_{k2} + \cdots + a_{in}\Delta_{kn}$ é o desenvolvimento do determinante de A', segundo o teorema de Laplace, ao longo da linha número k.

A partir do teorema de Laplace e deste seu corolário, podemos agora fazer o seguinte:

Consideremos um sistema de equações lineares da forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

Vamos escolher, arbitrariamente, uma das colunas da matriz simples do sistema, por exemplo a primeira (o raciocínio é análogo para qualquer outra). Calculamos o complemento algébrico de cada um desses elementos, multiplicamos a equação número i pelo complemento algébrico do elemento a_{i1}

$$\begin{cases} A_{11}a_{11}x_1 + A_{11}a_{12}x_2 + \cdots + A_{11}a_{1n}x_n = b_1A_{11} \\ A_{21}a_{21}x_1 + A_{21}a_{22}x_2 + \cdots + A_{21}a_{2n}x_n = b_2A_{21} \\ \vdots \\ A_{n1}a_{n1}x_1 + A_{n1}a_{n2}x_2 + \cdots + A_{n1}a_{nn}x_n = b_nA_{n1} \end{cases}$$

e somamos as várias equações assim obtidas:

$$\left(\sum_{i=1}^{n} a_{i1}A_{i1}\right)x_1 + \left(\sum_{i=1}^{n} a_{i2}A_{i1}\right)x_2 + \cdots + \left(\sum_{i=1}^{n} a_{in}A_{i1}\right)x_n = \sum_{i=1}^{n} b_iA_{i1}$$

Pelo Teorema de Laplace, sabemos que

$$\sum_{i1}^{n} a_{i1}A_{i1} = |A|$$

e, pelo Corolário 1-27, que

$$\sum_{i1}^{n} a_{ik}A_{i1} = 0$$

para cada $k \neq 1$. Por conseguinte, a igualdade anterior transforma-se em

$$|A|x_1 = \sum_{i=1}^{n} b_iA_{i1}$$

Por outro lado, a expressão

$$\sum_{i=1}^{n} b_iA_{i1}$$

corresponde ao desenvolvimento, pelo método de Laplace, ao longo da coluna número 1, do determinante de uma matriz que resulta da matriz simples do sistema, substituindo a primeira coluna pelos termos independentes do sistema:

$$\sum_{i=1}^{n} b_i A_{i1} = \begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Assim, chegámos à igualdade

$$|A|\, x_1 = \begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Nestas condições, se for $|A| \neq 0$, podemos obter o valor de x_1 pela fórmula

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}}{|A|}$$

Naturalmente, o mesmo raciocínio, aplicado a qualquer outra coluna da matriz simples, levará a conclusões similares:

$$x_j = \frac{\begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}}{|A|} \qquad (j = 1, 2, \ldots, n) \qquad \textbf{(regra de Cramer)}$$

Uma das vantagens da aplicação da regra de Cramer em relação a outros processos para a resolução de um sistema de equações lineares consiste na possibilidade de se calcular o valor de cada incógnita, independentemente da determinação das restantes.

Exemplo 1-28

Dado o sistema de equações lineares

$$\begin{cases} x + y - 2z = 3 \\ -x + 2y - z = 5 \\ x + z = 4 \end{cases}$$

usamos a regra de Cramer para calcular o valor de y:

Em primeiro lugar, para verificarmos que a regra de Cramer é, de facto, aplicável à resolução deste sistema de equações, vamos certificar-nos de que o determinante da sua matriz simples é não nulo:

$$\begin{vmatrix} 1 & 1 & -2 \\ -1 & 2 & -1 \\ 1 & 0 & 1 \end{vmatrix} = 2 - 1 + 4 + 1 = 6 \neq 0$$

Assim,

$$y = \frac{\begin{vmatrix} 1 & 3 & -2 \\ -1 & 5 & -1 \\ 1 & 4 & 1 \end{vmatrix}}{6} = \frac{5 - 3 + 8 + 10 + 3 + 4}{6} = \frac{27}{6} = \frac{9}{2} \blacksquare$$

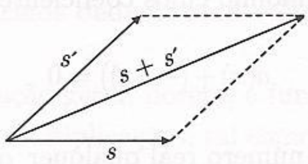
2 Espaços vectoriais

2-1 Generalidades

O objecto fundamental da Álgebra Linear é o estudo dos "espaços vectoriais" ou "espaços lineares". Para se compreender o alcance desse estudo, é necessário ter uma ideia dos métodos e dos propósitos da Álgebra em geral: pretende-se, no fundo, estudar, num plano abstracto, o que várias situações concretas distintas possam ter em comum, nomeadamente do ponto de vista das propriedades de operações definidas em certos conjuntos.

Consideremos, por exemplo, o conjunto S constituído por todos os segmentos orientados com origem num determinado ponto, fixo num plano. É bem sabido que, por exemplo, na Física, determinadas grandezas se costumam representar por meio de tais segmentos orientados.

Dados dois desses segmentos orientados, s e s', é costume construir-se a "soma" ou "resultante" dos dois, representada por $s+s'$, através da conhecida "regra do paralelogramo":



Podemos, portanto, dizer que no conjunto S está definida uma operação de "adição", representada pelo símbolo "+", e que, como se sabe, goza, entre outras, das seguintes propriedades:

a) $\forall s, s' \in S : s + s' = s' + s$

b) $\forall s, s', s'' \in S : (s + s') + s'' = s + (s' + s'')$

c) Representando por 0 o segmento reduzido a um só ponto:

$$\forall s \in S : s + 0 = s$$

d) Representando por $-s$ o segmento orientado com a mesma direcção e o mesmo comprimento que s, mas de sentido contrário:

$$s + (-s) = 0$$

Por outro lado, se tomarmos um número real qualquer α, podemos "multiplicar" cada segmento orientado s por α, como habitualmente, obtendo um novo segmento orientado αs. Valem as seguintes propriedades:

e) $\forall \alpha \in \mathbb{R}, \forall s, s' \in S : \alpha(s+s') = \alpha s + \alpha s'$

f) $\forall \alpha, \beta \in \mathbb{R}, \forall s, s' \in S : (\alpha+\beta)s = \alpha s + \beta s$

g) $\forall \alpha, \beta \in \mathbb{R}, \forall s \in S : \alpha(\beta s) = (\alpha\beta)s$

h) $\forall s \in S : 1s = s$

Veremos que, em muitas outras situações, se encontram propriedades inteiramente análogas a estas.

Tomemos, por exemplo, o conjunto $\mathbb{R}[x]$, constituído por todos os polinómios de coeficientes reais, de qualquer grau, com a variável x.

Dados dois polinómios, digamos $a(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ e $b(x) = b_0 + b_1x + b_2x^2 + + \cdots + b_mx^m$, é costume definir-se a sua "soma" como sendo o polinómio $a(x)+b(x) = (a_0+b_0) + + (a_1+b_1)x + (a_2+b_2)x^2 + \cdots$, onde se supõe que a parcela de menor grau foi prolongada, com coeficientes nulos, até ao grau da outra.

Esta operação de "adição" goza, como bem sabemos, das seguintes propriedades:

a) $\forall a(x), b(x) \in \mathbb{R}[x] : a(x)+b(x) = b(x)+a(x)$

b) $\forall a(x), b(x), c(x) \in \mathbb{R}[x] : (a(x)+b(x))+c(x) = a(x)+(b(x)+c(x))$

c) Designando por 0 o polinómio (do grau que se quiser) de coeficientes todos nulos:

$$\forall a(x) \in \mathbb{R}[x] : a(x)+0 = a(x)$$

d) Designando por $-a(x)$ o polinómio cujos coeficientes são os simétricos dos de $a(x)$:

$$a(x)+(-a(x)) = 0$$

Sabe-se também que, dado um número real qualquer α, se define $\alpha a(x) = (\alpha a_0) + (\alpha a_1)x + + \cdots + (\alpha a_n)x^n$ e que esta nova operação goza das seguintes propriedades:

e) $\forall \alpha \in \mathbb{R}, \forall a(x), b(x) \in \mathbb{R}[x] : \alpha(a(x)+b(x)) = \alpha a(x) + \alpha b(x)$

f) $\forall \alpha, \beta \in \mathbb{R}, \forall a(x) \in \mathbb{R}[x] : (\alpha+\beta)a(x) = \alpha a(x) + \beta a(x)$

g) $\forall \alpha, \beta \in \mathbb{R}, \forall a(x) \in \mathbb{R}[x] : \alpha(\beta a(x)) = (\alpha\beta)a(x)$

h) $\forall a(x) \in \mathbb{R}[x] : 1a(x) = a(x)$

Deve ainda salientar-se que as mesmas propriedades se verificariam se tomássemos as constantes α, β, etc., só no conjunto dos números racionais, por exemplo, em vez de as tomarmos no conjunto dos números reais.

Consideremos ainda o conjunto $\mathbb{C}$, dos números complexos. Como se sabe, dados os números complexos $z = a + bi$ e $z' = a' + b'i$, definimos a sua soma como sendo o complexo $z + z' = (a + a') + (b + b')i$, gozando esta operação de adição de diversas propriedades importantes, entre as quais as seguintes:

a) $\forall z, z' \in \mathbb{C} : z + z' = z' + z$

b) $\forall z, z', z'' \in \mathbb{C} : (z + z') + z'' = z + (z' + z'')$

c) Designando por 0 o complexo $0 = 0 + 0i$, tem-se: $\forall z \in \mathbb{C} : z + 0 = z$

d) Tomando, para cada complexo $z = a + bi$, o complexo $-z = -a - bi$, tem-se $z + (-z) = 0$.

Por sua vez, está também definida a multiplicação de um número complexo $z = a + bi$ por um outro número complexo ("constante") $\alpha = \alpha_1 + \alpha_2 i$:

$$\alpha z = (\alpha_1 a - \alpha_2 b) + (\alpha_1 b + \alpha_2 a)i$$

Ora, é fácil verificar que se tem as seguintes propriedades (que seriam igualmente válidas se tomássemos as "constantes" $\alpha, \beta, \ldots$ só no conjunto dos números reais, ou no conjunto dos números racionais, por exemplo):

e) $\forall \alpha \in \mathbb{C}, \forall z, z' \in \mathbb{C} : \alpha(z + z') = \alpha z + \alpha z'$

f) $\forall \alpha, \beta \in \mathbb{C}, \forall z \in \mathbb{C} : (\alpha + \beta)z = \alpha z + \beta z$

g) $\forall \alpha, \beta \in \mathbb{C}, \forall z \in \mathbb{C} : \alpha(\beta z) = (\alpha\beta)z$

h) $\forall z \in \mathbb{C} : 1z = z$

Estes e muitos outros exemplos, que poderíamos apresentar, levam-nos a considerar uma situação geral, abstracta, onde as mesmas propriedades se verifiquem, à partida, na certeza de que tudo quanto daí conseguirmos deduzir será válido em cada um dos exemplos considerados.

A fim de se chegar a essa situação geral, porém, é fundamental dizer alguma coisa acerca das "constantes" a utilizar para as multiplicações, tal como anteriormente encontrámos. Já nos dois últimos exemplos observámos que há uma certa latitude quanto ao conjunto a seleccionar para a escolha das tais constantes. Nuns casos, será adequado usar-se o conjunto dos números reais; noutros, o conjunto dos números racionais, noutros ainda, o conjunto dos números complexos. Muito sumariamente, convirá lidar com "constantes" com as quais se possam efectuar as operações aritméticas elementares (soma, subtracção, multiplicação e divisão por elementos não nulos) e em que essas operações gozem das propriedades mais habituais (comutativas, associativas, distributivas).

Num plano inteiramente abstracto, essa ideia corresponde à noção de "corpo", que passamos a definir:

Definição 2-1 Seja K um conjunto qualquer, no qual estejam definidas duas operações binárias, designadas por "adição" e "multiplicação" e representadas pela simbologia habitual. Diz-se que K, com essas operações, constitui um **corpo** quando se verificam as seguintes propriedades:

a) Propriedades da adição
  a1) Propriedade comutativa: $\forall \alpha, \beta \in K : \alpha + \beta = \beta + \alpha$
  a2) Propriedade associativa:
$$\forall \alpha, \beta, \gamma \in K : (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$$
  a3) Existência de elemento neutro: $\exists 0 \in K : 0 + \alpha = \alpha, \forall \alpha \in K$
  a4) Existência de simétricos: $\forall \alpha \in K, \exists -\alpha \in K : \alpha + (-\alpha) = 0$

b) Propriedades da multiplicação
  b1) Propriedade comutativa: $\forall \alpha, \beta \in K : \alpha\beta = \beta\alpha$
  b2) Propriedade associativa: $\forall \alpha, \beta, \gamma \in K : (\alpha\beta)\gamma = \alpha(\beta\gamma)$
  b3) Existência de elemento neutro: $\exists 1 \in K : 1\alpha = \alpha, \forall \alpha \in K$
  b4) Existência de simétricos: $\forall \alpha \in K \setminus \{0\}, \exists \frac{1}{\alpha} \in K : \alpha \frac{1}{\alpha} = 1$

c) Propriedade (distributiva) de ligação
$$\forall \alpha, \beta, \gamma \in K : (\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$$

Entre os exemplos mais vulgares de corpos contam-se: o conjunto $\mathbb{R}$, dos números reais, com as operações habituais; o conjunto $\mathbb{Q}$, dos números racionais, com as operações habituais; o conjunto $\mathbb{C}$, dos números complexos, com as operações habituais. Mas há muitos outros corpos. Assim, por exemplo, podemos construir um corpo no conjunto $K = \{0,1,2\}$, definindo as operações de adição e multiplicação através das seguintes tabelas:

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

×	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

A seguir, sempre com a anunciada preocupação de generalidade, falaremos frequentemente num corpo K qualquer. Mas na quase totalidade dos casos, poderá o leitor, caso isso lhe facilite a compreensão dos assuntos tratados, imaginar que nos referimos concretamente (e salvo menção especial) ao corpo dos números reais.

Com o que fica dito, estamos, finalmente, em condições de introduzir a definição de "espaço vectorial":

Definição 2-2 Seja $E=\left\{\vec{a},\vec{b},\vec{c},...,\vec{x},\vec{y},...\right\}$ um conjunto qualquer. Diz-se que introduzimos em E uma estrutura de **espaço vectorial** ou **espaço linear** sobre um dado corpo K, quando:

a) Definimos em E uma operação binária, chamada "adição" e representada pelo símbolo "+", que goze das seguintes propriedades:
   - a1) Propriedade comutativa: $\forall \vec{x},\vec{y}\in E:\vec{x}+\vec{y}=\vec{y}+\vec{x}$
   - a2) Propriedade associativa: $\forall \vec{x},\vec{y},\vec{z}\in E:(\vec{x}+\vec{y})+\vec{z}=\vec{x}+(\vec{y}+\vec{z})$
   - a3) Existência de elemento neutro: $\exists\vec{0}\in E:\vec{x}+\vec{0}=\vec{x},\ \forall\vec{x}\in E$
   - a4) Existência de elementos simétricos: $\forall\vec{x}\in E,\ \exists-\vec{x}\in E:\ \vec{x}+(-\vec{x})=\vec{0}$

b) Para cada $\alpha\in K$, define-se uma "multiplicação", que transforma cada elemento $\vec{x}\in E$ num outro elemento $\alpha\vec{x}\in E$, de tal modo que se verifiquem as seguintes propriedades:
   - a5) $\forall\alpha\in K,\ \forall\vec{x},\vec{y}\in E:\alpha(\vec{x}+\vec{y})=\alpha\vec{x}+\alpha\vec{y}$
   - a6) $\forall\alpha,\beta\in K,\forall\vec{x}\in E:(\alpha+\beta)\vec{x}=\alpha\vec{x}+\beta\vec{x}$
   - a7) $\forall\alpha,\beta\in K,\forall\vec{x}\in E:\alpha(\beta\vec{x})=(\alpha\beta)\vec{x}$
   - a8) $\forall\vec{x}\in E\ 1\vec{x}=\vec{x}$

Quando estas condições se verificam, os elementos de E (qualquer que seja a sua natureza) tomam o nome de **vectores**, enquanto os elementos do corpo K se chamam **escalares**.

Em muitos sectores da Matemática, surgem frequentemente espaços vectoriais. Ao longo do nosso estudo da Álgebra Linear encontraremos, mais adiante, diversos exemplos. Mas podemos, desde já, apresentar os seguintes:

Exemplos 2-3

a) Consideremos o conjunto $\mathbb{R}^2$, formado por todos os pares ordenados de números reais, e definamos nele as seguintes operações (onde α designa um número real qualquer):

$$(x_1,x_2)+(y_1,y_2)=(x_1+y_1,x_2+y_2)$$

$$\alpha(x_1,x_2)=(\alpha x_1,\alpha x_2)$$

Obtém-se, deste modo, um espaço vectorial sobre $\mathbb{R}$ (abreviadamente, um **espaço vectorial real**).

b) Analogamente, no conjunto $\mathbb{R}^3$, formado por todos os ternos ordenados de números reais, obtemos um espaço vectorial real quando definamos as seguintes operações (onde α designa um número real qualquer):

$$(x_1,x_2,x_3)+(y_1,y_2,y_3)=(x_1+y_1,x_2+y_2,x_3+y_3)$$
$$\alpha(x_1,x_2,x_3)=(\alpha x_1,\alpha x_2,\alpha x_3)$$

De modo geral, para cada número natural n, podemos considerar o conjunto $\mathbb{R}^n$, constituído por todas as listas (ordenadas) de n números reais, e definir um espaço vectorial real, com as seguintes operações:

$$(x_1,x_2,\ldots,x_n)+(y_1,y_2,\ldots,y_n)=(x_1+y_1,x_2+y_2,\ldots,x_n+y_n)$$
$$\alpha(x_1,x_2,\ldots,x_n)=(\alpha x_1,\alpha x_2,\ldots,\alpha x_n)$$

c) Para cada número natural n, obtemos um espaço vectorial real definindo operações, no conjunto $\mathbb{C}^n$ — formado por todas as listas ordenadas de n números complexos — através das seguintes igualdades, onde α designa um número real qualquer:

$$(x_1,x_2,\ldots,x_n)+(y_1,y_2,\ldots,y_n)=(x_1+y_1,x_2+y_2,\ldots,x_n+y_n)$$
$$\alpha(x_1,x_2,\ldots,x_n)=(\alpha x_1,\alpha x_2,\ldots,\alpha x_n)$$

Ao mesmo tempo, no conjunto $\mathbb{C}^n$, podemos definir um **espaço vectorial complexo** (isto é, um espaço vectorial sobre o corpo dos números complexos), definindo operações pelas mesmas igualdades anteriores, mas onde α designa agora um número complexo qualquer.

d) Dado um corpo K qualquer, podemos construir um espaço vectorial sobre K, para cada número natural n, definindo operações, no conjunto K^n, formado por todas as listas ordenadas de n elementos de K, através das seguintes igualdades, onde α designa um elemento qualquer de K:

$$(x_1,x_2,\ldots,x_n)+(y_1,y_2,\ldots,y_n)=(x_1+y_1,x_2+y_2,\ldots,x_n+y_n)$$
$$\alpha(x_1,x_2,\ldots,x_n)=(\alpha x_1,\alpha x_2,\ldots,\alpha x_n)$$

Os espaços vectoriais reais da forma $\mathbb{R}^n$, indicados nos exemplos a) e b)), bem como os espaços complexos da forma $\mathbb{C}^n$, do exemplo c)), são casos particulares destes espaços vectoriais da forma K^n.

É claro que, nestes conjuntos, poderíamos eventualmente definir outras operações que satisfizessem as mesmas propriedades, obtendo-se então diferentes espaços vectoriais com os mesmos conjuntos de vectores. No entanto, sempre que nos referirmos ao "espaço vectorial K^n sobre o corpo K^T", consideraremos, salvo indicação expressa em sentido contrário, que as operações são definidas como acabámos de indicar.

e) Seja E o conjunto de todas as sucessões $\left(u_n\right)_{n\in\mathbb{N}}$ de números reais, onde consideramos as operações usuais, ou seja,

$$\left(u_n\right)_{n\in\mathbb{N}}+\left(v_n\right)_{n\in\mathbb{N}}=\left(u_n+v_n\right)_{n\in\mathbb{N}}$$
$$\alpha\left(u_n\right)_{n\in\mathbb{N}}=\left(\alpha u_n\right)_{n\in\mathbb{N}}$$

Obtém-se assim um espaço vectorial real.

f) Seja F o conjunto de todas as funções reais definidas num certo domínio real D. Obtemos um espaço vectorial real quando consideramos em F as operações definidas do modo habitual:

$$(f+g)(x)=f(x)+g(x)$$
$$(\alpha f)(x)=\alpha f(x)$$

g) Se num conjunto com um só elemento, $E=\left\{\vec{0}\right\}$, definirmos as necessárias operações da única maneira possível, isto é,

$$\vec{0}+\vec{0}=\vec{0}$$
$$\alpha\vec{0}=\vec{0}$$

onde α designa um elemento qualquer de um corpo qualquer K, obtemos um espaço vectorial sobre K, chamado **espaço vectorial nulo.** ■

A partir das propriedades exigidas na definição de espaço vectorial, é possível deduzir outras, muito simples, como as que a seguir se indicam:

Teorema 2-4 Seja E um espaço vectorial qualquer, sobre um corpo K. Para vectores e escalares arbitrários, valem as seguintes igualdades:

a) $0\vec{x}=\vec{0}$
b) $\alpha\vec{0}=\vec{0}$
c) $(-\alpha)\vec{x}=-(\alpha\vec{x})$
d) $\alpha(-\vec{x})=-(\alpha\vec{x})$
e) $(\alpha-\beta)\vec{x}=\alpha\vec{x}-\beta\vec{x}$
f) $\alpha(\vec{x}-\vec{y})=\alpha\vec{x}-\alpha\vec{y}$

Demonstração:

a) Note-se que $0\vec{x}=(0+0)\vec{x}=0\vec{x}+0\vec{x}$. Se somarmos $-0\vec{x}$ ao primeiro e ao último membros das igualdades anteriores, obtemos $\vec{0}=0\vec{x}$.

b) Note-se que $\alpha\vec{0}=\alpha(\vec{0}+\vec{0})=\alpha\vec{0}+\alpha\vec{0}$. Se somarmos $-\alpha\vec{0}$ ao primeiro e ao último membros das igualdades anteriores, obtemos $\vec{0}=\alpha\vec{0}$.

c) Em virtude da igualdade a), estabelecida acima, vemos que $(-\alpha)\vec{x}+\alpha\vec{x}=(-\alpha+\alpha)\vec{x}=0\vec{x}=\vec{0}$, pelo que $(-\alpha)\vec{x}=-(\alpha\vec{x})$.
d) Resulta da igualdade b).
e) Tem-se $(\alpha-\beta)\vec{x}=(\alpha+(-\beta))\vec{x}=\alpha\vec{x}+(-\beta)\vec{x}=\alpha\vec{x}-\beta\vec{x}$, atendendo à igualdade c).
f) Resulta da igualdade d).

2-2 Combinações lineares

Num espaço vectorial, temos apenas os vectores, de um lado, e os escalares, do outro. Podemos multiplicar vectores por escalares, para obter novos vectores, e podemos somar os vectores uns com os outros. Nada mais podemos lá fazer! Isso justifica a seguinte definição:

Definição 2-5 Seja E um espaço vectorial sobre um corpo K. Tomando vectores $\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k\in E$ e escalares $\alpha_1,\alpha_2,\dots,\alpha_k\in K$, diz-se que o vector

$$\vec{v}=\alpha_1\vec{v}_1+\alpha_2\vec{v}_2+\dots+\alpha_k\vec{v}_k$$

é obtido como **combinação linear** dos vectores $\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k$, com os coeficientes $\alpha_1,\alpha_2,\dots,\alpha_k$.

Exemplos 2-6

a) O vector $(-8,7,5)$, do espaço vectorial real $\mathbb{R}^3$, é combinação linear dos vectores $(1,-1,2)$, $(2,1,1)$ e $(-1,3,1)$, uma vez que se tem, conforme facilmente se verifica,

$$(-8,7,5)=2(1,-1,2)-3(2,1,1)+4(-1,3,1)$$

b) No espaço vectorial real $\mathbb{R}[x]$ consideremos os vectores $\vec{v}_1=1+x-2x^2-x^3$, $\vec{v}_2=-1+3x+x^3$, $\vec{v}_3=1-x-x^2-x^3$, $\vec{v}_4=5-7x-4x^2-5x^3$. Vamos verificar se o vector $\vec{v}_4$ se pode escrever como combinação linear dos restantes.

Pondo

$$5-7x-4x^2-5x^3=\alpha\left(1+x-2x^2-x^3\right)+\beta\left(-1+3x+x^3\right)+\gamma\left(1-x-x^2-x^3\right)$$

obtemos, ao igualar os coeficientes dos polinómios que figuram nos dois membros da igualdade, o sistema de equações

$$\begin{cases} \alpha - \beta + \gamma = 5 \\ \alpha + 3\beta - \gamma = -7 \\ -2\alpha - \gamma = -4 \\ -\alpha + \beta - \gamma = -5 \end{cases}$$

Resolvendo este sistema, encontramos

$$\begin{cases} \beta = -1 - \alpha \\ \gamma = 4 - 2\alpha \end{cases}$$

com α arbitrário. Tomando, por exemplo, $\alpha = 1$, ficamos com

$$\begin{cases} \beta = -2 \\ \gamma = 2 \end{cases}$$

pelo que $\vec{v}_4 = \vec{v}_1 - 2\vec{v}_2 + 2\vec{v}_3$. ■

A possibilidade de escrevermos vectores de um certo espaço vectorial como combinações lineares de outros leva-nos a introduzir uma nova definição:

Definição 2-7 Consideremos, num espaço vectorial E, os vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ e os vectores $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_t$. Dizemos que a lista (ou "sistema") de vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ é **equivalente** à lista de vectores $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_t$, e representamos esse facto por $(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k) \approx (\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_t)$, quando cada um dos vectores $\vec{v}_i$ $(i = 1, 2, \ldots, k)$ se pode escrever como combinação linear de $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_t$ e, ao mesmo tempo, cada um dos vectores $\vec{w}_j$ $(j = 1, 2, \ldots, t)$ se pode escrever como combinação linear de $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$.

Exemplo 2-8

Consideremos, no espaço vectorial real $\mathbb{R}^3$, os vectores $\vec{v}_1 = (1,1,0)$, $\vec{v}_2 = (0,1,1)$, $\vec{v}_3 = (1,0,1)$ e os vectores $\vec{w}_1 = (1,-1,1)$, $\vec{w}_2 = (-1,1,1)$, $\vec{w}_3 = (0,0,1)$, $\vec{w}_4 = (1,1,-1)$.

É fácil ver que se tem $(\vec{v}_1, \vec{v}_2, \vec{v}_3) \approx (\vec{w}_1, \vec{w}_2, \vec{w}_3, \vec{w}_4)$, uma vez que, por exemplo,

$$\vec{v}_1 = \vec{w}_3 + \vec{w}_4 \qquad \vec{w}_1 = \frac{1}{2}\vec{v}_1 - \frac{1}{2}\vec{v}_2 + \frac{3}{2}\vec{v}_3$$

$$\vec{v}_2 = -\frac{1}{2}\vec{w}_1 + 2\vec{w}_3 + \frac{1}{2}\vec{w}_4 \qquad \vec{w}_2 = -\frac{1}{2}\vec{v}_1 + \frac{3}{2}\vec{v}_2 - \frac{1}{2}\vec{v}_3$$

$$\vec{v}_3 = \frac{1}{2}\vec{w}_1 + \vec{w}_3 + \frac{1}{2}\vec{w}_4 \qquad \vec{w}_3 = -\frac{1}{2}\vec{v}_1 + \frac{1}{2}\vec{v}_2 + \frac{1}{2}\vec{v}_3$$

$$\vec{w}_4 = \frac{3}{2}\vec{v}_1 - \frac{1}{2}\vec{v}_2 - \frac{1}{2}\vec{v}_3 \quad ■$$

O conceito de equivalência entre listas de vectores goza de algumas propriedades importantes, a saber:

Teorema 2-9

Num espaço vectorial, são válidas, para vectores arbitrários, as seguintes propriedades:

a) $(\vec{v}_1, \vec{v}_2, ..., \vec{v}_k) \approx (\vec{v}_1, \vec{v}_2, ..., \vec{v}_k)$ — propriedade reflexiva.

b) Se $(\vec{v}_1, \vec{v}_2, ..., \vec{v}_k) \approx (\vec{w}_1, \vec{w}_2, ..., \vec{w}_t)$, então $(\vec{w}_1, \vec{w}_2, ..., \vec{w}_t) \approx (\vec{v}_1, \vec{v}_2, ..., \vec{v}_k)$ — propriedade simétrica.

c) Se se tiver $(\vec{v}_1, \vec{v}_2, ..., \vec{v}_k) \approx (\vec{w}_1, \vec{w}_2, ..., \vec{w}_t)$ e $(\vec{w}_1, \vec{w}_2, ..., \vec{w}_t) \approx (\vec{z}_1, \vec{z}_2, ..., \vec{z}_s)$, então $(\vec{v}_1, \vec{v}_2, ..., \vec{v}_k) \approx (\vec{z}_1, \vec{z}_2, ..., \vec{z}_s)$ — propriedade transitiva.

d) $(\vec{v}_1, \vec{v}_2, ..., \vec{v}_i, ..., \vec{v}_k) \approx (\vec{v}_1, \vec{v}_2, ..., \xi\vec{v}_i, ..., \vec{v}_k)$, qualquer que seja o escalar não nulo ξ.

e) $(\vec{v}_1, \vec{v}_2, ..., \vec{v}_i, ..., \vec{v}_j, ..., \vec{v}_k) \approx (\vec{v}_1, \vec{v}_2, ..., \vec{v}_i, ..., \xi\ \vec{v}_i + \vec{v}_j, ..., \vec{v}_k)$ qualquer que seja o escalar não nulo ξ.

Demonstração:

a) Resulta imediatamente da definição, uma vez que, para cada índice i entre 1 e k, se tem

$$\vec{v}_i = 0\vec{v}_1 + 0\vec{v}_2 + \cdots + 0\vec{v}_{i-1} + 1\vec{v}_i + 0\vec{v}_{i+1} + \cdots + 0\vec{v}_k$$

b) Resulta trivialmente da definição.

c) Tomemos um vector $\vec{v}_i$ $(i = 1, 2, ..., k)$. Por hipótese, ter-se-á

$$\vec{v}_i = \alpha_1\vec{w}_1 + \alpha_2\vec{w}_2 + \cdots + \alpha_t\vec{w}_t = \sum_{j=1}^{t} \alpha_j\vec{w}_j$$

Mas, também por hipótese, teremos

$$\vec{w}_j = \sum_{r=1}^{s} \beta_{jr}\vec{z}_r \ (j = 1, 2, ..., t)$$

Substituindo na igualdade anterior, obtemos então

$$\vec{v}_i = \sum_{j=1}^{t} \alpha_j \left(\sum_{r=1}^{s} \beta_{jr}\vec{z}_r \right) = \sum_{j=1}^{t} \left(\sum_{r=1}^{s} \alpha_j\beta_{jr}\vec{z}_r \right) = \sum_{r=1}^{s} \left(\sum_{j=1}^{t} \alpha_j\beta_{jr}\vec{z}_r \right) = \sum_{r=1}^{s} \left(\sum_{j=1}^{t} \alpha_j\beta_{jr} \right) \vec{z}_r$$

o que mostra que cada vector $\vec{v}_i$ $(i = 1, 2, ..., k)$ é combinação linear dos vectores $(\vec{z}_1, \vec{z}_2, ..., \vec{z}_s)$. Analogamente se demonstra a recíproca.

d) Basta que notar que, sendo $\xi \neq 0$, podemos escrever

$$\vec{v}_i = 0\vec{v}_1 + 0\vec{v}_2 + \cdots + 0\vec{v}_{i-1} + \frac{1}{\xi}\xi\vec{v}_i + 0\vec{v}_{i+1} + \cdots + 0\vec{v}_k$$

e) Basta notar que

$$\vec{v}_i = 0\vec{v}_1 + 0\vec{v}_2 + \cdots - \xi \ \vec{v}_i + \cdots + 1\left(\xi\vec{v}_i + \vec{v}_j\right) + \cdots + 0\vec{v}_k$$

2-3 Dependência linear

Um outro conceito, da maior importância, prende-se com o número de diferentes maneiras de escrever um dado vector como combinação linear de outros. Esse conceito é traduzido pela próxima definição:

Definição 2-10 Os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$, de um espaço vectorial E, são **linearmente independentes** quando há apenas uma maneira de escrever o vector $\vec{0}$ como combinação linear deles, que é, como trivialmente se compreende, com os coeficientes todos nulos. Quando há mais de uma maneira de escrever o vector $\vec{0}$ como combinação linear dos vectores dados, estes são **linearmente dependentes**.

Resumidamente, os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ são linearmente independentes quando verificam a condição

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k = \vec{0} \Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$$

Pelo contrário, são linearmente dependentes quando existem escalares $\alpha_1, \alpha_2, ..., \alpha_k$, não todos nulos (embora alguns deles possam, evidentemente, ser nulos), tais que $\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k = \vec{0}$.

Exemplos 2-11

a) Os vectores $(1,1,1)$, $(1,1,0)$ e $(0,1,-1)$, do espaço vectorial real $\mathbb{R}^3$, são linearmente independentes. Com efeito, partindo de

$$\alpha(1,1,1) + \beta(1,1,0) + \gamma(0,1,-1) = (0,0,0)$$

obtemos

$$\begin{cases} \alpha + \beta = 0 \\ \alpha + \beta + \gamma = 0 \\ \alpha - \gamma = 0 \end{cases}$$

de onde resulta $\alpha = \beta = \gamma = 0$.

b) Os vectores $1+x-2x^2-x^3$, $-1+3x+x^3$, $1-x-x^2-x^3$, do espaço vectorial $\mathbb{R}[x]$, são linearmente dependentes.

Com efeito, partindo de

$$\alpha\left(1+x-2x^2-x^3\right)+\beta\left(-1+3x+x^3\right)+\gamma\left(1-x-x^2-x^3\right)=0$$

obtemos

$$\begin{cases}\alpha-\beta+\gamma=0\\ \alpha+3\beta-\gamma=0\\ -2\alpha-\gamma=0\\ -\alpha+\beta-\gamma=0\end{cases}$$

ou seja,

$$\begin{cases}\beta=-\alpha\\ \gamma=-2\alpha\end{cases}$$

com α arbitrário.

Dando diferentes valores a α, obtemos diversas soluções para o problema. ■

As designações de "dependência" e "independência" linear, aplicadas aos conceitos que acabámos de introduzir, são justificadas pela segunda parte do seguinte teorema:

Teorema 2-12 Seja E um espaço vectorial qualquer

a) Um vector $\vec{v}\in E$ forma, só por si, um sistema linearmente dependente se e somente se $\vec{v}=\vec{0}$.

b) Os vectores $\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k$, com $k>1$, são linearmente dependentes se e só se algum deles se pode obter como combinação linear dos restantes.

Demonstração:

a) Se se tiver $\vec{v}=\vec{0}$, então $1\vec{v}=1\vec{0}=\vec{0}$, sendo 1 um coeficiente não nulo, pelo que $\vec{v}$ fica, por definição, linearmente dependente.

Pelo contrário, se $\vec{v}\neq\vec{0}$, suponhamos que se tem $\xi\vec{v}=\vec{0}$, com $\xi\neq 0$; nesse caso, seria

$$\vec{v}=\frac{1}{\xi}(\xi\vec{v})=\frac{1}{\xi}\vec{0}=\vec{0}$$

contra a hipótese. Isso mostra que, para se ter $\xi\vec{v}=\vec{0}$, terá de ser $\xi=0$, e o vector $\vec{v}$ é linearmente independente.

b) Se os vectores $\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k$ são linearmente dependentes, existirão coeficientes não todos nulos, $\alpha_1,\alpha_2,\ldots,\alpha_k$, tais que $\alpha_1\vec{v}_1+\alpha_2\vec{v}_2+\cdots+\alpha_k\vec{v}_k=\vec{0}$.

Uma vez que a adição de vectores é comutativa, não há quebra de generalidade em se supor que, por exemplo, $\alpha_1 \neq 0$. Mas então, da igualdade

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k = \vec{0}$$

resulta

$$\vec{v}_1 = -\frac{\alpha_2}{\alpha_1}\vec{v}_2 - \cdots - \frac{\alpha_k}{\alpha_1}\vec{v}_k$$

Vemos, pois, que qualquer vector que esteja afectado de um coeficiente não nulo, na igualdade $\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k = \vec{0}$, se pode escrever à custa dos demais.

Reciprocamente, suponhamos que um dos vectores dados é combinação linear dos restantes, por exemplo, que

$$\vec{v}_k = \zeta_1\vec{v}_1 + \cdots + \zeta_{k-1}\vec{v}_{k-1}$$

Daqui resulta que

$$\zeta_1\vec{v}_1 - \cdots + \zeta_{k-1}\vec{v}_{k-1} - \vec{v}_k = \vec{0}$$

e, nesta igualdade, o vector $\vec{v}_k$ tem o coeficiente não nulo 1, o que mostra que os vectores considerados são linearmente dependentes.

A noção de dependência e independência linear goza ainda de uma importante propriedade, a saber:

Teorema 2-13 Consideremos um espaço vectorial E qualquer. Se os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ são dependentes, então os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k, \vec{v}_{k+1}, \vec{v}_{k+2}, ..., \vec{v}_{k+t}$ são também linearmente dependentes.

Demonstração:

Se os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ são dependentes, então existem coeficientes $\alpha_1, \alpha_2, ..., \alpha_k$, não todos nulos, tais que $\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k = \vec{0}$.

Mas então temos, trivialmente,

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k + 0\vec{v}_{k+1} + 0\vec{v}_{k+2} + \cdots + 0\vec{v}_{k+t} = \vec{0}$$

e entre os coeficientes $\alpha_1, \alpha_2, ..., \alpha_k$ haverá pelo menos um que é diferente de zero, pelo que os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k, \vec{v}_{k+1}, \vec{v}_{k+2}, ..., \vec{v}_{k+t}$ são linearmente dependentes.

Corolário 2-14 Qualquer parte de uma lista de vectores independentes é ainda independente.

Demonstração:
Resulta trivialmente do teorema anterior.

Corolário 2-15 Qualquer lista de vectores que inclua o vector nulo é dependente.

Demonstração:
Resulta trivialmente do teorema anterior, uma vez que o vector nulo, só por si, é linearmente dependente.

Corolário 2-16 Qualquer lista de vectores que inclua dois vectores iguais é dependente.

Demonstração:
Um par de vectores iguais, digamos $\vec{a}$ e $\vec{a}$, é linearmente dependente, já que podemos escrever

$$1\vec{a} - 1\vec{a} = \vec{0}$$

sendo os coeficientes não nulos.

Por consequência, qualquer lista de vectores que os inclua fica também dependente, em face do teorema anterior.

O Teorema 2-12 mostrou que, quando um sistema de vectores com mais do que um vector é linearmente dependente, algum deles se pode escrever à custa dos outros (isto é, "depende" dos outros). Mas, à partida, não se sabe ao certo qual dos vectores dados se pode obter à custa dos restantes, podendo até haver várias possibilidades de o fazer, como vimos. Há, porém, uma situação, muito útil na prática, em que podemos adiantar algo mais:

Teorema 2-17 Suponhamos que os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$, de um dado espaço vectorial, são linearmente independentes, e que os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k, \vec{w}$ são dependentes. Então, o vector $\vec{w}$ pode obter-se como combinação linear de $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$.

Demonstração:
Uma vez que os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k, \vec{w}$ são, por hipótese, dependentes, existirão coeficientes $\alpha_1, \alpha_2, ..., \alpha_k, \beta$, não todos nulos, tais que

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k + \beta\vec{w} = \vec{0}$$

Ora, nesta igualdade, o coeficiente β não pode ser 0, porque, nesse caso, algum dos coeficientes $\alpha_1, \alpha_2, ..., \alpha_k$ seria não nulo e teríamos

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_k \vec{v}_k + \beta \vec{w} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_k \vec{v}_k + 0\vec{w} = \vec{0}$$

ou seja,

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_k \vec{v}_k = \vec{0}$$

com algum dos coeficientes $\alpha_1, \alpha_2, ..., \alpha_k$ não nulo, o que contradiz a hipótese de os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ serem linearmente independentes.

Logo, ter-se-á $\beta \neq 0$, e poderemos concluir que $\vec{w}$ é combinação linear de $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$:

$$\vec{w} = -\frac{\alpha_1}{\beta} \vec{v}_1 - \frac{\alpha_2}{\beta} \vec{v}_2 - \cdots - \frac{\alpha_k}{\beta} \vec{v}_k$$

Finalmente, convém ainda observar que as definições de dependência e independência linear se podem generalizar do seguinte modo:

Teorema 2-18 Os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$, de um dado espaço vectorial, são linearmente independentes se e só se qualquer vector que se possa obter como combinação linear deles tem, nessa expressão, coeficientes únicos.

Demonstração:
A condição enunciada é, obviamente, uma condição suficiente: se cada combinação linear dos vectores dados tem coeficientes fixos, também, em particular, uma combinação linear desses vectores que vá dar $\vec{0}$ terá de ter coeficientes únicos.

Resta, então, provar que a condição é necessária. Para isso, partamos de duas combinações lineares que vão dar o mesmo resultado:

$$\vec{u} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \cdots + \lambda_k \vec{v}_k$$
$$\vec{u} = \mu_1 \vec{v}_1 + \mu_2 \vec{v}_2 + \cdots + \mu_k \vec{v}_k$$

Nestas condições, teremos

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \cdots + \lambda_k \vec{v}_k = \mu_1 \vec{v}_1 + \mu_2 \vec{v}_2 + \cdots + \mu_k \vec{v}_k$$

logo,

$$(\lambda_1 - \mu_1)\vec{v}_1 + (\lambda_2 - \mu_2)\vec{v}_2 + \cdots + (\lambda_k - \mu_k)\vec{v}_k = \vec{0}$$

Caso os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ sejam linearmente independentes, daqui resulta que

$$\begin{cases} \lambda_1 - \mu_1 = 0 \\ \lambda_2 - \mu_2 = 0 \\ \vdots \\ \lambda_k - \mu_k = 0 \end{cases} \quad \text{ou seja, que} \quad \begin{cases} \lambda_1 = \mu_1 \\ \lambda_2 = \mu_2 \\ \vdots \\ \lambda_k = \mu_k \end{cases}$$

Quando, num espaço vectorial, tomamos arbitrariamente um certo número de vectores, o mais natural é que se trate de um sistema linearmente dependente. No entanto, entre esses vectores poderá haver alguns que sejam independentes.

Mais precisamente, consideremos, num espaço vectorial E, um sistema de vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$, não todos nulos.

Supondo, por exemplo, que seja $\vec{v}_1 \neq \vec{0}$, o sistema de vectores formado apenas por $\vec{v}_1$ será linearmente independente, em virtude do Teorema 2-12.

Podemos imaginar duas situações distintas: ou bem que cada um dos vectores $\vec{v}_2, ..., \vec{v}_k$ é proporcional a $\vec{v}_1$, ou então existe um vector, suponhamos que seja $\vec{v}_2$ (uma vez que a ordem por que se encontram escritos é irrelevante), que não se pode obter à custa de $\vec{v}_1$, e, neste último caso, os vectores $\vec{v}_1, \vec{v}_2$ são independentes.

Seguidamente, podemos verificar se haverá um terceiro vector, entre os dados, que não se possa escrever à custa de $\vec{v}_1, \vec{v}_2$. Se, por exemplo, $\vec{v}_3$ estiver nessas condições, os vectores $\vec{v}_1, \vec{v}_2, \vec{v}_3$ serão ainda independentes.

Prosseguindo deste modo, chegaremos a um subsistema linearmente independente **maximal** do sistema dado, isto é, um subsistema do sistema dado que seja linearmente independente mas que se torne dependente se lhe acrescentarmos qualquer outro vector do sistema inicial.

Um subsistema linearmente independente maximal de um dado sistema de vectores goza de uma propriedade fundamental, a saber:

Teorema 2-19

Seja $\vec{v}_1, \vec{v}_2, ..., \vec{v}_r, \vec{v}_{r+1}, ..., \vec{v}_k$ um sistema de vectores, num espaço vectorial E, e suponhamos que os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_r$ constituem um subsistema linearmente independente maximal do sistema considerado. Então, tem-se

$$(\vec{v}_1, \vec{v}_2, ..., \vec{v}_r, \vec{v}_{r+1}, ..., \vec{v}_k) \approx (\vec{v}_1, \vec{v}_2, ..., \vec{v}_r)$$

Demonstração:

É claro que cada vector do sistema $\vec{v}_1, \vec{v}_2, ..., \vec{v}_r$ se pode obter como combinação linear dos vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_r, \vec{v}_{r+1}, ..., \vec{v}_k$ (bastando usar o coeficiente 1 para o vector que se pretenda obter e coeficientes nulos para todos os restantes). Falta, portanto, provar que cada um dos vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_r, \vec{v}_{r+1}, ..., \vec{v}_k$ se pode também escrever à custa de $\vec{v}_1, \vec{v}_2, ..., \vec{v}_r$.

Ora, tomando um vector $\vec{v}_i$, com $i > r$, o sistema de vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_r, \vec{v}_i$ tem mais vectores do que $\vec{v}_1, \vec{v}_2, ..., \vec{v}_r$, pelo que, como este é um subsistema independente maximal do sistema dado, os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_r, \vec{v}_i$ terão de ser linearmente dependentes.

Mas então, como $\vec{v}_1, \vec{v}_2, ..., \vec{v}_r$ são independentes, resulta do Teorema 2-17 que $\vec{v}_i$ é combinação linear de $\vec{v}_1, \vec{v}_2, ..., \vec{v}_r$.

Apresentaremos, a seguir, um importante teorema, que se prende com os conceitos de sistemas de vectores equivalentes e de independência linear. Trata-se de um resultado que, por si mesmo ou mediante os seus corolários, será de utilização permanente em tudo quanto se seguirá.

Teorema 2-20

Teorema de Steinitz

Sejam $\vec{u}_1, \vec{u}_2, ..., \vec{u}_p$ vectores linearmente independentes, de um certo espaço vectorial E. Suponhamos que cada um dos vectores $\vec{u}_i$ $(i = 1, 2, ..., p)$ é combinação linear de certos vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$, fixos em E. Então, é possível substituir p dos vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ pelos vectores $\vec{u}_1, \vec{u}_2, ..., \vec{u}_p$ e obter, desse modo, um sistema equivalente a $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$.

Demonstração:

Faremos a demonstração por indução em p:

a) No caso de ser $p = 1$, teremos um só vector $\vec{u}_1$, linearmente independente, portanto não nulo.

Escrevendo esse vector como combinação linear de $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$,

$$\vec{u}_1 = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_k \vec{v}_k$$

verificamos que algum dos coeficientes utilizados terá de ser diferente de zero (sob pena de se concluir que $\vec{u}_1 = \vec{0}$). Supondo, por exemplo, que seja $\alpha_1 \neq 0$, poderemos então escrever

$$\vec{v}_1 = \frac{1}{\alpha_1} \vec{u}_1 - \frac{\alpha_2}{\alpha_1} \vec{v}_2 - \cdots - \frac{\alpha_k}{\alpha_1} \vec{v}_k$$

Daqui resulta então imediatamente que

$$(\vec{v}_1, \vec{v}_2, ..., \vec{v}_k) \approx (\vec{u}_1, \vec{v}_2, ..., \vec{v}_k)$$

uma vez que os vectores comuns aos dois sistemas se podem facilmente escrever à custa uns dos outros.

b) Suponhamos, agora, que a afirmação feita é válida quando partimos de p vectores independentes, e consideremos os vectores independentes $\vec{u}_1, \vec{u}_2, ..., \vec{u}_p, \vec{u}_{p+1}$, cada um dos quais combinação linear de $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$.

Se desprezarmos o último dos vectores independentes considerados, obteremos o sistema, ainda independente, $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p$. Como cada um destes vectores é combinação linear de $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$, poderemos, de acordo com a hipótese de indução, substituir p dos vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ pelos vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p$ e obter ainda um sistema equivalente a $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$. Como a ordem dos vectores é irrelevante, suponhamos que se tem, por exemplo,

$$(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k) \approx (\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{v}_{p+1}, \ldots, \vec{v}_k)$$

O vector $\vec{u}_{p+1}$ é combinação linear de $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$; logo, será também combinação linear de $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{v}_{p+1}, \ldots, \vec{v}_k$: basta substituir, na expressão de $\vec{u}_{p+1}$ à custa de $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ cada um destes vectores pela sua expressão em função de $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{v}_{p+1}, \ldots, \vec{v}_k$. Obteremos, então,

$$\vec{u}_{p+1} = \lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \cdots + \lambda_p \vec{u}_p + \mu_{p+1} \vec{v}_{p+1} + \cdots + \mu_k \vec{v}_k$$

Ora, nesta igualdade, terá de haver algum coeficiente não nulo entre os coeficientes $\mu_{p+1}, \ldots, \mu_k$, já que, se todos estes coeficientes fossem iguais a zero, a igualdade anterior se transformaria em

$$\vec{u}_{p+1} = \lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \cdots + \lambda_p \vec{u}_p$$

o que contrariaria o facto de os vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}$ serem linearmente independentes.

Supondo, então, que, por exemplo, seja $\mu_{p+1} \neq 0$, poderemos, procedendo como no passo 1) acima, substituir o vector $\vec{v}_{p+1}$ pelo vector $\vec{u}_{p+1}$, obtendo um sistema equivalente a $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{v}_{p+1}, \ldots, \vec{v}_k$:

$$(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{v}_{p+1}, \ldots, \vec{v}_k) \approx (\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}, \ldots, \vec{v}_k)$$

Finalmente, de

$$(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k) \approx (\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{v}_{p+1}, \ldots, \vec{v}_k)$$

e de

$$(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{v}_{p+1}, \ldots, \vec{v}_k) \approx (\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}, \ldots, \vec{v}_k)$$

resulta, por transitividade,

$$(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k) \approx (\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}, \ldots, \vec{v}_k)$$

como queríamos demonstrar.

Do teorema de Steinitz resultam imediatamente alguns importantes corolários, que passamos a apresentar:

Corolário 2-21 Se p vectores linearmente independentes, de um espaço vectorial, se podem escrever como combinações lineares de k vectores quaisquer do espaço, então $p \leq k$.

Demonstração:
Resulta trivialmente da construção indicada no teorema de Steinitz.

Corolário 2-22 Dois sistemas equivalentes de vectores, ambos independentes, têm de ter o mesmo número de vectores.

Demonstração:
Suponhamos que se tem

$$\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_p\right) \approx \left(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_q\right)$$

sendo ambos os sistemas independentes.

Pelo Corolário 2-21, como cada um dos vectores $\vec{a}_i \; (i = 1, 2, \ldots, p)$ é combinação linear de $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_q$, ter-se-á $p \leq q$. Mas, do mesmo modo, como cada vector $\vec{b}_j \; (j = 1, 2, \ldots, q)$ é combinação linear de $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_p$, será $q \leq p$. Assim, concluímos que $p = q$, como se pretendia.

Corolário 2-23 Dois sistemas equivalentes com o mesmo número de vectores têm a mesma natureza, isto é, são ambos dependentes ou ambos independentes.

Demonstração:
Suponhamos que se tem $\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right) \approx \left(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m\right)$, mas que os dois sistemas de vectores são de natureza diferente: digamos que os vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ são independentes e os vectores $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m$ são dependentes.

É claro que, se todos os vectores $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m$ forem nulos, também todos os vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ serão nulos, dada a equivalência entre os dois sistemas, o que vai contra a hipótese. Consequentemente, nem todos os vectores $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m$ serão nulos, e poderemos, portanto, construir, tal como vimos anteriormente, um subsistema independente máximo de $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m$: será, por exemplo, um sistema de vectores $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_r$, com $r < m$ (já que imaginámos os vectores $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m$ dependentes).

Ora, se os vectores independentes $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ se podem escrever como combinações lineares de $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m$, também se poderão escrever como combinações lineares de $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_r$, uma vez que, pelo Teorema 2-19, se tem $\left(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m\right) \approx \left(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_r\right)$.

Mas então, de acordo com o Corolário 2-21, deveria ser $m \leq r < m$, o que é absurdo. Por conseguinte, os dois sistemas de vectores considerados têm de ter a mesma natureza.

Corolário 2-24 Dois subsistemas independentes maximais de um determinado sistema de vectores têm o mesmo número de vectores.

Demonstração:
Sejam $\vec{v}_{i_1}, \vec{v}_{i_2}, ..., \vec{v}_{i_r}$ e $\vec{v}_{j_1}, \vec{v}_{j_2}, ..., \vec{v}_{j_s}$ dois subsistemas independentes maximais do sistema de vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$.

De acordo com o Teorema 2-19, teremos

$$\left(\vec{v}_{i_1}, \vec{v}_{i_2}, ..., \vec{v}_{i_r}\right) \approx \left(\vec{v}_1, \vec{v}_2, ..., \vec{v}_k\right) \quad \text{e} \quad \left(\vec{v}_1, \vec{v}_2, ..., \vec{v}_k\right) \approx \left(\vec{v}_{j_1}, \vec{v}_{j_2}, ..., \vec{v}_{j_s}\right)$$

e, por transitividade,

$$\left(\vec{v}_{i_1}, \vec{v}_{i_2}, ..., \vec{v}_{i_r}\right) \approx \left(\vec{v}_{j_1}, \vec{v}_{j_2}, ..., \vec{v}_{j_s}\right)$$

Ora, como estes dois últimos sistemas de vectores são ambos independentes, resulta do Corolário 2-22 que se tem $r = s$.

O facto de o número de vectores de um subsistema independente maximal de um dado sistema de vectores ser fixo permite que se lhe atribua um nome:

Definição 2-25 Ao número máximo de vectores independentes escolhidos entre os vectores de um dado sistema de vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ chama-se **característica** deste sistema de vectores.

Finalmente, podemos ainda acrescentar o seguinte:

Corolário 2-26 Dois sistemas de vectores equivalentes têm a mesma característica.

Demonstração:
Resulta imediatamente do Teorema 2-19 e do Corolário 2-22.

2-4 Geradores. Bases

Consideremos, por exemplo, os vectores (1,1,0), (1,0,1) e (0,1,1), do espaço vectorial real $\mathbb{R}^3$. É fácil ver que qualquer vector (a,b,c) do espaço se pode obter como combinação linear dos vectores dados.

Com efeito, se considerarmos a igualdade

$$(a,b,c)=\alpha(1,1,0)+\beta(1,0,1)+\gamma(0,1,1)$$

vemos que ela se dará quando

$$\begin{cases}\alpha+\beta=a\\ \alpha+\gamma=b\\ \beta+\gamma=c\end{cases}$$

ora, este sistema de equações é sempre possível, quaisquer que sejam a, b e c, sendo a sua solução

$$\begin{cases}\alpha=\dfrac{1}{2}(a+b-c)\\ \beta=\dfrac{1}{2}(a-b+c)\\ \gamma=\dfrac{1}{2}(-a+b+c)\end{cases}$$

Esta propriedade dos três vectores dados é notável e vamos destacá-la através da seguinte definição:

Definição 2-27 Sejam $\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k$ vectores de um certo espaço vectorial E. Diz-se que esses vectores são **geradores** do espaço (ou que geram o espaço, ou que este é **gerado** por eles), facto que se representa por $E=\langle\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k\rangle$, quando qualquer vector do espaço se pode escrever como combinação linear de $\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k$.

De acordo com o que vimos anteriormente, teremos, por exemplo,

$$\mathbb{R}^3=\langle(1,1,0),(1,0,1),(0,1,1)\rangle$$

É fácil provar o seguinte:

Teorema 2-28 a) Se $E=\langle\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k\rangle$ e $(\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k)\approx(\vec{w}_1,\vec{w}_2,\dots,\vec{w}_t)$, então também $E=\langle\vec{w}_1,\vec{w}_2,\dots,\vec{w}_t\rangle$

b) Reciprocamente, se se tiver $E=\langle\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k\rangle=\langle\vec{w}_1,\vec{w}_2,\dots,\vec{w}_t\rangle$, então $(\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k)\approx(\vec{w}_1,\vec{w}_2,\dots,\vec{w}_t)$

Demonstração:

a) Seja $\vec{u}$ um vector qualquer de E. Como, por hipótese, se tem $E=\langle\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k\rangle$, o vector $\vec{u}$ é combinação linear de $\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k$:

$$\vec{u}=\xi_1\vec{v}_1+\xi_2\vec{v}_2+\cdots+\xi_k\vec{v}_k=\sum_{i=1}^{k}\xi_i\vec{v}_i$$

Por sua vez, por ser $(\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k)\approx(\vec{w}_1,\vec{w}_2,\dots,\vec{w}_t)$, cada vector $\vec{v}_i$ é combinação linear de $\vec{w}_1,\vec{w}_2,\dots,\vec{w}_t$:

$$\vec{v}_i=\sigma_{1i}\vec{w}_1+\sigma_{2i}\vec{w}_2+\cdots+\sigma_{ti}\vec{w}_t=\sum_{j=1}^{t}\sigma_{ji}\vec{w}_j\quad(i=1,2,\dots,k)$$

Substituindo na igualdade anterior, encontramos

$$\vec{u}=\sum_{i=1}^{k}\xi_i\vec{v}_i=\sum_{i=1}^{k}\xi_i\left(\sum_{j=1}^{t}\sigma_{ji}\vec{w}_j\right)=\sum_{j=1}^{t}\sum_{i=1}^{k}\xi_i\sigma_{ji}\vec{w}_j=\sum_{j=1}^{t}\left(\sum_{i=1}^{k}\xi_i\sigma_{ji}\right)\vec{w}_j$$

Verificamos, assim, que $\vec{u}$ é combinação linear de $\vec{w}_1,\vec{w}_2,\dots,\vec{w}_t$, e como $\vec{u}$ é qualquer, concluímos que $E=\langle\vec{w}_1,\vec{w}_2,\dots,\vec{w}_t\rangle$.

b) Resulta imediatamente das definições.

À definição anterior podemos acrescentar a seguinte:

Definição 2-29 Um espaço vectorial E diz-se **finitamente gerado** quando existe uma lista (finita) de vectores $\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k$ tais que $E=\langle\vec{v}_1,\vec{v}_2,\dots,\vec{v}_k\rangle$.

Exemplos 2-30

a) O espaço vectorial $\mathbb{R}^3$ é finitamente gerado, já que, como vimos acima, se tem, por exemplo, $\mathbb{R}^3=\langle(1,1,0),(1,0,1),(0,1,1)\rangle$.

b) O espaço $\mathbb{R}[x]$ não é finitamente gerado. Com efeito, suponhamos que se teria

$$\mathbb{R}[x]=\langle p_1(x),p_2(x),\dots,p_k(x)\rangle$$

e seja m o máximo dos graus dos polinómios $p_1(x), p_2(x), ..., p_k(x)$. Ao efectuarmos combinações lineares destes polinómios, o grau do polinómio obtido não pode ultrapassar m, pelo que um polinómio de grau superior a m não poderia ser obtido à custa dos polinómios considerados, o que contradiz a ideia de eles serem geradores do espaço. ■

Um tipo especial de geradores (caso existam) merece relevo especial: trata-se do caso em que os geradores sejam vectores linearmente independentes:

Definição 2-31 Chama-se **base** de um espaço vectorial E, finitamente gerado, a um sistema $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ de geradores linearmente independentes. Por convenção, o espaço nulo tem como base o conjunto vazio $\varnothing$.

Deve notar-se que no espaço vectorial nulo, $E = \{\vec{0}\}$, não há vectores independentes. No entanto, este espaço é finitamente gerado, já que $E = \{\vec{0}\} = \langle\vec{0}\rangle$.

Teorema 2-32 Todo o espaço finitamente gerado tem base.

Demonstração:
Em virtude do que se disse acima, bastará considerar um espaço vectorial E finitamente gerado não nulo. Suponhamos então que se tem $E = \langle \vec{v}_1, \vec{v}_2, ..., \vec{v}_k \rangle$: uma vez que E não é o espaço nulo, terá de haver, entre os geradores, algum vector diferente de $\vec{0}$.

Tomando, então, um subsistema independente maximal de $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$, digamos $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$, com $n \leq k$, sabemos (pelo Teorema 2-19) que se tem

$$(\vec{v}_1, \vec{v}_2, ..., \vec{v}_k) \approx (\vec{v}_1, \vec{v}_2, ..., \vec{v}_n)$$

e, pelo Teorema 2-28, $E = \langle \vec{v}_1, \vec{v}_2, ..., \vec{v}_n \rangle$. Como estes geradores são independentes, constituem uma base de E.

Por sua vez, do Corolário 2-22 resulta o seguinte:

Teorema 2-33 Duas bases de um mesmo espaço vectorial têm o mesmo número de vectores.

Demonstração:
Sejam $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$ e $\vec{f}_1, \vec{f}_2, ..., \vec{f}_m$ duas bases do espaço vectorial E. Por ser $E = \langle \vec{e}_1, \vec{e}_2, ..., \vec{e}_n \rangle = \langle \vec{f}_1, \vec{f}_2, ..., \vec{f}_m \rangle$, teremos $(\vec{e}_1, \vec{e}_2, ..., \vec{e}_n) \approx (\vec{f}_1, \vec{f}_2, ..., \vec{f}_m)$. Tratando-se de dois sistemas equivalentes e linearmente independentes, o Corolário 2-22 garante que $m = n$.

Pelo facto de ser o mesmo para todas as bases do espaço, o número de vectores de uma base é típico do próprio espaço e merece, por isso, uma designação apropriada:

Definição 2-34 Se E é um espaço vectorial finitamente gerado, o número de vectores de uma base qualquer de E chama-se **dimensão** do espaço e representa-se por $\dim(E)$. Um espaço finitamente gerado também se chama um espaço **de dimensão finita**, enquanto um espaço que não seja finitamente gerado tem **dimensão infinita**.

O conhecimento da dimensão de um dado espaço vectorial (finitamente gerado) reveste-se de alguma importância, em virtude das seguintes propriedades:

Teorema 2-35 Seja E um espaço vectorial de dimensão n.

a) Não há, no espaço E, sistemas de vectores independentes com mais de n vectores.
b) Qualquer sistema com n vectores independentes é uma base de E.
c) Qualquer sistema de n geradores de E é uma base de E.

Demonstração:

Uma vez que $\dim(E) = n$, vamos fixar no espaço uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$.

a) Sejam $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ vectores independentes em E. Por definição de base, cada um destes vectores é combinação linear dos vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$; estamos, portanto, nas condições do teorema de Steinitz.

De acordo com o Corolário 2-21, podemos concluir imediatamente que $p \leq n$.

b) Sejam $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ vectores independentes em E.

Por definição de base, sabemos que cada um destes vectores é combinação linear dos vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$.

Se, por outro lado, tomarmos um dos vectores da base, digamos $\vec{e}_i$ ($i = 1, 2, \ldots, n$), o sistema $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n, \vec{e}_i$ tem mais de n vectores, pelo que, pela parte a) do presente teorema, será linearmente dependente. Como os vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ são independentes, segue-se que $\vec{e}_i$ é combinação linear de $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ (Teorema 2-17).

Concluímos, assim, que $(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n) \approx (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$, pelo que, como os vectores da base são, por definição, geradores do espaço, também $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ gerarão o espaço E (Teorema 2-28). Formarão, portanto, uma base de E.

c) Sejam $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ geradores de E. Pelo Teorema 2-28, podemos concluir que $(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n) \approx (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$, e, pelo Corolário 2-23, os vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ serão independentes; logo, uma base de E.

Quando, num espaço vectorial E, de dimensão n, fixamos uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, cada vector $\vec{v}$ pode escrever-se na forma

$$\vec{v} = \nu_1\vec{e}_1 + \nu_2\vec{e}_2 + \cdots + \nu_n\vec{e}_n$$

uma vez que os vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ são geradores do espaço. Por outro lado, os coeficientes $(\nu_1, \nu_2, \ldots, \nu_n)$ são únicos, pelo facto de os vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ serem linearmente independentes.

Deste modo, cada vector $\vec{v}$ fica perfeitamente identificado pela lista de coeficientes $(\nu_1, \nu_2, \ldots, \nu_n)$, e, como é evidente, cada lista de coeficientes $(\zeta_1, \zeta_2, \ldots, \zeta_n)$ define o vector

$$\vec{z} = \zeta_1\vec{e}_1 + \zeta_2\vec{e}_2 + \cdots + \zeta_n\vec{e}_n$$

Para que não haja ambiguidade, é conveniente considerar os vectores da base escritos por uma determinada ordem. Com efeito, os vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, escritos por outra ordem qualquer, continuam a constituir uma base do espaço E, uma vez que tanto o conceito de independência linear como o facto de gerarem o espaço não dependem da ordem por que os vectores se apresentam. No entanto, duas bases constituídas pelos mesmos vectores, escritos por ordem diferente, considerar-se-ão, daqui para a frente, bases diferentes.

Definição 2-36

Seja $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ uma base (ordenada) do espaço vectorial E, de dimensão n, sobre o corpo K. Escrevendo $\vec{v} = \nu_1\vec{e}_1 + \nu_2\vec{e}_2 + \cdots + \nu_n\vec{e}_n$, os coeficientes da lista $(\nu_1, \nu_2, \ldots, \nu_n)$ tomam o nome de **componentes** do vector $\vec{v}$ em relação à base considerada. Representar-se-á esse facto por

$$\vec{v} \equiv_{(\vec{e}_i)} (\nu_1, \nu_2, \ldots, \nu_n)$$

ou, se a base estiver subentendida,

$$\vec{v} \equiv (\nu_1, \nu_2, \ldots, \nu_n)$$

As observações que antecedem esta definição permitem, desde logo, concluir o seguinte:

Teorema 2-37

Seja E um espaço vectorial, de dimensão n, sobre um corpo K, e seja $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ uma base fixa em E. A aplicação $\omega : E \to K^n$ que, a cada vector de E, faz corresponder a lista das suas componentes em relação à base considerada, é uma bijecção.

Exemplo 2-38

Consideremos, no espaço vectorial real $\mathbb{R}^3$, a base constituída pelos vectores (1,1,0), (1,0,1), (0,1,1), por esta ordem.

Para determinar as componentes de um vector, por exemplo o vector $\vec{v} = (2,-1,5)$, em relação à base considerada, teremos de escrever

$$(2,-1,5) = \alpha\,(1,1,0) + \beta\,(1,0,1) + \gamma\,(0,1,1)$$

de onde tiramos

$$\begin{cases} \alpha + \beta = 2 \\ \alpha + \gamma = -1 \\ \beta + \gamma = 5 \end{cases}, \quad \text{ou seja,} \quad \begin{cases} \alpha = -2 \\ \beta = 4 \\ \gamma = 1 \end{cases}$$

Por consequência: $\vec{v} \equiv (-2,4,1)$. ■

Importa também observar o seguinte:

Consideremos, novamente fixa, no espaço vectorial E, a base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, e suponhamos que, em relação a esta base, se tem

$$\vec{x} \equiv (x_1, x_2, \ldots, x_n) \;\text{ e }\; \vec{y} \equiv (y_1, y_2, \ldots, y_n)$$

Isto significa, então, que

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n \;\text{ e }\; \vec{y} = y_1\vec{e}_1 + y_2\vec{e}_2 + \cdots + y_n\vec{e}_n$$

Daqui resulta que

$$\begin{aligned} \vec{x} + \vec{y} &= \left(x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n\right) + \left(y_1\vec{e}_1 + y_2\vec{e}_2 + \cdots + y_n\vec{e}_n\right) = \\ &= (x_1 + y_1)\vec{e}_1 + (x_2 + y_2)\vec{e}_2 + \cdots + (x_n + y_n)\vec{e}_n \end{aligned}$$

e

$$\alpha\vec{x} = \alpha\left(x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n\right) = (\alpha x_1)\vec{e}_1 + (\alpha x_2)\vec{e}_2 + \cdots + (\alpha x_n)\vec{e}_n$$

Consequentemente, podemos escrever

$$\begin{aligned} \vec{x} + \vec{y} &\equiv (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \\ \alpha\vec{x} &\equiv (\alpha x_1, \alpha x_2, \ldots, \alpha x_n) \end{aligned}$$

Deve notar-se que, deste modo, conhecidas que sejam as componentes de certos vectores em relação a uma dada base, podemos imediatamente calcular as componentes, em relação à mesma base, da soma deles e dos seus produtos por escalares. Por outro lado, as regras para obter essas componentes não são mais nem menos do que as operações que considerámos nos espaços vectoriais da forma K^n. Mais adiante veremos uma interpretação mais profunda deste facto.

Na prática, trabalha-se muitas vezes com espaços vectoriais da forma K^n. Em cada um destes espaços, há uma base especial, chamada a respectiva **base canónica**, que é formada pelos vectores

$$(1,0,0,\ldots,0),\ (0,1,0,\ldots,0),\ (0,0,1,\ldots,0),\ \ldots,\ (0,0,0,\ldots,1)$$

por esta ordem.

Assim, por exemplo, a base canónica de K^2 é formada pelos vectores $(1,0)$, $(0,1)$; a base canónica de K^3 é formada pelos vectores $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, e assim sucessivamente.

É particularmente fácil determinar as componentes de um vector de K^n em relação à respectiva base canónica, já que, como é óbvio, para qualquer vector $\vec{x} = (x_1, x_2, \ldots, x_n) \in K^n$, se tem

$$(x_1, x_2, \ldots, x_n) = x_1(1,0,0,\ldots,0) + x_2(0,1,0,\ldots,0) + x_3(0,0,1,\ldots,0) + \cdots + x_n(0,0,0,\ldots,1)$$

e, por conseguinte,

$$\vec{x} \equiv (x_1, x_2, \ldots, x_n)$$

isto é, a lista das componentes do vector $\vec{x}$ em relação à base canónica coincide com a lista de elementos de K que constitui o próprio vector $\vec{x}$.

Finalmente, observaremos ainda que, do teorema de Steinitz, resulta uma outra propriedade, que usaremos mais tarde, em diversas construções:

Teorema 2-39

Sejam $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ vectores linearmente independentes de um espaço vectorial E, de dimensão $n > k$. Então, é possível construir uma base de E que inclua os vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$.

Demonstração:

Seja $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ uma base qualquer de E. Cada vector $\vec{v}_i$ $(i = 1, 2, \ldots, k)$ é combinação linear dos vectores da base, e como os vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ são, por hipótese, independentes, estamos nas condições de aplicar o teorema de Steinitz.

Consequentemente, será possível substituir k dos vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ pelos vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ e obter um sistema equivalente a $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$. Uma vez que a ordem é, para o efeito, irrelevante, admitamos que são substituídos os primeiros k vectores:

$$(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) \approx (\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k, \vec{e}_{k+1}, \vec{e}_{k+2}, \ldots, \vec{e}_n)$$

O facto de os dois sistemas serem equivalentes e de $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ serem geradores de E garante que $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k, \vec{e}_{k+1}, \vec{e}_{k+2}, \ldots, \vec{e}_n$ são também geradores de E (pelo Teorema 2-28); por outro lado, tratando-se de sistemas equivalentes e com o mesmo número de vectores, e sabendo-se que os vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ são linearmente independentes, o mesmo acontecerá aos vectores

$$\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k, \vec{e}_{k+1}, \vec{e}_{k+2}, \ldots, \vec{e}_n \text{ (pelo Corolário 2-23)}$$

Portanto, os vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k, \vec{e}_{k+1}, \vec{e}_{k+2}, \ldots, \vec{e}_n$ constituem uma base do espaço E, que inclui os vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$, como se pretendia.

3 | Subespaços vectoriais

3-1 | Generalidades

Quando lidamos com os vectores de um certo espaço vectorial E, nem sempre estamos interessados em considerá-los todos: poderemos estar interessados em trabalhar apenas com um determinado conjunto de vectores de E. No entanto, ao efectuarmos operações com vectores do conjunto escolhido poderemos obter resultados que não estejam já nesse conjunto; nessas circunstâncias, não poderíamos trabalhar exclusivamente com o conjunto desejado. Este tipo de considerações leva-nos a introduzir a seguinte definição:

Definição 3-1 Seja E um espaço vectorial sobre um corpo K, e seja F um subconjunto de E. Diz-se que F é um **subespaço vectorial** de E quando:

a) $F \neq \varnothing$

b) $\vec{x},\vec{y} \in F \Rightarrow \vec{x}+\vec{y} \in F$

c) $\vec{x} \in F \Rightarrow \alpha\vec{x} \in F, \ \forall\alpha \in K$

Para indicar que F é um subespaço vectorial de E, podemos escrever $F \leq E$.

Costumamos enunciar a propriedade indicada em b) dizendo que F é "fechado para a adição", e a propriedade indicada em c) dizendo que F é "fechado para a multiplicação por escalares".

Exemplos 3-2

a) Consideremos, no espaço vectorial real $\mathbb{R}^3$, o conjunto

$$F = \{(a,b,0) : a,b \in \mathbb{R}\}$$

É fácil verificar que $F \leq \mathbb{R}^3$:

Em primeiro ligar, F não é vazio, já que, por exemplo, $(2,1,0) \in F$.

Se $\vec{x},\vec{y} \in F$, será, digamos, $\vec{x} = (a,b,0)$, $\vec{y} = (c,d,0)$, e então

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

Finalmente, se $\vec{x} \in F$, será, digamos, $\vec{x} = (a,b,0)$, e então

$$\alpha\vec{x} = \alpha(a,b,0) = (\alpha a,\alpha b,\alpha 0) = (\alpha a,\alpha b,0) \in F$$

b) Consideremos, no espaço vectorial real $\mathbb{R}^4$ os conjuntos

$$G = \left\{(x_1,x_2,x_3,x_4) : 3x_1 - 5x_2 + 7x_3 + 2x_4 = 0\right\}$$

$$H = \left\{(x_1,x_2,x_3,x_4) : x_1 + 2x_2 - x_4 = 0,\ 5x_1 + 4x_3 - x_4 = 0\right\}$$

$$W = \left\{(x_1,x_2,x_3,x_4) : x_1 + x_2 - x_4 = 0,\ x_1 + x_3 - x_4 = 0,\ x_2 - x_3 + x_4 = 0\right\}$$

É fácil verificar que todos eles são subespaços vectoriais de $\mathbb{R}^4$.

c) Consideremos, no espaço vectorial real $\mathbb{R}[x]$, o conjunto $\mathbb{R}_2[x]$, formado pelos polinómios de coeficientes reais até ao segundo grau. Resulta imediatamente das definições que este conjunto constitui um subespaço vectorial de $\mathbb{R}[x]$.

Do mesmo modo, o conjunto $\mathbb{R}_k[x]$, formado pelos polinómios de coeficientes reais até ao grau k (onde k designa um número inteiro não negativo qualquer) constitui um subespaço vectorial de $\mathbb{R}[x]$.

Também o conjunto

$$S = \left\{p(x) \in \mathbb{R}[x] : p''(4) = p(4)\right\}$$

onde $p''(x)$ designa a segunda derivada do polinómio $p(x)$, é um subespaço vectorial de $\mathbb{R}[x]$.

d) Seja E o espaço vectorial formado pelas sucessões de números reais, com as operações usuais.

O conjunto das sucessões convergentes em $\mathbb{R}$, o conjunto das sucessões limitadas, o conjunto dos infinitésimos, são exemplos de subespaços vectoriais de E.

e) Seja E o espaço vectorial formado pelas funções reais definidas num certo intervalo aberto de $\mathbb{R}$.

O conjunto das funções limitadas nesse intervalo, o conjunto das funções contínuas nesse intervalo, o conjunto das funções diferenciáveis nesse intervalo, são exemplos de subespaços vectoriais de E.

f) Qualquer que seja o espaço vectorial E, o conjunto E constitui um subespaço vectorial de si mesmo (chamado o **subespaço impróprio**), e o conjunto $\left\{\vec{0}\right\}$ é também um subespaço vectorial de E (chamado o **subespaço nulo**). ■

Da definição anterior resultam imediatamente duas importantes propriedades dos subespaços vectoriais, a saber:

Teorema 3-3 Seja F um subespaço vectorial do espaço vectorial E, sobre o corpo K. Então:

a) $\vec{0} \in F$

b) $\vec{x} \in F \Rightarrow -\vec{x} \in F$

Demonstração:

a) Como F não é vazio (por definição de subespaço vectorial), tomemos um vector qualquer $\vec{u} \in F$.

Aplicando a propriedade c) da Definição 3-1, com $\alpha = 0$, obtemos

$$\vec{0} = 0\vec{u} \in F$$

b) Resulta imediatamente da Definição 3-1 c), tomando $\alpha = -1$.

Para se verificar que um determinado subconjunto de um espaço vectorial E é um subespaço de E, é necessário verificar as propriedades a) a c) da Definição 3-1. No entanto, estas propriedades podem reduzir-se apenas a duas, o que abrevia as necessárias verificações:

Teorema 3-4

Um subconjunto F, de um espaço vectorial E, sobre um corpo K, é um subespaço vectorial de E se e só se satisfaz as seguintes condições:

a′) $F \neq \varnothing$

b′) $\vec{x}, \vec{y} \in F \Rightarrow \alpha\vec{x} + \beta\vec{y} \in F, \forall \alpha, \beta \in K$

Demonstração:

Se F é um subespaço vectorial de E, então F é não vazio e é fechado para as operações de adição e de multiplicação por um escalar, pelo que, obviamente, satisfaz as condições indicadas.

Reciprocamente, suponhamos que F satisfaz as condições a′) e b′).

Uma vez que a condição a) da Definição 3-1 de subespaço vectorial coincide com a′), resta-nos comprovar que se verificam b) e c).

Ora, b) é um caso particular de b′), com $\alpha = \beta = 1$, enquanto c) é um caso particular de b′), com α qualquer e $\beta = 0$.

A designação de "subespaço vectorial" é justificada pela propriedade enunciada no teorema seguinte:

Teorema 3-5

Seja F um subespaço vectorial do espaço vectorial E, sobre o corpo K. Com as restrições das operações definidas em E, o conjunto F forma um espaço vectorial sobre K.

Demonstração:

As propriedades b) e c) da Definição 3-1 garantem que as operações de adição e multiplicação por um escalar, definidas em E, se podem restringir a F, ficando, portanto, a constituir operações em F.

A validade para F das propriedades que definem um espaço vectorial resulta então trivialmente das mesmas propriedades em E, bem como do Teorema 3-3.

O facto de um subespaço de um espaço vectorial ser, ele próprio, um espaço vectorial permite aplicar-lhe os conceitos estudados no capítulo anterior, e nomeadamente falar em geradores, base e dimensão de um subespaço de um espaço vectorial.

Em particular, tomemos, num espaço vectorial E, sobre um corpo K, um sistema de vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ e consideremos o conjunto de todas as possíveis combinações lineares destes vectores, isto é, o conjunto

$$L = \{\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k : \alpha_1, \alpha_2, ..., \alpha_k \in K\}$$

Podemos verificar facilmente que este conjunto constitui um subespaço vectorial de E. Com efeito, L não é vazio, por construção, e, tomando $\vec{x}, \vec{y} \in L$, teremos

$$\vec{x} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_k\vec{v}_k,\ \vec{y} = \mu_1\vec{v}_1 + \mu_2\vec{v}_2 + \cdots + \mu_k\vec{v}_k$$

pelo que, para quaisquer escalares α e β, se tem

$$\begin{aligned}\alpha\vec{x} + \beta\vec{y} &= \alpha(\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_k\vec{v}_k) + \beta(\mu_1\vec{v}_1 + \mu_2\vec{v}_2 + \cdots + \mu_k\vec{v}_k) = \\ &= (\alpha\lambda_1 + \beta\mu_1)\vec{v}_1 + (\alpha\lambda_2 + \beta\mu_2)\vec{v}_2 + \cdots + (\alpha\lambda_k + \beta\mu_k)\vec{v}_k \in L\end{aligned}$$

Pelo Teorema 3-4, L é um subespaço vectorial de E.

Note-se que, por construção, teremos

$$L = \langle \vec{v}_1, \vec{v}_2, ..., \vec{v}_k \rangle$$

De modo ainda mais geral, consideremos em E um subconjunto qualquer $X \neq \varnothing$ e tomemos o conjunto de todas as possíveis combinações lineares feitas com vectores pertencentes a X:

$$L = \left\{\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k : k \in \mathbb{N},\ \alpha_i \in K\ (i = 1, 2, ..., k),\ \vec{v}_i \in X\ (i = 1, 2, ..., k)\right\}$$

Por um raciocínio inteiramente semelhante ao anterior, verifica-se imediatamente que L é um subespaço vectorial de E.

Este subespaço, construído a partir do conjunto X, tem uma designação natural:

Definição 3-6 | Dado $\varnothing \neq X \subseteq E$, o subespaço vectorial L constituído por todas as combinações lineares feitas com vectores de X chama-se **subespaço gerado** por X, e representa-se por $\langle X \rangle$. Por convenção, tomamos $\langle \varnothing \rangle = \{\vec{0}\}$.

O subespaço gerado por um determinado conjunto goza das seguintes importantes propriedades:

Teorema 3-7 Seja E um espaço vectorial sobre um corpo K, e sejam X e Y subconjuntos quaisquer de E. Então:

a) $X \subseteq \langle X \rangle$

b) Se $F \leq E$ e $X \subseteq F$, então $\langle X \rangle \subseteq F$

c) $X = \langle X \rangle$ se e só se $X \leq E$

d) $\langle \vec{0} \rangle = \langle \{\vec{0}\} \rangle = \{\vec{0}\}$

e) $\langle E \rangle = E$

f) $\langle X \rangle$ é o único subespaço de E que verifica as propriedades indicadas em a) e b)

g) $X \subseteq Y \Rightarrow \langle X \rangle \subseteq \langle Y \rangle$

Demonstração:

a) Resulta imediatamente da construção: no caso de X ser vazio, a afirmação é trivial, e no caso de X não ser vazio basta notar que cada vector de X é combinação linear de si mesmo.

b) Resulta imediatamente da definição e do facto de F ser fechado para as operações.

c) Resulta imediatamente de b).

d) Resulta imediatamente de c).

e) Resulta imediatamente de c).

f) Suponhamos que um certo subespaço vectorial W, do espaço E, verifica as propriedades indicadas em a) e b).

Por se ter $X \subseteq W$ e aplicando a propriedade b), podemos concluir que $\langle X \rangle \subseteq W$. Mas, ao mesmo tempo, $X \subseteq \langle X \rangle$, e como W satisfaz a propriedade b), teremos $W \subseteq \langle X \rangle$. Por conseguinte, $W = \langle X \rangle$.

g) Resulta imediatamente da construção, já que toda a combinação linear de vectores de X se pode considerar uma combinação linear de vectores de Y.

As propriedades a), b) e f) do teorema anterior traduzem-se, geralmente, dizendo que o subespaço vectorial de E gerado por X é o mais pequeno subespaço vectorial de E que contém o conjunto X.

3-2 Construção de subespaços

Vamos agora ver que, a partir de subespaços vectoriais dados, num certo espaço vectorial, é possível construir novos subespaços:

Teorema 3-8

Seja E um espaço vectorial sobre um corpo K.

a) Se F e G são subespaços vectoriais de E, então $F \cap G$ é um subespaço vectorial de E.

b) De modo mais geral, se H_j $(j \in J)$ constitui uma família qualquer de subespaços vectoriais de E, então $H = \bigcap_{j \in J} H_j$ é um subespaço vectorial de E.

Demonstração:

Faremos a demonstração apenas para o caso geral:

Pelo facto de cada H_j ser um subespaço vectorial de E, teremos $\vec{0} \in H_j (j \in J)$, pelo que

$$\vec{0} \in H = \bigcap_{j \in J} H_j. \quad \text{Logo,} \quad H = \bigcap_{j \in J} H_j \neq \varnothing$$

Por outro lado, tomemos

$$\vec{x}, \vec{y} \in H = \bigcap_{j \in J} H_j: \text{ ter-se-á } \vec{x}, \vec{y} \in H_j \ (j \in J)$$

e, como cada H_j é um subespaço vectorial de E, podemos concluir que

$$\alpha\vec{x} + \beta\vec{y} \in H_j, \forall j \in J, \quad \text{pelo que} \quad \alpha\vec{x} + \beta\vec{y} \in H = \bigcap_{j \in J} H_j$$

Se tomarmos um subconjunto X qualquer, num espaço vectorial E, existem sempre subespaços vectoriais de E que contêm o conjunto X, quanto mais não seja, o subespaço impróprio E.

Consequentemente, podemos considerar o conjunto de todos os subespaços vectoriais de E que contêm X:

$$\{H \leq E : X \subseteq H\}$$

De acordo com o Teorema 3-8, a intersecção de todos estes subespaços é ainda um subespaço vectorial de E:

$$W = \bigcap_{X \subseteq H \leq E} H \leq E$$

Por construção, ter-se-á $X \subseteq W$, uma vez que os vectores de X são comuns a todos os subespaços intersectados.

Por outro lado, se F é um subespaço vectorial de E e contém X, então F é um dos subespaços H intersectados, e, por consequência, $W \subseteq F$.

Ora, de acordo com a propriedade enunciada no Teorema 3-7 f), podemos daqui concluir que

$$W = \bigcap_{X \subseteq H \leq E} H = \langle X \rangle$$

Quando se pensa na intersecção de conjuntos, é natural pensar-se também na união. Ora, a união de dois ou mais subespaços do espaço vectorial E não é, necessariamente, um subespaço de E.

Com efeito, se tomarmos, por exemplo, no espaço vectorial $\mathbb{R}^2$, os subespaços

$$F = \langle (1,0) \rangle,\ G = \langle (0,1) \rangle$$

verificamos que

$$(1,0) \in F \subseteq F \cup G,\ (0,1) \in G \subseteq F \cup G \quad \text{logo,} \quad (1,0),(0,1) \in F \cup G$$

mas

$$(1,1) = (1,0) + (0,1) \notin F \cup G \quad \text{uma vez que} \quad (1,1) \notin F, (1,1) \notin G$$

Consequentemente, o conjunto $F \cup G$ não é fechado para a adição; logo, não é um subespaço vectorial de $\mathbb{R}^2$.

Só em casos muito particulares (e triviais) é que a união de dois subespaços dá um subespaço:

Teorema 3-9 Sejam F e G subespaços vectoriais de um espaço vectorial E. A união dos dois subespaços é um subespaço vectorial de E se e só se $F \subseteq G$ ou $G \subseteq F$.

Demonstração:
Se $F \subseteq G$, então $F \cup G = G$, e se $G \subseteq F$, então $F \cup G = F$, pelo que, nesses casos, a união é um subespaço de E.

Reciprocamente, suponhamos, por redução ao absurdo, que $F \cup G$ é um subespaço vectorial de E, mas que nem F está contido em G, nem G está contido em F.

Se F não está contido em G, tomemos um vector $\vec{f} \in F$ tal que $\vec{f} \notin G$; e se G não está contido em F, tomemos um vector $\vec{g} \in G$ tal que $\vec{g} \notin F$.

Os vectores $\vec{f}$ e $\vec{g}$ pertencerão ambos a $F \cup G$, e como este conjunto é, por hipótese, um subespaço vectorial de E, é fechado para a adição, e teremos

$$\vec{s} = \vec{f} + \vec{g} \in F \cup G$$

Se for $\vec{s} \in F$, então, como F é um subespaço,

$$\vec{g} = \vec{s} - \vec{f} \in F$$

o que vai contra a hipótese.

Por outro lado, sendo $\vec{s} \in G$ e atendendo a que G é um subespaço, teremos

$$\vec{f} = \vec{s} - \vec{g} \in G$$

o que contraria igualmente a hipótese.

Temos, portanto, o absurdo que procurávamos.

No caso de lidarmos com a união de mais do que dois subespaços, a união pode dar um subespaço vectorial sem que haja qualquer relação de inclusão entre eles. Por exemplo, para cada número real m, consideremos, no espaço vectorial real $\mathbb{R}^2$, o subespaço $F_m = \langle(1,m)\rangle$ e ainda o subespaço $G = \langle(0,1)\rangle$. Não há qualquer relação de inclusão entre estes diversos subespaços, mas, como é fácil de verificar,

$$G \cup \left(\bigcup_{m \in \mathbb{R}} F_m \right) = \mathbb{R}^2$$

já que, dado $\vec{u} = (u_1, u_2) \in \mathbb{R}^2$, se for $u_1 = 0$, então

$$\vec{u} = (0, u_2) = u_2(0,1) \in G$$

e, se for $u_1 \neq 0$, então

$$\vec{u} = (u_1, u_2) = u_1 \left(1, \frac{u_2}{u_1} \right) \in F_{u_2/u_1}$$

Para termos ainda outro exemplo, consideremos o corpo $K = \{0,1\}$, com as operações definidas da única maneira possível, isto é, através das tabelas

+	0	1
0	0	1
1	1	0

×	0	1
0	0	0
1	0	1

e tomemos o espaço vectorial K^2, sobre K. O espaço K^2 será formado por quatro vectores, a saber:

$$K^2 = \{(0,0), (0,1), (1,0), (1,1)\}$$

É fácil verificar, directamente, que os conjuntos

$$F_1 = \{(0,0), (0,1)\}, \quad F_2 = \{(0,0), (1,0)\}, \quad F_3 = \{(0,0), (1,1)\}$$

são subespaços vectoriais de K^2. Embora não haja qualquer relação de inclusão entre eles, tem-se, obviamente,

$$F_1 \cup F_2 \cup F_3 = K^2$$

Vamos, seguidamente, estudar ainda mais um processo de construir subespaços à custa de outros:

Definição 3-10 Sejam X e Y dois subconjuntos de um certo espaço vectorial E. Designamos por $X+Y$ o conjunto dos vectores que se obtém somando cada vector de X com cada vector de Y, ou seja,

$$X+Y = \{\vec{x}+\vec{y} : \vec{x} \in X,\ \vec{y} \in Y\}$$

A definição anterior aplica-se a dois subconjuntos quaisquer de E, mas, de momento, o que pretendemos é aplicá-la a dois subespaços vectoriais. Podemos provar o seguinte:

Teorema 3-11 Sejam F e G dois subespaços de um certo espaço vectorial E. Então:

a) $F+G \leq E$

b) $F+G = \langle F \cup G \rangle$

c) $F+G = G$ se e só se $F \subseteq G$

Demonstração:

a) Atendendo a que F e G não são vazios, é claro que a sua soma também não fica vazia.

Tomemos então $\vec{u}, \vec{v} \in F+G$ e $\alpha, \beta \in K$. Por definição de $F+G$, teremos

$$\vec{u} = \vec{f}_1 + \vec{g}_1,\ \vec{v} = \vec{f}_2 + \vec{g}_2, \quad \text{com} \quad \vec{f}_1, \vec{f}_2 \in F, \vec{g}_1, \vec{g}_2 \in G$$

Consequentemente,

$$\alpha\vec{u}+\beta\vec{v}=\alpha\left(\vec{f}_1+\vec{g}_1\right)+\beta\left(\vec{f}_2+\vec{g}_2\right)=\left(\alpha\vec{f}_1+\beta\vec{f}_2\right)+\left(\alpha\vec{g}_1+\beta\vec{g}_2\right)$$

e, como F e G são subespaços, portanto fechados para as operações,

$$\alpha\vec{f}_1+\beta\vec{f}_2\in F,\ \alpha\vec{g}_1+\beta\vec{g}_2\in G$$

pelo que $\alpha\vec{u}+\beta\vec{v}\in F+G$ e $F+G$ é um subespaço vectorial de E.

b) Dado um vector $\vec{f}\in F$, tem-se $\vec{f}=\vec{f}+\vec{0}\in F+G$, pelo que $F\subseteq F+G$; analogamente, $G\subseteq F+G$ e, portanto, $F\cup G\subseteq F+G$.

Por outro lado, se W é um subespaço vectorial de E e $F\cup G\subseteq W$, então tanto os vectores de F como os de G pertencem a W; sendo este um subespaço, é fechado para a adição, pelo que $F+G\subseteq W$.

De acordo com o Teorema 3-7 f), podemos concluir que

$$F+G=\langle F\cup G\rangle$$

c) Suponhamos que $F+G=G$. Nestas condições, temos, evidentemente,

$$F\subseteq F\cup G\subseteq F+G=G$$

Reciprocamente, se $F\subseteq G$ então, por definição $F+G\subseteq G$ e, portanto,

$$G\subseteq F\cup G\subseteq F+G\subseteq G$$

pelo que $F+G=G$.

Em relação com o que vimos acima, podemos ainda provar um outro resultado acerca da soma de subespaços vectoriais:

Teorema 3-12 Sejam X e Y subconjuntos quaisquer de um certo espaço vectorial E. Então, tem-se: $\langle X\cup Y\rangle=\langle X\rangle+\langle Y\rangle$.

Em particular, se X e Y são finitos, podemos escrever:

$$\left\langle\vec{u}_1,\vec{u}_2,\dots,\vec{u}_p\right\rangle+\left\langle\vec{v}_1,\vec{v}_2,\dots,\vec{v}_q\right\rangle=\left\langle\vec{u}_1,\vec{u}_2,\dots,\vec{u}_p,\vec{v}_1,\vec{v}_2,\dots,\vec{v}_q\right\rangle$$

Demonstração:
Uma vez que $X\subseteq X\cup Y$ e $Y\subseteq X\cup Y$, teremos

$$\langle X\rangle\subseteq\langle X\cup Y\rangle \text{ e } \langle Y\rangle\subseteq\langle X\cup Y\rangle$$

e, automaticamente,

$$\langle X\rangle + \langle Y\rangle \subseteq \langle X \cup Y\rangle$$

Reciprocamente, consideremos um vector $\vec{z} \in \langle X \cup Y\rangle$. Por construção, $\vec{z}$ será uma combinação linear de vectores de $X \cup Y$, que se pode desdobrar em duas parcelas (eventualmente nulas) juntando numa as parcelas que usem vectores de X, e noutra as parcelas que usem vectores de Y. Assim, teremos $\vec{z} \in \langle X\rangle + \langle Y\rangle$.

3-3 Subespaços de dimensão finita

No caso de um espaço vectorial E ser finitamente gerado, todos os seus subespaços vectoriais gozam da mesma propriedade, conforme vamos ver já em seguida:

Teorema 3-13 Seja F um subespaço vectorial de um espaço vectorial finitamente gerado E. Então, F é finitamente gerado e $\dim(F) \leq \dim(E)$, dando-se aqui a igualdade se e só se $F = E$.

Demonstração:
Se F é o espaço nulo, então F é finitamente gerado: $F = \left\langle \vec{0} \right\rangle$.

No caso de F não ser o espaço nulo, podemos tomar em F um vector $\vec{f}_1 \neq \vec{0}$. Por ser diferente de zero, este vector será independente.

Se for $F = \left\langle \vec{f}_1 \right\rangle$, então F é finitamente gerado; caso contrário, e uma vez que $\left\langle \vec{f}_1 \right\rangle \subseteq F$, existirá um vector $\vec{f}_2 \in F$ tal que $\vec{f}_2 \notin \left\langle \vec{f}_1 \right\rangle$. Nestas circunstâncias, $\vec{f}_2$ não se poderá obter à custa de $\vec{f}_1$, pelo que os vectores $\vec{f}_1, \vec{f}_2$ serão independentes.

Se for $F = \left\langle \vec{f}_1, \vec{f}_2 \right\rangle$, então F é finitamente gerado; se não, existirá um vector $\vec{f}_3 \in F$ tal que $\vec{f}_3 \notin \left\langle \vec{f}_1, \vec{f}_2 \right\rangle$, e os vectores $\vec{f}_1, \vec{f}_2, \vec{f}_3$ serão independentes.

Prosseguindo este raciocínio, ou bem que chegamos a um sistema finito de geradores de F, ou bem que construímos um sistema de vectores independentes com um número cada vez maior de vectores.

Ora, se o espaço vectorial E tem dimensão finita n, o Teorema 2-35 assegura-nos que não é possível que exista nesse espaço um sistema de vectores independentes com mais de n vectores. Consequentemente, chegaremos forçosamente a um sistema finito de geradores para F.

Resulta então novamente do Teorema 2-35 que $\dim(F) \leq \dim(E)$, já que qualquer base de F é um sistema de vectores independentes de E.

Finalmente, se for $\dim(F) = \dim(E) = n$, uma base de F será um sistema de n vectores independentes em E; logo, uma base de E, pelo que $F = E$.

É claro que, quer um espaço vectorial tenha dimensão finita, quer não, podemos considerar nele subespaços vectoriais finitamente gerados.

Dos Teoremas 3-12 e 3-13 resulta imediatamente que se F e G são subespaços finitamente gerados de um certo espaço vectorial, também $F \cap G$ e $F + G$ são finitamente gerados. Com efeito, $F \cap G$ é um subespaço vectorial de F e de G, pelo que terá de ter uma dimensão não superior às destes; quanto a $F + G$, ele será gerado pela união de um sistema (que pode ser finito) de geradores de F com um sistema (que pode ser finito) de geradores de G.

Ora, há uma relação interessante entre as dimensões de F, G, $F \cap G$ e $F + G$, a saber:

Teorema 3-14

Teorema das dimensões

Sejam F e G subespaços vectoriais finitamente gerados de um certo espaço vectorial E. Tem-se:

$$\dim(F+G) = \dim(F) + \dim(G) - \dim(F \cap G)$$

Demonstração:

Comecemos por fixar uma base $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_r$ no subespaço $F \cap G$. De acordo com o Teorema 2-39 e atendendo a que $F \cap G \subseteq F$ e $F \cap G \subseteq G$, poderemos construir bases de F e G que incluam os vectores da base de $F \cap G$: sejam, por exemplo, $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_r, \vec{f}_1, \vec{f}_2, \ldots, \vec{f}_s$ uma base de F, e $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_r, \vec{g}_1, \vec{g}_2, \ldots, \vec{g}_t$ uma base de G.

Tal como vimos no Teorema 3-12, teremos, nestas circunstâncias,

$$F + G = \left\langle \vec{w}_1, \ldots, \vec{w}_r, \vec{f}_1, \ldots, \vec{f}_s, \vec{w}_1, \ldots, \vec{w}_r, \vec{g}_1, \ldots, \vec{g}_t \right\rangle = \left\langle \vec{w}_1, \ldots, \vec{w}_r, \vec{f}_1, \ldots, \vec{f}_s, \vec{g}_1, \ldots, \vec{g}_t \right\rangle$$

Se provarmos que os vectores $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_r, \vec{f}_1, \vec{f}_2, \ldots, \vec{f}_s, \vec{g}_1, \vec{g}_2, \ldots, \vec{g}_t$ são linearmente independentes, eles constituirão uma base de $F + G$. Esse facto completará a demonstração do teorema! Com efeito, a dimensão de $F + G$ será então igual a $r + s + t$, e, por outro lado,

$$\dim(F) + \dim(G) - \dim(F \cap G) = (r+s) + (r+t) - r = r + s + t$$

chegando-se à igualdade desejada.

Para vermos que os vectores $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_r, \vec{f}_1, \vec{f}_2, \ldots, \vec{f}_s, \vec{g}_1, \vec{g}_2, \ldots, \vec{g}_t$ são independentes, consideremos uma combinação linear nula

$$\lambda_1 \vec{w}_1 + \lambda_2 \vec{w}_2 + \cdots + \lambda_r \vec{w}_r + \mu_1 \vec{f}_1 + \mu_2 \vec{f}_2 + \cdots + \mu_s \vec{f}_s + \sigma_1 \vec{g}_1 + \sigma_2 \vec{g}_2 + \cdots + \sigma_t \vec{g}_t = \vec{0}$$

Tomemos o vector auxiliar

$$\vec{z} = \lambda_1 \vec{w}_1 + \lambda_2 \vec{w}_2 + \cdots + \lambda_r \vec{w}_r + \mu_1 \vec{f}_1 + \mu_2 \vec{f}_2 + \cdots + \mu_s \vec{f}_s$$

Por construção, ter-se-á $\vec{z} \in F$.

Mas da igualdade anterior resulta que

$$\vec{z} = -\sigma_1\vec{g}_1 - \sigma_2\vec{g}_2 - \cdots - \sigma_t\vec{g}_t \in G$$

Logo, $\vec{z} \in F \cap G$, e, como $\vec{w}_1, \vec{w}_2, ..., \vec{w}_r$ é uma base de $F \cap G$, poderemos escrever

$$\vec{z} = \eta_1\vec{w}_1 + \eta_2\vec{w}_2 + \cdots + \eta_r\vec{w}_r$$

Igualando as duas últimas expressões de $\vec{z}$, obtemos

$$\eta_1\vec{w}_1 + \eta_2\vec{w}_2 + \cdots + \eta_r\vec{w}_r = -\sigma_1\vec{g}_1 - \sigma_2\vec{g}_2 - \cdots - \sigma_t\vec{g}_t$$

ou seja,

$$\eta_1\vec{w}_1 + \eta_2\vec{w}_2 + \cdots + \eta_r\vec{w}_r + \sigma_1\vec{g}_1 + \sigma_2\vec{g}_2 + \cdots + \sigma_t\vec{g}_t = \vec{0}$$

Ora, os vectores $\vec{w}_1, \vec{w}_2, ..., \vec{w}_r, \vec{g}_1, \vec{g}_2, ..., \vec{g}_t$ são independentes, visto formarem uma base de G, pelo que podemos daqui concluir que

$$\eta_1 = \eta_2 = \cdots = \eta_r = \sigma_1 = \sigma_2 = \cdots = \sigma_t = 0$$

Nestas condições, a igualdade considerada inicialmente reduz-se apenas a

$$\lambda_1\vec{w}_1 + \lambda_2\vec{w}_2 + \cdots + \lambda_r\vec{w}_r + \mu_1\vec{f}_1 + \mu_2\vec{f}_2 + \cdots + \mu_s\vec{f}_s = \vec{0}$$

e, como os vectores $\vec{w}_1, \vec{w}_2, ..., \vec{w}_r, \vec{f}_1, \vec{f}_2, ..., \vec{f}_s$ também são independentes (por serem base de F), concluímos, finalmente, que

$$\lambda_1 = \lambda_2 = \cdots = \lambda_r = \mu_1 = \mu_2 = \cdots = \mu_s = 0$$

Do teorema das dimensões resulta, trivialmente, o seguinte corolário:

Corolário 3-15 Nas condições do teorema anterior, tem-se

$$\dim(F+G) = \dim(F) + \dim(G)$$

se e só se $F \cap G = \{\vec{0}\}$.

Demonstração:
De acordo com o teorema das dimensões, será

$$\dim(F+G)=\dim(F)+\dim(G)$$

se e só se $\dim(F\cap G)=0$, o que, por sua vez, equivale a $F\cap G=\{\vec{0}\}$.

A situação considerada no Corolário 3-15 é, como veremos, particularmente importante, pelo que é conveniente distingui-la, com uma designação apropriada, que, de resto, apresentaremos num enquadramento mais geral:

Definição 3-16 Sejam F e G dois subespaços vectoriais quaisquer de um espaço vectorial E. Diz-se que a soma $S=F+G$ é uma **soma directa**, o que se representa por $S=F\oplus G$, quando $F\cap G=\{\vec{0}\}$.

O facto de a soma de dois subespaços ser directa pode reconhecer-se de diversas maneiras. Na verdade, podemos, de imediato, demonstrar o seguinte:

Teorema 3-17 Sejam F e G dois subespaços vectoriais de um espaço vectorial E, e consideremos a sua soma $S=F+G$. São equivalentes as seguintes afirmações:

a) $S=F\oplus G$
b) $F\cap G=\{\vec{0}\}$
c) Cada vector de S se escreve de uma única maneira, como soma de um vector de F com um vector de G.
d) $\vec{0}$ escreve-se de uma única maneira, como soma de um vector de F com um vector de G.

Demonstração:
A equivalência entre a) e b) resulta da própria definição de soma directa.

Vamos provar que c) resulta de b):

Suponhamos que se tem

$$\vec{s}=\vec{f}+\vec{g}=\vec{f}'+\vec{g}', \text{ com } \vec{f},\vec{f}'\in F \text{ e } \vec{g},\vec{g}'\in G$$

Então,

$$\vec{d}=\vec{f}-\vec{f}'\in F$$

mas, ao mesmo tempo,

$$\vec{d}=\vec{g}'-\vec{g}\in G$$

Consequentemente, $\vec{d} \in F \cap G = \{\vec{0}\}$, ou seja, $\vec{d} = \vec{0}$, o que significa que $\vec{f} = \vec{f}'$ e $\vec{g}' = \vec{g}$.

Uma vez que d) é um caso particular de c), completaremos a demonstração provando que, de d), se deduz b):

Consideremos um vector $\vec{u} \in F \cap G$. Atendendo a que o vector $\vec{0}$ pertence a ambos os subespaços, podemos escrever

$$\vec{u} = \vec{u} + \vec{0} \in F + G, \ \vec{u} = \vec{0} + \vec{u} \in F + G$$

Dada a unicidade da decomposição de $\vec{u}$ como soma de um vector de F com um vector de G, concluímos que $\vec{u} = \vec{0}$. Como $\vec{u}$ é um vector arbitrário de $F \cap G$, podemos afirmar que $F \cap G = \{\vec{0}\}$.

No caso de lidarmos com subespaços de dimensão finita, podemos, de acordo com os resultados anteriores, acrescentar ainda o seguinte:

Teorema 3-18 Se F e G são subespaços vectoriais de dimensão finita de um espaço vectorial E, e $S = F + G$, então são equivalentes as seguintes afirmações:

a) $S = F \oplus G$

b) $\dim(F + G) = \dim(F) + \dim(G)$

c) Se $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_s$ é uma base de F e $\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_t$ é uma base de G, então $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_s, \vec{g}_1, \vec{g}_2, \ldots, \vec{g}_t$ é uma base de $F + G$.

Demonstração:

Resulta imediatamente do que foi visto anteriormente.

A noção de soma directa permite-nos, ainda, a introdução de um outro conceito, a saber:

Definição 3-19 Seja F um subespaço vectorial de um espaço vectorial E. Diz-se que um subespaço vectorial F^*, do espaço E, é um **subespaço complementar** de F quando se tem $E = F \oplus F^*$.

No caso de se lidar com um espaço vectorial de dimensão finita, é possível garantir, para cada um dos seus subespaços, a existência de um subespaço complementar, conforme veremos já no próximo teorema:

Teorema 3-20 Todo o subespaço de um espaço vectorial de dimensão finita tem pelo menos um subespaço complementar.

Demonstração:

Seja F um subespaço vectorial de um espaço vectorial E, de dimensão n. É claro que, se for $F = E$, o seu (único) complementar será o subespaço $F^* = \{\vec{0}\}$;

reciprocamente, o único complementar de $F=\{\vec{0}\}$ é $F^*=E$. Consideraremos então que o subespaço F não é E nem $\{\vec{0}\}$: o subespaço vectorial F terá também dimensão finita p, tal que $0<p<n$.

Fixemos em F uma base $\vec{f}_1,\vec{f}_2,\ldots,\vec{f}_p$. De acordo com o Teorema 2-39, podemos construir uma base de E que inclua estes vectores, por exemplo,

$$\vec{f}_1,\vec{f}_2,\ldots,\vec{f}_p,\vec{e}_{p+1},\vec{e}_{p+2},\ldots,\vec{e}_n$$

Tomando, então, o subespaço $F^*=\langle\vec{e}_{p+1},\vec{e}_{p+2},\ldots,\vec{e}_n\rangle$, resulta imediatamente do Teorema 3-18 que $E=F\oplus F^*$, ou seja, que F^* é um subespaço complementar de F em E.

A noção de soma directa pode alargar-se à soma de mais que dois subespaços, mediante a definição seguinte:

Definição 3-21 Sejam $F_1,F_2,\ldots,F_t$ subespaços de um espaço vectorial E, e seja

$$S=F_1+F_2+\cdots+F_t$$

Diz-se que esta soma é uma **soma directa**, o que se representa por

$$S=F_1\oplus F_2\oplus\cdots\oplus F_t$$

quando, para cada $i=1,2,\ldots,t$, se tem

$$F_i\cap\left(F_1+F_2+\cdots+F_{i-1}+F_{i+1}+\cdots+F_t\right)=F_i\cap\sum_{j\neq i}F_j=\{\vec{0}\}$$

A demonstração do seguinte teorema é inteiramente análoga às dos teoremas correspondentes, para o caso de dois subespaços, pelo que a omitiremos:

Definição 3-22 Sejam $F_1,F_2,\ldots,F_t$ subespaços de um espaço vectorial E, e seja

$$S=F_1+F_2+\cdots+F_t$$

São equivalentes as afirmações seguintes:

a) $S=F_1\oplus F_2\oplus\cdots\oplus F_t$

b) $F_i\cap\left(F_1+F_2+\cdots+F_{i-1}+F_{i+1}+\cdots+F_t\right)=\{\vec{0}\}$, $\forall i=1,2,\ldots,t$

c) Cada vector de S se escreve de uma única maneira como soma de vectores de cada um dos subespaços F_i $(i=1,2,\ldots,t)$.

d) O vector $\vec{0}$ escreve-se de uma única maneira como soma de vectores de cada um dos subespaços F_i $(i=1,2,\ldots,t)$.

No caso de cada um dos subespaços F_i $(i = 1,2,...,t)$ considerados ter dimensão finita, as afirmações anteriores são ainda equivalentes a:

e) $\dim(S) = \dim(F_1) + \dim(F_2) + \cdots + \dim(F_t)$

f) A justaposição de bases dos diferentes subespaços F_i $(i = 1,2,...,t)$ dá uma base de S.

3-4 Congruências e espaços-quociente

Quando, num conjunto A qualquer, definimos uma relação de equivalência R, isto é, uma relação que verifica as propriedades reflexiva, simétrica e transitiva, podemos associar, a cada elemento $a \in A$, a respectiva "classe de equivalência", que é o conjunto dos elementos de A que estão em relação com a:

$$[a]_R = \{x \in A : xRa\}$$

É bem sabido que o conjunto A fica partido em classes de equivalência (que são conjuntos não vazios, disjuntos dois a dois, e cuja união reproduz o conjunto A), que constituem o chamado "conjunto-quociente" de A por R:

$$\frac{A}{R} = \{[a]_R : a \in A\}$$

Num espaço vectorial, há um tipo particular de relações de equivalência, definidas à custa dos respectivos subespaços, que vamos estudar, começando por apresentar esta definição:

Definição 3-23 Seja F um subespaço vectorial de um espaço vectorial qualquer E. Chama-se **congruência associada** a F à relação R_F definida, em E, do seguinte modo:

$$\vec{x}R_F\vec{y} \text{ se e só se } \vec{x} - \vec{y} \in F$$

As relações da forma R_F são relações de equivalência e gozam ainda de outras propriedades notáveis:

Teorema 3-24 Seja F um subespaço vectorial de um espaço vectorial E. A relação R_F satisfaz as seguintes propriedades:

a) $\vec{x}R_F\vec{x}$, $\forall \vec{x} \in E$ (propriedade reflexiva)

b) Se $\vec{x}R_F\vec{y}$, então $\vec{y}R_F\vec{x}$ (propriedade simétrica)

c) Se $\vec{x}R_F\vec{y}$ e $\vec{y}R_F\vec{z}$, então $\vec{x}R_F\vec{z}$ (propriedade transitiva)

d) Se $\vec{x}R_F\vec{y}$ e $\vec{x}'R_F\vec{y}'$, então $\vec{x}+\vec{x}'R_F\vec{y}+\vec{y}'$
e) Se $\vec{x}R_F\vec{y}$, então $\alpha\vec{x}R_F\alpha\vec{y}$, qualquer que seja o escalar α.

Demonstração:
a) Tem-se, evidentemente, $\vec{x}-\vec{x}=\vec{0}\in F$, pelo que $\vec{x}R_F\vec{x}$.
b) Se $\vec{x}R_F\vec{y}$, tem-se $\vec{x}-\vec{y}\in F$. Mas, então,

$$\vec{y}-\vec{x}=-(\vec{x}-\vec{y})\in F$$

pelo que também $\vec{y}R_F\vec{x}$.
c) Supondo que $\vec{x}R_F\vec{y}$ e $\vec{y}R_F\vec{z}$, então $\vec{x}-\vec{y}\in F$ e $\vec{y}-\vec{z}\in F$. Mas, então,

$$\vec{x}-\vec{z}=(\vec{x}-\vec{y})+(\vec{y}-\vec{z})\in F$$

pelo que $\vec{x}R_F\vec{z}$.
d) Partindo de $\vec{x}R_F\vec{y}$ e $\vec{x}'R_F\vec{y}'$, temos $\vec{x}-\vec{y}\in F$ e $\vec{x}'-\vec{y}'\in F$. Consequentemente,

$$(\vec{x}+\vec{x}')-(\vec{y}+\vec{y}')=(\vec{x}-\vec{y})+(\vec{x}'-\vec{y}')\in F$$

pelo que $\vec{x}+\vec{x}'R_F\vec{y}+\vec{y}'$.
e) Sendo $\vec{x}R_F\vec{y}$, temos $\vec{x}-\vec{y}\in F$, portanto

$$\alpha\vec{x}-\alpha\vec{y}=\alpha(\vec{x}-\vec{y})\in F \quad \text{e, assim,} \quad \alpha\vec{x}R_F\alpha\vec{y}.$$

As propriedades a) a c) do Teorema 3-24 dizem-nos que a relação R_F é uma relação de equivalência. Sabemos, portanto, que cada vector de E define uma determinada classe de equivalência para esta relação. Podemos identificar com toda a precisão as classes de equivalência, graças à próxima proposição:

Teorema 3-25 Nas condições do teorema anterior, a classe de cada vector $\vec{u}\in E$ é

$$[\vec{u}]_{R_F}=\vec{u}+F=\{\vec{u}+\vec{x}:\vec{x}\in F\}$$

Demonstração:
Dado um vector $\vec{y}=\vec{u}+\vec{x}$, com $\vec{x}\in F$, temos, evidentemente,

$$\vec{y}-\vec{x}=\vec{u}\in F$$

pelo que $\vec{y}R_F\vec{x}$, e, por conseguinte, $\vec{y}\in[\vec{u}]_{R_F}$.

Reciprocamente, se $\vec{y}\in[\vec{u}]_{R_F}$, será $\vec{y}R_F\vec{x}$, ou seja, $\vec{y}-\vec{x}\in F$. Escrevendo $\vec{y}-\vec{x}=\vec{u}$, obtemos $\vec{y}=\vec{u}+\vec{x}$, com $\vec{x}\in F$, pelo que $\vec{y}\in\vec{u}+F$.

Do teorema anterior resulta, em particular, que

$$[\vec{0}]_{R_F} = \vec{0} + F = F$$

O conjunto-quociente de E pela relação R_F representa-se, abreviadamente, por $\dfrac{E}{R_F} = \dfrac{E}{F}$. Temos, portanto,

$$\frac{E}{F} = \left\{F,\ \vec{a} + F,\ \vec{b} + F,\ \dots\right\} = \{\vec{x} + F : \vec{x} \in F\}$$

Definição 3-26 Uma relação de equivalência, definida num espaço vectorial E, que satisfaça as propriedades d) e e) do Teorema 3-24, toma o nome de **relação de congruência**.

Os resultados anteriores mostram, portanto, que a cada subespaço vectorial de um espaço vectorial E podemos associar uma relação de congruência, coincidindo o subespaço em causa com a classe de equivalência do vector nulo. Ora, reciprocamente, podemos mostrar que toda a relação de congruência é precisamente desse tipo. É o que enunciamos já a seguir:

Teorema 3-27 Seja R uma relação de congruência num espaço vectorial E, isto é, uma relação que verifique as propriedades a) a e) do Teorema 3-24. Então:

a) $H = [\vec{0}]_R$ é um subespaço vectorial de E.

b) $R_H = R$.

Demonstração:

a) Tem-se $H \neq \varnothing$, porque, pela propriedade reflexiva, $\vec{0}R\vec{0}$; logo, $\vec{0} \in H$.

Tomemos, então, $\vec{x}, \vec{y} \in H$ e os escalares α e β. Por definição de "classe de equivalência", ter-se-á $\vec{x}R\vec{0}$ e $\vec{y}R\vec{0}$.

Da propriedade e) resulta, então, que

$$\alpha\vec{x}R\alpha\vec{0} \quad \text{e} \quad \beta\vec{y}R\beta\vec{0} \quad \text{isto é,} \quad \alpha\vec{x}R\vec{0} \quad \text{e} \quad \beta\vec{y}R\vec{0}$$

mas, então, da propriedade d) resulta que

$$(\alpha\vec{x} + \beta\vec{y})R\left(\vec{0} + \vec{0}\right) \quad \text{ou seja,} \quad (\alpha\vec{x} + \beta\vec{y})R\vec{0}$$

Portanto, $\alpha\vec{x} + \beta\vec{y} \in H$, o que garante que H é um subespaço vectorial de E.

b) Suponhamos que se tem $\vec{x}R\vec{y}$.

Pela propriedade reflexiva, sabemos que se tem também $-\vec{y}R-\vec{y}$, e, utilizando a propriedade d), obtemos

$$(\vec{x}-\vec{y})R(\vec{y}-\vec{y}) \quad \text{ou seja,} \quad (\vec{x}-\vec{y})R\vec{0}$$

pelo que $\vec{x}-\vec{y}\in H$ e, por definição, $\vec{x}R_H\vec{y}$.

Reciprocamente, se $\vec{x}R_H\vec{y}$, temos $\vec{x}-\vec{y}\in H$, ou seja, $(\vec{x}-\vec{y})R\vec{0}$.

Ora, como $\vec{y}R\vec{y}$, resulta da propriedade d) que $\vec{x}R\vec{y}$.

O que mais nos interessa salientar é que as propriedades d) e e) das relações de congruência nos permitem introduzir no conjunto-quociente

$$\frac{E}{F}=\{\vec{x}+F:\vec{x}\in F\}$$

operações que lhe confiram a estrutura de espaço vectorial.

Para esse efeito, definimos a adição e a multiplicação por um escalar da seguinte maneira:

$$(\vec{x}+F)+(\vec{y}+F)=(\vec{x}+\vec{y})+F \qquad \alpha(\vec{x}+F)=(\alpha\vec{x})+F$$

Convém observar que se pretende definir operações com as classes de equivalência (ou, neste caso, de congruência), mas estas são definidas à custa de representantes. É, por isso, necessário verificar se as definições apresentadas são coerentes, isto é, se conduzem aos mesmos resultados se, para as mesmas classes, usarmos representantes diferentes.

Ora, suponhamos que se tem

$$\vec{x}+F=\vec{x}'+F \quad \text{e} \quad \vec{y}+F=\vec{y}'+F$$

Isto significa que

$$\vec{x}-\vec{x}'\in F \text{ e } \vec{y}-\vec{y}'\in F \quad \text{pelo que} \quad (\vec{x}+\vec{x}')-(\vec{y}+\vec{y}')=(\vec{x}-\vec{x}')+(\vec{y}-\vec{y}')\in F$$

e, portanto,

$$(\vec{x}+\vec{x}')R_F(\vec{y}+\vec{y}') \quad \text{ou seja,} \quad (\vec{x}+\vec{x}')+F=(\vec{y}+\vec{y}')+F$$

Isto mostra que a maneira como definimos a operação de adição de classes é coerente: o resultado não se altera quando mudamos os representantes das classes.

Analogamente, de

$$\vec{x}+F=\vec{x}'+F \quad \text{resulta} \quad \vec{x}-\vec{x}'\in F$$

logo,

$$\alpha\vec{x}-\alpha\vec{x}'=\alpha\left(\vec{x}-\vec{x}'\right)\in F \quad \text{pelo que} \quad \left(\alpha\vec{x}\right)+F=\left(\alpha\vec{x}'\right)+F$$

É agora muito fácil verificar que, com as operações que acabamos de definir, se obtém, de facto, um espaço vectorial:

Teorema 3-28 Seja F um subespaço de um espaço vectorial E, sobre um corpo K. Com as operações definidas acima no conjunto-quociente E/F, obtém-se um espaço vectorial sobre K.

Demonstração:
Trata-se, evidentemente, de verificar as oito propriedades que caracterizam a estrutura de espaço vectorial:

a) $(\vec{x}+F)+(\vec{y}+F)=(\vec{x}+\vec{y})+F=(\vec{y}+\vec{x})+F=(\vec{y}+F)+(\vec{x}+F)$

b) $$\begin{aligned}\big((\vec{x}+F)+(\vec{y}+F)\big)+(\vec{z}+F)&=\big((\vec{x}+\vec{y})+F\big)+(\vec{z}+F)=\\&=\big((\vec{x}+\vec{y})+\vec{z}\big)+F=\big(\vec{x}+(\vec{y}+\vec{z})\big)+F=\\&=(\vec{x}+F)+\big((\vec{y}+\vec{z})+F\big)=\\&=(\vec{x}+F)+\big((\vec{y}+F)+(\vec{z}+F)\big)\end{aligned}$$

c) $(\vec{x}+F)+\left(\vec{0}+F\right)=\left(\vec{x}+\vec{0}\right)+F=\vec{x}+F$

d) $(\vec{x}+F)+(-\vec{x}+F)=(\vec{x}-\vec{x})+F=\vec{0}+F$

e) $$\begin{aligned}\alpha\big((\vec{x}+F)+(\vec{y}+F)\big)&=\alpha\big((\vec{x}+\vec{y})+F\big)=\big(\alpha(\vec{x}+\vec{y})\big)+F=\\&=(\alpha\vec{x}+\alpha\vec{y})+F=(\alpha\vec{x}+F)+(\alpha\vec{y}+F)=\alpha(\vec{x}+F)+\alpha(\vec{y}+F)\end{aligned}$$

f) $$\begin{aligned}(\alpha+\beta)(\vec{x}+F)&=\big((\alpha+\beta)\vec{x}\big)+F=(\alpha\vec{x}+\beta\vec{x})+F=\\&=(\alpha\vec{x}+F)+(\beta\vec{x}+F)=\alpha(\vec{x}+F)+\beta(\vec{x}+F)\end{aligned}$$

g) $\alpha\big(\beta(\vec{x}+F)\big)=\alpha\big((\beta\vec{x})+F\big)=\alpha(\beta\vec{x})+F=\big((\alpha\beta)\vec{x}\big)+F=(\alpha\beta)(\vec{x}+F)$

h) $1(\vec{x}+F)=(1\vec{x})+F=\vec{x}+F$.

Na construção do espaço-quociente, há duas situações particulares que convém examinar. Em primeiro lugar, consideremos o quociente E/E.

Deve observar-se que a relação R_E, definida como vimos acima, é, neste caso, a relação universal, já que, quaisquer que sejam $\vec{x},\vec{y}\in E$, se tem, obviamente,

$$\vec{x}-\vec{y}\in E \quad \text{ou seja,} \quad \vec{x}R_E\vec{y}$$

Por consequência, o espaço-quociente E/E tem uma classe só:

$$\frac{E}{E}=\left\{\vec{0}+E\right\}=\{E\}$$

Trata-se, portanto, de um espaço vectorial com um só vector, isto é, do espaço vectorial nulo.

Por outro lado, examinemos o quociente $\dfrac{E}{\{\vec{0}\}}$:

Para se ter

$$\vec{x} R_{\{\vec{0}\}} \vec{y}$$

terá de ser

$$\vec{x} - \vec{y} \in \{\vec{0}\} \quad \text{ou seja,} \quad \vec{x} = \vec{y}$$

Quer isto dizer que, no quociente $E/_{\{\vec{0}\}}$, cada classe é constituída por um só vector: a classe de cada vector $\vec{u}$ é $\vec{u} + \{\vec{0}\} = \{\vec{u}\}$. Consequentemente, sendo

$$E = \{\vec{0},\ \vec{a},\ \vec{b},\ \vec{c}, \ldots\}$$

será

$$\frac{E}{\{\vec{0}\}} = \left\{\{\vec{0}\}, \{\vec{a}\}, \{\vec{b}\}, \{\vec{c}\}, \ldots\right\}$$

Atendendo ao modo como se definiram as operações no quociente, será

$$\{\vec{x}\} + \{\vec{y}\} = \{\vec{x} + \vec{y}\} \ \text{ e } \ \alpha\{\vec{x}\} = \{\alpha\vec{x}\}$$

pelo que podemos identificar o espaço-quociente $E/_{\{\vec{0}\}}$ com o próprio espaço E.

Sucede muitas vezes que o espaço vectorial quociente tem propriedades interessantes, quando o mesmo não acontecia com o espaço inicialmente considerado. Por outro lado, certas propriedades do espaço dado transmitem-se aos seus quocientes, como, por exemplo, a seguinte:

Teorema 3-29 Seja F um subespaço de um espaço vectorial E. Se o espaço E tem dimensão finita, então o espaço-quociente $E/_F$ também tem dimensão finita, sendo

$$\dim\left(\frac{E}{F}\right) = \dim(E) - \dim(F)$$

Demonstração:

Suponhamos que E tem dimensão n; nesse caso, F terá dimensão $p \leq n$.

Uma vez que o caso de ser $p = 0$ é trivial, em virtude do que vimos acima, suponhamos que p é positivo e tomemos uma base $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_p$, do subespaço F.

Por aplicação do teorema de Steinitz, podemos construir uma base de E que inclua os vectores da base escolhida em F: $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_p, \vec{e}_{p+1}, \vec{e}_{p+2}, \ldots, \vec{e}_n$.

Dada uma classe qualquer $\vec{x}+F \in E/F$, podemos escrever o seu representante $\vec{x}$ como combinação linear dos vectores da base de E, e, então,

$$\vec{x}+F=\left(\lambda_1\vec{f}_1+\lambda_2\vec{f}_2+\cdots+\lambda_p\vec{f}_p+\mu_{p+1}\vec{e}_{p+1}+\mu_{p+2}\vec{e}_{p+2}+\cdots+\mu_n\vec{e}_n\right)+F=$$

$$=\lambda_1\left(\vec{f}_1+F\right)+\lambda_2\left(\vec{f}_2+F\right)+\cdots+\lambda_p\left(\vec{f}_p+F\right)+\mu_{p+1}\left(\vec{e}_{p+1}+F\right)+$$

$$+\mu_{p+2}\left(\vec{e}_{p+2}+F\right)+\cdots+\mu_n\left(\vec{e}_n+F\right)=$$

$$=\lambda_1\left(\vec{0}+F\right)+\lambda_2\left(\vec{0}+F\right)+\cdots+\lambda_p\left(\vec{0}+F\right)+\mu_{p+1}\left(\vec{e}_{p+1}+F\right)+$$

$$+\mu_{p+2}\left(\vec{e}_{p+2}+F\right)+\cdots+\mu_n\left(\vec{e}_n+F\right)=$$

$$=\mu_{p+1}\left(\vec{e}_{p+1}+F\right)+\mu_{p+2}\left(\vec{e}_{p+2}+F\right)+\cdots+\mu_n\left(\vec{e}_n+F\right)=$$

Isto mostra que se tem

$$\frac{E}{F}=\left\langle \vec{e}_{p+1}+F, \vec{e}_{p+2}+F, \ldots, \vec{e}_n+F\right\rangle$$

Se provarmos que os vectores $\vec{e}_{p+1}+F, \vec{e}_{p+2}+F, \ldots, \vec{e}_n+F$ são linearmente independentes em E/F, então estes vectores constituirão uma base do espaço-quociente e a demonstração ficará terminada, já que, nessas condições,

$$\dim\left(\frac{E}{F}\right)=n-p=\dim(E)-\dim(F)$$

Ora, para vermos que os vectores $\vec{e}_{p+1}+F,\ \vec{e}_{p+2}+F, \ldots,\ \vec{e}_n+F$ são, de facto, independentes, partamos de uma sua combinação linear nula:

$$\xi_{p+1}\left(\vec{e}_{p+1}+F\right)+\xi_{p+2}\left(\vec{e}_{p+2}+F\right)+\cdots+\xi_n\left(\vec{e}_n+F\right)=\vec{0}+F$$

Daqui resulta, sucessivamente, que

$$\left(\xi_{p+1}\vec{e}_{p+1}+\xi_{p+2}\vec{e}_{p+2}+\cdots+\xi_n\vec{e}_n\right)+F=\vec{0}+F$$

$$\left(\xi_{p+1}\vec{e}_{p+1}+\xi_{p+2}\vec{e}_{p+2}+\cdots+\xi_n\vec{e}_n\right)-\vec{0}\in F$$

$$\xi_{p+1}\vec{e}_{p+1}+\xi_{p+2}\vec{e}_{p+2}+\cdots+\xi_n\vec{e}_n\in F$$

$$\xi_{p+1}\vec{e}_{p+1}+\xi_{p+2}\vec{e}_{p+2}+\cdots+\xi_n\vec{e}_n=\upsilon_1\vec{f}_1+\upsilon_2\vec{f}_2+\cdots+\upsilon_p\vec{f}_p$$

$$-\upsilon_1\vec{f}_1-\upsilon_2\vec{f}_2-\cdots-\upsilon_p\vec{f}_p+\xi_{p+1}\vec{e}_{p+1}+\xi_{p+2}\vec{e}_{p+2}+\cdots+\xi_n\vec{e}_n=\vec{0}$$

Ora, como os vectores $\vec{f}_1, \vec{f}_2, ..., \vec{f}_p, \vec{e}_{p+1}, \vec{e}_{p+2}, ..., \vec{e}_n$ constituem uma base do espaço E, são linearmente independentes, pelo que podemos concluir que

$$v_1 = v_2 = \cdots = v_p = \xi_{p+1} = \xi_{p+2} = \cdots = \xi_n = 0$$

O facto de se ter

$$\xi_{p+1} = \xi_{p+2} = \cdots = \xi_n = 0$$

mostra que os vectores $\vec{e}_{p+1} + F,\ \vec{e}_{p+2} + F, ...,\ \vec{e}_n + F$ são linearmente independentes em E/F e termina a demonstração.

Convém observar que, sendo verdade que todo o quociente de um espaço de dimensão finita tem também dimensão finita, pode acontecer que um quociente de um espaço vectorial de dimensão infinita tenha, por sua vez, dimensão finita.

Com efeito, consideremos, por exemplo, o espaço vectorial $\mathbb{R}[x]$ e o seu subespaço vectorial

$$W = \left\langle x^2, x^3, x^4, ..., x^n, ... \right\rangle$$

Dado um polinómio qualquer,

$$a(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_m x^m$$

a sua classe no quociente $\mathbb{R}[x]/W$ será

$$\begin{aligned} a(x) + W &= \left(a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_m x^m\right) + W = \\ &= (a_0 + W) + (a_1 x + W) + \left(a_2 x^2 + a_3 x^3 + \cdots + a_m x^m + W\right) = \\ &= a_0 (1 + W) + a_1 (x + W) + (0 + W) = \\ &= a_0 (1 + W) + a_1 (x + W) \end{aligned}$$

atendendo a que $a_2 x^2 + a_3 x^3 + \cdots + a_m x^m \in W$.

Por conseguinte, teremos

$$\frac{\mathbb{R}[x]}{W} = \langle 1 + W,\ x + W \rangle$$

e o espaço vectorial $\mathbb{R}[x]/W$ é finitamente gerado. É, de resto, fácil verificar que os dois geradores encontrados são linearmente independentes em $\mathbb{R}[x]/W$, pelo que

$$\dim\left(\frac{\mathbb{R}[x]}{W}\right) = 2$$

4 Aplicações lineares

4-1 Generalidades

Sucede, por vezes, que precisamos de lidar, num mesmo problema, com vectores de mais do que um espaço vectorial. Nesses casos, é, muitas vezes, necessário dispor de um instrumento que nos permita passar de um espaço para outro, e o instrumento adequado a esse fim é, naturalmente, uma aplicação.

Mas os vectores de um espaço vectorial não se devem imaginar desligados uns dos outros, já que um espaço vectorial não é um mero conjunto de vectores. Há, com efeito, uma certa estrutura, que, por assim dizer, liga os vectores entre si: por exemplo, certos vectores serão obtidos como combinações lineares de outros.

Por essa razão, não têm interesse todas as aplicações entre espaços vectoriais, mas apenas aquelas que, de alguma maneira, tomem em consideração essas relações entre os vectores.

As aplicações em causa serão as "aplicações lineares", que passamos a definir:

Definição 4-1 Sejam E e E' espaços vectoriais sobre um certo corpo K. Diz-se que uma aplicação $\varphi : E \to E'$ é uma **aplicação linear** quando verifica as seguintes propriedades, válidas para vectores arbitrários $\vec{x}, \vec{y} \in E$ e para qualquer escalar $\alpha \in K$:

a) $\varphi(\vec{x}+\vec{y}) = \varphi(\vec{x}) + \varphi(\vec{y})$

b) $\varphi(\alpha\vec{x}) = \alpha\varphi(\vec{x})$

Note-se que só se definem aplicações lineares entre dois espaços vectoriais sobre um mesmo corpo. Na verdade, na propriedade b) desta definição, o escalar α aplica-se, no primeiro membro, a um vector de E, e, no segundo membro, a um vector de E', o que exige, de facto, que o corpo dos escalares seja o mesmo para os dois espaços.

Exemplos 4-2

a) Consideremos a aplicação $\varphi : \mathbb{R}^3 \to \mathbb{R}^2$ definida por

$$\varphi(x,y,z) = (3x - 7z,\ x - 5y + 8z)$$

É fácil verificar que se trata de uma aplicação linear. Na verdade, tem-se:

$$\begin{aligned}
\varphi\big((x,y,z)+(x',y',z')\big) &= \varphi\big(x+x',\, y+y',z+z'\big) = \\
&= \big(3(x+x') - 7(z+z'),\, (x+x') - 5(y+y') + 8(z+z')\big) = \\
&= \big(3x + 3x' - 7z - 7z',\, x + x' - 5y - 5y' + 8z + 8z'\big) = \\
&= \big(3x - 7z,\, x - 5y + 8z\big) + \big(3x' - 7z',\, x' - 5y' + 8z'\big) = \\
&= \varphi(x,y,z) + \varphi(x',y',z')
\end{aligned}$$

$$\begin{aligned}
\varphi\big(\alpha(x,y,z)\big) &= \varphi(\alpha x,\alpha y,\alpha z) = \big(3(\alpha x) - 7(\alpha z),\ (\alpha x) - 5(\alpha y) + 8(\alpha z)\big) = \\
&= \big(\alpha(3x - 7z),\ \alpha(x - 5y + 8z)\big) = \\
&= \alpha(3x - 7z,\ x - 5y + 8z) = \alpha\varphi(x,y,z)
\end{aligned}$$

b) Analogamente se pode verificar que a aplicação $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$\varphi(x,y,z) = (x - y + z,\ -x + y + z,\ x + y - z)$$

é uma aplicação linear.

c) A aplicação $\varphi : \mathbb{R}[x] \to \mathbb{R}[x]$ definida por $\varphi\big(p(x)\big) = p'(x)$, onde $p'(x)$ designa a derivada do polinómio $p(x)$, é uma aplicação linear. Esta afirmação resulta trivialmente de propriedades bem conhecidas das derivadas.

Analogamente, a aplicação $\psi : \mathbb{R}[x] \to \mathbb{R}[x]$ definida por

$$\psi\big(p(x)\big) = p''(x) - 3p'(x) + 2p(x)$$

é uma aplicação linear.

d) Seja S o espaço vectorial real formado pelas sucessões convergentes de números reais, com as operações usuais. Atendendo a propriedades bem conhecidas, vemos imediatamente que a aplicação $\zeta : S \to \mathbb{R}$ em que $\zeta\big((u_n)_{n\in\mathbb{N}}\big) = \lim u_n$ é uma aplicação linear.

e) Seja F o espaço vectorial real formado pelas funções reais de variável real definidas e contínuas num certo intervalo $[a,b]$, com as operações usuais. A aplicação $\tau : F \to \mathbb{R}$ definida por

$$\tau(f) = \int_a^b f(x)dx$$

é uma aplicação linear.

f) Seja E um espaço vectorial qualquer sobre um corpo K. Dado um escalar $\lambda \in K$, resulta imediatamente da definição de espaço vectorial que a aplicação $\eta : E \rightarrow E$ definida por $\eta(\vec{x}) = \lambda\vec{x}$ é uma aplicação linear (chamada uma **homotetia**).

g) A aplicação identidade $1_E : E \rightarrow E$, em que $1_E(\vec{x}) = \vec{x}$, é uma aplicação linear, seja qual for o espaço vectorial E.

h) A aplicação nula $0 : E \rightarrow E'$, definida por $0(\vec{x}) = \vec{0}$, para qualquer vector $\vec{x} \in E$, é uma aplicação linear, quaisquer que sejam os espaços vectoriais considerados.

i) Seja F um subespaço vectorial de um espaço vectorial E arbitrário, e consideremos a aplicação $\pi_F : E \rightarrow E/F$, definida por $\pi_F(\vec{x}) = \vec{x} + F$, para cada vector $\vec{x} \in E$. Resulta imediatamente do modo como foram definidas as operações no espaço-quociente que π_F é uma aplicação linear (chamada a **aplicação linear canónica** ou **natural** de E para E/F).

j) Seja E um espaço vectorial qualquer e suponhamos que F e G são subespaços vectoriais de E, tais que $E = F \oplus G$. Consideremos as aplicações

$$\rho_F : E \rightarrow F \quad \text{e} \quad \rho_G : E \rightarrow G$$

definidas conforme passamos a expor:

Uma vez que se tem $E = F \oplus G$, cada vector $\vec{x} \in E$ tem uma expressão única na forma $\vec{x} = \vec{f}_{\vec{x}} + \vec{g}_{\vec{x}}$, com $\vec{f}_{\vec{x}} \in F$ e $\vec{g}_{\vec{x}} \in G$. Nestas condições, pomos

$$\rho_F(\vec{x}) = \vec{f}_{\vec{x}} \quad \text{e} \quad \rho_G(\vec{x}) = \vec{g}_{\vec{x}}$$

É muito fácil verificar que ρ_F e ρ_G são aplicações lineares (chamadas **projecções** de E, sobre F e sobre G, respectivamente). ■

Os exemplos anteriores mostram, em particular, que o conceito de "aplicação linear" aparece, com bastante naturalidade, em muitas áreas da Matemática. As derivadas, os limites, os integrais, fornecem-nos exemplos de aplicações lineares.

Como é natural, o facto de uma aplicação linear tomar em consideração a estrutura do espaço vectorial (diz-se, por vezes, que uma aplicação linear é aquela que é "compatível" com as operações do espaço vectorial) garante que ela vai satisfazer algumas propriedades interessantes em relação às noções previamente estudadas. Em primeiro lugar, podemos provar as seguintes propriedades elementares:

Teorema 4-3

Sejam E e E' espaços vectoriais sobre um certo corpo K, e seja $\varphi : E \rightarrow E'$ uma aplicação linear. Então:

a) $\varphi(\vec{0}) = \vec{0}$

b) $\varphi(-\vec{x}) = -\varphi(\vec{x})$

c) $\varphi(\vec{x} - \vec{y}) = \varphi(\vec{x}) - \varphi(\vec{y})$

Demonstração:

a) $\varphi(\vec{0}) = \varphi(0\,\vec{0}) = 0\varphi(\vec{0}) = \vec{0}$

b) Tem-se $\varphi(-\vec{x}) + \varphi(\vec{x}) = \varphi(-\vec{x} + \vec{x}) = \varphi(\vec{0}) = \vec{0}$, pelo que $\varphi(-\vec{x}) = -\varphi(\vec{x})$

c) $\varphi(\vec{x} - \vec{y}) = \varphi(\vec{x} + (-\vec{y})) = \varphi(\vec{x}) + \varphi(-\vec{y}) = \varphi(\vec{x}) - \varphi(\vec{y})$

Por outro lado, quando se trata de verificar se uma dada aplicação é ou não linear, as duas propriedades consignadas na Definição 4-1 podem resumir-se numa só, a saber:

Teorema 4-4 Sejam E e E' espaços vectoriais sobre um certo corpo K, e seja $\varphi : E \to E'$ uma aplicação. A aplicação φ é uma aplicação linear se e só se, para vectores arbitrários de E e escalares quaisquer, se tem:

$$\varphi(\alpha\vec{x}+\beta\vec{y}) = \alpha\varphi(\vec{x})+\beta\varphi(\vec{y})$$

Demonstração:
Supondo que φ é uma aplicação linear, tem-se imediatamente

$$\varphi(\alpha\vec{x}+\beta\vec{y}) = \varphi(\alpha\vec{x})+\varphi(\beta\vec{y}) = \alpha\varphi(\vec{x})+\beta\varphi(\vec{y})$$

Reciprocamente, se φ satisfaz esta última igualdade, obtêm-se as propriedades que definem uma aplicação linear, do seguinte modo: a primeira propriedade resulta de se tomar, nesta igualdade, $\alpha=\beta=1$; a segunda resulta de se tomar $\beta=0$, com α qualquer.

O teorema seguinte exprime o comportamento das aplicações lineares em relação às noções de dependência e de independência lineares:

Teorema 4-5
a) Uma aplicação linear φ, entre dois espaços vectoriais, transforma vectores linearmente dependentes em vectores linearmente dependentes, isto é, se os vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in E$ são linearmente dependentes, então também os vectores $\varphi(\vec{v}_1), \varphi(\vec{v}_2), \ldots, \varphi(\vec{v}_k) \in E'$ o são.
b) Uma aplicação linear transforma vectores linearmente independentes em vectores linearmente independentes se e só se é injectiva.

Demonstração:
a) Supondo que os vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in E$ são dependentes, existirão escalares $\xi_1, \xi_2, \ldots, \xi_k$, não todos nulos, tais que

$$\xi_1\vec{v}_1 + \xi_2\vec{v}_2 + \cdots + \xi_k\vec{v}_k = \vec{0}$$

Teremos, então,

$$\xi_1\varphi(\vec{v}_1) + \xi_2\varphi(\vec{v}_2) + \cdots + \xi_k\varphi(\vec{v}_k) = \varphi(\xi_1\vec{v}_1 + \xi_2\vec{v}_2 + \cdots + \xi_k\vec{v}_k) = \varphi(\vec{0}) = \vec{0}$$

e, como os escalares $\xi_1, \xi_2, \ldots, \xi_k$ são os mesmos, concluímos que os vectores $\varphi(\vec{v}_1), \varphi(\vec{v}_2), \ldots, \varphi(\vec{v}_k)$ são linearmente dependentes.

b) Suponhamos que a aplicação linear φ é injectiva e tomemos vectores linearmente independentes $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k \in E$.

Partindo de

$$\xi_1\varphi(\vec{v}_1) + \xi_2\varphi(\vec{v}_2) + \cdots + \xi_k\varphi(\vec{v}_k) = \vec{0}$$

obtemos

$$\varphi(\xi_1\vec{v}_1 + \xi_2\vec{v}_2 + \cdots + \xi_k\vec{v}_k) = \vec{0} = \varphi(\vec{0})$$

e da injectividade de φ resulta

$$\xi_1\vec{v}_1 + \xi_2\vec{v}_2 + \cdots + \xi_k\vec{v}_k = \vec{0}$$

Como os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ são independentes, podemos concluir que os coeficientes, nesta combinação linear, são todos nulos e, portanto, os vectores $\varphi(\vec{v}_1), \varphi(\vec{v}_2), ..., \varphi(\vec{v}_k)$ são também independentes.

Por outro lado, suponhamos que a aplicação linear φ não é injectiva: existirão então vectores distintos $\vec{a}$ e $\vec{b}$ tais que $\varphi(\vec{a}) = \varphi(\vec{b})$. Nestas condições, o vector $\vec{c} = \vec{a} - \vec{b}$ é independente, visto que $\vec{c} = \vec{a} - \vec{b} \neq \vec{0}$, mas

$$\varphi(\vec{c}) = \varphi(\vec{a} - \vec{b}) = \varphi(\vec{a}) - \varphi(\vec{b}) = \vec{0}$$

é um vector linearmente dependente.

As aplicações lineares injectivas merecem uma atenção especial, e atribuir-lhes-emos um nome próprio:

Definição 4-6

Uma aplicação linear $\varphi : E \to E'$ chama-se:

a) Um **monomorfismo** quando é injectiva;
b) Um **epimorfismo** quando é sobrejectiva;
c) Um **isomorfismo** quando é bijectiva;
d) Um **endomorfismo** quando $E = E'$;
e) Um **automorfismo** quando é um endomorfismo e um isomorfismo.

Analisaremos, em seguida, o comportamento das aplicações lineares perante os subespaços vectoriais:

Teorema 4-7

Dada uma aplicação linear $\varphi : E \to E'$, sejam F um subespaço vectorial de E e F' um subespaço vectorial de E'. Então:

a) $\varphi(F) = \{\varphi(\vec{x}) : \vec{x} \in F\}$ é um subespaço vectorial de E'.
b) $\varphi^{-1}(F') = \{\vec{x} \in E : \varphi(\vec{x}) \in F'\}$ é um subespaço vectorial de E.

Demonstração:

Trata-se, apenas, de verificar, em cada caso, as propriedades que definem um subespaço vectorial:

a) Sendo F um subespaço vectorial de E, então F não é vazio, pelo que $\varphi(F)$ também não é vazio.

Tomando $\vec{u}',\vec{v}' \in \varphi(F)$ e escalares quaisquer α e β, teremos, por definição de $\varphi(F)$, $\vec{u}' = \varphi(\vec{u})$ e $\vec{v}' = \varphi(\vec{v})$, com $\vec{u},\vec{v} \in F$. Mas, então, uma vez que $\alpha\vec{u} + \beta\vec{v} \in F$, vemos que

$$\alpha\vec{u}' + \beta\vec{v}' = \alpha\varphi(\vec{u}) + \beta\varphi(\vec{v}) = \varphi(\alpha\vec{u} + \beta\vec{v}) \in \varphi(F)$$

Portanto, $\varphi(F)$ é um subespaço vectorial de E'.

b) O facto de ser $\varphi(\vec{0}) = \vec{0} \in F'$ mostra que $\vec{0} \in \varphi^{-1}(F')$, pelo que $\varphi^{-1}(F') \neq \varnothing$.

Por outro lado, tomemos $\vec{x},\vec{y} \in \varphi^{-1}(F')$ e escalares quaisquer α e β. Por definição de $\varphi^{-1}(F')$, teremos $\varphi(\vec{x}),\varphi(\vec{y}) \in F'$, e, então,

$$\varphi(\alpha\vec{x} + \beta\vec{y}) = \alpha\varphi(\vec{x}) + \beta\varphi(\vec{y}) \in F'$$

pelo que $\alpha\vec{x} + \beta\vec{y} \in \varphi^{-1}(F')$, o que prova que $\varphi^{-1}(F') \leq E$.

Do teorema anterior resultam dois casos particulares importantes:
Consideremos uma aplicação linear $\varphi : E \to E'$.
Sabemos que $E \leq E$, pelo que, imediatamente, $\varphi(E) \leq E'$.
Por outro lado, sabemos também que $\{\vec{0}\} \leq E'$, e, por consequência,

$$\varphi^{-1}(\{\vec{0}\}) = \varphi^{-1}(\vec{0}) = \{\vec{x} \in E : \varphi(\vec{x}) = \vec{0}\} \leq E$$

Estes subespaços, de E e de E', respectivamente, têm significado especial, pelo que vamos atribuir-lhes designações especiais:

Definição 4-8

Seja $\varphi : E \to E'$ uma aplicação linear qualquer. O subespaço $\varphi(E) = \operatorname{Im}(E)$, do espaço vectorial E', é o **espaço-imagem** ou **espaço característico** de φ, e o subespaço $\varphi^{-1}(\{\vec{0}\}) = \varphi^{-1}(\vec{0}) = \operatorname{Nuc}(\varphi)$ é o **núcleo** ou **espaço de nulidade** de φ. Quando o espaço-imagem de φ é finitamente gerado, a sua dimensão chama-se **característica** de φ: $c_\varphi = \dim(\operatorname{Im}(\varphi))$; por sua vez, quando o núcleo é finitamente gerado, a sua dimensão é a **nulidade** de φ: $n_\varphi = \dim(\operatorname{Nuc}(\varphi))$.

Alguns autores designam o núcleo por $\operatorname{Ker}(\varphi)$, abreviatura da palavra inglesa "kernel", com o mesmo significado.

O estudo do núcleo é da máxima importância para o conhecimento das propriedades de uma dada aplicação linear, em face do seguinte:

Teorema 4-9 Seja $\varphi : E \to E'$ uma aplicação linear qualquer.

a) Dado um vector $\vec{a} \in E$ e $\vec{a}' = \varphi(\vec{a}) \in E'$, tem-se

$$\varphi^{-1}(\vec{a}') = \{\vec{x} \in E : \varphi(\vec{x}) = \vec{a}'\} = \vec{a} + \mathrm{Nuc}(\varphi)$$

b) A aplicação linear φ é um monomorfismo se e só se $\mathrm{Nuc}(\varphi) = \{\vec{0}\}$.

Demonstração:

a) Tomando $\vec{z} \in \vec{a} + \mathrm{Nuc}(\varphi)$, teremos $\vec{z} = \vec{a} + \vec{w}$, para algum $\vec{w} \in \mathrm{Nuc}(\varphi)$. Nestas circunstâncias, temos

$$\varphi(\vec{z}) = \varphi(\vec{a} + \vec{w}) = \varphi(\vec{a}) + \varphi(\vec{z}) = \vec{a}' + \vec{0} = \vec{a}'$$

pelo que $\vec{z} \in \varphi^{-1}(\vec{a}')$.

Reciprocamente, se for $\vec{z} \in \varphi^{-1}(\vec{a}')$, então $\varphi(\vec{z}) = \vec{a}' = \varphi(\vec{a})$, pelo que, sucessivamente,

$$\begin{aligned} \varphi(\vec{z}) - \varphi(\vec{a}) &= \vec{0} \\ \varphi(\vec{z} - \vec{a}) &= \vec{0} \\ \vec{z} - \vec{a} &= \vec{n} \in \mathrm{Nuc}(\varphi) \\ \vec{z} &= \vec{a} + \vec{n} \in \vec{a} + \mathrm{Nuc}(\varphi) \end{aligned}$$

b) Resulta imediatamente da alínea anterior.

Em relação ao espaço-imagem, podemos adiantar o seguinte:

Teorema 4-10 Seja $\varphi : E \to E'$ uma aplicação linear qualquer, e seja F um subespaço vectorial de E. Então, se $F = \langle X \rangle$, tem-se $\varphi(F) = \langle \varphi(X) \rangle$. Em particular, se F é finitamente gerado, também $\varphi(F)$ é finitamente gerado, tendo-se, então,

$$\dim(\varphi(F)) \le \dim(F)$$

Demonstração:

Se $F = \langle X \rangle$, então cada vector de F é combinação linear de alguns vectores de X. Escrevendo $\vec{f} = \vartheta_1\vec{x}_1 + \vartheta_2\vec{x}_2 + \cdots + \vartheta_m\vec{x}_m$, obtemos

$$\varphi(\vec{f}) = \varphi(\vartheta_1\vec{x}_1 + \vartheta_2\vec{x}_2 + \cdots + \vartheta_m\vec{x}_m) = \vartheta_1\varphi(\vec{x}_1) + \vartheta_2\varphi(\vec{x}_2) + \cdots + \vartheta_m\varphi(\vec{x}_m)$$

Atendendo a que, evidentemente, $\varphi(X) \subseteq \varphi(F)$, podemos concluir que $\varphi(F) = \langle \varphi(X) \rangle$. As restantes afirmações resultam trivialmente daqui.

O teorema seguinte, de grande importância, relaciona as dimensões do espaço-imagem e do núcleo de uma aplicação linear, quando elas são finitas:

Teorema 4-11 Seja $\varphi : E \to E'$ uma aplicação linear qualquer. Se o espaço E tem dimensão finita, então: $\dim(E) = n_\varphi + c_\varphi = \dim(\mathrm{Nuc}(\varphi)) + \dim(\mathrm{Im}(\varphi))$.

Demonstração:
Seja $\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_s$ uma base de $\mathrm{Nuc}(\varphi)$, e, por aplicação do teorema de Steinitz, construamos uma base de E que inclua estes vectores, digamos

$$\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_s, \vec{e}_{s+1}, \vec{e}_{s+2}, \ldots, \vec{e}_n$$

Uma vez que $E = \langle \vec{z}_1, \vec{z}_2, \ldots, \vec{z}_s, \vec{e}_{s+1}, \vec{e}_{s+2}, \ldots, \vec{e}_n \rangle$, teremos, de acordo com o Teorema 4-10,

$$\begin{aligned}\varphi(E) &= \langle \varphi(\vec{z}_1), \varphi(\vec{z}_2), \ldots, \varphi(\vec{z}_s), \varphi(\vec{e}_{s+1}), \varphi(\vec{e}_{s+2}), \ldots, \varphi(\vec{e}_n) \rangle = \\ &= \langle \vec{0}, \vec{0}, \ldots, \vec{0}, \varphi(\vec{e}_{s+1}), \varphi(\vec{e}_{s+2}), \ldots, \varphi(\vec{e}_n) \rangle = \\ &= \langle \varphi(\vec{e}_{s+1}), \varphi(\vec{e}_{s+2}), \ldots, \varphi(\vec{e}_n) \rangle\end{aligned}$$

Se provarmos que os vectores $\varphi(\vec{e}_{s+1}), \varphi(\vec{e}_{s+2}), \ldots, \varphi(\vec{e}_n)$ são linearmente independentes, eles constituirão uma base do espaço imagem $\varphi(E)$, e, então, a demonstração ficará feita, uma vez que se terá

$$\dim(\varphi(E)) = n - s = \dim(E) - \dim(\mathrm{Nuc}(\varphi))$$

o que equivale à igualdade procurada.

Para estudarmos, então, a independência dos vectores

$$\varphi(\vec{e}_{s+1}), \varphi(\vec{e}_{s+2}), \ldots, \varphi(\vec{e}_n)$$

consideremos a igualdade

$$\varepsilon_{s+1}\varphi(\vec{e}_{s+1}) + \varepsilon_{s+2}\varphi(\vec{e}_{s+2}) + \cdots + \varepsilon_n\varphi(\vec{e}_n) = \vec{0}$$

Desta igualdade concluímos, sucessivamente, que

$$\varphi(\varepsilon_{s+1}\vec{e}_{s+1} + \varepsilon_{s+2}\vec{e}_{s+2} + \cdots + \varepsilon_n\vec{e}_n) = \vec{0}$$

$$\varepsilon_{s+1}\vec{e}_{s+1} + \varepsilon_{s+2}\vec{e}_{s+2} + \cdots + \varepsilon_n\vec{e}_n = \vec{0}$$

$$\varepsilon_{s+1}\vec{e}_{s+1} + \varepsilon_{s+2}\vec{e}_{s+2} + \cdots + \varepsilon_n\vec{e}_n \in \mathrm{Nuc}(\varphi)$$

$$\varepsilon_{s+1}\vec{e}_{s+1}+\varepsilon_{s+2}\vec{e}_{s+2}+\dots+\varepsilon_n\vec{e}_n=\delta_1\vec{z}_1+\delta_2\vec{z}_2+\dots+\delta_s\vec{z}_s$$
$$-\delta_1\vec{z}_1-\delta_2\vec{z}_2-\dots-\delta_s\vec{z}_s+\varepsilon_{s+1}\vec{e}_{s+1}+\varepsilon_{s+2}\vec{e}_{s+2}+\dots+\varepsilon_n\vec{e}_n=\vec{0}$$

Uma vez que os vectores $\vec{z}_1,\vec{z}_2,\dots,\vec{z}_s,\vec{e}_{s+1},\vec{e}_{s+2},\dots,\vec{e}_n$ são linearmente independentes (visto constituírem uma base de E), podemos inferir daqui que

$$\varepsilon_{s+1}=\varepsilon_{s+2}=\dots=\varepsilon_n=\delta_1=\delta_2=\dots=\delta_s=0$$

Sendo $\varepsilon_{s+1}=\varepsilon_{s+2}=\dots=\varepsilon_n=0$, os vectores $\varphi(\vec{e}_{s+1}),\varphi(\vec{e}_{s+2}),\dots,\varphi(\vec{e}_n)$ são linearmente independentes, como se pretendia.

Do teorema anterior resulta, por sua vez, um corolário bastante importante:

Corolário 4-12 Seja $\varphi: E\to E'$ uma aplicação linear qualquer. Se os espaços E e E' têm ambos a dimensão finita n, então φ é um monomorfismo se e só se φ é um epimorfismo. Em particular, um endomorfismo num espaço de dimensão finita é um monomorfismo se e só se é um epimorfismo.

Demonstração:

Suponhamos que φ é um monomorfismo. Sabemos, então, que $\mathrm{Nuc}(\varphi)=\{\vec{0}\}$, pelo que $n_\varphi=0$.

Da igualdade $\dim(E)=n_\varphi+c_\varphi$, demonstrada no Teorema 4-11, resulta então que $n=\dim(E)=n_\varphi+c_\varphi=0+c_\varphi=c_\varphi$. Mas, sendo $\varphi(E)$ um subespaço vectorial de E' e tendo a dimensão n, ter-se-á $\varphi(E)=E'$, pelo que φ é um epimorfismo.

Reciprocamente, se φ é um epimorfismo, então $\varphi(E)=E'$, pelo que $c_\varphi=\dim(E')=n$. Consequentemente, resulta do teorema anterior que

$$n=\dim(E)=n_\varphi+c_\varphi=n_\varphi+n$$

pelo que $n_\varphi=0$. Mas, nestas condições, será $\mathrm{Nuc}(\varphi)=\{\vec{0}\}$, e, portanto, φ é um monomorfismo.

4-2 Operações com aplicações lineares

À custa de aplicações lineares dadas, podemos, em certas condições, construir novas aplicações lineares. Vamos estudar, nesta secção, alguns dos métodos de construção de aplicações lineares.

Em primeiro lugar, vamos ver que, em determinadas condições, se pode definir a "soma" de duas aplicações lineares:

Definição 4-13 Sejam E e E' dois espaços vectoriais sobre um certo corpo K, e consideremos duas aplicações lineares, φ e ψ, de E para E'. Definimos a soma de φ com ψ como sendo a aplicação $\varphi+\psi : E \to E'$ em que, para qualquer vector $\vec{x} \in E$, se faz $(\varphi+\psi)(\vec{x}) = \varphi(\vec{x}) + \psi(\vec{x})$.

É fácil verificar que a soma de duas aplicações lineares é ainda uma aplicação linear:

Teorema 4-14 Nas condições da Definição 4-13, a aplicação $\varphi+\psi$ é uma aplicação linear.

Demonstração:
Dados vectores quaisquer $\vec{x}, \vec{y} \in E$, e escalares $\alpha, \beta \in K$, igualmente arbitrários, vemos que

$$\begin{aligned}(\varphi+\psi)(\alpha\vec{x}+\beta\vec{y}) &= \varphi(\alpha\vec{x}+\beta\vec{y}) + \psi(\alpha\vec{x}+\beta\vec{y}) = \\ &= \alpha\varphi(\vec{x}) + \beta\varphi(\vec{y}) + \alpha\psi(\vec{x}) + \beta\psi(\vec{y}) = \\ &= \alpha\big(\varphi(\vec{x}) + \psi(\vec{x})\big) + \beta\big(\varphi(\vec{y}) + \psi(\vec{y})\big) = \\ &= \alpha(\varphi+\psi)(\vec{x}) + \beta(\varphi+\psi)(\vec{y})\end{aligned}$$

o que mostra que $\varphi+\psi$ é, de facto, uma aplicação linear.

A operação de adição de aplicações lineares, entre dois dados espaços vectoriais, que acabamos de definir, goza de algumas propriedades notáveis, que resumimos no próximo teorema:

Teorema 4-15 Sejam E e E' dois espaços vectoriais sobre um certo corpo K. Para quaisquer aplicações φ, ψ e θ, de E para E', tem-se:

a) $\varphi+\psi = \psi+\varphi$
b) $(\varphi+\psi)+\theta = \varphi+(\psi+\theta)$
c) $\varphi+0 = \varphi$, onde 0 designa a aplicação linear nula, tal como foi definida no Exemplo 4-2 h).
d) A aplicação $-\varphi : E \to E'$, definida por $(-\varphi)(\vec{x}) = -\varphi(\vec{x})$, é uma aplicação linear, tendo-se $\varphi+(-\varphi) = 0$.

Demonstração:
a) Para um vector $\vec{x} \in E$ qualquer, tem-se

$$(\varphi+\psi)(\vec{x}) = \varphi(\vec{x}) + \psi(\vec{x}) = \psi(\vec{x}) + \varphi(\vec{x}) = (\psi+\varphi)(\vec{x})$$

b) Para um vector $\vec{x} \in E$ qualquer, tem-se

$$((\varphi+\psi)+\theta)(\vec{x}) = (\varphi+\psi)(\vec{x})+\theta(\vec{x}) = (\varphi(\vec{x})+\psi(\vec{x}))+\theta(\vec{x}) =$$
$$= \varphi(\vec{x})+(\psi(\vec{x})+\theta(\vec{x})) = \varphi(\vec{x})+(\psi+\theta)(\vec{x}) = (\varphi+(\psi+\theta))(\vec{x})$$

c) Para um vector $\vec{x} \in E$ qualquer, tem-se

$$(\varphi+0)(\vec{x}) = \varphi(\vec{x})+0(\vec{x}) = \varphi(\vec{x})+\vec{0} = \varphi(\vec{x})$$

d) Para vermos que a aplicação $-\varphi$ é linear, basta reparar que

$$(-\varphi)(\alpha\vec{x}+\beta\vec{y}) = -\varphi(\alpha\vec{x}+\beta\vec{y}) = -(\alpha\varphi(\vec{x})+\beta\varphi(\vec{y})) =$$
$$= -\alpha\varphi(\vec{x})-\beta\varphi(\vec{y}) = \alpha(-\varphi(\vec{x}))+\beta(-\varphi(\vec{y})) =$$
$$= \alpha(-\varphi)(\vec{x})+\beta(-\varphi)(\vec{y})$$

A igualdade $\varphi+(-\varphi)=0$ resulta imediatamente das definições.

Estudaremos agora uma outra maneira de se construírem aplicações lineares à custa de aplicações lineares dadas:

Definição 4-16 Seja $\varphi : E \to E'$ uma aplicação linear entre dois espaços vectoriais quaisquer. Dado um escalar arbitrário $\lambda \in K$, definimos uma nova aplicação $\lambda\varphi : E \to E'$, pondo, para qualquer vector $\vec{x} \in E$, $(\lambda\varphi)(\vec{x}) = \lambda\varphi(\vec{x})$.

As aplicações da forma $\lambda\varphi$, para um escalar arbitrário $\lambda \in K$, são novas aplicações lineares, como vamos ver:

Teorema 4-17 Nas condições da Definição 4-16, a aplicação $\lambda\varphi : E \to E'$ é uma aplicação linear.

Demonstração:
Dados vectores quaisquer $\vec{x}, \vec{y} \in E$ e escalares $\alpha, \beta \in K$, igualmente arbitrários, vemos que

$$(\lambda\varphi)(\alpha\vec{x}+\beta\vec{y}) = \lambda\varphi(\alpha\vec{x}+\beta\vec{y}) = \lambda(\alpha\varphi(\vec{x})+\beta\varphi(\vec{y})) =$$
$$= \lambda\alpha\varphi(\vec{x})+\lambda\beta\varphi(\vec{y}) = \alpha\lambda\varphi(\vec{x})+\beta\lambda\varphi(\vec{y}) = \alpha(\lambda\varphi)(\vec{x})+\beta(\lambda\varphi)(\vec{y})$$

Também esta nova construção goza de propriedades interessantes:

Teorema 4-18 Sejam E e E' dois espaços vectoriais sobre um certo corpo K. Para quaisquer aplicações φ e ψ, de E para E', e para quaisquer escalares $\lambda, \mu \in K$, tem-se:
a) $\lambda(\varphi+\psi)=\lambda\varphi+\lambda\psi$
b) $(\lambda+\mu)\varphi=\lambda\varphi+\mu\varphi$
c) $\lambda(\mu\varphi)=(\lambda\mu)\varphi$
d) $1\varphi=\varphi$

Demonstração:
a) Para um vector $\vec{x} \in E$ qualquer, tem-se

$$\begin{aligned}(\lambda(\varphi+\psi))(\vec{x}) &= \lambda(\varphi+\psi)(\vec{x}) = \lambda(\varphi(\vec{x})+\psi(\vec{x})) = \lambda\varphi(\vec{x})+\lambda\psi(\vec{x}) = \\ &= (\lambda\varphi)(\vec{x})+(\lambda\psi)(\vec{x}) = (\lambda\varphi+\lambda\psi)(\vec{x})\end{aligned}$$

b) Para um vector $\vec{x} \in E$ qualquer, tem-se

$$\begin{aligned}((\lambda+\mu)\varphi)(\vec{x}) &= (\lambda+\mu)\varphi(\vec{x}) = \lambda\varphi(\vec{x})+\mu\varphi(\vec{x}) = \\ &= (\lambda\varphi)(\vec{x})+(\mu\varphi)(\vec{x}) = (\lambda\varphi+\mu\varphi)(\vec{x})\end{aligned}$$

c) Para um vector $\vec{x} \in E$ qualquer, tem-se

$$(\lambda(\mu\varphi))(\vec{x}) = \lambda(\mu\varphi)(\vec{x}) = \lambda(\mu\varphi(\vec{x})) = (\lambda\mu)\varphi(\vec{x}) = ((\lambda\mu)\varphi)(\vec{x})$$

d) Resulta trivialmente da definição.

As considerações anteriores podem ser interpretadas de uma maneira que, no nosso contexto, se reveste do máximo interesse.

Com efeito, se designarmos por $L(E, E')$ o conjunto de todas as aplicações lineares do espaço vectorial E para o espaço vectorial E', as construções que acabámos de estudar correspondem a introduzir, neste conjunto, uma operação binária de "adição" e, para cada escalar $\lambda \in K$, uma "multiplicação" por esse escalar. Tendo em consideração os Teoremas 4-15 e 4-18, podemos imediatamente concluir o seguinte:

Teorema 4-19 Com as operações de adição e multiplicação por um escalar definidas acima, o conjunto $L(E, E')$, formado por todas as aplicações lineares do espaço vectorial E para o espaço vectorial E', constitui um espaço vectorial sobre o corpo K.

Demonstração:
Resulta imediatamente da definição de espaço vectorial e dos Teoremas 4-15 e 4-18.

No caso de lidarmos com espaços vectoriais de dimensão finita, podemos assegurar que o novo espaço vectorial $L(E,E')$ tem também dimensão finita. Embora este facto resulte com grande facilidade de matérias que apresentaremos no próximo capítulo, pode ter interesse apresentar, desde já, uma demonstração directa deste facto, pelo que enunciamos o seguinte teorema:

Teorema 4-20 Se os espaços vectoriais E e E', sobre o corpo K, têm ambos dimensão finita, então o espaço vectorial $L(E,E')$ tem também dimensão finita, sendo

$$\dim(L(E,E')) = \dim(E) \times \dim(E')$$

Demonstração:
Suponhamos que os espaços vectoriais E e E' têm ambos dimensão finita, e fixemos uma base em cada um: seja $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ uma base de E, e $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_p$ uma base de E'.

Para cada par (i,j), com $j = 1,2,\ldots,n$ e $i = 1,2,\ldots,p$, definimos uma aplicação

$$\zeta_{ij} : E \to E'$$

do seguinte modo:

Dado um vector $\vec{x} \in E$, ele escreve-se, de modo único, como combinação linear dos vectores da base fixa em E, ou seja, na forma

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n$$

Note-se que, por uma questão de clareza e facilidade de referência, representámos as componentes do vector $\vec{x}$, em relação à base considerada, por x_j (com $j = 1,2,\ldots,n$), em vez de pelas letras gregas que temos vindo a usar para designar os escalares. Fá-lo-emos muitas vezes.

Poremos, então,

$$\zeta_{ij}(\vec{x}) = x_j\vec{e}'_i$$

É fácil verificar que as pn aplicações assim definidas são aplicações lineares. Com efeito, tomemos vectores quaisquer $\vec{x}, \vec{y} \in E$ e escalares arbitrários $\alpha, \beta \in K$.

Teremos, então,

$$\begin{aligned}\zeta_{ij}(\alpha\vec{x} + \beta\vec{y}) &= (\alpha x_j + \beta y_j)\vec{e}'_i = \alpha x_j\vec{e}'_i + \beta y_j\vec{e}'_i = \\ &= \alpha(x_j\vec{e}'_i) + \beta(y_j\vec{e}'_i) = \alpha\zeta_{ij}(\vec{x}) + \beta\zeta_{ij}(\vec{y})\end{aligned}$$

Vamos provar que as aplicações ζ_{ij} constituem uma base do espaço vectorial $L(E,E')$, o que termina a demonstração.

Comecemos por verificar que as aplicações em causa geram o espaço. Para isso, tomemos uma aplicação linear qualquer $\varphi \in L(E,E')$.

Para cada vector $\vec{e}_j$, da base de E, consideremos $\varphi(\vec{e}_j) \in E'$. Podemos escrever este vector à custa da base fixa em E':

$$\varphi(\vec{e}_j) = \sum_{i=1}^{p} a_{ij}\vec{e}'_i$$

Para cada vector $\vec{x} = \sum_{j=1}^{n} x_j\vec{e}_j$, teremos, atendendo às propriedades das aplicações lineares,

$$\varphi(\vec{x}) = \varphi\left(\sum_{j=1}^{n} x_j\vec{e}_j\right) = \sum_{j=1}^{n} x_j\varphi(\vec{e}_j) = \sum_{j=1}^{n} x_j\left(\sum_{i=1}^{p} a_{ij}\vec{e}'_i\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{p} a_{ij}x_j\vec{e}'_i\right) =$$

$$= \sum_{i=1}^{p}\left(\sum_{j=1}^{n} a_{ij}x_j\vec{e}'_i\right) = \sum_{i=1}^{p}\left(\sum_{j=1}^{n} a_{ij}\zeta_{ij}(\vec{x})\right) = \sum_{i=1}^{p}\left(\sum_{j=1}^{n}(a_{ij}\zeta_{ij})(\vec{x})\right) = \left(\sum_{i=1}^{p}\left(\sum_{j=1}^{n} a_{ij}\zeta_{ij}\right)\right)(\vec{x})$$

Como $\vec{x}$ é arbitrário, podemos concluir que

$$\varphi = \sum_{i=1}^{p}\sum_{j=1}^{n} a_{ij}\zeta_{ij}$$

o que significa que as aplicações ζ_{ij} geram o espaço vectorial $L(E,E')$.

Resta apenas verificar que estes geradores são linearmente independentes.

Ora, suponhamos que

$$\sum_{i=1}^{p}\sum_{j=1}^{n} a_{ij}\zeta_{ij} = 0$$

isto é, que, para qualquer vector $\vec{x} \in E$,

$$\left(\sum_{i=1}^{p}\sum_{j=1}^{n} a_{ij}\zeta_{ij}\right)(\vec{x}) = \vec{0}$$

Sendo esta igualdade válida para qualquer vector de E, continuará a verificar-se se substituirmos $\vec{x}$ por cada um dos vectores da base fixa em E.

Repare-se que, se tomarmos $\vec{x} = \vec{e}_k$ (para $k = 1,2,\ldots,n$), teremos

$$x_j = \begin{cases} 1 \ , \ \text{se} \ j = k \\ 0 \ , \ \text{se} \ j \neq k \end{cases} \qquad (j = 1,2,\ldots,n)$$

Teremos, então:

$$\left(\sum_{i=1}^{p}\sum_{j=1}^{n} a_{ij}\zeta_{ij}\right)(\vec{e}_k) = \vec{0}$$

$$\sum_{i=1}^{p}\sum_{j=1}^{n} a_{ij}\zeta_{ij}\left(\vec{e}_k\right) = \vec{0}$$

$$\sum_{i=1}^{p}\sum_{j=1}^{n} a_{ij}x_j\vec{e}'_i = \vec{0}$$

Atendendo ao que vimos acima, quanto às componentes do vector $\vec{x} = \vec{e}_k$, bastará, neste somatório, aproveitar as parcelas em que $j = k$, uma vez que todas as outras se anulam; como, para essas parcelas, se tem $x_k = 1$, obteremos

$$\sum_{i=1}^{p} a_{ik}\vec{e}'_i = 0$$

Ora, os vectores $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_p$ formam uma base de E', pelo que são linearmente independentes. Por conseguinte, da igualdade anterior podemos concluir que se tem

$$a_{ik} = 0 \quad (i = 1,2,\ldots,p)$$

Como k é também um número arbitrário, entre 1 e n, verificamos que os coeficientes a_{ij} são todos nulos, o que prova que os vectores ζ_{ij} do espaço vectorial $L(E,E')$ são linearmente independentes, como queríamos demonstrar.

Estudaremos ainda uma outra construção de aplicações lineares a partir de aplicações lineares dadas. Trata-se, neste caso, da composição de aplicações:

Teorema 4-21 Sejam $\varphi : E \to E'$ e $\psi : E' \to E''$ aplicações lineares, entre espaços vectoriais sobre um certo corpo K. Então, a aplicação $\psi \circ \varphi : E \to E''$ é ainda uma aplicação linear.

Demonstração:
Atendendo à maneira como se define a composição de duas aplicações e também às propriedades das aplicações lineares, temos, para vectores quaisquer $\vec{x}, \vec{y} \in E$ e escalares arbitrários $\alpha, \beta \in K$:

$$(\psi \circ \varphi)(\alpha\vec{x} + \beta\vec{y}) = \psi(\varphi(\alpha\vec{x} + \beta\vec{y})) = \psi(\alpha\varphi(\vec{x}) + \beta\varphi(\vec{y})) =$$
$$= \alpha\psi(\varphi(\vec{x})) + \beta\psi(\varphi(\vec{y})) = \alpha(\psi \circ \varphi)(\vec{x}) + \beta(\psi \circ \varphi)(\vec{y})$$

Fica assim provado que $\psi \circ \varphi$ é, de facto, uma aplicação linear.

Para além das propriedades específicas da composição de aplicações, podemos ainda encontrar outras, que relacionam a composição com as operações de adição e de multiplicação por um escalar, anteriormente definidas:

Teorema 4-22 Sejam E, E', E'' e E''' espaços vectoriais sobre um certo corpo K. Então:
a) $\theta \circ (\psi \circ \varphi) = (\theta \circ \psi) \circ \varphi, \ \forall \varphi \in L(E, E'), \ \psi \in L(E', E''), \ \theta \in L(E'', E''')$
b) $\psi \circ (\varphi + \varphi') = (\psi \circ \varphi) + (\psi \circ \varphi'), \ \forall \varphi, \varphi' \in L(E, E'), \ \psi \in L(E', E'')$
c) $(\psi + \psi') \circ \varphi = (\psi \circ \varphi) + (\psi' \circ \varphi), \ \forall \varphi \in L(E, E'), \ \psi, \ \psi' \in L(E', E'')$
d) $\psi \circ (\lambda\varphi) = (\lambda\psi) \circ \varphi = \lambda(\psi \circ \varphi), \ \forall \varphi \in L(E, E'), \ \psi \in L(E', E''), \ \forall \lambda \in K$

Demonstração:
Tomando um vector arbitrário $\vec{x} \in E$, temos, atendendo às definições e às propriedades das aplicações lineares:

a) $$(\theta \circ (\psi \circ \varphi))(\vec{x}) = \theta((\psi \circ \varphi)(\vec{x})) = \theta(\psi(\varphi(\vec{x}))) =$$
$$= (\theta \circ \psi)(\varphi(\vec{x})) = ((\theta \circ \psi) \circ \varphi)(\vec{x})$$

b) $$(\psi \circ (\varphi + \varphi'))(\vec{x}) = \psi((\varphi + \varphi')(\vec{x})) = \psi(\varphi(\vec{x}) + \varphi'(\vec{x})) =$$
$$= \psi(\varphi(\vec{x})) + \psi(\varphi'(\vec{x})) = (\psi \circ \varphi)(\vec{x}) + (\psi \circ \varphi')(\vec{x}) =$$
$$= ((\psi \circ \varphi) + (\psi \circ \varphi'))(\vec{x})$$

c) $$((\psi + \psi') \circ \varphi)(\vec{x}) = (\psi + \psi')(\varphi(\vec{x})) = \psi(\varphi(\vec{x})) + \psi'(\varphi(\vec{x})) =$$
$$= (\psi \circ \varphi)(\vec{x}) + (\psi' \circ \varphi)(\vec{x}) = (\psi \circ \varphi)(\vec{x}) + (\psi' \circ \varphi)(\vec{x}) =$$
$$= ((\psi \circ \varphi) + (\psi' \circ \varphi))(\vec{x})$$

d) $\big(\psi\circ(\lambda\varphi)\big)(\vec{x})=\psi\big((\lambda\varphi)(\vec{x})\big)=\psi\big(\lambda\varphi(\vec{x})\big)=\lambda\psi\big(\varphi(\vec{x})\big)=$
$=\lambda(\psi\circ\varphi)(\vec{x})=\big(\lambda(\psi\circ\varphi)\big)(\vec{x})$

$\big((\lambda\psi)\circ\varphi\big)(\vec{x})=(\lambda\psi)\big(\varphi(\vec{x})\big)=\lambda\psi\big(\varphi(\vec{x})\big)=\lambda(\psi\circ\varphi)(\vec{x})=\big(\lambda(\psi\circ\varphi)\big)(\vec{x})$

4-3 Isomorfismos

Entre as aplicações lineares, merecem atenção especial os isomorfismos. A existência de um isomorfismo entre dois espaços vectoriais assegura que eles partilham um grande número de propriedades.

A fim de nos referirmos mais facilmente aos casos em que existe um isomorfismo entre dois espaços vectoriais, introduzimos a seguinte designação:

Definição 4-23 Quando existe um isomorfismo φ de um espaço vectorial E para um espaço vectorial E', dizemos que E é **isomorfo** a E'. Indica-se esse facto escrevendo $E\cong E'$.

Exemplo 4-24

Seja E um espaço vectorial de dimensão n sobre um corpo K. Tem-se então $E\cong K^n$.

Com efeito, vimos, no Teorema 2-37, que, fixando em E uma base $\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n$, a aplicação $\omega\ :E\to K^n$, que associa a cada vector de E a lista das suas componentes, na base considerada, é uma aplicação bijectiva. Por outro lado, as observações que se seguem aos Exemplos 2-38 mostram que a aplicação ω é uma aplicação linear; portanto, é um isomorfismo. ■

Esta noção de isomorfismo comporta-se como uma relação de equivalência, em qualquer conjunto de espaços vectoriais. Com efeito, são válidas as seguintes propriedades:

Teorema 4-25 Para espaços vectoriais quaisquer, tem-se:

a) $E\cong E$
b) Se $E\cong E'$, então $E'\cong E$
c) Se $E\cong E'$ e $E'\cong E''$, então $E\cong E''$

Demonstração:

a) Vimos já que a aplicação identidade $1_E:E\to E$ é uma aplicação linear, e, como é, evidentemente, uma bijecção, é um isomorfismo de E para E.

b) Se se tem $E \cong E'$, é porque existe um isomorfismo $\varphi : E \to E'$.

Ora, sendo φ uma aplicação bijectiva, existirá a respectiva aplicação inversa $\varphi^{-1} : E' \to E$. Vamos ver que esta aplicação φ^{-1} é uma aplicação linear, o que prova imediatamente que se trata de um isomorfismo e que, por consequência, $E' \cong E$.

Tomando vectores quaisquer $\vec{x}', \vec{y}' \in E'$, sabemos que, pelo facto de φ ser sobrejectiva, existirão vectores $\vec{x}, \vec{y} \in E$ tais que $\varphi(\vec{x}) = \vec{x}', \varphi(\vec{y}) = \vec{y}'$. Nessas condições, considerando também escalares arbitrários $\alpha, \beta \in K$, e atendendo a que φ verifica as propriedades das aplicações lineares, temos:

$$\varphi^{-1}\left(\alpha\vec{x}' + \beta\vec{y}'\right) = \varphi^{-1}\left(\alpha\varphi(\vec{x}) + \beta\varphi(\vec{y})\right) = \varphi^{-1}\left(\varphi(\alpha\vec{x} + \beta\vec{y})\right) =$$
$$= \left(\varphi^{-1} \circ \varphi\right)\left(\alpha\vec{x} + \beta\vec{y}\right) = 1_E\left(\alpha\vec{x} + \beta\vec{y}\right) = \alpha\vec{x} + \beta\vec{y} = \alpha\varphi^{-1}\left(\vec{x}'\right) + \beta\varphi^{-1}\left(\vec{y}'\right)$$

c) Sendo $\varphi : E \to E'$ e $\psi : E' \to E''$ isomorfismos, portanto aplicações lineares bijectivas, a aplicação composta $\psi \circ \varphi : E \to E''$ é também uma aplicação linear e igualmente bijectiva, ou seja, um isomorfismo.

Conforme acima indicámos, dois espaços isomorfos têm muitas propriedades comuns. Em particular, quando se lida com espaços vectoriais de dimensão finita, podemos provar o seguinte:

Teorema 4-26 Sejam E e E' espaços vectoriais sobre um corpo K.

a) Se $E \cong E'$ e E tem dimensão n, então também E' tem dimensão n.

b) Se E e E' têm ambos dimensão n, então $E \cong E'$.

Demonstração:

a) Seja $\varphi : E \to E'$ um isomorfismo, e tomemos em E uma base $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$. Pelo facto de se ter $E = \langle \vec{e}_1, \vec{e}_2, ..., \vec{e}_n \rangle$ e atendendo a que φ é sobrejectiva, sabemos que $E' = \varphi(E) = \langle \varphi(\vec{e}_1), \varphi(\vec{e}_2), ..., \varphi(\vec{e}_n) \rangle$.

Por outro lado, sendo os vectores $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$ linearmente independentes e sendo φ um monomorfismo, os vectores $\varphi(\vec{e}_1), \varphi(\vec{e}_2), ..., \varphi(\vec{e}_n)$ serão também independentes. Logo, constituirão uma base de E' e teremos $\dim(E') = n$.

b) Se $\dim(E) = \dim(E') = n$, então, tal como se viu no Exemplo 4-24, sabemos que $E \cong K^n$ e $E' \cong K^n$.

Ora, tendo em conta o que se provou no Teorema 4-25, de $E \cong K^n$ e $K^n \cong E'$ resulta imediatamente $E \cong E'$.

4-4 Propriedade universal do quociente

Consideremos um espaço vectorial E, sobre um corpo K, e um seu subespaço vectorial F. Vimos, no Exemplo 4-2 i), que a aplicação $\pi_F : E \to E/F$, definida, para cada vector $\vec{x} \in E$, por $\pi_F(\vec{x}) = \vec{x} + F$, é uma aplicação linear, a aplicação linear canónica ou natural de E para E/F.

Por construção e atendendo ao que sabemos das classes de E/F, esta aplicação π_F é um epimorfismo, que designaremos, daqui para a frente, por **epimorfismo canónico** ou **natural**.

O epimorfismo canónico goza de uma propriedade da máxima importância, que passamos a apresentar:

Teorema 4-27 **Propriedade universal do quociente**

Seja F um subespaço de um espaço vectorial E. Seja, por sua vez, $\varphi : E \to E'$ uma aplicação linear tal que $F \subseteq \mathrm{Nuc}(\varphi)$. Então, existe uma aplicação linear $\varphi^* : E/F \to E'$ tal que $\varphi^* o \pi_F = \varphi$, isto é, tal que seja comutativo o seguinte diagrama:

$$\begin{tikzcd} E \arrow[r, "\pi_F"] \arrow[d, "\varphi"'] & E/F \arrow[dl, "\varphi^*"] \\ E' & \end{tikzcd}$$

A aplicação φ^* é um monomorfismo se e só se $F = \mathrm{Nuc}(\varphi)$, e um epimorfismo se e só se φ é um epimorfismo.

Demonstração:

Definimos $\varphi^* : E/F \to E'$ escrevendo $\varphi^*(\vec{x} + F) = \varphi(\vec{x})$, para cada vector $\vec{x} \in E$. Uma vez que a imagem de cada classe $\vec{x} + F$ é calculada a partir do seu representante $\vec{x}$, temos necessidade de nos assegurarmos de que a definição é coerente, isto é, de que a imagem não se altera se mudarmos o representante considerado.

Ora, sendo $\vec{x} + F = \vec{y} + F$, teremos, por definição, $\vec{x} - \vec{y} \in F \subseteq \mathrm{Nuc}(\varphi)$. Mas, então, $\varphi(\vec{x} - \vec{y}) = \vec{0}$, ou seja, $\varphi(\vec{x}) - \varphi(\vec{y}) = \vec{0}$, ou ainda $\varphi(\vec{x}) = \varphi(\vec{y})$, pelo que $\varphi^*(\vec{x} + F) = \varphi(\vec{x}) = \varphi(\vec{y}) = \varphi^*(\vec{y} + F)$.

Posto isto, é extremamente fácil verificar que a aplicação φ^* é uma aplicação linear:

$$\varphi^*\big(\alpha(\vec{x}+F)+\beta(\vec{y}+F)\big)=\varphi^*\big((\alpha\vec{x}+\beta\vec{y})+F\big)=$$
$$=\varphi(\alpha\vec{x}+\beta\vec{y})=\alpha\varphi(\vec{x})+\beta\varphi(\vec{y})=\alpha\varphi^*(\vec{x}+F)+\beta\varphi^*(\vec{y}+F)$$

Resulta agora imediatamente da definição que $\varphi^*\circ\pi_F=\varphi$, uma vez que, para um vector arbitrário tomado em E, se tem

$$(\varphi^*\circ\pi_F)(\vec{x})=\varphi^*(\pi_F(\vec{x}))=\varphi^*(\vec{x}+F)=\varphi(\vec{x})$$

Por sua vez, se φ^* é um epimorfismo, então a igualdade $\varphi^*\circ\pi_F=\varphi$ mostra que φ é a composição de dois epimorfismos; portanto, ele próprio, um epimorfismo. Reciprocamente, se φ é um epimorfismo, dado um vector qualquer $\vec{u}\in E'$, existirá um vector $\vec{u}\in E$ tal que $\varphi(\vec{u})=\vec{u}$, e, nessas condições,

$$\vec{u}'=\varphi(\vec{u})=(\varphi^*\circ\pi_F)(\vec{u})=\varphi^*(\pi_F(\vec{u}))=\varphi^*(\vec{u}+F)$$

o que, dada a arbitrariedade de $\vec{u}\in E'$, prova que φ^* é um epimorfismo. Finalmente, observemos que

$$\mathrm{Nuc}(\varphi^*)=\left\{\vec{x}+F\in\frac{E}{F}:\varphi^*(\vec{x}+F)=\vec{0}\right\}=\left\{\vec{x}+F\in\frac{E}{F}:\varphi(\vec{x})=\vec{0}\right\}=$$
$$=\left\{\vec{x}+F\in\frac{E}{F}:\vec{x}\in\mathrm{Nuc}(\varphi)\right\}=\frac{\mathrm{Nuc}(\varphi)}{F}$$

Assim, φ^* é um monomorfismo se e só se o seu núcleo se reduz ao vector nulo, isto é, se e só se

$$\frac{\mathrm{Nuc}(\varphi)}{F}=\left\{\vec{0}+F\right\}=\{F\}$$

ou seja, se e só se $F=Nuc(\varphi)$.

Da propriedade universal resultam diversas conclusões notáveis, uma das mais importantes das quais é o chamado "teorema do homomorfismo", que apresentamos já a seguir. A palavra "homomorfismo" pode considerar-se sinónima de "aplicação linear" quando se estudam espaços vectoriais, sendo a designação utilizada num contexto mais abstracto no estudo de estruturas algébricas quaisquer.

Teorema 4-28 **Teorema do homomorfismo**

Seja $\varphi : E \rightarrow E'$ uma aplicação linear qualquer, entre dois espaços vectoriais. Então, $\dfrac{E}{\mathrm{Nuc}(\varphi)} \cong \varphi(E)$. Em particular, se φ é um epimorfismo, tem-se

$$\frac{E}{\mathrm{Nuc}(\varphi)} \cong E'$$

Demonstração:

Consideremos o epimorfismo canónico

$$\pi_{\mathrm{Nuc}(\varphi)} : E \rightarrow \frac{E}{\mathrm{Nuc}(\varphi)}$$

Aplicando a propriedade universal do quociente à restrição $\varphi : E \rightarrow \varphi(E)$, construímos uma aplicação linear φ^* que torna comutativo o diagrama

$$\begin{tikzcd} E \arrow[r, "\pi_{\mathrm{Nuc}(\varphi)}"] \arrow[d, "\varphi"'] & \dfrac{E}{\mathrm{Nuc}(\varphi)} \arrow[dl, "\varphi^*"] \\ \varphi(E) & \end{tikzcd}$$

De acordo com a parte final do Teorema 4-29, podemos afirmar imediatamente que φ^* é um isomorfismo, pelo que

$$\frac{E}{\mathrm{Nuc}(\varphi)} \cong \varphi(E)$$

como se pretendia.

Por sua vez, o teorema do homomorfismo permitir-nos-á demonstrar os chamados "teoremas de isomorfismo":

Corolário 4-29 **1.º teorema de isomorfismo**

Sejam F e G subespaços de um espaço vectorial E. Tem-se:

$$\frac{F+G}{F} \cong \frac{G}{F \cap G}$$

Demonstração:

Considerando o epimorfismo canónico $\pi_F : E \to E/_F$, tomemos a sua restrição

$$\pi_F^o : F+G \to \frac{F+G}{F}.$$

Tomemos igualmente a aplicação $\tau : G \to F+G$, em que, para cada vector $\vec{g} \in G$, pomos $\tau(\vec{g}) = \vec{0} + \vec{g} = \vec{g}$; esta aplicação τ é manifestamente uma aplicação linear.

Podemos, nestas condições, fazer a composição

$$G \xrightarrow{\ \tau\ } F+G \xrightarrow{\ \pi_F^o\ } \frac{F+G}{F}$$

A aplicação $\pi_F^o \circ \tau$ é, como sabemos, uma aplicação linear, e tem-se

$$\begin{aligned}\mathrm{Nuc}\left(\pi_F^o \circ \tau\right) &= \left\{\vec{x} \in G : \left(\pi_F^o \circ \tau\right)(\vec{x}) = \vec{0} + F\right\} = \\ &= \left\{\vec{x} \in G : \pi_F^o\left(\tau(\vec{x})\right) = F\right\} = \left\{\vec{x} \in G : \vec{x} + F = F\right\} = \\ &= \left\{\vec{x} \in G : \vec{x} \in F\right\} = F \cap G\end{aligned}$$

Por outro lado, dado um vector qualquer $\vec{w} \in F+G$, poderemos escrever

$$\vec{w} = \vec{f} + \vec{g}, \quad \text{com} \quad \vec{f} \in F \ \text{ e } \ \vec{g} \in G$$

pelo que

$$\vec{w} + F = \left(\vec{f} + \vec{g}\right) + F = \vec{g} + F = \tau(\vec{g}) + F = \pi_F^o\left(\tau(\vec{g})\right) = \left(\pi_F^o \circ \tau\right)(\vec{g})$$

pelo que $\pi_F^o \circ \tau$ é um epimorfismo.

Mas então resulta do teorema do homomorfismo que

$$\frac{G}{\mathrm{Nuc}\left(\pi_F^o \circ \tau\right)} \cong \frac{F+G}{F}$$

ou seja, que

$$\frac{G}{F\cap G} \cong \frac{F+G}{F}$$

Corolário 4-30 **2.º teorema de isomorfismo**

Sejam F e G subespaços de um espaço vectorial E, tais que $F \subseteq G$. Então, $G/_F$ é um subespaço vectorial de $E/_F$, e tem-se

$$\frac{E/_F}{G/_F} \cong \frac{E}{G}$$

Demonstração:

Considerando o epimorfismo canónico $\pi_F : E \to E/_F$, vemos imediatamente que

$$G/_F = \pi_F(G)$$

e, como G é um subespaço vectorial de E, resulta daqui que $G/_F$ é um subespaço vectorial de $E/_F$.

Consideremos agora a aplicação

$$\sigma : E/_F \to E/_G$$

definida, para cada classe $\vec{x}+F \in E/_F$, por $\sigma(\vec{x}+F) = \vec{x}+G$.

Vamos ver que esta definição é coerente, isto é, que cada classe de $E/_F$ vai ter uma e uma só imagem em $E/_G$, seja qual for o representante tomado para a definir.

Com efeito, se for $\vec{x}+F = \vec{y}+F$, teremos, por definição,

$$\vec{x}-\vec{y} \in F \subseteq G$$

ou seja, $\vec{x}-\vec{y} \in G$, pelo que

$$\sigma(\vec{x}+F) = \vec{x}+G = \vec{y}+G = \sigma(\vec{y}+F)$$

É agora evidente que σ é uma aplicação linear, em virtude do modo como se fazem as operações num e noutro espaços. Mais do que isso, σ é obviamente um epimorfismo.

Calculemos o núcleo de σ:

$$\mathrm{Nuc}(\sigma) = \left\{\vec{x} + F \in {E}/{F} : \sigma(\vec{x} + F) = \vec{0} + G\right\} =$$
$$= \left\{\vec{x} + F \in {E}/{F} : \vec{x} + G = G\right\} = \left\{\vec{x} + F \in {E}/{F} : \vec{x} \in G\right\} = {G}/{F}$$

Aplicando então o teorema do homomorfismo, concluímos que

$$\frac{{E}/{F}}{{G}/{F}} = \frac{{E}/{F}}{Nuc(\sigma)} \cong \frac{E}{G}$$

como se pretendia.

5 Matrizes

5-1 Matriz de uma aplicação linear

Consideremos uma aplicação linear $\varphi : E \to E'$ entre dois espaços vectoriais. Suponhamos que o espaço E tem dimensão finita n e que o espaço E' tem dimensão finita p, e fixemos em E uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, e em E' uma base $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_p$.

Para cada vector da base de E, podemos calcular a respectiva imagem:

$$\begin{aligned} \vec{e}_1 &\mapsto \varphi(\vec{e}_1) \\ \vec{e}_2 &\mapsto \varphi(\vec{e}_2) \\ &\vdots \\ \vec{e}_n &\mapsto \varphi(\vec{e}_n) \end{aligned}$$

Uma vez que os vectores $\varphi(\vec{e}_1), \varphi(\vec{e}_2), \ldots, \varphi(\vec{e}_n)$ pertencem ao espaço vectorial E', poderemos escrevê-los como combinações lineares dos vectores da base fixa em E':

$$\begin{aligned} \varphi(\vec{e}_1) &= a_{11}\vec{e}'_1 + a_{21}\vec{e}'_2 + \cdots + a_{p1}\vec{e}'_p \\ \varphi(\vec{e}_2) &= a_{12}\vec{e}'_1 + a_{22}\vec{e}'_2 + \cdots + a_{p2}\vec{e}'_p \\ &\vdots \\ \varphi(\vec{e}_n) &= a_{1n}\vec{e}'_1 + a_{2n}\vec{e}'_2 + \cdots + a_{pn}\vec{e}'_p \end{aligned}$$

Com as componentes assim obtidas, construímos uma matriz do tipo $p \times n$, ou seja, com p linhas e n colunas:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}$$

Esta construção leva-nos à seguinte definição:

Definição 5-1 Seja $\varphi : E \to E'$ uma aplicação linear entre dois espaços vectoriais. Fixando no espaço E uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, e no espaço E' uma base $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_p$, a matriz construída pelo processo indicado acima é chamada **matriz da aplicação linear** φ em relação às bases consideradas. Representá-la-emos por

$$M\left(\varphi\ ;(\vec{e}_i)_{i=1,2,\ldots,n}\ ,(\vec{e}'_j)_{j=1,2,\ldots,p}\right)$$

As colunas desta matriz são constituídas pelas componentes dos vectores da base fixa em E, em relação à base fixa em E'.

A utilidade da matriz de uma aplicação linear, entre espaços de dimensão finita, resulta do facto de, a partir dela, ser possível calcular facilmente a imagem de um vector qualquer do domínio. Efectivamente, podemos demonstrar o seguinte:

Teorema 5-2 Nas condições anteriores, seja $\vec{x} \underset{(\vec{e}_i)}{\equiv} (x_1, x_2, \ldots, x_n)$ e seja

$$M\left(\varphi\ ;(\vec{e}_i)_{i=1,2,\ldots,n}\ ,(\vec{e}'_j)_{j=1,2,\ldots,p}\right) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}$$

Então, as componentes de $\varphi(\vec{x})$ em relação à base fixa em E' são:

$$\left(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n,\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \ldots, a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n\right)$$

Demonstração:
Atendendo às construções anteriores, vemos que

$$\varphi(\vec{x}) = \varphi\left(\sum_{i=1}^{n} x_i\vec{e}_i\right) = \sum_{i=1}^{n} x_i\varphi(\vec{e}_i) = \sum_{i=1}^{n} x_i\left(\sum_{j=1}^{p} a_{ji}\vec{e}'_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{ji}x_i\vec{e}'_j =$$

$$= \sum_{j=1}^{p}\sum_{i=1}^{n} a_{ji}x_i\vec{e}'_j = \sum_{j=1}^{p}\left(\sum_{i=1}^{n} a_{ji}x_i\right)\vec{e}'_j$$

As componentes de $\varphi(\vec{x})$ são, portanto, as indicadas no enunciado.

Vemos, assim, que as componentes de $\varphi(\vec{x})$ se obtêm a partir das componentes de $\vec{x}$ e dos elementos da matriz

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}$$

É habitual escreverem-se as componentes de um vector em relação a uma base fixa como uma coluna. Podemos, por isso, pôr

$$\vec{x} \equiv \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad \varphi(\vec{x}) \equiv \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n \end{bmatrix}$$

Mais ainda, é costume dizer-se que esta coluna

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n \end{bmatrix}$$

resulta de se "multiplicar" a matriz

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix} \quad \text{pela coluna} \quad \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

escrevendo-se

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Exemplo 5-3

Consideremos a aplicação linear $\delta : \mathbb{R}_2[x] \to \mathbb{R}_1[x]$ definida por $\delta(p(x)) = p'(x)$, onde $p'(x)$ designa a derivada de $p(x)$. Vamos fixar no espaço vectorial $\mathbb{R}_2[x]$ a base constituída pelos vectores

$$1-x,\ 1+x-x^2,\ 1+x+2x^2$$

e no espaço $\mathbb{R}_1[x]$ a base $1+2x,\ 2+3x$.

Vamos construir a matriz de δ em relação às bases consideradas, e, à custa dessa matriz, calcular a imagem do vector $2-3x+5x^2$.

Começamos por calcular as imagens dos vectores da base fixa em $\mathbb{R}_2[x]$:

$$\begin{array}{ccc} 1-x & \mapsto & -1 \\ 1+x-x^2 & \mapsto & 1-2x \\ 1+x+2x^2 & \mapsto & 1+4x \end{array}$$

Seguidamente, calculamos as componentes destas imagens em relação à base fixa em $\mathbb{R}_1[x]$:

$$\begin{aligned} -1 &= 3(1+2x) - 2(2+3x) \\ 1-2x &= -7(1+2x) + 4(2+3x) \\ 1+4x &= 5(1+2x) - 2(2+3x) \end{aligned}$$

Consequentemente, a matriz de δ em relação às bases consideradas é

$$A = \begin{bmatrix} 3 & -7 & 5 \\ -2 & 4 & -2 \end{bmatrix}$$

Para calcularmos agora a imagem de $2-3x+5x^2$, começamos por determinar as componentes deste vector em relação à base escolhida em $\mathbb{R}_2[x]$: obtemos

$$2-3x+5x^2 = \frac{5}{2}(1-x) - 2(1+x-x^2) + \frac{3}{2}(1+x+2x^2)$$

A partir da matriz A e da coluna

$$\begin{bmatrix} 5/2 \\ -2 \\ 3/2 \end{bmatrix}$$

obtemos as componentes da imagem de $2-3x+5x^2$ fazendo a "multiplicação"

$$\begin{bmatrix} 3 & -7 & 5 \\ -2 & 4 & -2 \end{bmatrix} \begin{bmatrix} 5/2 \\ -2 \\ 3/2 \end{bmatrix} = \begin{bmatrix} 29 \\ -16 \end{bmatrix}$$

Quer isto dizer que

$$\delta\left(2-3x+5x^2\right) = 29(1+2x) - 16(2+3x) = -3+10x$$

Ora, $-3+10x$ é, efectivamente, a derivada de $2-3x+5x^2$. ■

Vemos, portanto, que, dados dois espaços vectoriais, E e E', ambos de dimensões finitas n e p, respectivamente, e uma aplicação linear $\varphi : E \to E'$, fixando uma base em cada um dos espaços, podemos associar a φ uma matriz do tipo $p \times n$. Reciprocamente, desde que as bases estejam fixas, qualquer matriz do tipo $p \times n$ define uma aplicação linear de E para E', como veremos na proposição seguinte:

Teorema 5-4

Sejam E e E' espaços vectoriais, de dimensões finitas n e p, respectivamente. Fixemos em E uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, e em E' uma base $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_p$. Então, fazendo corresponder a cada aplicação linear $\varphi \in L\left(E, E'\right)$ a sua matriz em relação às bases consideradas, obtemos uma aplicação bijectiva entre o conjunto $L\left(E, E'\right)$ e o conjunto $\mathrm{M}_{p\times n}(K)$, formado por todas as matrizes do tipo $p \times n$, feitas com elementos do corpo K.

Demonstração:

Começaremos por provar o seguinte lema:

Lema 5-5

Nas condições do enunciado e tomando em E' vectores arbitrários $\vec{w}'_1, \vec{w}'_2, \ldots, \vec{w}'_n$, existe uma e uma só aplicação linear $\zeta : E \to E'$ tal que $\zeta\left(\vec{e}_j\right) = \vec{w}'_j$ (para $j = 1, \ldots, n$).

Demonstração:

Com efeito, definamos ζ do seguinte modo: como cada vector $\vec{x} \in E$ se pode escrever como combinação linear da base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ fixa no espaço E, tomando $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n$, fazemos $\zeta(\vec{x}) = x_1\vec{w}'_1 + x_2\vec{w}'_2 + \cdots + x_n\vec{w}'_n$.

É fácil verificar que ζ é uma aplicação linear, bastando observar que, com uma notação natural, se tem

$$\begin{aligned}\zeta\left(\alpha\vec{x}+\beta\vec{y}\right) &= \left(\alpha x_1+\beta y_1\right)\vec{w}_1'+\left(\alpha x_2+\beta y_2\right)\vec{w}_2'+\cdots+\left(\alpha x_n+\beta y_n\right)\vec{w}_n' = \\ &= \alpha x_1\vec{w}_1'+\beta y_1\vec{w}_1'+\alpha x_2\vec{w}_2'+\beta y_2\vec{w}_2'+\cdots+\alpha x_n\vec{w}_n'+\beta y_n\vec{w}_n' = \\ &= \left(\alpha x_1\vec{w}_1'+\alpha x_2\vec{w}_2'+\cdots+\alpha x_n\vec{w}_n'\right)+\left(\beta y_1\vec{w}_1'+\beta y_2\vec{w}_2'+\cdots+\beta y_n\vec{w}_n'\right) = \\ &= \alpha\left(x_1\vec{w}_1'+x_2\vec{w}_2'+\cdots+x_n\vec{w}_n'\right)+\beta\left(y_1\vec{w}_1'+y_2\vec{w}_2'+\cdots+y_n\vec{w}_n'\right) = \\ &= \alpha\zeta\left(\vec{x}\right)+\beta\zeta\left(\vec{y}\right)\end{aligned}$$

Por outro lado, resulta da própria construção de ζ que se tem $\zeta\left(\vec{e}_j\right)=\vec{w}_j'$, para cada $j=1,\dots,n$.

A unicidade de ζ resulta imediatamente das propriedades das aplicações lineares. Se houvesse uma outra aplicação linear $\vartheta : E \to E'$ tal que $\vartheta\left(\vec{e}_j\right)=\vec{w}_j'$, para cada vector da base fixa em E, então

$$\begin{aligned}\vartheta\left(\vec{x}\right) &= \vartheta\left(x_1\vec{e}_1+x_2\vec{e}_2+\cdots+x_n\vec{e}_n\right)=x_1\vartheta\left(\vec{e}_1\right)+x_2\vartheta\left(\vec{e}_2\right)+\cdots+\vartheta\left(\vec{e}_n\right)= \\ &= x_1\vec{w}_1'+x_2\vec{w}_2'+\cdots+x_n\vec{w}_n'=\zeta\left(\vec{x}\right)\end{aligned}$$

Assim, teríamos $\vartheta=\zeta$, o que prova a unicidade de ζ.

Podemos, agora, regressar à demonstração do Teorema 5-4:

Estando as bases de E e de E' fixas, vimos já de que maneira podemos associar, a cada aplicação linear $\varphi\in L\left(E,E'\right)$, a sua matriz em relação a essas bases.

Obtemos assim uma aplicação

$$\Theta : L\left(E,E'\right)\to \mathrm{M}_{p\times n}\left(K\right)$$

na qual

$$\Theta(\varphi)=M\left(\varphi\,;\left(\vec{e}_i\right)_{i=1,2,\dots,n},\left(\vec{e}_j'\right)_{j=1,2,\dots,p}\right)$$

Reciprocamente, dada uma matriz $A\in\mathrm{M}_{p\times n}\left(K\right)$, associamos a cada coluna de A um vector de E': $\vec{w}_j'=\sum_{i=1}^{p}a_{ij}\vec{e}_i'\ \left(j=1,2,\dots,n\right)$. De acordo com o Lema 5-5, estes vectores $\vec{w}_1',\vec{w}_2',\dots,\vec{w}_n'$, assim construídos, definem uma aplicação linear $\zeta_A : E\to E'$.

Por este processo, obtemos uma aplicação linear $\Omega : \mathrm{M}_{p\times n}\left(K\right)\to L\left(E,E'\right)$, na qual $\Omega\left(A\right)=\zeta_A$.

Se verificarmos que se tem $\Omega = \Theta^{-1}$, fica demonstrado que Θ e Ω são aplicações bijectivas.

Ora, por construção, a matriz de ζ_A é formada pelas componentes dos vectores $\zeta_A(\vec{e}_j)$ em relação à base de E'. Como, por construção, se tem

$$\zeta_A(\vec{e}_j) = \vec{w}'_j = \sum_{i=1}^{p} a_{ij}\vec{e}'_i$$

trata-se exactamente da matriz A.

Reciprocamente, sendo

$$B = M\left(\varphi\,;(\vec{e}_i)_{i=1,2,\ldots,n},(\vec{e}'_j)_{j=1,2,\ldots,p}\right)$$

a aplicação ζ_B, construída pelo processo descrito anteriormente é φ, uma vez que os vectores cujas componentes se encontram nas colunas de B são $\varphi(\vec{e}_j)$ (com $j = 1,2,\ldots,n$).

Fica, assim, terminada a demonstração.

No capítulo anterior, vimos que, a partir de aplicações lineares dadas, podemos construir outras. Vamos agora ver que há relações simples entre as respectivas matrizes.

Teorema 5-6 Sejam $\varphi : E \to E'$ e $\psi : E \to E'$ aplicações lineares entre dois espaços vectoriais de dimensões finitas. Fixemos em E uma base $\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n$, e em E' uma base $\vec{e}'_1,\vec{e}'_2,\ldots,\vec{e}'_p$, e suponhamos que, em relação a estas bases, a aplicação φ tem a matriz $A = [a_{ij}]$ $(i = 1,2,\ldots,p;\ j = 1,2,\ldots,n)$ e a aplicação ψ tem a matriz $A = [b_{ij}]$ $(i = 1,2,\ldots,p;\ j = 1,2,\ldots,n)$. Então, em relação às bases consideradas:

a) A matriz de $\varphi + \psi$, é

$$C = [a_{ij} + b_{ij}]\ (i = 1,2,\ldots,p;\ j = 1,2,\ldots,n)$$

b) Para cada escalar $\lambda \in K$, a matriz de $\lambda\varphi$ é

$$D = [\lambda a_{ij}]\ (i = 1,2,\ldots,p;\ j = 1,2,\ldots,n)$$

Demonstração:

a) Sabemos que, para cada vector $\vec{e}_j$, da base de E, se tem

$$(\varphi+\psi)(\vec{e}_j) = \varphi(\vec{e}_j) + \psi(\vec{e}_j) = \sum_{i=1}^{p} a_{ij}\vec{e}'_i + \sum_{i=1}^{p} b_{ij}\vec{e}'_i = \sum_{i=1}^{p}(a_{ij} + b_{ij})\vec{e}'_i$$

o que mostra que a matriz de $\varphi+\psi$, nas mesmas bases, é

$$C=\left[a_{ij}+b_{ij}\right]$$

b) Sabemos que, para cada vector $\vec{e}_j$, da base de E, se tem

$$(\lambda\varphi)\left(\vec{e}_j\right)=\lambda\varphi\left(\vec{e}_j\right)=\lambda\sum_{i=1}^{p}a_{ij}\vec{e}'_i=\sum_{i=1}^{p}\lambda a_{ij}\vec{e}'_i$$

o que mostra que a matriz de $\lambda\varphi$, nas mesmas bases, é $D=\left[\lambda a_{ij}\right]$.

As construções anteriores suscitam, de um modo muito natural, as seguintes definições:

Definição 5-7 No conjunto $\mathrm{M}_{p\times n}(K)$, formado pelas matrizes do tipo $p\times n$ com elementos do corpo K, definimos operações de "adição" e de "multiplicação por um escalar" $\alpha\in K$ do seguinte modo:

a) Sejam $A,B\in\mathrm{M}_{p\times n}(K)$. Sendo

$$A=\left[a_{ij}\right] \text{ e } B=\left[b_{ij}\right], \quad \text{com} \quad i=1,2,\dots,p \text{ e } j=1,2,\dots,n$$

a matriz $A+B$ é a matriz de $\mathrm{M}_{p\times n}(K)$ definida por

$$A+B=\left[a_{ij}+b_{ij}\right], \quad \text{para} \quad i=1,2,\dots,p \text{ e } j=1,2,\dots,n$$

b) Seja $A\in\mathrm{M}_{p\times n}(K)$ e seja $\alpha\in K$. Sendo

$$A=\left[a_{ij}\right], \quad \text{com} \quad i=1,2,\dots,p \text{ e } j=1,2,\dots,n$$

a matriz αA é a matriz de $\mathrm{M}_{p\times n}(K)$ definida por

$$\alpha A=\left[\alpha a_{ij}\right], \quad \text{para} \quad i=1,2,\dots,p \text{ e } j=1,2,\dots,n$$

É muito fácil transportar para as operações que acabámos de definir com matrizes as propriedades que conhecemos para as correspondentes operações definidas com aplicações lineares, beneficiando da existência de uma bijecção entre matrizes e aplicações lineares, que encontrámos no Teorema 5-4. Mais concretamente, podemos enunciar o próximo teorema.

Teorema 5-8 Consideremos o conjunto $\mathbf{M}_{p\times n}(K)$, formado pelas matrizes do tipo $p\times n$ com elementos do corpo K. Para matrizes arbitrárias, de $\mathbf{M}_{p\times n}(K)$, e escalares também quaisquer, são válidas as seguintes propriedades:

a) $A+B=B+A$

b) $(A+B)+C=A+(B+C)$

c) Sendo $0=\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$ tem-se $A+0=A$

d) Para $A=\left[a_{ij}\right]$, pomos $-A=\left[-a_{ij}\right]$, e tem-se $A+(-A)=0$

e) $\alpha(A+B)=\alpha A+\alpha B$

f) $(\alpha+\beta)A=\alpha A+\beta A$

g) $\alpha(\beta A)=(\alpha\beta)A$

h) $1A=A$

Demonstração:

Escolhendo espaços vectoriais E e E', de dimensões finitas n e p, respectivamente, e fixando bases em ambos os espaços, sabemos, pelo Teorema 5-4, que existe uma aplicação bijectiva $\Theta: L(E,E') \to \mathbf{M}_{p\times n}(K)$, que faz corresponder a cada aplicação linear a sua matriz em relação às bases consideradas.

As propriedades enunciadas, acerca das operações em $\mathbf{M}_{p\times n}(K)$, resultam então imediatamente das propriedades análogas conhecidas para as operações correspondentes definidas em $L(E,E')$, em face do que se viu no Teorema 5-6.

Por exemplo, para se provar que, dadas matrizes quaisquer $A,B \in \mathbf{M}_{p\times n}(K)$, se tem $A+B=B+A$, procedemos do seguinte modo:
Tomamos aplicações lineares $\varphi,\psi \in L(E,E')$ tais que $\Theta(\varphi)=A$, $\Theta(\psi)=B$. Pelo Teorema 5-6, sabemos que $\Theta(\varphi+\psi)=A+B$ e $\Theta(\psi+\varphi)=B+A$, pelo que

$$A+B=\Theta(\varphi+\psi)=\Theta(\psi+\varphi)=B+A$$

Em particular, resulta do modo como é construída a matriz de uma aplicação linear, em relação a certas bases, que a matriz da aplicação nula é precisamente a matriz

$$0=\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

do tipo $p\times n$, e que, se $\Theta(\varphi)=A$, então $\Theta(-\varphi)=-A$.

Por outro lado, o teorema anterior pode ser reformulado de um modo mais sintético, como se indica na primeira parte do corolário seguinte:

Corolário 5-9 Nas condições do teorema anterior, tem-se:

a) Com as operações acima definidas, $\mathrm{M}_{p\times n}(K)$ constitui um espaço vectorial sobre o corpo K.

b) O espaço vectorial $\mathrm{M}_{p\times n}(K)$ tem a dimensão pn.

c) Tomando espaços vectoriais E e E', de dimensões finitas n e p, respectivamente, $\mathrm{M}_{p\times n}(K) \cong L(E,E')$, pelo que $L(E,E')$ tem a dimensão pn.

Demonstração:

a) Resulta trivialmente do Teorema 5-8 e da definição de espaço vectorial.

b) Consideremos as matrizes $E_{ij} \in \mathrm{M}_{p\times n}(K)$, em que, para cada $i = 1,2,\dots,p$ e cada $j = 1,2,\dots,n$, se tem $E_{ij} = [e_{kt}]$, onde

$$e_{kt} = \begin{cases} 1, & \text{se } \; k = i \; \text{ e } \; t = j \\ 0, & \text{se } \; k \neq i \; \text{ ou } \; t \neq j \end{cases}$$

Resulta imediatamente da maneira como as operações foram definidas que cada matriz $A = [a_{ij}] \in \mathrm{M}_{p\times n}(K)$ tem a expressão única

$$A = \sum_{i=1}^{p} \sum_{j=1}^{n} a_{ij} E_{ij}$$

o que revela que as pn matrizes E_{ij} formam uma base de $\mathrm{M}_{p\times n}(K)$.

c) Tendo em atenção os teoremas anteriores, e fixando bases nos espaços vectoriais E e E', a aplicação

$$\Theta : L(E,E') \to \mathrm{M}_{p\times n}(K)$$

definida como acima, é um isomorfismo.

Há ainda, como vimos oportunamente, uma outra maneira de construir aplicações lineares a partir doutras: a composição. Importa-nos ver qual a matriz que corresponde à composição de duas aplicações lineares, cujas matrizes, em relação a certas bases, sejam conhecidas. Ora, podemos afirmar o seguinte:

Teorema 5-10 Sejam $\psi : E \to E'$ e $\varphi : E' \to E''$ aplicações lineares entre espaços de dimensão finita. Consideremos fixa em E a base $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$, em E' a base $\vec{e}'_1, \vec{e}'_2, ..., \vec{e}'_p$, e em E'', a base $\vec{e}''_1, \vec{e}''_2, ..., \vec{e}''_q$, e suponhamos que, em relação às bases consideradas, as aplicações φ e ψ são definidas pelas matrizes

$$A = [a_{ij}] \text{ e } B = [b_{jk}]$$

respectivamente. Então, sempre em relação às bases que fixámos, a aplicação linear $\varphi \circ \psi : E \to E''$ é definida por uma matriz

$$C = [c_{jk}] \ (i = 1, 2, ..., q;\ k = 1, 2, ..., n), \quad \text{na qual} \quad c_{ik} = \sum_{j=1}^{p} a_{ij} b_{jk}$$

Demonstração:

Atendendo à definição de matriz de uma aplicação linear, para construirmos a matriz C precisamos de calcular as imagens, por meio de $\varphi \circ \psi$, dos vectores da base escolhida em E. Temos, então:

$$(\varphi \circ \psi)(\vec{e}_k) = \varphi(\psi(\vec{e}_k)) = \varphi\left(\sum_{j=1}^{p} b_{jk} \vec{e}'_j\right) = \sum_{j=1}^{p} b_{jk} \varphi(\vec{e}'_j) =$$

$$= \sum_{j=1}^{p} b_{jk} \left(\sum_{i=1}^{k} a_{ij} \vec{e}''_i\right) = \sum_{j=1}^{p} \sum_{i=1}^{k} b_{jk} a_{ij} \vec{e}''_i = \sum_{i=1}^{k} \sum_{j=1}^{p} a_{ij} b_{jk} \vec{e}''_i = \sum_{i=1}^{k} \left(\sum_{j=1}^{p} a_{ij} b_{jk}\right) \vec{e}''_i$$

Estas igualdades mostram precisamente que a matriz C, correspondente à aplicação composta $\varphi \circ \psi$, é constituída pelos elementos

$$c_{ik} = \sum_{j=1}^{p} a_{ij} b_{jk}$$

Podemos, agora, introduzir a seguinte definição:

Definição 5-11 a) Dadas as matrizes

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \text{ e } \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

dos tipos $1\times n$ e $n\times 1$, respectivamente, define-se o seu **produto** como sendo a seguinte matriz do tipo 1×1:

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1b_1 + a_2b_2 + \cdots + a_nb_n \end{bmatrix} = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

b) Dadas as matrizes $A \in \mathrm{M}_{q\times p}(K)$ e $B \in \mathrm{M}_{p\times n}(K)$, define-se o seu **produto** como sendo a matriz $AB \in \mathrm{M}_{q\times n}(K)$ que se obtém multiplicando cada linha de A por cada coluna de B, tal como foi definido na alínea a), isto é, a matriz AB, que é formada pelos elementos c_{ik} ($i = 1,2,...,q;\ k = 1,2,...,n$), com

$$c_{ik} = \sum_{j=1}^{p} a_{ij}b_{jk}$$

Deve salientar-se que, de acordo com esta definição de produto de matrizes, a multiplicação de uma matriz por outra é possível se e só se o número de colunas do primeiro factor coincide com o número de linhas do segundo factor. Note-se ainda que o produto de uma matriz do tipo $q\times p$ por uma matriz do tipo $p\times n$ é uma matriz do tipo $q\times n$, isto é, que a matriz produto tem tantas linhas como o primeiro factor e tantas colunas como o segundo.

Perante a definição anterior, podemos imediatamente reformular o Teorema 5-10, enunciando-o da seguinte maneira:

Teorema 5-12 Se $\psi : E \to E'$ e $\varphi : E' \to E''$ são aplicações lineares entre espaços de dimensão finita, definidas, em relação a certas bases fixas nos vários espaços vectoriais, pelas matrizes A e B, respectivamente, então, sempre em relação às bases escolhidas, a aplicação linear $\varphi \circ \psi : E \to E''$ é definida pela matriz AB.

É fácil provar propriedades desta operação de multiplicação de matrizes, bem como propriedades que a relacionem com as operações definidas anteriormente:

Teorema 5-13 São válidas as seguintes propriedades:

a) $(AB)C = A(BC),\ \forall A \in \mathbf{M}_{q\times p}(K),\ B \in \mathbf{M}_{p\times n}(K),\ C \in \mathbf{M}_{n\times s}(K)$

b) $(A + A')B = AB + A'B,\ \forall A, A' \in \mathbf{M}_{q\times p}(K),\ B \in \mathbf{M}_{p\times n}(K)$

c) $A(B + B') = AB + AB',\ \forall A \in \mathbf{M}_{q\times p}(K),\ B, B' \in \mathbf{M}_{p\times n}(K)$

d) $(\alpha A)B = A(\alpha B) = \alpha(AB),\ \forall A \in \mathbf{M}_{q\times p}(K),\ B \in \mathbf{M}_{p\times n}(K),\ \forall \alpha \in K$

e) $|AB| = |A||B|,\ \forall A, B \in \mathbf{M}_{n\times n}(K)$

Demonstração:

a) Consideremos espaços vectoriais E, E', E'' e E''', sobre o corpo K, de dimensões s, n, p e q, respectivamente, e consideremos fixa em cada um deles uma base. Nestas condições, as matrizes A, B e C definem aplicações lineares, $\varphi : E'' \to E'''$, $\psi : E' \to E''$ e $\theta : E \to E'$, respectivamente.

Sabemos que a composição de aplicações é associativa, ou seja, que

$$(\varphi \circ \psi) \circ \theta = \varphi \circ (\psi \circ \theta)$$

e também que à composição de aplicações lineares corresponde o produto das respectivas matrizes, desde que as bases estejam fixas. Por consequência, podemos concluir, da igualdade anterior, que

$$(AB)C = A(BC)$$

b) Consideremos espaços vectoriais E, E' e E'', sobre o corpo K, de dimensões n, p e q, respectivamente, e consideremos fixa em cada um deles uma base. Nestas condições, as matrizes A, A' e B definem aplicações lineares, $\varphi : E' \to E''$, $\varphi' : E' \to E''$ e $\psi : E \to E'$, respectivamente.

Sabemos que se tem

$$(\varphi + \varphi') \circ \psi = \varphi \circ \psi + \varphi' \circ \psi$$

e também que à composição de aplicações lineares corresponde o produto das respectivas matrizes, e à soma de aplicações lineares corresponde a soma das respectivas matrizes, desde que as bases estejam fixas. Por consequência, podemos concluir, da igualdade anterior, que

$$(A + A')B = AB + A'B$$

c) A justificação é análoga à da propriedade anterior.

d) Consideremos espaços vectoriais E, E' e E'', sobre o corpo K, de dimensões n, p e q, respectivamente, e consideremos fixa em cada um deles uma base. Nestas condições, as matrizes A e B definem aplicações lineares, $\varphi : E' \to E''$ e $\psi : E \to E'$, respectivamente.

Sabemos que se tem

$$(\alpha\varphi)\circ\psi = \varphi\circ(\alpha\psi) = \alpha(\varphi\circ\psi)$$

e também que à composição de aplicações lineares corresponde o produto das respectivas matrizes, e ao produto de uma aplicação linear por um escalar corresponde o produto da respectiva matriz por esse escalar, desde que as bases estejam fixas. Por consequência, podemos concluir, da igualdade anterior, que

$$(\alpha A)B = A(\alpha B) = \alpha(AB)$$

e) Consideremos as matrizes

$$A = [a_{ij}] \text{ e } B = [b_{jk}]$$

com I, j e k tomando valores de 1 a n. O produto AB será uma matriz $C = [c_{ik}]$, na qual

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

Vamos, então, calcular o determinante da matriz AB. Por definição, será

$$|AB| = \sum (-1)^{\sigma} c_{1k_1} c_{2k_2} \cdots c_{nk_n}$$

onde $k_1, k_2, \ldots, k_n$ percorre todas as possíveis permutações dos números $1, 2, \ldots, n$, e o expoente σ é 0 ou 1 consoante a permutação $k_1, k_2, \ldots, k_n$ é par ou ímpar.

Temos, então,

$$|AB| = \sum (-1)^{\sigma} c_{1k_1} c_{2k_2} \cdots c_{nk_n} = \sum (-1)^{\sigma} \left(\sum_{j_1=1}^{n} a_{1j_1} b_{j_1 k_1} \right) c_{2k_2} \cdots c_{nk_n} =$$

$$= \sum \sum_{j_1=1}^{n} (-1)^{\sigma} a_{1j_1} b_{j_1 k_1} c_{2k_2} \cdots c_{nk_n} = \sum \sum_{j_1=1}^{n} (-1)^{\sigma} a_{1j_1} b_{j_1 k_1} \left(\sum_{j_2=1}^{n} a_{2j_2} b_{j_2 k_2} \right) \cdots c_{nk_n} =$$

$$= \sum \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} (-1)^{\sigma} a_{1j_1} b_{j_1 k_1} a_{2j_2} b_{j_2 k_2} \cdots c_{nk_n} =$$

$$= \sum \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_n}^{n} (-1)^{\sigma} a_{1j_1} b_{j_1 k_1} a_{2j_2} b_{j_2 k_2} \cdots a_{nj_n} b_{j_n k_n} =$$

$$= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_n=1}^{n} \left(\sum (-1)^{\sigma} b_{j_1 k_1} b_{j_2 k_2} \cdots b_{j_n k_n} \right) a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

Ora, se comparamos a expressão

$$\sum (-1)^{\sigma} b_{j_1 k_1} b_{j_2 k_2} \cdots b_{j_n k_n}$$

onde $k_1, k_2, \ldots, k_n$ percorre todas as possíveis permutações dos números $1, 2, \ldots, n$, e o expoente σ é 0 ou 1 consoante a permutação $k_1, k_2, \ldots, k_n$ é par ou ímpar, com a expressão geral de um determinante verificamos que a expressão em causa corresponde ao determinante de uma matriz $B_{j_1 j_2 \cdots j_n}$ cuja primeira linha é a linha j_1 de B, cuja segunda linha é a linha j_2 de B, e assim sucessivamente.

Sempre que haja dois índices iguais entre $j_1, j_2, \ldots, j_n$, a matriz $B_{j_1 j_2 \cdots j_n}$ tem duas linhas iguais e o seu determinante é nulo. Por consequência, na expressão que obtivemos para o determinante de AB, podemos desprezar as parcelas em que haja índices $j_1, j_2, \ldots, j_n$ repetidos e aproveitar apenas aquelas em que os índices $j_1, j_2, \ldots, j_n$ formem uma permutação dos números $1, 2, \ldots, n$. Ora, neste caso, a matriz $B_{j_1 j_2 \cdots j_n}$ terá todas as linhas de B, embora por outra ordem.

Se colocarmos as linhas por ordem para obtermos a matriz B, o determinante ficará inalterado se a permutação $j_1, j_2, \ldots, j_n$ for par, e mudará de sinal se a permutação $j_1, j_2, \ldots, j_n$ for ímpar. Podemos, então, escrever

$$|AB| = \sum (-1)^{\tau} |B| a_{1j_1} a_{2j_2} \cdots a_{nj_n} = |B| \sum (-1)^{\tau} a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

onde $j_1, j_2, ..., j_n$ percorre todas as permutações dos números $1,2,...,n$, e o expoente τ toma o valor 0 ou 1 conforme a permutação $j_1, j_2, ..., j_n$ é par ou ímpar.

Mas, por definição, teremos, nestas condições,

$$\sum (-1)^{\tau} a_{1j_1} a_{2j_2} \cdots a_{nj_n} = |A|$$

pelo que, finalmente, $|AB| = |B|\,|A|$, como pretendíamos provar.

Para certos fins, há ainda a necessidade de introduzir uma outra construção, que podemos relacionar com as operações definidas acima. Trata-se da noção de "transposição" de uma matriz, que passamos a apresentar:

Definição 5-14 Dada uma matriz A qualquer, do tipo $p \times n$, chama-se **matriz transposta** de A àquela cujas linhas são as colunas de A, e vice-versa. Representa-se a matriz transposta de A por A^T.

Assim, sendo

$$A = \left[a_{ij}\right] \text{ (com } i = 1,2,...,p;\ j = 1,2,...,n)$$

a matriz A^T será uma matriz do tipo $n \times p$, podendo-se escrever

$$A^T = \left[a'_{ji}\right] \text{ (com } i = 1,2,...,p;\ j = 1,2,...,n)$$

sendo $a'_{ji} = a_{ij}$, para cada i e cada j.

Exemplo 5-15

Sendo

$$A = \begin{bmatrix} 3 & 5 & 2 \\ 9 & -2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 6 & 7 & 3 \\ 9 & -1 & 7 \\ -8 & 1 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 2 \\ 5 \\ 1 \\ 9 \end{bmatrix}$$

teremos

$$A^T = \begin{bmatrix} 3 & 9 \\ 5 & -2 \\ 2 & 1 \end{bmatrix}, \quad B^T = \begin{bmatrix} 6 & 9 & -8 \\ 7 & -1 & 1 \\ 3 & 7 & 2 \end{bmatrix}, \quad C^T = \begin{bmatrix} 2 & 5 & 1 & 9 \end{bmatrix} \ \blacksquare$$

É fácil demonstrar algumas propriedades elementares da transposição de matrizes:

Teorema 5-16 Para matrizes arbitrárias e para um escalar qualquer, são válidas (desde que façam sentido) as igualdades seguintes:

a) $(A+B)^T = A^T + B^T$

b) $(\alpha A)^T = \alpha A^T$

c) $(AB)^T = B^T A^T$

d) $\left(A^T\right)^T = A$

Demonstração:

a) Sendo $A=[a_{ij}]$ e $B=[b_{ij}]$, temos $A+B=[s_{ij}]$, com $s_{ij}=a_{ij}+b_{ij}$. Consequentemente, a transposta de $A+B$ será uma matriz $(A+B)^T = [s'_{ji}]$, onde $s'_{ji} = s_{ij} = a_{ij} + b_{ij}$.

Por sua vez, teremos $A^T = [a'_{ji}]$, com $a'_{ji} = a_{ij}$, e $B^T = [b'_{ji}]$ com $b'_{ji} = b_{ij}$. Portanto, $A^T + B^T = [a'_{ji} + b'_{ji}]$, e temos $a'_{ji} + b'_{ji} = a_{ij} + b_{ij} = s_{ij} = s'_{ji}$.

Assim, as matrizes $(A+B)^T$ e $A^T + B^T$, para além de serem do mesmo tipo, são formadas pelos mesmos elementos.

b) Sendo $A=[a_{ij}]$, temos $\alpha A = [\alpha a_{ij}]$, e, portanto, $(\alpha A)^T = [c'_{ji}]$, com $c'_{ji} = \alpha a_{ij}$. Por sua vez, será $A^T = [a'_{ji}]$ com $a'_{ji} = a_{ij}$, pelo que $\alpha A^T = [\alpha a'_{ji}]$. Ora, $\alpha a'_{ji} = \alpha a_{ij} = c'_{ji}$.

c) Tomemos $A=[a_{ij}]$ e $B=[b_{jk}]$. Teremos, então, $AB=[d_{ik}]$, onde

$$d_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

Por consequência, $(AB)^T = [d'_{ki}]$, onde

$$d'_{ki} = d_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

Por outro lado, teremos $A^T = [a'_{ji}]$, com $a'_{ji} = a_{ij}$, e $B^T = [b'_{kj}]$, com $b'_{kj} = b_{jk}$. Assim sendo, $B^T A^T = [d''_{ki}]$, onde

$$d''_{ki} = \sum_{j=1}^{n} b'_{kj} a'_{ji} = \sum_{j=1}^{n} b_{jk} a_{ij} = d_{ik} = d'_{ki}$$

d) Resulta trivialmente da definição.

5-2 Matrizes invertíveis

Consideremos uma aplicação linear qualquer $\varphi : E \to E'$ entre dois espaços vectoriais. Suponhamos que o espaço E tem dimensão finita n e que o espaço E' tem dimensão finita p, e fixemos em E uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ e em E' uma base $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_p$. Nestas condições, a aplicação linear φ será definida por uma determinada matriz A, do tipo $p \times n$.

Consideremos a aplicação identidade $1_E : E \to E$, que é, como sabemos, uma aplicação linear. Vamos determinar a matriz desta aplicação, em relação à base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, fixa no espaço E, considerado quer como espaço de partida quer como espaço de chegada de 1_E. Para isso, basta seguir a definição de matriz de uma aplicação linear e observar que

$$\begin{gathered} 1_E(\vec{e}_1) = \vec{e}_1 = 1\vec{e}_1 + 0\vec{e}_2 + \cdots + 0\vec{e}_n \\ 1_E(\vec{e}_2) = \vec{e}_2 = 0\vec{e}_1 + 1\vec{e}_2 + \cdots + 0\vec{e}_n \\ \vdots \\ 1_E(\vec{e}_n) = \vec{e}_n = 0\vec{e}_1 + 0\vec{e}_2 + \cdots + 1\vec{e}_n \end{gathered}$$

A matriz de 1_E em relação à base considerada será, portanto,

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Esta matriz, do tipo $n \times n$, vai desempenhar um papel importante, pelo que vamos atribuir-lhe uma designação especial:

Definição 5-17 Para cada número natural n, a matriz

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

toma o nome de **matriz identidade** de ordem n.

Esta matriz identidade vai gozar de uma propriedade fundamental, a saber:

Teorema 5-18 Dada uma matriz qualquer $A \in \mathrm{M}_{p\times n}(K)$, tem-se

$$I_p A = A = A I_n$$

Demonstração:

Dada a matriz A, do tipo $p \times n$, consideremos um espaço vectorial E, de dimensão finita n, e um espaço vectorial E', de dimensão finita p, sobre o corpo K. Fixando em E uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, e em E', uma base $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_p$, a matriz A definirá, como vimos, uma aplicação linear $\varphi : E \to E'$.

Ora, sabemos que se tem

$$1_{E'} \circ \varphi = \varphi = \varphi \circ 1_E$$

e também que à composição de aplicações lineares corresponde o produto das respectivas matrizes, desde que as bases estejam fixas.

Por outro lado, atendendo a que a matriz de uma aplicação identidade em relação a uma certa base é, como vimos acima, a respectiva matriz identidade, podemos concluir, da igualdade anterior, que

$$I_p A = A = A I_n$$

O teorema anterior revela que, desde que a multiplicação se possa fazer, a matriz identidade, seja de que ordem for, se comporta como um "elemento neutro", ou seja, o produto da matriz identidade por qualquer outra, tanto do lado esquerdo como do lado direito, não afecta essa outra.

Estas ideias sugerem que se introduza a seguinte definição:

Definição 5-19 Seja A uma matriz qualquer. Diz-se que a matriz A é **invertível** quando existe uma matriz A' tal que $AA' = A'A = I_n$, para algum número natural n.

É fácil verificar o seguinte:

Teorema 5-20 Seja A uma matriz invertível. Então,

a) A é uma matriz quadrada (isto é, tem tantas linhas como colunas).

b) A matriz A', referida na definição, é única.

Demonstração:

a) Sendo A invertível, existirá, por definição, uma matriz A' tal que $AA' = A'A = I_n$.

Sabe-se que o número de linhas da matriz produto AA' vai ser igual ao número de linhas de A. Portanto, para que o produto AA' vá dar a matriz I_n, o número de linhas de A tem de ser n. Mas o número de colunas de $A'A$ coincide com o número de colunas de A, pelo que, sendo $A'A = I_n$, o número de colunas de A será igualmente n.

b) Suponhamos que as matrizes A' e A'' verificam ambas a definição anterior, isto é, que

$$AA' = A'A = I_n \quad \text{e} \quad AA'' = A''A = I_n$$

Atendendo à propriedade associativa da multiplicação de matrizes, sabemos que se tem

$$(A'A)A'' = A'(AA'')$$

De acordo com as hipóteses, obtemos, a partir desta igualdade,

$$I_nA'' = A'I_n, \quad \text{logo} \quad A'' = A'$$

A unicidade da matriz A', referida na definição anterior, que acabamos de comprovar, permite-lhe dar-lhe um nome. Temos, então:

Definição 5-21 Seja A uma matriz invertível. À matriz A', referida na definição, chama-se **matriz inversa** de A e representa-se por A^{-1}. Tem-se, portanto,

$$AA^{-1} = A^{-1}A = I_n$$

É claro que existem matrizes invertíveis, porque, por exemplo, resulta trivialmente da definição que $I_n^{-1} = I_n$. No entanto, nem todas as matrizes são invertíveis: na verdade, basta considerar a matriz nula, do tipo $n \times n$ e observar que, seja qual for a matriz B, do mesmo formato, se tem

$$0B = B0 = 0 \neq I_n$$

Pode, pois, perguntar-se como poderemos distinguir as matrizes invertíveis das não invertíveis, e, para as matrizes efectivamente invertíveis, como poderemos determinar a respectiva matriz inversa. Vamos passar a responder a estas questões:

Teorema 5-22 Seja A uma matriz do tipo $n \times n$. Sejam E e E' dois espaços vectoriais, ambos de dimensão n, nos quais fixámos, respectivamente, as bases $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ e $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_n$, e seja $\varphi : E \to E'$ a aplicação linear definida pela matriz A em relação a essas bases. Então, a matriz A é invertível se e só se φ é um isomorfismo.

Demonstração:
Suponhamos que A é invertível, e seja $\psi : E' \to E$ a aplicação linear definida, em relação às bases consideradas, pela matriz A^{-1}. Sabemos que ao produto das matrizes corresponde a composição das respectivas aplicações lineares, desde que as bases estejam fixas. Consequentemente, a matriz $A^{-1}A$ é a matriz da aplicação linear $\psi \circ \varphi$ em relação à base $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$, do espaço E, e a matriz AA^{-1} é a matriz da aplicação linear $\varphi \circ \psi$ em relação à base $\vec{e}'_1, \vec{e}'_2, ..., \vec{e}'_n$, do espaço E'.

Mas $A^{-1}A = I_n$, e a matriz I_n é a matriz da aplicação 1_E em relação à base $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$, pelo que $\psi \circ \varphi = 1_E$; analogamente, tem-se $AA^{-1} = I_n$, e podemos concluir que $\varphi \circ \psi = 1_{E'}$.

Ora, das igualdades $\psi \circ \varphi = 1_E$ e $\varphi \circ \psi = 1_{E'}$ resulta que $\psi = \varphi^{-1}$, pelo que φ é um isomorfismo.

Reciprocamente, se φ é um isomorfismo, então podemos tomar o isomorfismo inverso $\psi = \varphi^{-1}$. Designando por A' a matriz de φ^{-1}, resulta das igualdades $\varphi^{-1} \circ \varphi = 1_E$ e $\varphi \circ \varphi^{-1} = 1_{E'}$ que se tem $A'A = I_n$ e $AA' = I_n$, pelo que A é invertível.

O Teorema 5-22 responde já, no plano teórico, às questões formuladas antes do seu enunciado: são invertíveis as matrizes associadas a isomorfismos, e a matriz inversa de uma matriz invertível é a que representa, em relação às mesmas bases, o isomorfismo inverso. Na prática, porém, há maneiras mais simples de verificar se uma dada matriz é ou não invertível, e de obter, nesse caso, a sua inversa. Vejamos como:

Teorema 5-23 Seja A uma matriz do tipo $n \times n$. A matriz A é invertível se e só se o seu determinante é não nulo.

Demonstração:
Seja E um espaço vectorial de dimensão n, com uma base fixa $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$. Seja $\zeta : E \to E$ o endomorfismo de E definido pela matriz A em relação a essa base.

Sabemos, pelo teorema anterior, que A é invertível se e só se ζ é um isomorfismo, isto é, se e só se

$$\mathrm{Nuc}(\zeta) = \{\vec{0}\}$$

Ora, representando cada vector $\vec{x}$ de E pela coluna

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

das suas componentes, o núcleo de ζ é constituído pelos vectores cujas componentes satisfaçam o sistema de equações $AX = 0$.

Para ser $\mathrm{Nuc}(\zeta) = \{\vec{0}\}$, este sistema de equações tem de ser possível e determinado; logo, necessariamente, $|A| \neq 0$. Reciprocamente, sendo $|A| \neq 0$, o sistema de equações $AX = 0$ será um sistema de Cramer; portanto, possível e determinado, pelo que $\mathrm{Nuc}(\zeta) = \{\vec{0}\}$.

No que se refere, por sua vez, à busca da inversa de uma dada matriz invertível, comecemos por observar o seguinte:

Teorema 5-24 Seja A uma matriz do tipo $n \times n$. Se existir uma matriz B, do tipo $n \times n$, tal que $AB = I_n$, então A é invertível e $B = A^{-1}$.

Demonstração:

Suponhamos que se tem $AB = I_n$. Consideremos um espaço vectorial E, de dimensão n, com uma base fixa $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$. Sejam φ e ψ os endomorfismos de E definidos pelas matrizes A e B, respectivamente, em relação à base considerada.

Pelo facto de se ter $AB = I_n$, podemos concluir que $\varphi \circ \psi = 1_E$.

Tomando, então, um vector arbitrário $\vec{x} \in E$, podemos escrever

$$\vec{x} = 1_E(\vec{x}) = (\varphi \circ \psi)(\vec{x}) = \varphi(\psi(\vec{x}))$$

o que, dada a arbitrariedade de $\vec{x}$, mostra que a aplicação φ é sobrejectiva.

Tratando-se de um endomorfismo de um espaço vectorial de dimensão finita, daí resulta que φ é mesmo um isomorfismo, pelo que A é invertível.

Resulta então imediatamente que $B = A^{-1}$.

Segue-se, do Teorema 5-24, que, para procurar a inversa de uma matriz invertível A, bastará procurar uma matriz B que verifique a igualdade $AB = I_n$.

Para procurarmos uma tal matriz, poderíamos proceder coluna por coluna: o produto de A pela coluna número j de B tem de dar a coluna número j da matriz identidade:

$$A \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Para encontrar cada coluna de B, teríamos, portanto, de resolver um sistema de equações cuja matriz simples seria A e cujos termos independentes seriam constituídos por uma das colunas da matriz identidade. Ora, sendo a matriz simples comum aos n sistemas de equações a resolver, podemos resolvê-los simultaneamente.

Por exemplo, para procurarmos a inversa da matriz

$$A=\begin{bmatrix}1&1&0\\0&1&1\\1&0&1\end{bmatrix}$$

teríamos de resolver os três sistemas de equações cujas matrizes ampliadas seriam

$$\left[\begin{array}{ccc|c}1&1&0&1\\0&1&1&0\\1&0&1&0\end{array}\right],\quad \left[\begin{array}{ccc|c}1&1&0&0\\0&1&1&1\\1&0&1&0\end{array}\right]\quad \text{e}\quad \left[\begin{array}{ccc|c}1&1&0&0\\0&1&1&0\\1&0&1&1\end{array}\right]$$

que se podem resolver simultaneamente, usando o método de Gauss para condensar a matriz ampliada

$$\left[\begin{array}{ccc|ccc}1&1&0&1&0&0\\0&1&1&0&1&0\\1&0&1&0&0&1\end{array}\right]$$

Teremos, sucessivamente,

$$\left[\begin{array}{ccc|ccc}1&1&0&1&0&0\\0&1&1&0&1&0\\1&0&1&0&0&1\end{array}\right]\to\left[\begin{array}{ccc|ccc}1&1&0&1&0&0\\0&1&1&0&1&0\\0&-1&1&-1&0&1\end{array}\right]\to\left[\begin{array}{ccc|ccc}1&0&-1&1&-1&0\\0&1&1&0&1&0\\0&0&2&-1&1&1\end{array}\right]\to$$

$$\to\left[\begin{array}{ccc|ccc}1&0&-1&1&-1&0\\0&1&1&0&1&0\\0&0&1&-\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\end{array}\right]\to\left[\begin{array}{ccc|ccc}1&0&0&\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}\\0&1&0&\frac{1}{2}&\frac{1}{2}&-\frac{1}{2}\\0&0&1&-\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\end{array}\right]$$

Os três sistemas de equações considerados ficam assim resolvidos, e temos

$$A^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

O processo exposto é, evidentemente, um processo geral, que podemos resumir do seguinte modo: para se procurar a inversa de uma matriz invertível A escrevemos a matriz ampliada $\left[A \;\middle|\; I_n\right]$ e utilizamos o método de Gauss para proceder à sua condensação. Obter-se-á, no final, a matriz $\left[I_n \;\middle|\; A^{-1}\right]$, desde que, ao longo da condensação, não se efectuem trocas de colunas, o que é sempre possível quando se parte de uma matriz invertível. Resumidamente:

$$\left[A \;\middle|\; I_n\right] \xrightarrow[\text{condensar sem trocar colunas}]{} \left[I_n \;\middle|\; A^{-1}\right]$$

Há ainda um outro modo de determinar a matriz inversa de uma dada matriz invertível, que se apoia no teorema de Laplace. Passamos a expô-lo:

Seja A uma matriz qualquer, do tipo $n \times n$. Para cada elemento a_{ij}, da matriz A, consideremos o respectivo complemento algébrico Δ_{ij}.

Sabe-se, precisamente pelo teorema de Laplace, que o determinante de A se pode obter à custa dos complementos algébricos de uma dada linha ou coluna, através das expressões

$$|A| = \sum_{i=1}^{n} a_{ij}\Delta_{ij} = \sum_{j=1}^{n} a_{ij}\Delta_{ij}$$

Recordamos ainda o Corolário 1-27, segundo o qual, dada uma matriz A do tipo $n \times n$, se tem

$$\sum_{i=1}^{n} a_{ik}\Delta_{ij} = 0, \forall k \neq j \quad \text{e} \quad \sum_{j=1}^{n} a_{kj}\Delta_{ij} = 0, \forall k \neq i$$

ou seja, a soma dos produtos dos elementos de uma linha (resp.: coluna) de A pelos complementos algébricos de uma outra linha (resp.: coluna) dá sempre 0.

Podemos, portanto, afirmar que

$$\sum_{i=1}^{n} a_{ik}\Delta_{ij} = \begin{cases} |A|, & \text{se } k = j \\ 0\ , & \text{se } k \neq j \end{cases}$$

Em face disto, se tomarmos a matriz $\tilde{A} = [\Delta_{ij}]$ $(i, j = 1, 2, \ldots, n)$, resulta imediatamente da forma como se multiplicam matrizes que se tem

$$A\tilde{A}^T = \tilde{A}^T A = \begin{bmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & |A| \end{bmatrix} = |A|\, I_n$$

A matriz $\tilde{A}^T$ tem um nome especial, que passamos a indicar:

Definição 5-25 Nas condições anteriores, a matriz $\tilde{A}$ é chamada **matriz complementar** de A. A matriz $\tilde{A}^T$, por sua vez, é a **matriz adjunta** de A.

Supondo, agora, que a matriz A é invertível, sabemos, pelo Teorema 5-23, que o seu determinante não é nulo, pelo que podemos voltar a escrever a igualdade

$$A\tilde{A}^T = \tilde{A}^T A = |A| \; I_n \quad \text{na forma} \quad A\frac{\tilde{A}^T}{|A|} = \frac{\tilde{A}^T}{|A|}A = I_n$$

Ora, estas igualdades significam, pela própria definição de matriz inversa, que se tem

$$A^{-1} = \frac{\tilde{A}^T}{|A|}$$

Assim obtemos uma fórmula para a determinação da matriz inversa.

Finalmente, podemos ainda registar mais algumas propriedades das matrizes invertíveis:

Teorema 5-26 Para matrizes invertíveis quaisquer, de ordem n, tem-se:

a) $(AB)^{-1} = B^{-1}A^{-1}$

b) $(A^{-1})^{-1} = A$

c) $(A^T)^{-1} = (A^{-1})^T$

d) $|A^{-1}| = \dfrac{1}{|A|}$

Demonstração:

a) Basta observar que

$$(AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$

b) Resulta trivialmente da definição.

c) Basta observar que

$$A^T\left(A^{-1}\right)^T = \left(A^{-1}A\right)^T = I_n^T = I_n$$

d) De $AA^{-1} = I_n$ resulta que $\left|AA^{-1}\right| = \left|I_n\right| = 1$. Atendendo a que o determinante do produto de duas matrizes quadradas coincide com o produto dos determinantes dos factores, obtemos $|A|\left|A^{-1}\right| = 1$, pelo que

$$\left|A^{-1}\right| = \frac{1}{|A|}$$

5-3 Mudanças de base

Consideremos um espaço vectorial E, de dimensão finita n. Sabemos que a aplicação identidade $1_E : E \to E$ é uma aplicação linear.

Fixemos no espaço E, considerado domínio de 1_E, uma base $\vec{e}_1', \vec{e}_2', ..., \vec{e}_n'$, e no espaço E, espaço de chegada de 1_E, uma base $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$, e consideremos a matriz

$$P = M\left(1_E;\ (\vec{e}_i'),\ (\vec{e}_i)\right)$$

Sejam, por sua vez,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad X' = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix}$$

as colunas das componentes de um vector arbitrário $\vec{x}$, em relação às bases $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$ e $\vec{e}_1', \vec{e}_2', ..., \vec{e}_n'$, respectivamente.

Uma vez que se tem

$$1_E(\vec{x}) = \vec{x}$$

teremos então, em termos de componentes,

$$PX' = X$$

A matriz P relaciona, portanto, as componentes de cada vector em relação às duas bases consideradas. Em face disso, vamos atribuir-lhe uma designação especial, a saber:

Definição 5-27 Nas condições descritas acima, a matriz P toma o nome de **matriz de mudança de base**.

Deve observar-se que, por definição de matriz de uma aplicação linear, a matriz de mudança de base P, sendo a matriz da aplicação identidade, na situação

$$\begin{array}{cccc} 1_E : & E & \rightarrow & E \\ & (\vec{e}'_j) & & (\vec{e}_j) \end{array}$$

é construída da seguinte maneira: nas colunas de P figuram as componentes das imagens dos vectores $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_n$ (as quais são os próprios vectores $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_n$, visto lidarmos com a aplicação identidade) em relação à base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$. Assim, para se escrever a matriz de mudança de base, basta escrever os vectores de uma das bases em função da outra.

Em face do que vimos acima, é também fácil observar o seguinte:

Teorema 5-28 Uma matriz de mudança de base é invertível.

Demonstração:

Uma matriz de mudança de base é a matriz correspondente à aplicação identidade em relação a certas bases escolhidas num espaço vectorial. Ora, a aplicação identidade é um isomorfismo, pelo que a respectiva matriz é invertível.

Mas o mais interessante é que o recíproco do teorema anterior é igualmente válido:

Teorema 5-29 Toda a matriz invertível pode ser considerada uma matriz de mudança de base.

Demonstração:

Seja P uma matriz invertível qualquer, do tipo $n \times n$. Tomemos um espaço vectorial E qualquer, de dimensão n, e fixemos nele uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$.

Sabemos já que, desde que as bases estejam fixas, toda a matriz define uma aplicação linear, entre espaços vectoriais de dimensões apropriadas. Sendo assim, consideremos a aplicação linear $\xi : E \to E$, definida, em relação à base considerada, pela matriz P, ou seja, tal que

$$P = M\big(\xi;\ (\vec{e}_j),\ (\vec{e}_j)\big)$$

Uma vez que P é invertível, a aplicação ξ será um automorfismo. Por consequência, os vectores $\vec{e}'_1 = \xi(\vec{e}_1), \vec{e}'_2 = \xi(\vec{e}_2), \ldots, \vec{e}'_n = \xi(\vec{e}_n)$ constituem uma nova base de E.

Consideremos então a aplicação identidade em E, na seguinte situação:

$$\begin{array}{cccc} 1_E: & E & \to & E \\ & (\vec{e}'_j) & & (\vec{e}_j) \end{array}$$

As colunas da matriz desta aplicação em relação às bases indicadas são constituídas pelas componentes dos vectores $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_n$ em relação à base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$. Ora, por construção, essas colunas são precisamente as mesmas que as da matriz P: nestas figuram as componentes, em relação à base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, das imagens, por meio da aplicação linear ξ, dos vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, imagens essas que são os vectores $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_n$. Logo, a matriz da identidade, nas condições descritas, é P, que fica assim considerada matriz de mudança de base.

As matrizes de mudança de base vão desempenhar um papel importante num problema específico:

Consideremos uma aplicação linear $\varphi : E \to E'$, entre dois espaços lineares de dimensão finita. Suponhamos que se fixa em E uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, e em E', uma base $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_p$. Nessas condições, a aplicação linear φ será definida por uma determinada matriz A. Mas se fixarmos outras bases, digamos $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_n$ em E e $\vec{f}'_1, \vec{f}'_2, \ldots, \vec{f}'_p$ em E', a mesma aplicação linear φ será definida por uma outra matriz, B. Pretendemos, precisamente, relacionar as matrizes A e B.

Para o efeito, consideremos o diagrama

$$\begin{tikzcd}
\underset{(\vec{e}_j)}{E} \arrow[r, "\varphi", "A"'] & \underset{(\vec{e}'_i)}{E'} \arrow[d, "Q"', "1_{E'}"] \\
\underset{(\vec{e}_j)}{E} \arrow[u, "1_E", "P"'] \arrow[r, "B", "\varphi"'] & \underset{(\vec{f}'_i)}{E'}
\end{tikzcd}$$

Tem-se, como é evidente,

$$1_{E'} \circ \varphi \circ 1_E = \varphi$$

pelo que, atendendo a que à composição de aplicações lineares corresponde o produto das respectivas matrizes, obtemos

$$B = QAP$$

Deve ainda observar-se que as matrizes de mudança de base P e Q são matrizes invertíveis, pelo que a igualdade anterior é equivalente a

$$A = Q^{-1}BP^{-1}$$

Vemos, assim, que as matrizes de uma mesma aplicação linear em relação a bases diferentes se relacionam entre si através da multiplicação por matrizes de mudança de base. Introduzimos, a este propósito, uma nova definição:

Definição 5-30 Sejam A e B duas matrizes quaisquer. Diz-se que A é **equivalente** a B quando existem matrizes invertíveis P e Q, tais que $B = QAP$. Quando A é equivalente a B, representamos esse facto por $A \sim B$.

Relacionando esta definição com o que apresentámos acima, podemos agora enunciar o seguinte:

Teorema 5-31 Uma matriz A é equivalente a uma matriz B se e só se A e B podem representar uma mesma aplicação linear em relação a determinadas bases.

Demonstração:
Vimos já que quando A e B representam uma mesma aplicação linear, a matriz A é equivalente à matriz B.

Reciprocamente, suponhamos que A é equivalente a B e consideremos uma aplicação linear $\varphi: E \to G$ definida pela matriz A em relação a certa base $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ do espaço vectorial E e a certa base $\vec{g}_1, \vec{g}_2, \dots, \vec{g}_p$ do espaço vectorial G.

Sendo A equivalente a B, existirão matrizes invertíveis P e Q, tais que $B = QAP$. Ora, pelo Teorema 5-29, as matrizes P e Q podem ser consideradas matrizes de mudança de base. Mais ainda, resulta da demonstração do Teorema 5-29 que, sendo P uma matriz invertível, existirá em E uma base $\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_n$, tal que P é a matriz da aplicação identidade na situação

$$\begin{array}{ccc} 1_E: & E & \to & E' \\ & (\vec{e}'_j) & & (\vec{e}_j) \end{array}$$

Analogamente, existirá em G uma base $\vec{g}'_1, \vec{g}'_2, \dots, \vec{g}'_p$, tal que Q^{-1} seja a matriz da aplicação identidade, na situação

$$\begin{array}{rccc} 1_G: & G & \to & G \\ & (\vec{g}'_i) & & (\vec{g}_i) \end{array}$$

pelo que Q será a matriz da aplicação identidade na situação

$$\begin{array}{rccc} 1_G: & G & \to & G \\ & (\vec{g}_i) & & (\vec{g}'_i) \end{array}$$

Nestas condições, se tomarmos o diagrama

$$\begin{tikzcd} \underset{(\vec{e}_j)}{E} \arrow[r, "\varphi"', "A"'] & \underset{(\vec{g}_i)}{G} \arrow[d, "Q"', "1_G"] \\ \underset{(\vec{e}'_j)}{E} \arrow[u, "1_E", "P"'] \arrow[r, "\varphi"'] & \underset{(\vec{g}'_i)}{G} \end{tikzcd}$$

verificamos que a matriz da aplicação identidade, na situação

$$\begin{array}{rccc} \varphi: & E & \to & G \\ & (\vec{e}'_j) & & (\vec{g}_i) \end{array}$$

é precisamente $QAP = B$, pelo que B representa, em relação às bases que construímos, a mesma aplicação linear que A.

Do teorema anterior resulta trivialmente que a relação de "equivalência" entre duas matrizes é, na verdade, uma relação de equivalência. Mas podemos também fazer a demonstração deste facto directamente a partir da definição:

Teorema 5-32 A relação de "equivalência" entre matrizes é uma relação de equivalência no conjunto $\mathrm{M}_{p\times n}(K)$, isto é, são válidas, para matrizes quaisquer, as seguintes propriedades:

a) $A \sim A$.

b) Se $A \sim B$, então $B \sim A$.

c) Se $A \sim B$ e $B \sim C$, então $A \sim C$.

Por outro lado, uma matriz equivalente a uma matriz invertível é também invertível.

Demonstração:

a) Para vermos que $A \sim A$, basta reparar que se tem $A = I_n A I_n$.

b) Sendo $A \sim B$, teremos $B = QAP$, onde P e Q são matrizes invertíveis. Daí resulta que $A = Q^{-1}BP^{-1}$, pelo que $B \sim A$.

c) De $A \sim B$ e $B \sim C$, resulta, respectivamente, $B = QAP$ e $C = RBS$, sendo P, Q, R e S matrizes invertíveis. Consequentemente,

$$C = RBS = R(QAP)S = (RQ)A(PS)$$

e as matrizes RQ e PS são invertíveis, pelo que $A \sim C$.

Quanto à última afirmação, suponhamos que A é uma matriz invertível e que $A \sim B$. Ter-se-á, por definição, $B = QAP$, com P e Q invertíveis, pelo que $B^{-1} = (QAP)^{-1} = P^{-1}A^{-1}Q^{-1}$ e B é também invertível.

Quando se trata de problemas de mudança de base, há um caso particular importante, que vamos examinar.

Consideremos o diagrama

$$\begin{tikzcd}
\underset{(\vec{e}_j)}{E} \arrow[rr, "\varphi", "A"'] & & \underset{(\vec{e}_i)}{E} \arrow[dd, "Q"', "1_{E'}"] \\
 & & \\
\underset{(\vec{f}_j)}{E} \arrow[uu, "1_E", "P"'] \arrow[rr, "\varphi"'] & & \underset{(\vec{f}_i)}{E}
\end{tikzcd}$$

Nestas circunstâncias, atendendo a que a aplicação identidade é inversa de si mesma, teremos $Q = P^{-1}$, pelo que

$$B = P^{-1}AP$$

Este caso particular merece também algum relevo, para o que introduzimos uma designação específica:

Definição 5-33 Diz-se que uma matriz A é **semelhante** a uma matriz B, quando existe uma matriz invertível P tal que $B = P^{-1}AP$. Quando A é semelhante a B, escrevemos $A \approx B$.

Tratando-se a semelhança de um caso particular de equivalência entre matrizes, as respectivas propriedades deduzem-se imediatamente:

Teorema 5-34 Uma matriz A é semelhante a uma matriz B se e só se ambas podem representar um mesmo endomorfismo de um espaço vectorial em relação a diferentes bases.

Demonstração:
Resulta imediatamente das construções anteriores.

Teorema 5-35 A relação de "semelhança" entre matrizes é uma relação de equivalência no conjunto $\mathrm{M}_{n\times n}(K)$, isto é, para matrizes quaisquer, são válidas as propriedades:

a) $A \approx A$.
b) Se $A \approx B$, então $B \approx A$.
c) Se $A \approx B$ e $B \approx C$, então $A \approx C$.

Demonstração:
Resulta imediatamente do teorema anterior.

5-4 | Sistemas de equações e subespaços vectoriais

Consideremos um sistema de p equações e n incógnitas, por exemplo

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n = b_p \end{cases}$$

Atendendo ao modo como se definiu a multiplicação de matrizes, este sistema pode ser escrito, mais abreviadamente, na forma matricial.

Com efeito, considerando as matrizes

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix} \quad \text{e} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

o sistema de equações anteriores equivale à igualdade matricial

$$AX = B$$

Definição 5-36 Um sistema de equações da forma $AX = 0$ diz-se um **sistema homogéneo**.

Os sistemas de equações homogéneos gozam de propriedades notáveis, e podemos provar a seguinte proposição:

Teorema 5-37 Seja $AX = 0$ um sistema homogéneo com p equações e n incógnitas. Esse sistema é sempre possível, e o conjunto das suas soluções é um subespaço vectorial do espaço vectorial K^n.

Demonstração:
O sistema de equações $AX = 0$ é sempre possível, visto que a matriz $X = 0$ satisfaz, trivialmente, o sistema. Por consequência, o conjunto S das suas soluções não é vazio.

Cada elemento de S é uma matriz-coluna com n elementos de K, pelo que pode considerar-se um elemento de K^n.

Se tomarmos dois vectores $X_1, X_2 \in S$, teremos $AX_1 = 0$ e $AX_2 = 0$, pelo que, para escalares α e β quaisquer, se tem:

$$A(\alpha X_1 + \beta X_2) = \alpha AX_1 + \beta AX_2 = \alpha 0 + \beta 0 = 0$$

Portanto, $\alpha X_1 + \beta X_2$ é ainda uma solução do sistema dado, isto é,

$$\alpha X_1 + \beta X_2 \in S$$

o que prova que S é um subespaço vectorial de K^n.

Interessa-nos, agora, verificar que se pode também provar a afirmação recíproca do teorema anterior, a saber:

Teorema 5-38 Seja F um subespaço vectorial qualquer de um espaço vectorial E, de dimensão n, sobre o corpo K. Existe um sistema de equações lineares, homogéneo, com n incógnitas, cujas soluções são precisamente as componentes dos vectores de F em relação a uma base fixa em E.

Demonstração:
Fixemos em E uma base $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$, e, ao mesmo tempo, consideremos um subespaço complementar de F em E: seja, por exemplo, $E = F \oplus F'$.

Conforme se viu no Exemplo 4-2 j), a projecção de E sobre F' é uma aplicação linear:

$$\begin{aligned}\rho_{F'}: \quad E = F \oplus F' \quad &\to \quad F' \\ \vec{x} = \vec{f} + \vec{f}' \quad &\mapsto \quad \vec{f}'\end{aligned}$$

O núcleo de $\rho_{F'}$ é

$$\text{Nuc}\left(\rho_{F'}\right) = \left\{\vec{x} = \vec{f} + \vec{f}' : \vec{f}' = \vec{0}\right\} = \left\{\vec{x} = \vec{f} + \vec{0} : \vec{f} \in F\right\} = F$$

Ora, se fixarmos em E a base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, e em F', uma base qualquer, a aplicação linear $\rho_{F'}$ será definida por uma certa matriz A. Se representarmos cada vector $\vec{x} \in E$ pela coluna

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

das suas componentes em relação à base escolhida em E, e tomarmos em consideração o modo como, a partir de A e de X, se obtêm as componentes da imagem de cada vector de E, temos

$$F = \text{Nuc}\left(\rho_{F'}\right) = \left\{\vec{x} \in E : \rho_{F'}(\vec{x}) = \vec{0}\right\} = \left\{\vec{x} \in E : AX = 0\right\}$$

Por conseguinte, F é formado pelos vectores cujas componentes em relação à base considerada são as soluções do sistema de equações $AX = 0$.

5-5 Característica de uma matriz

Consideremos uma matriz qualquer, do tipo $p \times n$, com elementos de um certo corpo K. Cada coluna de A é uma lista de p elementos de K, ou seja, um elemento do espaço vectorial K^p, e cada linha de A é um elemento do espaço vectorial K^n.

Designemos esses vectores por:

$$\vec{C}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{p1} \end{bmatrix}, \quad \vec{C}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{p2} \end{bmatrix}, \quad \ldots, \quad \vec{C}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{pn} \end{bmatrix}$$

e

$$\vec{L}_1 = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix}, \ \vec{L}_2 = \begin{bmatrix} a_{21} & a_{22} & \cdots & a_{2n} \end{bmatrix}, \ldots, \ \vec{L}_p = \begin{bmatrix} a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}$$

Para nos podermos referir mais facilmente ao assunto, introduzimos ainda a seguinte definição:

Definição 5-39 Dada uma matriz A, do tipo $p \times n$, chama-se **espaço das colunas** (resp.: **espaço das linhas**) de A ao subespaço de K^p (resp.: de K^n) gerado pelos vectores $\vec{C}_1, \vec{C}_2, \ldots, \vec{C}_n$ (resp.: pelos vectores $\vec{L}_1, \vec{L}_2, \ldots, \vec{L}_p$). Designaremos o espaço das colunas de A por $C(A)$, e o espaço das linhas por $L(A)$. A dimensão de $C(A)$ (resp.: de $L(A)$) chama-se **característica de coluna** (resp.: **característica de linha**) de A e representa-se por $r_c(A)$ (resp.: $r_l(A)$).

Resulta imediatamente da definição que se tem $r_c(A) = r_l(A^T)$.

Para podermos determinar, com facilidade, a característica de linha e a característica de coluna de uma dada matriz, começaremos por provar o seguinte:

Teorema 5-40 Seja A uma matriz do tipo $p \times n$, e sejam P e Q matrizes invertíveis, dos tipos $n \times n$ e $p \times p$, respectivamente. Tem-se

$$r_c(A) = r_c(QA) = r_c(AP), \quad r_l(A) = r_l(QA) = r_l(AP)$$

Demonstração:

Consideremos os espaços vectoriais K^n e K^p, e suponhamos fixas bases $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ e $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_p$, respectivamente: a matriz A definirá, em relação a estas bases, uma aplicação linear $\psi : K^n \to K^p$.

As colunas de A são constituídas pelas componentes das imagens dos vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, em relação à base $\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_p$, e essas imagens dos vectores da base de K^n geram o espaço imagem de ψ, pelo que $r_c(A) = \dim\left(\psi\left(K^n\right)\right) = c_\psi$.

Ora, sendo as matrizes P e Q invertíveis, a matriz A é equivalente a qualquer das matrizes $QA = QAI_n$ e $AP = I_pAP$. Consequentemente, as matrizes QA e PA poderão representar a mesma aplicação linear que A, ou seja, a aplicação ψ, em relação a diferentes bases de K^n e K^p, pelo que

$$r_c(QA) = c_\psi = r_c(A) = r_c(AP)$$

Por sua vez, se P e Q são invertíveis, também P^T e Q^T são invertíveis, pelo que podemos escrever

$$r_l(QA) = r_c\left((QA)^T\right) = r_c\left(A^T Q^T\right) = r_c\left(A^T\right) = r_l(A)$$

$$r_l(AP) = r_c\left((AP)^T\right) = r_c\left(P^T A^T\right) = r_c\left(A^T\right) = r_l(A)$$

Podemos agora observar o seguinte:

Lema 5-41 A característica de linha e a característica de coluna de uma matriz da forma

$$D = \begin{bmatrix} d_{11} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d_{rr} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

com $d_{ii} \neq 0$ $(i = 1, 2, \ldots, r)$ são ambas iguais a r.

Demonstração:
A afirmação feita resulta imediatamente do facto de quer as linhas quer as colunas não nulas de D serem vectores da base canónica do espaço vectorial K^n ou K^p.

Vamos seguidamente verificar que a característica de linha de uma matriz qualquer coincide sempre com a respectiva característica de coluna, e simultaneamente encontraremos um método fácil para a calcular.

Teorema 5-42 Dada uma matriz A qualquer, tem-se sempre $r_c(A) = r_l(A)$.

Demonstração:
Consideremos a matriz A, do tipo $p \times n$.

Vamos estudar o efeito da multiplicação de A por determinadas matrizes invertíveis.

Comecemos por tomar a matriz

$$M_{kk}(\alpha) = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & \alpha & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix},$$

que resulta da matriz identidade substituindo por um escalar $\alpha \neq 0$ o elemento que se encontra na posição kk. Podemos verificar que a matriz $M_{kk}(\alpha)$ é invertível calculando o respectivo determinante: $|M_{kk}(\alpha)| = \alpha \neq 0$.

Se escolhermos a matriz $M_{kk}(\alpha)$ do tipo $p \times p$, poderemos efectuar a multiplicação $M_{kk}(\alpha)A$. Ora, é fácil verificar directamente que

$$M_{kk}(\alpha)A = \begin{bmatrix} a_{11} & a_{12} & & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha a_{k1} & \alpha a_{k2} & \cdots & \alpha a_{kn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}$$

isto é, que o efeito de multiplicar $M_{kk}(\alpha)$ por A é o de multiplicar os elementos da linha k de A pelo escalar α.

Por sua vez, se tomarmos $M_{kk}(\alpha)$ do tipo $n \times n$, obtemos

$$AM_{kk}(\alpha) = \begin{bmatrix} a_{11} & a_{12} & \cdots & \alpha & a_{1k} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \alpha & a_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & \alpha & a_{pk} & \cdots & a_{pn} \end{bmatrix}$$

isto é, o efeito de multiplicar A por $M_{kk}(\alpha)$ é o de multiplicar os elementos da coluna k de A pelo escalar α.

Consideremos agora a matriz $M_{kt}(\alpha)$, que resulta da matriz identidade substituindo por um escalar α qualquer o elemento que se encontra na posição kt (com $k \neq t$). Esquematicamente, supondo $k < t$,

$$M_{kt}(\alpha) = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & \alpha & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

A matriz $M_{kk}(\alpha)$ é invertível, visto que o seu determinante nos dá $|M_{kt}(\alpha)| = 1 \neq 0$.

Podemos verificar, efectuando directamente a operação, que a matriz $M_{kk}(\alpha)A$ resulta de A somando aos elementos da linha k os elementos da linha t multiplicados por α, mantendo-se as demais linhas. O resultado de multiplicar A por $M_{kk}(\alpha)$ é análogo, mas sobre as colunas de A.

Por sua vez, através das transformações referidas, é possível proceder à troca de posição entre duas linhas ou entre duas colunas. Exemplifiquemos, para o caso das linhas, de forma esquemática, designando por (1) a transformação que consiste em multiplicar uma linha por -1, e por (2) a transformação que consiste em somar uma linha com outra, ambas englobadas pelas construções anteriores:

$$\begin{bmatrix} \vdots \\ L_k \\ \vdots \\ L_t \\ \vdots \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} \vdots \\ L_k \\ \vdots \\ L_k + L_t \\ \vdots \end{bmatrix} \xrightarrow{(1)} \begin{bmatrix} \vdots \\ L_k \\ \vdots \\ -L_k - L_t \\ \vdots \end{bmatrix} \xrightarrow{(2)}$$

$$\rightarrow \begin{bmatrix} \vdots \\ -L_t \\ \vdots \\ -L_k - L_t \\ \vdots \end{bmatrix} \xrightarrow{(1)} \begin{bmatrix} \vdots \\ L_t \\ \vdots \\ -L_k - L_t \\ \vdots \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} \vdots \\ L_t \\ \vdots \\ -L_k \\ \vdots \end{bmatrix} \xrightarrow{(1)} \begin{bmatrix} \vdots \\ L_t \\ \vdots \\ L_k \\ \vdots \end{bmatrix}$$

Vemos, assim, que, através da multiplicação de uma dada matriz por matrizes invertíveis, ou do lado direito, ou do lado esquerdo, é possível obter as seguintes transformações:

a) Trocar entre si as posições de duas linhas (ou de duas colunas) da matriz.
b) Multiplicar uma linha ou uma coluna da matriz por um escalar não nulo.
c) Somar a uma linha (resp.: coluna) uma outra linha (resp.: coluna), multiplicada por um escalar.

De acordo com o que se viu no Teorema 5-41, conclui-se daí que essas transformações não afectam nem a característica de linha, nem a característica de coluna de A.

As transformações descritas são, para linhas e para colunas, as que se usam na aplicação do método de Gauss, que estudámos a propósito da resolução de sistemas de equações ou do cálculo de determinantes. Ora, a aplicação do método de Gauss permite-nos, a partir de uma matriz A qualquer, chegar a uma matriz da forma

$$D=\begin{bmatrix} d_{11} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d_{rr} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

pelo que $r_c(A)=r_c(D)$ e $r_l(A)=r_l(D)$.

Finalmente, sabemos, pelo Lema 5-41, que $r_c(D)=r_l(D)=r$, pelo que

$$r_c(A)=r_c(D)=r=r_l(D)=r_l(A)$$

O facto de a característica de linha de uma matriz coincidir com a respectiva característica de coluna dispensa-nos de especificar se se trata de uma ou de outra. Assim:

Definição 5-43 O valor comum da característica de linha e da característica de coluna de uma matriz A designa-se por **característica** de A e representa-se por $r(A)$.

O conhecimento da característica de uma matriz dá-nos valiosas informações acerca dessa matriz, nomeadamente a seguinte:

Teorema 5-44 Uma matriz quadrada A, do tipo $n \times n$, é invertível se e só se $r(A) = n$.

Demonstração:
Seja φ uma aplicação linear definida pela matriz A em relação a certas bases.
Se A é invertível, então φ é um isomorfismo, pelo que $r(A) = c_\varphi = n$. Reciprocamente, se $r(A) = n$, então, através da aplicação do método de Gauss, verificamos que A é uma matriz equivalente à matriz identidade I_n, que é invertível, pelo que A é também invertível.

A noção de característica auxilia-nos também na discussão de um sistema de equações lineares:

Teorema 5-45 Um sistema de equações lineares $AX = B$, com p equações e n incógnitas, é possível se e só se a característica da sua matriz simples coincidir com a característica da respectiva matriz ampliada, isto é, $r(A) = r\left(\left[A \mid B\right]\right)$. Sendo possível, será determinado se e só se $r(A)$ coincidir ainda com o número de colunas de A.

Demonstração:
Resulta imediatamente da aplicação do método de Gauss.

6 Valores e vectores próprios

6-1 Vectores próprios

Consideremos o endomorfismo do espaço vectorial real $\mathbb{R}^3$ definido, em relação à base canónica, pela matriz

$$A = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Para calcularmos a imagem do vector $\vec{u} = (1,1,1)$ basta calcular o produto

$$\begin{bmatrix} 1 & -1 & 3 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

pelo que $\varphi(\vec{u}) = 3\vec{u}$.

Analogamente, seja E o espaço vectorial formado pelas funções infinitamente diferenciáveis num certo intervalo, com as operações usuais, e consideremos o endomorfismo $\delta : E \to E$, em que $\delta(f(x)) = f'(x)$, a derivada de $f(x)$.

Se tomarmos a função definida por $f(x) = e^{5x}$, verificamos que

$$\delta(f(x)) = \delta\left(e^{5x}\right) = \left(e^{5x}\right)' = 5e^{5x} = 5f(x)$$

Casos como os que acabámos de examinar, em que a imagem de um certo vector, através de um endomorfismo, é proporcional ao vector inicial, revestem-se de uma grande importância. Para os tratar, começamos por introduzir designações adequadas:

Definição 6-1

Seja φ um endomorfismo de um espaço vectorial E sobre um corpo K. Um vector $\vec{u} \in E$ diz-se um **vector próprio** (ou **vector característico**) de φ associado ao **valor próprio** (ou **valor característico**) $\lambda \in K$ quando:

a) $\vec{u} \neq \vec{0}$

b) $\varphi(\vec{u}) = \lambda\vec{u}$

Ao conjunto dos valores próprios de um dado endomorfismo costuma chamar-se **espectro** desse endomorfismo.

Naturalmente, dado um endomorfismo de um certo espaço vectorial, apresenta-se o problema de saber se existem ou não vectores próprios para esse endomorfismo, e, em caso afirmativo, haverá que determiná-los. Vamos ser capazes de responder a estas questões de uma forma eficaz no caso de o espaço vectorial considerado ter dimensão finita. O próximo teorema esclarece já o assunto:

Teorema 6-2

Seja E um espaço vectorial de dimensão n e seja φ um endomorfismo de E. Fixemos uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ em E, e seja A a matriz que representa o endomorfismo φ em relação a essa base. Nestas condições:

a) Um escalar $\lambda \in K$ é valor próprio de φ se e só se λ é raiz da equação $|A - \lambda I_n| = 0$.

b) Um vector $\vec{u} \in E$ é vector próprio de φ, associado a um valor próprio λ, se e só se as componentes de $\vec{u}$ em relação à base considerada são uma solução não nula do sistema de equações lineares $(A - \lambda I_n)X = 0$.

Demonstração:

Por definição, $\lambda \in K$ é valor próprio de φ se e só se existir um vector $\vec{u} \neq \vec{0}$ tal que $\varphi(\vec{u}) = \lambda\vec{u}$.

A igualdade $\varphi(\vec{u}) = \lambda\vec{u}$ pode traduzir-se na forma matricial, desde que representemos o vector $\vec{u}$ pela coluna

$$U = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

por $AU = \lambda U$.

Ora, a igualdade $AU = \lambda U$ equivale a $AU - \lambda U = 0$ ou, ainda, sucessivamente, a

$$AU - \lambda I_n U = 0$$
$$(A - \lambda I_n)U = 0$$

Portanto, podemos afirmar que λ é um valor próprio de φ se e só se existir uma coluna $U \neq 0$ tal que $(A - \lambda I_n)U = 0$, isto é, se e só se o sistema de equações

$$(A - \lambda I_n)X = 0$$

tiver uma solução não nula.

Ora, este sistema de equações é homogéneo, pelo que admite sempre a solução nula. Consequentemente, admitirá soluções não nulas se e só se for um sistema indeterminado, o que equivale a dizer que o determinante da respectiva matriz simples seja igual a zero: $|A - \lambda I_n| = 0$.

Os vectores próprios associados a λ serão precisamente as soluções não nulas daquele sistema de equações.

Exemplo 6-3

Consideremos o endomorfismo do espaço vectorial real $\mathbb{R}^3$ definido, em relação à base canónica, pela matriz

$$A = \begin{bmatrix} 1 & -2 & 5 \\ 2 & 1 & 1 \\ 0 & 3 & 1 \end{bmatrix}$$

Para calcular os valores próprios desse endomorfismo, calculamos

$$|A - \lambda I_3| = \begin{vmatrix} 1-\lambda & -2 & 5 \\ 2 & 1-\lambda & 1 \\ 0 & 3 & 1-\lambda \end{vmatrix} = (1-\lambda)^3 + 30 - 3(1-\lambda) + 4(1-\lambda) = -\lambda^3 + 3\lambda^2 - 4\lambda + 32$$

Igualando a zero, temos

$$-\lambda^3 + 3\lambda^2 - 4\lambda + 32 = 0 \quad \text{ou} \quad (\lambda - 4)(-\lambda^2 - \lambda - 8) = 0$$

A única raiz desta equação é $\lambda = 4$, que será, portanto, o único valor próprio do endomorfismo em causa. Para procurarmos os respectivos vectores próprios, teremos de resolver o sistema de equações

$$\begin{bmatrix} 1-4 & -2 & 5 \\ 2 & 1-4 & 1 \\ 0 & 3 & 1-4 \end{bmatrix} X = 0 \quad \text{ou seja,} \quad \begin{bmatrix} -3 & -2 & 5 \\ 2 & -3 & 1 \\ 0 & 3 & -3 \end{bmatrix} X = 0$$

Usando, para o efeito, o método de Gauss, obtemos:

$$\begin{bmatrix} -3 & -2 & 5 \\ 2 & -3 & 1 \\ 0 & 3 & -3 \end{bmatrix} \to \begin{bmatrix} -3 & -2 & 5 \\ 0 & -13/3 & 13/3 \\ 0 & 3 & -3 \end{bmatrix} \to \begin{bmatrix} -3 & -2 & 5 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \to \begin{bmatrix} -3 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

A solução do sistema obtido,

$$\begin{cases} x_1 - x_3 = 0 \\ x_2 - x_3 = 0 \end{cases} \quad \text{é} \quad \begin{cases} x_1 = x_3 \\ x_2 = x_3 \end{cases}$$

com x_3 qualquer. Os vectores próprios do endomorfismo dado, associados ao valor próprio 4, serão, portanto, os vectores não nulos da forma $(x_3, x_3, x_3) = x_3(1,1,1)$. ■

Conforme vimos no teorema anterior, os vectores próprios de um dado endomorfismo de um espaço vectorial E de dimensão finita aparecem como as soluções não nulas de um certo sistema homogéneo de equações lineares. Ora, sabemos que o conjunto de todas as soluções (nulas ou não) de um sistema homogéneo de equações lineares constitui um subespaço vectorial de E.

Em termos gerais, apresentamos o seguinte resultado:

Teorema 6-4

Seja φ um endomorfismo de um espaço vectorial qualquer E. Para cada escalar $\gamma \in K$, o conjunto $U_\gamma = \{\vec{x} \in E : \varphi(\vec{x}) = \gamma\vec{x}\}$ é um subespaço vectorial de E. O escalar γ é um valor próprio de φ se e só se $U_\gamma \neq \{\vec{0}\}$.

Demonstração:

É claro que

$$\varphi(\vec{0}) = \vec{0} = \gamma\vec{0}$$

pelo que $\vec{0} \in U_\gamma$ e $U_\gamma \neq \varnothing$.

Tomando, então, vectores quaisquer $\vec{x}, \vec{y} \in U_\gamma$ e escalares α, β, igualmente arbitrários, vemos que

$$\varphi(\alpha\vec{x} + \beta\vec{y}) = \alpha\varphi(\vec{x}) + \beta\varphi(\vec{y}) = \alpha\gamma\vec{x} + \beta\gamma\vec{y} = \gamma(\alpha\vec{x} + \beta\vec{y})$$

o que mostra que U_γ é, efectivamente, um subespaço vectorial de E.

A última afirmação do enunciado resulta imediatamente das definições.

Estes subespaços U_γ têm uma designação apropriada, a saber:

Definição 6-5 Nas condições anteriores e sendo γ um valor próprio do endomorfismo φ, o subespaço U_γ toma o nome de **subespaço próprio** (ou subespaço **característico**) associado a γ. Quando U_γ tem dimensão finita, a essa dimensão chama-se **multiplicidade geométrica** de γ; representá-la-emos por $m_g(\gamma)$.

É necessário ter um certo cuidado para não confundir a designação de "subespaço próprio associado a um certo valor próprio" com a ideia de "subespaço próprio" no sentido de um subespaço que não coincide com todo o espaço vectorial (já que este é o seu "subespaço impróprio").

Voltando ao caso de um espaço vectorial de dimensão finita, vimos que os valores próprios de um endomorfismo aparecem como raízes de uma equação polinomial. Ora, é bem sabido que, dado um polinómio, uma mesma raiz pode, por assim dizer, aparecer diversas vezes, isto é, sendo ε uma raiz de um dado polinómio $p(\lambda)$, então $p(\lambda)$ pode ser divisível por $(\lambda-\varepsilon)^k$, com $k>1$. Introduzimos, a este propósito, as seguintes designações:

Definição 6-6 Seja E um espaço vectorial de dimensão n, e seja φ um endomorfismo de E. Seja A a matriz que representa o endomorfismo φ em relação a certa base de E. Ao polinómio $p(\lambda)=|A-\lambda I_n|$ é chamado **polinómio característico** de A. A multiplicidade de cada raiz λ_0, isto é, de cada valor próprio de φ enquanto raiz do polinómio característico, é a **multiplicidade algébrica** de λ_0; representamo-la por $m_a(\lambda_0)$.

Deve reparar-se que a definição anterior se refere expressamente ao polinómio característico de uma matriz. No entanto, podemos provar o seguinte:

Teorema 6-7 Duas matrizes semelhantes têm o mesmo polinómio característico. Por outras palavras, duas matrizes que representem o mesmo endomorfismo, de um espaço vectorial de dimensão finita, têm o mesmo polinómio característico.

Demonstração:
Se tivermos duas matrizes quadradas, A e B, semelhantes, será, por definição, $B=P^{-1}AP$, para alguma matriz invertível P.

Por conseguinte, podemos escrever

$$\begin{aligned}|B-\lambda\ \ I_n| &= |P^{-1}AP-\lambda I_n| = |P^{-1}AP-\lambda P^{-1}I_nP| = \\ &= |P^{-1}(A-\lambda I_n)P| = |P^{-1}|\,|A-\lambda I_n|\,|P| = \frac{1}{|P|}\,|A-\lambda I_n|\,|P| = |A-\lambda I_n|\end{aligned}$$

Vemos, portanto, que B tem o mesmo polinómio característico que A.

O teorema anterior mostra que os polinómios característicos de todas as matrizes que representam um certo endomorfismo, de um espaço de dimensão finita, são iguais, o que dá sentido à seguinte definição:

Definição 6-8 Seja E um espaço vectorial de dimensão n, e seja φ um endomorfismo de E. Seja A a matriz que representa o endomorfismo φ em relação a certa base de E. O polinómio $p(\lambda)=|A-\lambda I_n|$ é chamado **polinómio característico** de φ.

Há uma relação interessante entre as duas multiplicidades, algébrica e geométrica, de um mesmo valor próprio, a saber:

Teorema 6-9 Seja λ_0 um valor próprio de um endomorfismo φ de um espaço vectorial E, de dimensão finita. Tem-se: $m_g(\lambda_0)\le m_a(\lambda_0)$.

Demonstração:

Seja $U=\langle \vec{u}_1,\vec{u}_2,\ldots,\vec{u}_m\rangle$ o subespaço próprio associado ao valor próprio λ_0 do endomorfismo φ. Utilizando o teorema de Steinitz, podemos construir uma base de E que inclua os vectores da base de U, digamos $\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_m,\vec{e}_{m+1},\ldots,\vec{e}_n$.

Atendendo a que os vectores $\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_m,\vec{e}_{m+1},\ldots,\vec{e}_n$ são vectores próprios de φ associados ao valor próprio λ_0, sabemos que

$$\begin{gathered}\varphi(\vec{u}_1)=\lambda_0\vec{u}_1\\ \varphi(\vec{u}_2)=\lambda_0\vec{u}_2\\ \vdots\\ \varphi(\vec{u}_m)=\lambda_0\vec{u}_m\end{gathered}$$

Por consequência, a matriz de φ em relação à base $\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_m,\vec{e}_{m+1},\ldots,\vec{e}_n$ será da forma

$$A=\begin{bmatrix}\lambda_0 & 0 & \cdots & 0 & b_{1\,m+1} & \cdots & b_{1n}\\ 0 & \lambda_0 & \cdots & 0 & b_{2\,m+1} & \cdots & b_{2n}\\ \vdots & \vdots & & \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & \lambda_0 & b_{m\,m+1} & \cdots & b_{mn}\\ 0 & 0 & \cdots & 0 & b_{m+1\,m+1} & \cdots & b_{m+1\,n}\\ \vdots & \vdots & & \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 0 & b_{n\,m+1} & \cdots & b_{nn}\end{bmatrix}$$

Segundo vimos, o polinómio característico de φ é sempre o mesmo, seja qual for a base utilizada. Assim, se construirmos o polinómio característico a partir desta matriz A, obtemos

$$A-\lambda I_n=\begin{bmatrix}\lambda_0-\lambda & 0 & \cdots & 0 & b_{1\,m+1} & \cdots & b_{1n}\\ 0 & \lambda_0-\lambda & \cdots & 0 & b_{2\,m+1} & \cdots & b_{2n}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & \lambda_0-\lambda & b_{m\,m+1} & \cdots & b_{mn}\\ 0 & 0 & \cdots & 0 & b_{m+1\,m+1}-\lambda & \cdots & b_{m+1\,n}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & 0 & b_{n\,m+1} & \cdots & b_{nn}-\lambda\end{bmatrix}$$

pelo que

$$|A-\lambda I_n|=(\lambda_0-\lambda)^m\begin{vmatrix}b_{m+1\,m+1}-\lambda & \cdots & b_{m+1\,n}\\ \vdots & \vdots & \vdots\\ b_{n\,m+1} & \cdots & b_{nn}\end{vmatrix}$$

resultando $m_a(\lambda_0)\geq m=m_g(\lambda_0)$.

Até ao momento, temos vindo a examinar os vectores próprios associados a cada valor próprio de um dado endomorfismo. Podemos agora confrontar vectores próprios associados a valores próprios distintos:

Teorema 6-10 Sejam $\vec{u}_1,\vec{u}_2,\dots,\vec{u}_k$ vectores próprios de um endomorfismo φ, de um espaço vectorial E, associados a valores próprios distintos $\lambda_1,\lambda_2,\dots,\lambda_k$. Então, os vectores $\vec{u}_1,\vec{u}_2,\dots,\vec{u}_k$ são linearmente independentes.

Demonstração:

Faremos a demonstração por indução em k.

No caso de ser $k=1$, o resultado é trivial, visto que um único vector próprio, sendo não nulo (por definição), é linearmente independente.

Suponhamos então que o resultado é válido para $k-1$ vectores e analisemos o caso de termos k vectores ao todo. Se os vectores $\vec{u}_1,\vec{u}_2,\dots,\vec{u}_k$ fossem linearmente dependentes, algum deles seria combinação linear dos restantes, não havendo perda de generalidade em se admitir que seja $\vec{u}_1=\alpha_2\vec{u}_2+\cdots+\alpha_k\vec{u}_k$.

Nestas condições,

$$\varphi(\vec{u}_1)=\varphi(\alpha_2\vec{u}_2+\cdots+\alpha_k\vec{u}_k)=\alpha_2\varphi(\vec{u}_2)+\cdots+\alpha_k\varphi(\vec{u}_k)$$

Como os vectores dados são vectores próprios de φ, esta igualdade transforma-se em

$$\lambda_1\vec{u}_1 = \alpha_2\lambda_2\vec{u}_2 + \cdots + \alpha_k\lambda_k\vec{u}_k$$

Se substituirmos $\vec{u}_1$ pela sua expressão $\vec{u}_1 = \alpha_2\vec{u}_2 + \cdots + \alpha_k\vec{u}_k$, obtemos

$$\lambda_1\left(\alpha_2\vec{u}_2 + \cdots + \alpha_k\vec{u}_k\right) = \alpha_2\lambda_2\vec{u}_2 + \cdots + \alpha_k\lambda_k\vec{u}_k$$

ou seja

$$\alpha_2\lambda_1\vec{u}_2 + \cdots + \alpha_k\lambda_1\vec{u}_k = \alpha_2\lambda_2\vec{u}_2 + \cdots + \alpha_k\lambda_k\vec{u}_k$$

ou ainda

$$\alpha_2\left(\lambda_1 - \lambda_2\right)\vec{u}_2 + \cdots + \alpha_k\left(\lambda_1 - \lambda_k\right)\vec{u}_k = \vec{0}$$

Ora, os vectores $\vec{u}_2, \ldots, \vec{u}_k$ são apenas $k-1$ vectores próprios de φ, associados a valores próprios distintos, pelo que, por hipótese de indução, são linearmente independentes. Consequentemente, podemos concluir que, na última igualdade, todos os coeficientes são nulos, ou seja, que

$$\begin{cases} \alpha_2\left(\lambda_1 - \lambda_2\right) = 0 \\ \vdots \\ \alpha_k\left(\lambda_1 - \lambda_k\right) = 0 \end{cases}$$

Como os valores próprios $\lambda_1, \lambda_2, \ldots, \lambda_k$ são, por hipótese, todos distintos entre si, sabemos que $\lambda_1 - \lambda_i \neq 0$ (para $i = 2, \ldots, k$), pelo que podemos concluir que

$$\alpha_2 = \cdots = \alpha_k = 0$$

pelo que

$$\vec{u}_1 = \alpha_2\vec{u}_2 + \cdots + \alpha_k\vec{u}_k = 0\vec{u}_2 + \cdots + 0\vec{u}_k = \vec{0}$$

o que é absurdo, visto o vector $\vec{u}_1$ ser um vector próprio.

Foi, portanto, absurdo supor que os vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$ fossem linearmente dependentes.

O facto de os vectores próprios associados a vectores próprios distintos serem linearmente independentes garante o seguinte:

Corolário 6-11 Sejam $\lambda_1, \lambda_2, ..., \lambda_k$ valores próprios distintos de um endomorfismo φ. Então, a soma $U_{\lambda_1} + U_{\lambda_2} + \cdots + U_{\lambda_k}$ é uma soma directa.

Demonstração:
Resulta imediatamente do facto de, se juntarmos bases das várias parcelas, obtermos uma base da soma.

Por sua vez, a circunstância de a soma $U_{\lambda_1} \oplus U_{\lambda_2} \oplus \cdots \oplus U_{\lambda_k}$ ser uma soma directa, quando os vectores próprios considerados são distintos, leva-nos a introduzir a seguinte definição:

Definição 6-12 Um endomorfismo φ, de um espaço vectorial E de dimensão finita, diz-se **diagonalizável** quando existe em E uma base formada por vectores próprios de φ.

Resulta imediatamente das definições e dos resultados anteriores o seguinte:

Teorema 6-13 Seja φ um endomorfismo de um espaço vectorial E, de dimensão finita n, e suponhamos que $\lambda_1, \lambda_2, ..., \lambda_k$ são todos os valores próprios distintos de φ. Então, φ é diagonalizável se e só se $m_g(\lambda_1) + m_g(\lambda_2) + \cdots + m_g(\lambda_k) = n$.

Demonstração:
Resulta imediatamente das proposições anteriores.

A designação "diagonalizável" justifica-se pela seguinte propriedade:

Teorema 6-14 Seja φ um endomorfismo de um espaço vectorial E, de dimensão finita n. O endomorfismo φ é diagonalizável se e só se existe em E uma base em relação à qual a matriz de φ é uma matriz diagonal.

Demonstração:
Seja $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$ uma base de E, constituída por vectores próprios de φ. Escrevendo então

$$\begin{aligned} \varphi(\vec{e}_1) &= \lambda_1 \vec{e}_1 \\ \varphi(\vec{e}_2) &= \lambda_1 \vec{e}_2 \\ &\vdots \\ \varphi(\vec{e}_n) &= \lambda_1 \vec{e}_n \end{aligned}$$

onde os valores próprios $\lambda_1, \lambda_2, \ldots, \lambda_n$ podem, naturalmente, ser ou não distintos entre si, vemos que a matriz de φ, em relação à base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, é

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Reciprocamente, resulta da definição de matriz de uma aplicação linear que, se a matriz de φ, em relação a certa base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, é

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

então tem-se $\varphi(\vec{e}_1) = \lambda_1 \vec{e}_1$, $\varphi(\vec{e}_2) = \lambda_1 \vec{e}_2, \ldots, \varphi(\vec{e}_n) = \lambda_1 \vec{e}_n$, pelo que a base é formada por vectores próprios de φ, e φ é diagonalizável.

Repare-se que um endomorfismo de um dado espaço vectorial E, de dimensão finita, definido por uma matriz A, em relação a certa base fixa no espaço, é diagonalizável se e só se existe uma matriz de mudança de base (ou seja, uma matriz invertível) P, tal que $P^{-1}AP$ é uma matriz diagonal.

6-2 A forma canónica de Jordan

Para muitos efeitos, especialmente de carácter computacional, no cálculo matricial é vantajoso obter matrizes tão simples quanto possível, equivalentes ou mesmo semelhantes àquelas de que dispomos.

É fácil encontrar exemplos de endomorfismos não diagonalizáveis de um dado espaço vectorial.

Seja, por exemplo, φ o endomorfismo do espaço vectorial real $\mathbb{R}^3$ definido, em relação à base canónica do espaço, pela matriz

$$J = \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

O único valor próprio de φ é 5, com multiplicidade algébrica 3. Mas a matriz

$$J-5I_3=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$$

tem característica 1, pelo que o subespaço próprio U_5, de $\mathbb{R}^3$, associado ao valor próprio 5, tem a dimensão 2, e, por consequência, φ não é diagonalizável.

Ora, sucede que todo o endomorfismo de um espaço vectorial complexo admite uma representação matricial bastante simples, conhecida como a "forma canónica de Jordan", que passaremos a apresentar.

Começamos por introduzir uma definição:

Definição 6-15 Dado um corpo K qualquer e um elemento fixo $b \in K$, chama-se **bloco de Jordan** (associado a b) a uma matriz B, do tipo $m \times m$, em que

$$b_{ii}=b \qquad (i=1,2,\ldots,m)$$

$$b_{i,i+1}=1 \qquad (i=1,2,\ldots,m-1)$$

$$b_{ij}=0 \qquad \text{(em todos os outros casos)}$$

Esquematicamente:

$$B=\begin{bmatrix} b & 1 & 0 & 0 & \cdots & 0 & 0\\ 0 & b & 1 & 0 & \cdots & 0 & 0\\ 0 & 0 & b & 1 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & 0 & \cdots & b & 1\\ 0 & 0 & 0 & 0 & \cdots & 0 & b\end{bmatrix}$$

Os blocos de Jordan têm comportamentos especiais, que importa salientar e que reunimos no seguinte teorema:

Teorema 6-16 Seja B um bloco de Jordan, e φ o endomorfismo do espaço vectorial $E=K^m$, sobre o corpo K, definido por B, em relação a certa base $\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_m$. São válidas as seguintes propriedades (designando por 1_E a aplicação identidade de E):

a) $(\varphi-b1_E)(\vec{e}_i)=\vec{e}_{i-1} \quad (i=2,3,\ldots,m)$

b) $(\varphi-b1_E)(\vec{e}_1)=\vec{0}$

c) $(\varphi-b1_E)^m=0$

d) $(\varphi-b1_E)^k\neq 0 \quad (k=1,2,\ldots,m-1)$

Demonstração:
A matriz de $\varphi - b1_E$ em relação à base considerada é

$$B - bI_m = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Daí resultam imediatamente as afirmações a) e b).

Por outro lado, de a) e b) resulta que, aplicando sucessivamente $\varphi - b1_E$ ao vector $\vec{e}_m$, obtemos

$$\vec{e}_m \mapsto \vec{e}_{m-1} \mapsto \vec{e}_{m-2} \mapsto \cdots \mapsto \vec{e}_3 \mapsto \vec{e}_2 \mapsto \vec{e}_1 \mapsto \vec{0}$$

de onde se concluem trivialmente c) e d).

Em termos matriciais, o teorema anterior significa, obviamente, que, sendo B o bloco de Jordan em causa, a característica da matriz $(B - bI_m)^k$ vai baixando, de $m-1$ a 0, à medida que k aumenta de 1 a m.

Deve ainda observar-se que, se um bloco de Jordan B representa um endomorfismo de um espaço vectorial em relação a certa base, a primeira coluna de B mostra que o primeiro vector da base considerada é um vector próprio do endomorfismo em causa.

Apresentamos agora o tipo de matriz que vai generalizar as matrizes diagonais, provando, em seguida, que todo o endomorfismo de um espaço vectorial complexo pode ser representado, em relação a uma base convenientemente escolhida, por uma matriz deste tipo.

Definição 6-17 Uma matriz A, do tipo $n \times n$, tem a **forma canónica de Jordan** quando consiste na justaposição, canto a canto, de blocos de Jordan ao longo da diagonal principal, sendo nulos todos os demais elementos de A.

Mais precisamente, quando existem números naturais

$$1 = r_0 \leq r_1 \leq r_2 \leq r_3 \leq \cdots \leq r_k \leq r_{k+1} = n$$

tais que cada submatriz da matriz A, definida por

$$B_t = [a_{ij}]$$
$$(t = 0,1,\ldots,k;\ \ i = r_t, r_t + 1, r_t + 2, \ldots, r_{t+1};\ \ j = r_t, r_t + 1, r_t + 2, \ldots, r_{t+1})$$

é um bloco de Jordan.

Por exemplo, a matriz

$$\left[\begin{array}{ccc|cc|c|cc} a & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & b & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & c & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & d & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & d \end{array}\right]$$

tem a forma canónica de Jordan.

O Teorema 6-16 mostra que, se esta matriz definir um endomorfismo η de $\mathbb{C}^8$ em relação a certa base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_8$, se terá

$$\begin{aligned} &(\eta - a1_E)(\vec{e}_1) = \vec{0} \\ &(\eta - a1_E)(\vec{e}_2) = \vec{e}_1 \quad \text{formando a cadeia} \quad \vec{e}_3 \mapsto \vec{e}_2 \mapsto \vec{e}_1 \mapsto \vec{0}; \\ &(\eta - a1_E)(\vec{e}_3) = \vec{e}_2 \end{aligned}$$

por sua vez,

$$\left.\begin{aligned} &(\eta - b1_E)(\vec{e}_4) = \vec{0} \\ &(\eta - b1_E)(\vec{e}_5) = \vec{e}_4 \end{aligned}\right. \quad \text{formando a cadeia} \quad \vec{e}_5 \mapsto \vec{e}_4 \mapsto \vec{0};$$

seguidamente,

$$(\eta - c1_E)(\vec{e}_6) = \vec{0} \quad \text{formando a cadeia} \quad \vec{e}_6 \mapsto \vec{0};$$

e, finalmente,

$$\left.\begin{aligned} &(\eta - d1_E)(\vec{e}_7) = \vec{0} \\ &(\eta - b1_E)(\vec{e}_8) = \vec{e}_7 \end{aligned}\right. \quad \text{formando a cadeia} \quad \vec{e}_8 \mapsto \vec{e}_7 \mapsto \vec{0}.$$

Os vectores $\vec{e}_1, \vec{e}_4, \vec{e}_6, \vec{e}_7$ são vectores próprios de η associados aos valores próprios (que podem não ser todos distintos) a, b, c e d, respectivamente.

Verificamos, portanto, que cada bloco de Jordan, associado a um determinado valor próprio λ, de um endomorfismo φ, corresponde, na base considerada, a uma cadeia da forma

$$\vec{e}_p \mapsto \vec{e}_{p-1} \mapsto \vec{e}_{p-2} \mapsto \cdots \mapsto \vec{e}_3 \mapsto \vec{e}_2 \mapsto \vec{e}_1 \mapsto \vec{0}$$

Consequentemente, procurar matrizes na forma canónica de Jordan para um endomorfismo φ de um certo espaço vectorial equivale a procurar cadeias deste tipo para cada endomorfismo $\varphi - \lambda 1_E$, onde λ percorre os diversos valores próprios de φ.

Podemos introduzir nomenclatura apropriada:

Definição 6-18 Dado um endomorfismo φ de um espaço vectorial complexo E, e designando por λ um dos valores próprios de φ, diz-se que os vectores não nulos $\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_p$ formam uma **cadeia de Jordan**, associada a φ e λ, quando se tem

$$\varphi(\vec{z}_1) = \lambda \vec{z}_1,\ \varphi(\vec{z}_2) = \vec{z}_1 + \lambda \vec{z}_2,\ \varphi(\vec{z}_3) = \vec{z}_2 + \lambda \vec{z}_3,\ \varphi(\vec{z}_p) = \vec{z}_{p-1} + \lambda \vec{z}_p$$

ou seja, para o endomorfismo $\varphi - \lambda 1_E$, será

$$\vec{e}_p \mapsto \vec{e}_{p-1} \mapsto \vec{e}_{p-2} \mapsto \cdots \mapsto \vec{e}_3 \mapsto \vec{e}_2 \mapsto \vec{e}_1 \mapsto \vec{0}$$

As cadeias de Jordan obedecem à seguinte propriedade:

Teorema 6-19 Nas condições da Definição 6-18, os vectores $\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_p$ são linearmente independentes. Pondo $Z = \langle \vec{z}_1, \vec{z}_2, \cdots, \vec{z}_p \rangle$, podemos restringir φ ao subespaço Z, e tem-se

$$\mathrm{Nuc}(\varphi) \cap Z = \begin{cases} \{\vec{0}\}, & \text{se } \lambda \neq 0 \\ \langle \vec{z}_1 \rangle, & \text{se } \lambda = 0 \end{cases}$$

Demonstração:

a) Usaremos o método de indução para provar que os vectores $\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_p$ são linearmente independentes.

A afirmação é evidentemente verdadeira quando $p = 1$, dado que, por definição, o vector $\vec{z}_1$ é não nulo.

Suponhamos agora que a afirmação é válida quando a cadeia de Jordan é formada por $p-1$ vectores, e observemos desde já que os vectores $\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_{p-1}$ formam, só por si, uma cadeia de Jordan, pelo que são linearmente independentes.

Para provar que os vectores $\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_p$ são linearmente independentes, distinguiremos dois casos:

a1) $\lambda = 0$

Consideremos a igualdade

$$\alpha_1 \vec{z}_1 + \alpha_2 \vec{z}_2 + \cdots + \alpha_p \vec{z}_p = \vec{0}$$

Aplicando o endomorfismo φ, obtemos

$$\varphi\left(\alpha_1\vec{z}_1 + \alpha_2\vec{z}_2 + \cdots + \alpha_p\vec{z}_p\right) = \varphi\left(\vec{0}\right) = \vec{0}$$

ou

$$\alpha_1\varphi\left(\vec{z}_1\right) + \alpha_2\varphi\left(\vec{z}_2\right) + \cdots + \alpha_p\varphi\left(\vec{z}_p\right) = \vec{0}$$

Ora, por definição de cadeia de Jordan e não esquecendo que estamos a supor $\lambda = 0$, temos

$$\varphi\left(\vec{z}_1\right) = 0\,\vec{z}_1 = \vec{0}, \quad \varphi\left(\vec{z}_2\right) = \vec{z}_1 + 0\vec{z}_2 = \vec{z}_1, \ldots, \varphi\left(\vec{z}_p\right) = \vec{z}_{p-1} + 0\vec{z}_p = \vec{z}_{p-1}$$

pelo que, substituindo na igualdade anterior, ficamos com

$$\alpha_1\vec{0} + \alpha_2\vec{z}_1 + \cdots + \alpha_p\vec{z}_{p-1} = \vec{0}$$
$$\alpha_2\vec{z}_1 + \cdots + \alpha_p\vec{z}_{p-1} = \vec{0}$$

e, como os vectores $\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_{p-1}$ são linearmente independentes, conclui-se daqui que

$$\alpha_2 = \cdots = \alpha_p = 0$$

Mas então, resulta da igualdade inicialmente considerada que

$$\alpha_1\vec{z}_1 = \vec{0}$$

pelo que, como o vector $\vec{z}_1$ é não nulo, se tem também $\alpha_1 = 0$.

a2) $\lambda \neq 0$

Neste caso, consideremos o endomorfismo $\varphi - \lambda\, 1_E$ e observemos que

$$\left(\varphi - \lambda 1_E\right)\left(\vec{z}_1\right) = \varphi\left(\vec{z}_1\right) - \lambda\vec{z}_1 = \lambda\vec{z}_1 - \lambda\vec{z}_1 = \vec{0}$$
$$\left(\varphi - \lambda 1_E\right)\left(\vec{z}_2\right) = \varphi\left(\vec{z}_2\right) - \lambda\vec{z}_2 = \vec{z}_1 + \lambda\vec{z}_2 - \lambda\vec{z}_2 = \vec{z}_1$$
$$\vdots$$
$$\left(\varphi - \lambda 1_E\right)\left(\vec{z}_p\right) = \varphi\left(\vec{z}_p\right) - \lambda\vec{z}_p = \vec{z}_{p-1} + \lambda\vec{z}_p - \lambda\vec{z}_p = \vec{z}_{p-1}$$

ou seja, a cadeia $\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_p$ é ainda uma cadeia de Jordan para o endomorfismo $\varphi - \lambda 1_E$, desta vez associada ao escalar 0. Por conseguinte, enquadra-se no primeiro caso, e os vectores $\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_p$ são linearmente independentes.

b) Resulta da própria definição de cadeia de Jordan que $\varphi(Z) \subseteq Z$, pelo que podemos considerar a restrição de φ a Z.

c) Estudemos, finalmente, a intersecção $\mathrm{Nuc}(\varphi) \cap Z$.

c1) $\lambda = 0$

Seja $\vec{v} \in \mathrm{Nuc}(\varphi) \cap Z$.

Por ser $\vec{v} \in \mathrm{Nuc}(\varphi)$, será, evidentemente, $\varphi(\vec{v}) = \vec{0}$.

Por ser $\vec{v} \in Z$, poderemos escrever $\vec{v} = \xi_1\vec{z}_1 + \xi_2\vec{z}_2 + \cdots + \xi_p\vec{z}_p$.

Consequentemente,

$$\varphi\left(\xi_1\vec{z}_1 + \xi_2\vec{z}_2 + \cdots + \xi_p\vec{z}_p\right) = \varphi(\vec{v}) = \vec{0}$$

ou

$$\xi_1\varphi(\vec{z}_1) + \xi_2\varphi(\vec{z}_2) + \cdots + \xi_p\varphi(\vec{z}_p) = \vec{0}$$

ou ainda

$$\xi_1\vec{0} + \xi_2\vec{z}_1 + \cdots + \xi_p\vec{z}_{p-1} = \vec{0}$$

Ora, como os vectores $\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_{p-1}$ são linearmente independentes, conclui-se daqui que $\xi_2 = \cdots = \xi_p = 0$, e, por conseguinte, será $\vec{v} = \xi_1\vec{z}_1$. Ora, é claro que também $\vec{z}_1 \in \mathrm{Nuc}(\varphi) \cap Z$, pelo que $\mathrm{Nuc}(\varphi) \cap Z = \langle \vec{z}_1 \rangle$.

c2) $\lambda \neq 0$

Tal como no caso anterior, se tomarmos $\vec{v} \in Nuc(\varphi) \cap Z$, teremos $\varphi(\vec{v}) = \vec{0}$ e também $\vec{v} = \xi_1\vec{z}_1 + \xi_2\vec{z}_2 + \cdots + \xi_p\vec{z}_p$.

Mais uma vez poderemos escrever

$$\varphi\left(\xi_1\vec{z}_1 + \xi_2\vec{z}_2 + \cdots + \xi_p\vec{z}_p\right) = \varphi(\vec{v}) = \vec{0}$$

ou seja

$$\xi_1\varphi(\vec{z}_1) + \xi_2\varphi(\vec{z}_2) + \cdots + \xi_p\varphi(\vec{z}_p) = \vec{0}$$

ou ainda

$$\xi_1(\lambda\vec{z}_1) + \xi_2(\vec{z}_1 + \lambda\vec{z}_2) + \cdots + \xi_p(\vec{z}_{p-1} + \lambda\vec{z}_p) = \vec{0}$$

Podemos escrever esta última igualdade na forma

$$(\lambda\xi_1 + \xi_2)\vec{z}_1 + \cdots + (\lambda\xi_{p-1} + \xi_p)\vec{z}_{p-1} + \lambda\xi_p\vec{z}_p = \vec{0}$$

e, como os vectores $\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_p$ são linearmente independentes, concluímos que

$$\begin{cases} \lambda\xi_1 + \xi_2 = 0 \\ \vdots \\ \lambda\xi_{p-1} + \xi_p = 0 \\ \lambda\xi_p = 0 \end{cases}$$

Atendendo a que $\lambda \neq 0$, e começando precisamente pela última destas igualdades, obtemos

$$\begin{cases} \xi_1 = 0 \\ \xi_2 = 0 \\ \vdots \\ \xi_p = 0 \end{cases}$$

e, portanto, $\vec{v} = \vec{0}$. Consequentemente, $\mathrm{Nuc}(\varphi) \cap Z = \{\vec{0}\}$.

O teorema fundamental que nos propomos demonstrar é o seguinte:

Teorema 6-20 Seja φ um endomorfismo de um espaço vectorial complexo E, de dimensão n. Existe em E uma base $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$, em relação à qual a matriz de φ tem a forma canónica de Jordan. Essa base toma o nome de **base de Jordan** para φ.

Demonstração:

Designaremos por A a matriz de φ em relação a uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m$ arbitrariamente fixa em E, e distinguiremos dois casos:

a) φ não tem o valor próprio 0

Uma vez que E é um espaço complexo, φ terá pelo menos um valor próprio λ, e sabemos que, se λ é valor próprio de φ, então 0 é valor próprio de $\varphi - \lambda 1_E$.

Por outro lado, se existir uma base de Jordan para $\varphi - \lambda 1_E$, a matriz C deste endomorfismo nessa base terá a forma canónica de Jordan, e sendo P a matriz de mudança de base, teremos $C = P^{-1}(A - \lambda I_n)P$.

Ora, daqui resulta que

$$C = P^{-1}AP - \lambda I_n$$
$$P^{-1}AP = C + \lambda I_n$$

Atente-se a que $C + \lambda I_n$ tem, por definição, a forma canónica de Jordan, pelo que a nova base considerada é também uma base de Jordan para φ.

Vemos, assim, que o problema fica resolvido, neste primeiro caso, desde que se garanta a existência de uma base de Jordan para $\varphi - \lambda 1_E$, que tem o valor próprio 0, ou, de modo mais geral, para um endomorfismo que admita o valor próprio 0. Podemos, portanto, passar ao segundo caso:

b) φ tem o valor próprio 0

Faremos a demonstração por indução em n.

O teorema é trivial quando $n = 1$, já que qualquer matriz do tipo 1×1 é um bloco de Jordan, e está, portanto, na forma canónica de Jordan.

Suponhamos então que a tese é verdadeira para todos os espaços vectoriais de dimensão menor que n.

Como φ tem o valor próprio 0, será $r(A) = r < n$.

Designemos por φ^* a restrição de φ a $\varphi(E)$. Uma vez que

$$\dim(\varphi(E)) = c_\varphi = r(A) = r < n$$

a tese é verdadeira em $\varphi(E)$, por hipótese de indução, e existe então em $\varphi(E)$ uma base de Jordan, $\vec{v}_1^*, \vec{v}_2^*, \ldots, \vec{v}_r^*$, para φ^*.

Seja $W = \varphi(E) \cap \mathrm{Nuc}(\varphi) = \mathrm{Nuc}(\varphi^*)$ e seja ainda J^* a matriz de φ^* em relação à base de Jordan considerada. Como J^* tem a forma canónica de Jordan, a nulidade de φ^* é igual ao número de linhas nulas em J^*: se se tem $\dim(W) = s$, então há s linhas nulas em J^*.

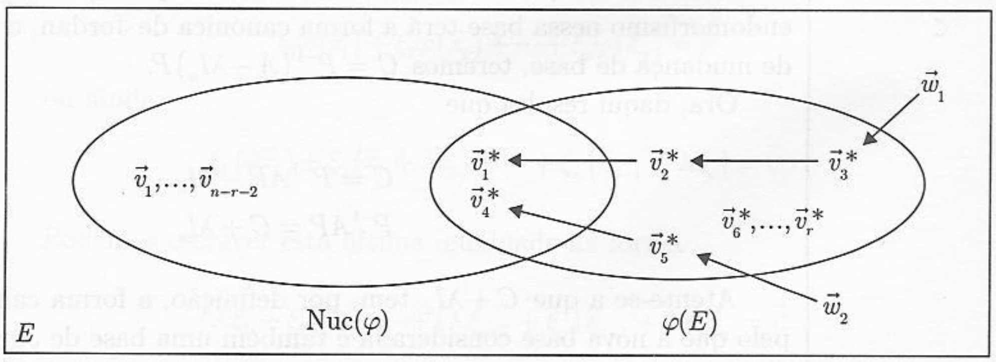
Por sua vez, J^* tem tantas linhas nulas quantos os blocos de Jordan associados ao valor próprio 0 (visto que em cada um destes e só aí há uma e uma só linha de zeros, que é a última).

Sendo λ um valor próprio de φ^*, e $\vec{v}_i^*$ um vector próprio associado a λ, podemos, por aplicação sucessiva de $\varphi^* - \lambda 1_E$, construir uma cadeia

$$\vec{v}_{j-1}^* \mapsto \cdots \mapsto \vec{v}_{i+1}^* \mapsto \vec{v}_i^* \mapsto \vec{0}, \quad \text{onde} \quad (\varphi^* - \lambda 1_E)(\vec{v}_j^*) \neq \vec{v}_{j-1}^*$$

Ora, há tantos blocos associados ao valor próprio 0 quantas as cadeias deste tipo em que $\vec{v}_i^* \in W$; ou seja, há exactamente s desses blocos.

A figura seguinte ilustra uma possível situação do tipo que acabámos de descrever, com $s = 2$:



Cada vector $\vec{v}_{j-1}^*$ está em $\varphi(E)$, pelo que, para algum vector $\vec{w} \in E$, se terá $\vec{v}_{j-1}^* = \varphi(\vec{w})$.

Obtemos deste modo os vectores $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_s$, em E.

Finalmente, como $\mathrm{Nuc}(\varphi)$ tem dimensão $n-r$ e os s vectores $\vec{v}_i^*$ estão em $\mathrm{Nuc}(\varphi)$, podemos, usando o teorema de Steinitz, procurar em $\mathrm{Nuc}(\varphi)$ vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-r-s}$ tais que $\vec{v}_1^*, \vec{v}_2^*, \ldots, \vec{v}_r^*, \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-r-s}$ seja uma base de $\mathrm{Nuc}(\varphi)$.

Vamos então ver que os vectores

$$\vec{v}_1^*, \vec{v}_2^*, \ldots, \vec{v}_r^*, \vec{w}_1, \vec{w}_2, \ldots, \vec{w}_s, \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-r-s}$$

se podem reordenar por forma a obter-se uma base de Jordan para φ.

Para isso, basta colocar cada $\vec{w}_i$ a seguir a um determinado $\vec{v}_p^*$, sempre que $\varphi(\vec{w}_i) = \vec{v}_p^*$.

O número total de vectores obtidos é $r + s + (n - r - s) = n$, pelo que, para que formem base para E, bastará que sejam linearmente independentes.

Tomemos então

$$\alpha_1 \vec{v}_1^* + \alpha_2 \vec{v}_2^* + \cdots + \alpha_r \vec{v}_r^* + \beta_1 \vec{w}_1 + \beta_2 \vec{w}_2 + \cdots + \beta_s \vec{w}_s + \\ + \gamma_1 \vec{v}_1 + \gamma_2 \vec{v}_2 + \cdots + \gamma_{n-r-s} \vec{v}_{n-r-s} = \vec{0}$$

Aplicando φ:

$$\alpha_1 \varphi(\vec{v}_1^*) + \cdots + \alpha_r \varphi(\vec{v}_r^*) + \beta_1 \varphi(\vec{w}_1) + \cdots + \beta_s \varphi(\vec{w}_s) + \\ + \gamma_1 \varphi(\vec{v}_1) + \cdots + \gamma_{n-r-s} \varphi(\vec{v}_{n-r-s}) = \vec{0}$$

Como $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-r-s} \in \mathrm{Nuc}(\varphi)$, esta igualdade reduz-se a

$$\alpha_1 \varphi(\vec{v}_1^*) + \cdots + \alpha_r \varphi(\vec{v}_r^*) + \beta_1 \varphi(\vec{w}_1) + \cdots + \beta_s \varphi(\vec{w}_s) = \vec{0}$$

Cada vector $\varphi(\vec{v}_i^*)\,(i = 1, 2, \ldots, r)$ só pode ser $\lambda \vec{v}_i^*$ ou $\lambda \vec{v}_i^* + \vec{v}_{i-1}^*$, pelo que

$$\alpha_1 \varphi(\vec{v}_1^*) + \cdots + \alpha_r \varphi(\vec{v}_r^*)$$

é uma combinação linear de alguns vectores de

$$\{\vec{v}_1^*, \vec{v}_2^*, \ldots, \vec{v}_r^*\}$$

que não iniciam qualquer das cadeias consideradas.

Por outro lado, os vectores $\varphi(\vec{w}_1), \ldots, \varphi(\vec{w}_s)$ são, entre os vectores $\vec{v}_1^*, \vec{v}_2^*, \ldots, \vec{v}_r^*$, precisamente aqueles que iniciam essas cadeias.

Daqui resulta que não há vectores de $\vec{v}_1^*, \vec{v}_2^*, \ldots, \vec{v}_r^*$ comuns às expressões de

$$\alpha_1\varphi\left(\vec{v}_1^*\right)+\cdots+\alpha_r\varphi\left(\vec{v}_r^*\right) \quad \text{e de} \quad \beta_1\varphi(\vec{w}_1)+\cdots+\beta_s\varphi(\vec{w}_s)$$

Dado que $\vec{v}_1^*, \vec{v}_2^*, \ldots, \vec{v}_r^*$ são linearmente independentes, concluímos que

$$\beta_1 = \beta_2 = \cdots = \beta_s = 0$$

e obtemos então

$$\alpha_1\vec{v}_1^* + \alpha_2\vec{v}_2^* + \cdots + \alpha_r\vec{v}_r^* + \gamma_1\vec{v}_1 + \gamma_2\vec{v}_2 + \cdots + \gamma_{n-r-s}\vec{v}_{n-r-s} = \vec{0}$$

Tomemos

$$\vec{z} = \alpha_1\vec{v}_1^* + \alpha_2\vec{v}_2^* + \cdots + \alpha_r\vec{v}_r^* = -\gamma_1\vec{v}_1 - \gamma_2\vec{v}_2 - \cdots - \gamma_{n-r-s}\vec{v}_{n-r-s}$$

Vemos que $\vec{z} \in \varphi(E) \cap \mathrm{Nuc}(\varphi) = W$, pelo que $\vec{z}$ é combinação linear dos vectores de $\left\{\vec{v}_1^*, \vec{v}_2^*, \ldots, \vec{v}_r^*\right\}$, em número de s, que formam a base de W.

Ora, foi a partir destes que, com o teorema de Steinitz, se obtiveram $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-r-s}$, o que nos permite, de imediato, concluir que

$$\gamma_1 = \gamma_2 = \cdots = \gamma_{n-r-s} = 0$$

e, logo, que

$$\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$$

O teorema anterior é ainda completado pelo seguinte:

Corolário 6-21 Nas condições do teorema anterior, a matriz de φ que tem a forma canónica de Jordan é única, excepto no que respeita à ordenação dos seus blocos de Jordan.

Demonstração:

Se J é a matriz de φ em relação a certa base, e está na forma canónica de Jordan, os blocos que nela figuram ficam completamente determinados pelas características das matrizes $\left(A - \lambda I_n\right)^k$ para cada valor próprio λ de φ e para cada natural k.

Independentemente da demonstração da existência de uma base de Jordan para um endomorfismo qualquer, é conveniente, na prática, dispormos de um algoritmo que permita determiná-la, bem como à matriz J, do endomorfismo em causa, na forma canónica de Jordan.

Comecemos por estudar alguns exemplos:

Exemplo 6-22

Consideremos a matriz

$$J=\begin{bmatrix} -i & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -i & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

que define um certo endomorfismo ψ do espaço vectorial complexo $E=\mathbb{C}^8$ em relação a certa base, $\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_8$.

Se considerarmos $\psi+i1_E$, obtemos

$$\vec{e}_3 \mapsto \vec{e}_2 \mapsto \vec{e}_1 \mapsto \vec{0}$$
$$\vec{e}_5 \mapsto \vec{e}_4 \mapsto \vec{0}$$

Por outro lado, $\psi+i1_E$ transforma $\vec{e}_6,\vec{e}_7,\vec{e}_8$ em vectores independentes (o mesmo acontecendo a qualquer potência de $\psi+i1_E$, já que a restrição deste endomorfismo a $\langle\vec{e}_6,\vec{e}_7,\vec{e}_8\rangle$ tem a matriz

$$\begin{bmatrix} 2+i & 0 & 0 \\ 0 & 5+i & 1 \\ 0 & 0 & 5+i \end{bmatrix}$$

de característica 3).

Vemos então que:

$$\mathrm{Nuc}(\psi+i1_E)=\langle\vec{e}_1,\vec{e}_4\rangle$$
$$\mathrm{Nuc}\left((\psi+i1_E)^2\right)=\langle\vec{e}_1,\vec{e}_2,\vec{e}_4,\vec{e}_5\rangle$$
$$\mathrm{Nuc}\left((\psi+i1_E)^3\right)=\langle\vec{e}_1,\vec{e}_2,\vec{e}_3,\vec{e}_4,\vec{e}_5\rangle=\mathrm{Nuc}\left((\psi+i1_E)^k\right),\quad \forall k>3$$

Analogamente,

$$\mathrm{Nuc}\left((\psi-21_E)^k\right)=\langle\vec{e}_6\rangle,\quad \forall k\geq 1$$

e

$$\mathrm{Nuc}(\psi-51_E)=\langle\vec{e}_7\rangle$$
$$\mathrm{Nuc}\left((\psi-51_E)^k\right)=\langle\vec{e}_7,\vec{e}_8\rangle,\quad \forall k>1 \ \blacksquare$$

Exemplo 6-23

Suponhamos que J é uma matriz do tipo 9×9, na forma canónica de Jordan, que verifica as seguintes condições:

$$r\left(\left(J-3iI_9\right)^k\right)=\begin{cases}7, & \text{se } k=1\\ 5, & \text{se } k=2\\ 4, & \text{se } k\geq 3\end{cases} \qquad r\left(\left(J+I_9\right)^k\right)=\begin{cases}6, & \text{se } k=1\\ 5, & \text{se } k\geq 2\end{cases}$$

Podemos, a partir destas propriedades, reconstituir os blocos de Jordan de J, raciocinando do seguinte modo, onde designamos por θ o endomorfismo do espaço complexo $\mathbb{C}^9$ definido, em relação a certa base $\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_9$, pela matriz J:

$$\begin{aligned} r\left(J-3iI_9\right)=7 \quad &\Rightarrow \quad \dim\left(\mathrm{Nuc}\left(\theta-3i1_E\right)\right)=2\\ &\Rightarrow \quad 3i \text{ é valor próprio de } \theta\text{, com multiplicidade geométrica 2}\\ &\Rightarrow \quad 3i \text{ dá lugar a 2 blocos de Jordan}\end{aligned}$$

Ao mesmo tempo, $\theta-3i1_E$ vai anular dois vectores, $\vec{e}_r$ e $\vec{e}_s$, da base considerada (que serão vectores próprios de θ associados ao valor próprio $3i$).

Como $r\left(\left(J-3iI_9\right)^2\right)=5$, será $\dim\left(\mathrm{Nuc}\left(\theta-3i1_E\right)^2\right)=4$, pelo que, aplicando $\theta-3i1_E$, teremos já

$$\vec{e}_{r+1}\mapsto\vec{e}_r\mapsto\vec{0}$$
$$\vec{e}_{s+1}\mapsto\vec{e}_s\mapsto\vec{0}$$

Finalmente, como $r\left(\left(J-3iI_9\right)^3\right)=4$, será então $\dim\left(\mathrm{Nuc}\left(\theta-3i1_E\right)^3\right)=5$ e haverá mais um vector numa das cadeias anteriores ($\vec{e}_{r+2}$ ou $\vec{e}_{s+2}$).

Consequentemente, os dois blocos de Jordan associados a $3i$ serão:

$$J_1=\begin{bmatrix}3i & 1 & 0\\ 0 & 3i & 1\\ 0 & 0 & 3i\end{bmatrix} \quad \text{e} \quad J_2=\begin{bmatrix}3i & 1\\ 0 & 3i\end{bmatrix}$$

Por outro lado,

$$\begin{aligned} r\left(J+I_9\right)=6 \quad &\Rightarrow \quad \dim\left(\mathrm{Nuc}\left(\theta+1_E\right)\right)=3\\ &\Rightarrow \quad -1 \text{ é valor próprio de } \theta\text{, com multiplicidade geométrica 3}\\ &\Rightarrow \quad -1 \text{ dá lugar a 3 blocos de Jordan}\end{aligned}$$

Como $r\left(\left(J+I_9\right)^2\right)=5$, será $\dim\left(\mathrm{Nuc}\left(\theta+1_E\right)^3\right)=4$, pelo que teremos cadeias do tipo

$$\vec{a}\mapsto\vec{b}\mapsto\vec{0}$$
$$\vec{c}\mapsto\vec{0}$$
$$\vec{d}\mapsto\vec{0}$$

e os blocos de Jordan associados a -1 serão

$$J_3 = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}, \quad J_4 = [-1] \quad \text{e} \quad J_5 = [-1]$$

Assim, a matriz J será constituída pelos blocos de Jordan J_1, J_2, J_3, J_4, J_5, dispostos ao longo da diagonal, por uma ordem qualquer.

Por exemplo:

$$J = \begin{bmatrix} 3i & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3i & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3i & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3i & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \blacksquare$$

Os raciocínios apresentados nestes exemplos têm, evidentemente, carácter geral, e permitem, portanto, a partir de qualquer matriz A obter a respectiva forma canónica de Jordan, bastando para o efeito, para cada valor próprio λ do endomorfismo φ definido por A em relação a certa base, calcular o valor de

$$r\left((A - \lambda I_n)^k\right) \quad (k \geq 1)$$

Em particular, deve observar-se que, para cada valor próprio λ, a multiplicidade algébrica de λ indica o número de vezes que este valor figura na diagonal da matriz, na forma canónica de Jordan, ao passo que a respectiva multiplicidade geométrica nos diz o número de blocos de Jordan que estão associados a λ.

Finalmente, indicaremos, através de um exemplo, como se poderá obter uma base de Jordan para um certo endomorfismo, seguindo os métodos que se encontram nas demonstrações dos teoremas anteriores.

Exemplo 6-24

Seja A uma matriz do tipo $n \times n$, e ξ o endomorfismo de um certo espaço vectorial definido por A; para um valor próprio λ, suponhamos que se tem

$$\dim\left(\mathrm{Nuc}(\xi - \lambda I_n)^k\right) = \begin{cases} 4, & \text{se } k = 1 \\ 7, & \text{se } k = 2 \\ 8, & \text{se } k \geq 3 \end{cases}$$

Neste caso, uma base de Jordan terá 4 cadeias, da forma

$$\vec{b}_3 \mapsto \vec{b}_2 \mapsto \vec{b}_1 \mapsto \vec{0}$$
$$\vec{b}_5 \mapsto \vec{b}_4 \mapsto \vec{0}$$
$$\vec{b}_7 \mapsto \vec{b}_6 \mapsto \vec{0}$$
$$\vec{b}_8 \mapsto \vec{0}$$

Seja $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_8$ uma base de $\mathrm{Nuc}(\xi - \lambda I_n)^3$.
Ter-se-á

$$\dim\Big((\xi - \lambda 1_E)^2 \big(\mathrm{Nuc}(\xi - \lambda I_n)^3\big)\Big) = 1$$

pelo que algum dos vectores $\vec{v}_i$ verificará a igualdade $(\xi - \lambda 1_E)^2(\vec{v}_i) \neq \vec{0}$. Seja $\vec{b}_3$ um vector nessas condições.

Pondo $\vec{b}_2 = (\xi - \lambda 1_E)(\vec{b}_3)$ e $\vec{b}_1 = (\xi - \lambda 1_E)(\vec{b}_2)$, obtemos a primeira cadeia.

Como $\vec{b}_1, \vec{b}_2 \in \mathrm{Nuc}(\xi - \lambda 1_E)^2$, construímos uma base

$$\vec{b}_1, \vec{b}_2, \vec{w}_1, \ldots, \vec{w}_5$$

para $\mathrm{Nuc}(\xi - \lambda 1_E)^2$.
Por ser

$$\dim\Big((\xi - \lambda 1_E)\big(\mathrm{Nuc}(\xi - \lambda I_n)^2\big)\Big) = 3$$

haverá dois vectores, $\vec{w}_i$ e $\vec{w}_j$, tais que $\vec{b}_1, (\xi - \lambda 1_E)(\vec{w}_i), (\xi - \lambda 1_E)(\vec{w}_j)$ são linearmente independentes.
Sejam

$$\vec{b}_5 = \vec{w}_i, \quad \vec{b}_4 = (\xi - \lambda 1_E)(\vec{b}_5)$$
$$\vec{b}_7 = \vec{w}_j, \quad \vec{b}_6 = (\xi - \lambda 1_E)(\vec{b}_7)$$

Os vectores $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_7$ são linearmente independentes.

Finalmente obtemos uma base $\vec{b}_1, \vec{b}_4, \vec{b}_6, \vec{b}_8$ para $\mathrm{Nuc}(\xi - \lambda 1_E)$ e assim completamos a parte da base de Jordan correspondente ao valor próprio λ. ■

Na prática estes cálculos são bastante morosos, especialmente para matrizes com um elevado número de linhas e colunas. Para as matrizes mais pequenas, porém, podem abreviar-se, se partirmos já da forma canónica de Jordan, mais fácil de obter.

Com efeito, se uma matriz A tem, por exemplo, a forma canónica de Jordan

$$\begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

para se obter uma base de Jordan deveremos procurar três vectores, $\vec{u}_1, \vec{u}_2, \vec{u}_3$, em que $\vec{u}_1$ e $\vec{u}_3$ serão vectores próprios do endomorfismo ψ definido por A em relação a certa base, associados aos valores próprios 3 e 4, respectivamente. Quanto ao vector $\vec{u}_2$, ele satisfará a equação

$$\psi(\vec{u}_2) = \vec{u}_1 + 3\vec{u}_2$$

que não é, em geral, difícil de resolver.

Analogamente, se A tem, por exemplo, a forma canónica de Jordan

$$\begin{bmatrix} -2 & 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

deveremos procurar uma base $\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4, \vec{u}_5$, na qual $\vec{u}_1$ e $\vec{u}_3$ são vectores próprios do endomorfismo ψ associados aos valores próprios -2 e 5, respectivamente, enquanto os restantes satisfazem as condições

$$\psi(\vec{u}_2) = \vec{u}_1 - 2\vec{u}_2$$
$$\psi(\vec{u}_4) = \vec{u}_3 + 5\vec{u}_4$$
$$\psi(\vec{u}_5) = \vec{u}_4 + 5\vec{u}_5$$

7 Produtos internos

7-1 Generalidades

No presente capítulo, veremos de que maneira é possível enriquecer a estrutura de um espaço vectorial com a introdução de um novo conceito, o de "produto interno". Por razões que exporemos, a seu tempo, a noção de produto interno é definida só para espaços vectoriais reais ou complexos.

Começamos precisamente por apresentar a definição, examinando, em primeiro lugar, o caso dos espaços reais:

Definição 7-1 Seja E um espaço vectorial real. Chama-se **produto interno** (ou **produto escalar**) em E a qualquer aplicação de $E \times E$ para $\mathbb{R}$, na qual designaremos a imagem de cada par $(\vec{x}, \vec{y})$, de vectores de E, por $\vec{x} \bullet \vec{y}$, que satisfaça as seguintes propriedades, para vectores e escalares quaisquer:

a) a1) $(\vec{x} + \vec{x}') \bullet \vec{y} = \vec{x} \bullet \vec{y} + \vec{x}' \bullet \vec{y}$

a2) $\vec{x} \bullet (\vec{y} + \vec{y}') = \vec{x} \bullet \vec{y} + \vec{x} \bullet \vec{y}'$

b) b1) $(\alpha \vec{x}) \bullet \vec{y} = \alpha (\vec{x} \bullet \vec{y})$

b2) $\vec{x} \bullet (\alpha \vec{y}) = \alpha (\vec{x} \bullet \vec{y})$

c) $\vec{x} \bullet \vec{y} = \vec{y} \bullet \vec{x}$

d) d1) $\vec{x} \bullet \vec{x} \geq 0$

d2) Se $\vec{x} \bullet \vec{x} = 0$, então $\vec{x} = \vec{0}$

A uma aplicação de $E \times E$ para $\mathbb{R}$ chama-se, geralmente, **forma**; por obedecer às propriedades a1) e a2) e b1) e b2), diz-se que um produto interno num espaço real é uma forma **bilinear**; por satisfazer a propriedade 3), será **simétrica**, e por verificar d1) e d2) será **definida positiva**. Assim, resumidamente, um produto interno num espaço vectorial é uma forma bilinear simétrica definida positiva.

No caso de lidarmos com um espaço vectorial complexo, a definição de produto interno sofre duas ligeiras alterações, como vamos indicar já a seguir:

Definição 7-2

Seja E um espaço vectorial complexo. Chama-se **produto interno** (ou **produto escalar**) em E a qualquer aplicação de $E \times E$ para $\mathbb{C}$, na qual designaremos a imagem de cada par $(\vec{x}, \vec{y})$, de vectores de E, por $\vec{x} \bullet \vec{y}$, que satisfaça as seguintes propriedades, para vectores e escalares quaisquer:

a) a1) $(\vec{x} + \vec{x}') \bullet \vec{y} = \vec{x} \bullet \vec{y} + \vec{x}' \bullet \vec{y}$

a2) $\vec{x} \bullet (\vec{y} + \vec{y}') = \vec{x} \bullet \vec{y} + \vec{x} \bullet \vec{y}'$

b) b1) $(\alpha\vec{x}) \bullet \vec{y} = \alpha(\vec{x} \bullet \vec{y})$

b2) $\vec{x} \bullet (\alpha\vec{y}) = \overline{\alpha}(\vec{x} \bullet \vec{y})$

c) $\vec{x} \bullet \vec{y} = \overline{\vec{y} \bullet \vec{x}}$

d) d1) $\vec{x} \bullet \vec{x} \geq 0$

d2) Se $\vec{x} \bullet \vec{x} = 0$, então $\vec{x} = \vec{0}$

Como habitualmente, sendo $z = a + bi$ um número complexo, designamos por $\overline{z}$ o seu conjugado, $\overline{z} = a - bi$.

A uma aplicação de $E \times E$ para $\mathbb{C}$ chama-se, geralmente, **forma**; por obedecer às propriedades a1) e a2) e b1) e b2), diz-se que um produto interno num espaço real é uma forma **sesquilinear**; por satisfazer a propriedade 3), será **hermítica**, e por verificar d1) e d2) será **definida positiva**. Assim, resumidamente, um produto interno num espaço vectorial é uma forma sesquilinear hermítica definida positiva.

Deve notar-se que um número complexo z coincide com o seu conjugado se e só se z é um número real. Por esse motivo, no caso de se tratar de um espaço vectorial real, as igualdades $\vec{x} \bullet (\alpha\vec{y}) = \overline{\alpha}(\vec{x} \bullet \vec{y})$ e $\vec{x} \bullet \vec{y} = \overline{\vec{y} \bullet \vec{x}}$ equivalem, respectivamente, a $\vec{x} \bullet (\alpha\vec{y}) = \alpha(\vec{x} \bullet \vec{y})$ e $\vec{x} \bullet \vec{y} = \vec{y} \bullet \vec{x}$. Sendo assim, usaremos muitas vezes as primeiras para abranger, simultaneamente, as duas possibilidades (espaço real ou espaço complexo). É o que faremos já na demonstração da seguinte propriedade:

Teorema 7-3

Seja E um espaço vectorial real ou complexo, onde se definiu um produto interno. Para qualquer vector $\vec{x} \in E$, tem-se $\vec{x} \bullet \vec{x} \in [0, +\infty[$.

Demonstração:

Pela propriedade c), aplicada ao caso particular em que $\vec{y} = \vec{x}$, vemos que $\vec{x} \bullet \vec{x} = \overline{\vec{x} \bullet \vec{x}}$, pelo que $\vec{x} \bullet \vec{x} \in \mathbb{R}$. A propriedade d1) garante que $\vec{x} \bullet \vec{x}$ não é negativo.

O facto de $\vec{x} \bullet \vec{x}$ ser sempre um número real não negativo dá sentido a uma nova definição, associada à definição de produto interno:

Definição 7-4 Seja E um espaço vectorial real ou complexo, onde se definiu um produto interno. Para cada vector $\vec{x} \in E$, chama-se **norma** de $\vec{x}$ ao número real não negativo $\|\vec{x}\| = \sqrt{\vec{x} \bullet \vec{x}}$.

Deve também observar-se que, na definição de produto interno, as propriedades a2) e b2) são, na verdade, supérfluas. Com efeito, a partir de c) e de a1), podemos verificar que

$$\vec{x} \bullet (\vec{y} + \vec{y}') = \overline{(\vec{y} + \vec{y}') \bullet \vec{x}} = \overline{\vec{y} \bullet \vec{x} + \vec{y}' \bullet \vec{x}} = \overline{\vec{y} \bullet \vec{x}} + \overline{\vec{y}' \bullet \vec{x}} = \vec{x} \bullet \vec{y} + \vec{x} \bullet \vec{y}'$$

e, a partir de c) e de b1), que

$$\vec{x} \bullet (\alpha \vec{y}) = \overline{(\alpha \vec{y}) \bullet \vec{x}} = \overline{\alpha (\vec{y} \bullet \vec{x})} = \overline{\alpha}\ \overline{\vec{y} \bullet \vec{x}} = \overline{\alpha} (\vec{x} \bullet \vec{y})$$

A razão pela qual não definimos produtos internos em espaços vectoriais sobre outros corpos (por exemplo, sobre corpos com um número finito de elementos) reside fundamentalmente na aplicabilidade da propriedade d1). Com efeito, na maioria dos corpos não está definida uma relação de ordem, pelo que não teria sentido falar em escalares maiores ou menores que 0. Por sua vez, a definição de "norma" exige que todo o escalar não negativo tenha raiz quadrada no corpo considerado, o que exclui também algumas possibilidades, como, por exemplo, a utilização do corpo dos racionais.

Exemplos 7-5

a) Consideremos, no espaço vectorial real $\mathbb{R}^2$, a aplicação de duas variáveis definida, para vectores quaisquer, $\vec{x} = (x_1, x_2)$, $\vec{y} = (y_1, y_2)$, por

$$\vec{x} \bullet \vec{y} = 5x_1y_1 - 2x_1y_2 - 2x_2y_1 + x_2y_2$$

É fácil verificar que se trata de um produto interno, examinando as diversas propriedades:

a1) $$\begin{aligned}(\vec{x} + \vec{x}') \bullet \vec{y} &= 5(x_1 + x_1')y_1 - 2(x_1 + x_1')y_2 - 2(x_2 + x_2')y_1 + x_2y_2 = \\ &= 5x_1y_1 + 5x_1'y_1 - 2x_1y_2 - 2x_1'y_2 - 2x_2y_1 - 2x_2'y_1 + x_2y_2 + x_2'y_2 = \\ &= (5x_1y_1 - 2x_1y_2 - 2x_2y_1 + x_2y_2) + (5x_1y_1 - 2x_1y_2 - 2x_2y_1 + x_2y_2) = \\ &= \vec{x} \bullet \vec{y} + \vec{x}' \bullet \vec{y}\end{aligned}$$

a2) $$\begin{aligned}(\alpha \vec{x}) \bullet \vec{y} &= 5(\alpha x_1)y_1 - 2(\alpha x_1)y_2 - 2(\alpha x_2)y_1 + (\alpha x_2)y_2 = \\ &= \alpha(5x_1y_1 - 2x_1y_2 - 2x_2y_1 + x_2y_2) = \alpha(\vec{x} \bullet \vec{y})\end{aligned}$$

a3) $\vec{y} \bullet \vec{x} = 5y_1x_1 - 2y_1x_2 - 2y_2x_1 + y_2x_2 = 5x_1y_1 - 2x_1y_2 - 2x_2y_1 + x_2y_2 = \vec{x} \bullet \vec{y}$

a4) 1) $\vec{x}\bullet\vec{x}=5x_1^2-2x_1x_2-2x_2x_1+x_2^2=5x_1^2-4x_1x_2+x_2^2=$
$=x_1^2+(2x_1-x_2)^2\geq 0$

2) $\vec{x}\bullet\vec{x}=x_1^2+(2x_1-x_2)^2=0\Rightarrow\begin{cases}x_1=0\\2x_1-x_2=0\end{cases}\Rightarrow\begin{cases}x_1=0\\x_2=0\end{cases}$

b) Analogamente se verifica que qualquer das seguintes expressões define um produto interno no espaço $\mathbb{R}^2$

$$\vec{x}\bullet\vec{y}=x_1y_1+x_1y_2+x_2y_1+2x_2y_2$$
$$\vec{x}\bullet\vec{y}=10x_1y_1+3x_1y_2+3x_2y_1+5x_2y_2$$
$$\vec{x}\bullet\vec{y}=x_1y_1+x_2y_2$$

Este último toma o nome de **produto interno canónico** em $\mathbb{R}^2$.

c) As expressões

$$\vec{x}\bullet\vec{y}=3x_1y_1+x_1y_2+x_1y_3+x_2y_1+2x_2y_2+x_2y_3+x_3y_1+x_3y_2+2x_3y_3$$
$$\vec{x}\bullet\vec{y}=5x_1y_1+2x_1y_2-x_1y_3+2x_2y_1+2x_2y_2-3x_2y_3-x_3y_1-3x_3y_2+12x_3y_3$$
$$\vec{x}\bullet\vec{y}=x_1y_1+x_2y_2+x_3y_3$$

definem produtos internos em $\mathbb{R}^3$, sendo o último deles o produto interno canónico. A verificação é semelhante à anterior.

De modo geral, em qualquer espaço $\mathbb{R}^n$, com n natural, a expressão

$$\vec{x}\bullet\vec{y}=x_1y_1+x_2y_2+\cdots+x_ny_n$$

define o produto interno canónico.

d) No espaço vectorial E, constituído pelas funções reais de variável real contínuas num certo intervalo $[a,b]$, com as operações usuais, podemos definir um produto interno por

$$f\bullet g=\int_a^b f(x)g(x)dx$$

A verificação é feita utilizando propriedades conhecidas dos integrais. ■

Podemos agora passar à demonstração de algumas propriedades dos produtos internos e das normas:

Teorema 7-6 | Seja E um espaço vectorial real, ou complexo, onde se definiu um produto interno. Para vectores arbitrários de E, são válidas as seguintes propriedades:

a) $\|\vec{x}\|=0$ se e só se $\vec{x}=\vec{0}$
b) $\|\alpha\vec{x}\|=|\alpha|\|\vec{x}\|$

c) $|\vec{x} \bullet \vec{y}| \leq \|\vec{x}\| \|\vec{y}\|$ **(desigualdade de Schwarz)**
d) Dá-se a igualdade, na desigualdade de Schwarz, se e só se os vectores $\vec{x}$ e $\vec{y}$ são linearmente dependentes.
e) $\|\vec{x}+\vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ **(desigualdade triangular)**
f) Dá-se a igualdade, na desigualdade triangular, se e só se um dos vectores se pode obter à custa do outro, por multiplicação por um escalar real não negativo.

Demonstração:

a) Atendendo à definição de norma, teremos $\|\vec{x}\| = 0$ se e só se $\vec{x} \bullet \vec{x} = 0$.

Ora, pelas propriedades de um produto interno, vemos que

$$\vec{0} \bullet \vec{0} = (0\vec{0}) \bullet \vec{0} = 0(\vec{0} \bullet \vec{0}) = 0$$

pelo que $\|\vec{0}\| = 0$. Com a propriedade d2) da definição de produto interno, a nossa afirmação fica provada.

b) Temos

$$\|\alpha\vec{x}\| = \sqrt{(\alpha\vec{x}) \bullet (\alpha\vec{x})} = \sqrt{\alpha\overline{\alpha}(\vec{x} \bullet \vec{x})} = \sqrt{|\alpha|^2 (\vec{x} \bullet \vec{x})} = |\alpha| \sqrt{\vec{x} \bullet \vec{x}} = |\alpha| \|\vec{x}\|$$

c) A desigualdade é trivial quando algum dos vectores considerados é nulo (visto ambos os membros darem zero), pelo que podemos, no que se segue, supor que lidamos com vectores não nulos. Atendendo à propriedade d1) da Definição 7-2, podemos escrever, para qualquer escalar γ:

$$(\gamma\vec{x} + \vec{y}) \bullet (\gamma\vec{x} + \vec{y}) \geq 0$$

Mas

$$\begin{aligned}(\gamma\vec{x} + \vec{y}) \bullet (\gamma\vec{x} + \vec{y}) &= \gamma\overline{\gamma}(\vec{x} \bullet \vec{x}) + \gamma(\vec{x} \bullet \vec{y}) + \overline{\gamma}(\vec{y} \bullet \vec{x}) + \vec{y} \bullet \vec{y} = \\ &= \gamma\overline{\gamma}\|\vec{x}\|^2 + \gamma(\vec{x} \bullet \vec{y}) + \overline{\gamma}(\vec{y} \bullet \vec{x}) + \|\vec{y}\|^2\end{aligned}$$

Podemos, portanto, escrever, para qualquer escalar γ:

$$\gamma\overline{\gamma}\|\vec{x}\|^2 + \gamma(\vec{x} \bullet \vec{y}) + \overline{\gamma}(\vec{y} \bullet \vec{x}) + \|\vec{y}\|^2 \geq 0$$

Uma vez que γ é arbitrário e tendo em consideração que $\vec{x} \neq \vec{0}$, tomaremos

$$\gamma = -\frac{\vec{y} \bullet \vec{x}}{\|\vec{x}\|^2}$$

sendo, então,

$$\overline{\gamma} = \overline{\left(-\frac{\vec{y}\cdot\vec{x}}{\|\vec{x}\|^2}\right)} = -\frac{\overline{\vec{y}\cdot\vec{x}}}{\|\vec{x}\|^2} = -\frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|^2}$$

Substituindo na desigualdade anterior, obtemos

$$\left(-\frac{\vec{y}\cdot\vec{x}}{\|\vec{x}\|^2}\right)\left(-\frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|^2}\right)\|\vec{x}\|^2 - \frac{\vec{y}\cdot\vec{x}}{\|\vec{x}\|^2}(\vec{x}\cdot\vec{y}) - \frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|^2}(\vec{y}\cdot\vec{x}) + \|\vec{y}\|^2 \geq 0$$

Simplificando, chegamos a

$$-\frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|^2}(\vec{y}\cdot\vec{x}) + \|\vec{y}\|^2 \geq 0 \quad \text{ou} \quad (\vec{x}\cdot\vec{y})(\vec{y}\cdot\vec{x}) \leq \|\vec{x}\|^2\ \|\vec{y}\|^2$$

Atendendo a que $\vec{y}\cdot\vec{x} = \overline{\vec{x}\cdot\vec{y}}$, será $(\vec{x}\cdot\vec{y})(\vec{y}\cdot\vec{x}) = |\vec{x}\cdot\vec{y}|^2$, pelo que

$$|\vec{x}\cdot\vec{y}|^2 \leq \|\vec{x}\|^2\ \|\vec{y}\|^2$$

Uma vez que tanto o módulo de um escalar como as normas dos vectores são números reais não negativos, podemos daqui deduzir, finalmente, que

$$|\vec{x}\cdot\vec{y}| \leq \|\vec{x}\|\ \|\vec{y}\|$$

d) Supondo que se tenha, por exemplo, $\vec{x} = \xi\vec{y}$, vemos que

$$|\vec{x}\cdot\vec{y}| = |\xi\vec{y}\cdot\vec{y}| = |\xi|\ |\vec{y}\cdot\vec{y}| = |\xi|\cdot\|\vec{y}\|^2$$
$$\|\vec{x}\|\ \|\vec{y}\| = \|\xi\vec{y}\|\ \|\vec{y}\| = |\xi|\ \|\vec{y}\|\ \|\vec{y}\| = |\xi|\ \|\vec{y}\|$$

Reciprocamente, suponhamos que se dá a igualdade na desigualdade de Schwarz com vectores não nulos (já que os vectores nulos acarretam automaticamente a dependência linear). Seguindo os cálculos efectuados acima, isso significa que

$$\left\|-\frac{\vec{y}\cdot\vec{x}}{\|\vec{x}\|^2}\vec{x} + \vec{y}\right\|^2 = -\frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|^2}(\vec{y}\cdot\vec{x}) + \|\vec{y}\|^2 = 0$$

pelo que

$$-\frac{\vec{y}\cdot\vec{x}}{\|\vec{x}\|^2}\vec{x} + \vec{y} = \vec{0} \quad \text{ou seja,} \quad \vec{y} = \frac{\vec{y}\cdot\vec{x}}{\|\vec{x}\|^2}\vec{x}$$

e) Usaremos a desigualdade de Schwarz para demonstrar a desigualdade triangular:

$$\begin{aligned}\|\vec{x}+\vec{y}\|^2 &= (\vec{x}+\vec{y})\bullet(\vec{x}+\vec{y}) = \|\vec{x}\|^2 + \vec{x}\bullet\vec{y} + \vec{y}\bullet\vec{x} + \|\vec{y}\|^2 = \\ &= \|\vec{x}\|^2 + \vec{x}\bullet\vec{y} + \overline{\vec{x}\bullet\vec{y}} + \|\vec{y}\|^2 = \|\vec{x}\|^2 + 2\,\mathrm{Re}(\vec{x}\bullet\vec{y}) + \|\vec{y}\|^2 \leq \\ &\leq \|\vec{x}\|^2 + 2|\vec{x}\bullet\vec{y}| + \|\vec{y}\|^2 \leq \|\vec{x}\|^2 + 2\|\vec{x}\|\ \|\vec{y}\| + \|\vec{y}\|^2 = \\ &= \left(\|\vec{x}\| + \|\vec{y}\|\right)^2\end{aligned}$$

De $\|\vec{x}+\vec{y}\|^2 \leq \left(\|\vec{x}\|+\|\vec{y}\|\right)^2$ e atendendo a que as bases são não negativas, resulta que $\|\vec{x}+\vec{y}\| \leq \|\vec{x}\|+\|\vec{y}\|$.

f) Para que seja $\|\vec{x}+\vec{y}\| = \|\vec{x}\|+\|\vec{y}\|$, é necessário e suficiente que

$$\mathrm{Re}(\vec{x}\bullet\vec{y}) = \|\vec{x}\|\ \|\vec{y}\|$$

Ora, uma vez que $\mathrm{Re}(\vec{x}\bullet\vec{y}) \leq |\vec{x}\bullet\vec{y}| \leq \|\vec{x}\|\ \|\vec{y}\|$, a igualdade anterior mostra que $\mathrm{Re}(\vec{x}\bullet\vec{y}) = |\vec{x}\bullet\vec{y}| = \|\vec{x}\|\ \|\vec{y}\| \geq 0$.

Pelo que se viu na propriedade e), os vectores $\vec{x}$ e $\vec{y}$ serão dependentes, e, pondo $\vec{x} = \xi\vec{y}$, será $\vec{x}\bullet\vec{y} = \xi\|\vec{y}\|^2$ e a igualdade $\mathrm{Re}(\vec{x}\bullet\vec{y}) = |\vec{x}\bullet\vec{y}|$ mostra que $\vec{x}\bullet\vec{y}$ é um número real não negativo. Logo, será também $\xi \geq 0$.

Reciprocamente, sendo $\vec{x} = \xi\vec{y}$, com $\xi \geq 0$, tem-se

$$\vec{x}\bullet\vec{y} = \xi\|\vec{y}\|^2 = \mathrm{Re}(\vec{x}\bullet\vec{y}) = \xi\|\vec{y}\|\ \|\vec{y}\| = \|\vec{x}\|\ \|\vec{y}\|$$

o que termina a demonstração.

A desigualdade de Schwarz é muito importante, em diversas aplicações, e dela podemos retirar, em certos casos, um conceito importante. Com efeito, supondo que o espaço vectorial em causa é um espaço real, e tomando vectores $\vec{x}, \vec{y} \neq \vec{0}$, de

$$|\vec{x}\bullet\vec{y}| \leq \|\vec{x}\|\ \|\vec{y}\|$$

podemos deduzir (recordando que vectores não nulos têm normas diferentes de zero e que as suas normas serão números positivos) a desigualdade

$$\frac{|\vec{x}\bullet\vec{y}|}{\|\vec{x}\|\ \|\vec{y}\|} \leq 1 \quad \text{ou seja,} \quad -1 \leq \frac{\vec{x}\bullet\vec{y}}{\|\vec{x}\|\ \|\vec{y}\|} \leq 1$$

Ora, nestas condições e atendendo ao conhecido comportamento da função co-seno, podemos afirmar que existe um e um só valor $\theta \in [0,\pi]$ tal que

$$\cos\ \theta = \frac{\vec{x}\bullet\vec{y}}{\|\vec{x}\|\ \|\vec{y}\|}$$

Estes factos permitem-nos, então, introduzir a seguinte definição:

Definição 7-7 Seja E um espaço vectorial real, onde se definiu um produto interno. Dados dois vectores, $\vec{x},\vec{y} \neq \vec{0}$, chama-se **ângulo** de $\vec{x}$ com $\vec{y}$, ao valor $\theta \in [0,\pi]$ tal que

$$\cos\ \theta = \frac{\vec{x} \bullet \vec{y}}{\|\vec{x}\|\,\|\vec{y}\|}$$

Representando o ângulo de $\vec{x}$ com $\vec{y}$ por $\angle(\vec{x},\vec{y})$, podemos, portanto, escrever

$$\angle(\vec{x},\vec{y}) = \arccos\frac{\vec{x} \bullet \vec{y}}{\|\vec{x}\|\,\|\vec{y}\|}$$

Deve notar-se que, de acordo com esta definição, o conceito de ângulo só fica definido num espaço real, não num espaço complexo. Por outro lado, só dois vectores não nulos definem um ângulo: o vector nulo não forma ângulo com qualquer outro, nem consigo mesmo.

Podemos demonstrar algumas propriedades dos ângulos:

Teorema 7-8 Seja E um espaço vectorial real, com um produto interno fixo. São válidas as seguintes propriedades, para vectores não nulos arbitrários:

a) $\angle(\vec{x},\vec{x}) = 0$

b) $\angle(\vec{x},\vec{y}) = \angle(\vec{y},\vec{x})$

c) $\angle(\vec{x},\vec{y}) = \angle(\alpha\vec{x},\beta\vec{y})$, desde que α e β sejam números do mesmo sinal.

d) $\angle(\vec{x},\vec{y}) = \pi - \angle(\alpha\vec{x},\beta\vec{y})$, se α e β são números de sinais contrários.

Demonstração:

a) $\angle(\vec{x},\vec{x}) = \arccos\dfrac{\vec{x} \bullet \vec{x}}{\|\vec{x}\|\,\|\vec{x}\|} = \arccos\dfrac{\|\vec{x}\|^2}{\|\vec{x}\|^2} = \arccos 1 = 0$

b) $\angle(\vec{x},\vec{y}) = \arccos\dfrac{\vec{x} \bullet \vec{y}}{\|\vec{x}\|\,\|\vec{y}\|} = \arccos\dfrac{\vec{y} \bullet \vec{x}}{\|\vec{y}\|\,\|\vec{x}\|} = \angle(\vec{y},\vec{x})$

c) $\angle(\alpha\vec{x},\beta\vec{y}) = \arccos\dfrac{(\alpha\vec{x}) \bullet (\beta\vec{y})}{\|\alpha\vec{x}\|\,\|\beta\vec{y}\|} = \arccos\dfrac{\alpha\beta(\vec{x} \bullet \vec{y})}{|\alpha|\,\|\vec{x}\|\,|\beta|\,\|\vec{y}\|} =$

$$= \arccos\frac{\alpha\beta(\vec{x} \bullet \vec{y})}{\alpha\|\vec{x}\|\,\beta\|\vec{y}\|} = \arccos\frac{\vec{x} \bullet \vec{y}}{\|\vec{x}\|\,\|\vec{y}\|} = \angle(\vec{x},\vec{y})$$

d) $$\angle(\alpha\vec{x},\beta\vec{y}) = \arccos\frac{(\alpha\ \vec{x})\bullet(\beta\ \vec{y})}{\|\alpha\ \ \vec{x}\|\,\|\beta\ \ \vec{y}\|} = \arccos\frac{\alpha\beta(\vec{x}\bullet\vec{y})}{|\alpha|\,\|\vec{x}\|\,|\beta|\,\|\vec{y}\|} =$$
$$= \arccos\frac{\alpha\beta(\vec{x}\bullet\vec{y})}{-\alpha\beta\|\vec{x}\|\,\|\vec{y}\|} = \arccos\left(-\frac{\vec{x}\bullet\vec{y}}{\|\vec{x}\|\,\|\vec{y}\|}\right) =$$
$$= \pi - \arccos\frac{\vec{x}\bullet\vec{y}}{\|\vec{x}\|\,\|\vec{y}\|} = \pi - \angle(\vec{x},\vec{y})$$

Da definição de ângulo entre dois vectores resulta imediatamente que, para vectores não nulos num espaço real, se tem $\angle(\vec{x},\vec{y}) = \pi/2$ se e só se $\vec{x}\bullet\vec{y} = 0$. Este facto sugere que se introduza uma definição mais geral:

Definição 7-9 Seja E um espaço vectorial, real ou complexo, com um produto interno. Diz-se que um vector $\vec{x}$ é **ortogonal** a um vector $\vec{y}$ quando $\vec{x}\bullet\vec{y} = 0$. Representa-se esse facto por $\vec{x}\perp\vec{y}$.

Observe-se que o conceito de ortogonalidade fica, assim, definido tanto num espaço real como num espaço complexo, e para vectores arbitrários, incluindo o vector nulo. Podemos demonstrar o seguinte:

Teorema 7-10 Seja E um espaço vectorial, real ou complexo, com um produto interno. São válidas as seguintes propriedades, para vectores arbitrários:

a) Se $\vec{x}\perp\vec{y}$, então $\vec{y}\perp\vec{x}$.
b) $\vec{0}\perp\vec{x}$.
c) $\vec{x}\perp\vec{x}$ se e só se $\vec{x} = \vec{0}$.
d) Se $\vec{x}\perp\vec{y}$, então $\vec{x}\perp\lambda\vec{y}$, qualquer que seja o escalar λ.

Demonstração:
Qualquer das propriedades resulta trivialmente das definições.

O conceito de ortogonalidade reveste-se de grande importância. Para podermos enunciar mais facilmente algumas proposições, introduzimos ainda a seguinte nomenclatura:

Definição 7-11 Seja E um espaço vectorial, real ou complexo, com um produto interno. Diz-se que os vectores $\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k$ formam um **sistema ortogonal** quando cada um dos vectores considerados é ortogonal a cada um dos outros, ou seja, quando $\vec{v}_i\bullet\vec{v}_j = 0$, sempre que $i \neq j$.

Um vector $\vec{u}$ diz-se **unitário** ou **normado** quando $\|\vec{u}\| = 1$. Os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ formam um **sistema ortonormado** quando se trata de um sistema ortogonal e formado por vectores unitários. Assim, o sistema de vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ é ortonormado quando, para quaisquer $i, j = 1, 2, ..., k$, se tem

$$\vec{v}_i \bullet \vec{v}_j = \begin{cases} 0, & \text{se } i \neq j \\ 1, & \text{se } i = j \end{cases}$$

Podemos relacionar a ortogonalidade com a noção de independência linear, através do seguinte resultado:

Teorema 7-12 Todo o sistema ortogonal formado por vectores não nulos é linearmente independente. Em particular, todo o sistema ortonormado é linearmente independente.

Demonstração:
Seja $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ um sistema ortogonal formado por vectores não nulos, e consideremos a igualdade

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k = \vec{0}$$

Para um índice arbitrário $i = 1, 2, ..., k$, teremos, então,

$$(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k) \bullet \vec{v}_i = \vec{0} \bullet \vec{v}_i = 0$$

ou seja,

$$\alpha_1(\vec{v}_1 \bullet \vec{v}_i) + \alpha_2(\vec{v}_2 \bullet \vec{v}_i) + \cdots + \alpha_k(\vec{v}_k \bullet \vec{v}_i) = 0$$

Ora, uma vez que os vectores considerados são ortogonais dois a dois, o produto interno $\vec{v}_j \bullet \vec{v}_i$ é nulo sempre que $j \neq i$, pelo que a igualdade anterior se reduz a

$$\alpha_i(\vec{v}_i \bullet \vec{v}_i) = 0$$

Sendo $\vec{v}_i \neq \vec{0}$, por hipótese, também $\vec{v}_i \bullet \vec{v}_i \neq 0$, e podemos então concluir que $\alpha_i = 0$. Como i era arbitrário, isto quer dizer que todos os coeficientes $\alpha_1, \alpha_2, ..., \alpha_k$ são iguais a zero e os vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ são linearmente independentes.

7-2 Produtos internos em espaços de dimensão finita

No caso particular de se trabalhar com um espaço de dimensão finita, o cálculo do produto interno de dois vectores fica muito simplificado. Para nos referirmos mais facilmente a este tipo de espaços, introduzimos, desde já, as seguintes denominações:

Definição 7-13 Um espaço vectorial real (resp.: complexo), com um produto interno fixo e de dimensão finita, toma o nome de **espaço euclidiano** (resp.: **espaço unitário**).

Seja então E um espaço euclidiano ou unitário, e fixemos em E uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$.

Dados dois vectores arbitrários $\vec{x}, \vec{y} \in E$, vamos procurar calcular o seu produto interno. Tendo em conta que qualquer vector de E se pode escrever como combinação linear dos vectores da base, e atendendo também às propriedades do produto interno, vemos que:

$$\vec{x} \bullet \vec{y} = \left(\sum_{i=1}^{n} x_i \vec{e}_i\right) \bullet \vec{y} = \sum_{i=1}^{n} x_i \left(\vec{e}_i \bullet \vec{y}\right) = \sum_{i=1}^{n} x_i \left[\vec{e}_i \bullet \left(\sum_{j=1}^{n} y_j \vec{e}_j\right)\right] =$$

$$= \sum_{i=1}^{n} x_i \left[\sum_{j=1}^{n} \overline{y_j} \left(\vec{e}_i \bullet \vec{e}_j\right)\right] = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \overline{y_j} \left(\vec{e}_i \bullet \vec{e}_j\right) = \sum_{i,j=1}^{n} \left(\vec{e}_i \bullet \vec{e}_j\right) x_i \overline{y_j}$$

Designando os escalares $\vec{e}_i \bullet \vec{e}_j$ ($i, j = 1, 2, \ldots, n$) por $\vec{e}_i \bullet \vec{e}_j = g_{ij}$, a igualdade anterior reveste a forma

$$\vec{x} \bullet \vec{y} = \sum_{i,j=1}^{n} g_{ij} x_i \overline{y_j}$$

Esta expressão é polinomial; portanto, sempre fácil de calcular. Mais ainda, se considerarmos a matriz quadrada $G = [\, g_{ij} \,]$ e as colunas

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

constituídas pelas componentes de $\vec{x}$ e $\vec{y}$ em relação à base considerada, a igualdade anterior é equivalente a

$$\vec{x} \bullet \vec{y} = X^T G \overline{Y}$$

onde $\overline{Y}$ designa, naturalmente, a coluna formada pelos conjugados dos elementos de Y.

No caso real, as igualdades anteriores simplificam-se ainda mais, visto o conjugado de um número real coincidir com o próprio número:

$$\vec{x} \bullet \vec{y} = \sum_{i,j=1}^{n} g_{ij} x_i y_j = X^T G Y$$

A matriz quadrada G, formada pelos produtos internos dos vectores da base escolhida, uns pelos outros, desempenha, portanto, um papel muito importante no cálculo do produto interno, pelo que merece uma designação especial:

Definição 7-14 | Nas condições anteriores, a matriz $G = [\vec{e}_i \bullet \vec{e}_j]_{i,j=1,2,\ldots,n}$ toma o nome de **matriz da métrica** do produto interno definido no espaço em relação à base considerada.

Note-se que os elementos da matriz da métrica são precisamente os coeficientes da expressão polinomial que acima encontrámos para o produto interno.

Assim, por exemplo, consideremos, no espaço vectorial real $\mathbb{R}^3$, o produto interno definido por

$$\vec{x} \bullet \vec{y} = 4x_1y_1 - x_1y_2 + x_1y_3 - x_2y_1 + 3x_2y_2 + 2x_2y_3 + x_3y_1 + 2x_3y_2 + 6x_3y_3$$

Se fixarmos no espaço a base canónica, cada vector $\vec{u} = (u_1, u_2, u_3)$ terá justamente as componentes (u_1, u_2, u_3), pelo que a expressão anterior se pode considerar escrita em relação à base canónica. A respectiva matriz da métrica será, portanto, formada pelos coeficientes da expressão:

$$G = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 2 \\ 1 & 2 & 6 \end{bmatrix}$$

Se pretendermos, por exemplo, calcular o produto interno do vector $\vec{x} = (3,2,1)$ pelo vector $\vec{y} = (-1,3,1)$, deveremos simplesmente calcular o produto matricial

$$\vec{x} \bullet \vec{y} = \begin{bmatrix} 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 2 \\ 1 & 2 & 6 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} -6 \\ 12 \\ 11 \end{bmatrix} = 17$$

Naturalmente, a matriz da métrica depende não só do produto interno considerado, mas também da base que escolhemos. Podemos, sem dificuldade, relacionar a matriz da métrica de um produto interno em relação a certa base, com a matriz da métrica do mesmo produto interno em relação a outra base:

Teorema 7-15

Seja E um espaço euclidiano ou unitário. Seja G a matriz da métrica em relação a uma base, $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$, de E, e G' a matriz da métrica em relação a uma outra base, $\vec{e}'_1, \vec{e}'_2, ..., \vec{e}'_n$. Seja P a matriz de mudança de base, isto é, a matriz cujas colunas são formadas pelas componentes dos vectores $\vec{e}'_1, \vec{e}'_2, ..., \vec{e}'_n$ em relação à base $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$. Então, tem-se $G' = P^T G \overline{P}$ (ou, no caso real, $G' = P^T G P$).

Demonstração:

Os elementos de G' obtêm-se calculando $\vec{e}'_i \bullet \vec{e}'_j$ (para todos os valores de i e j). Ora, as componentes de $\vec{e}'_i$ e de $\vec{e}'_j$ em relação à base $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$ encontram-se nas colunas i e j da matriz P, respectivamente.

Portanto, para calcular $\vec{e}'_i \bullet \vec{e}'_j$, usaremos essas colunas de P e a matriz da métrica G, fazendo o produto

$$\begin{bmatrix} p_{1i} & p_{2i} & \cdots & p_{ni} \end{bmatrix} G \overline{\begin{bmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{nj} \end{bmatrix}}$$

Como os índices i e j variam de 1 a n, estes cálculos correspondem precisamente a efectuarmos o produto de matrizes $P^T G \overline{P}$.

Consideremos agora, num espaço vectorial finitamente gerado, um sistema ortonormado de vectores. Pelo facto de serem independentes, os vectores desse sistema podem fazer parte de uma base do espaço. Mais do que isso, podemos mesmo construir uma base que seja, ela própria, um sistema ortonormado. Vamos ver como, começando por apresentar uma definição e um lema:

Definição 7-16

Seja $\vec{u} \neq \vec{0}$ um vector de um espaço vectorial, real ou complexo, com um produto interno. Chama-se **versor** de $\vec{u}$ ao vector

$$\frac{1}{\|\vec{u}\|}\vec{u} = \frac{\vec{u}}{\|\vec{u}\|}$$

Lema 7-17

O versor de um vector não nulo, num espaço vectorial, real ou complexo, com um produto interno, é sempre um vector unitário.

Demonstração:

Aplicando as propriedades já provadas acerca da norma, vemos que

$$\left\|\frac{\vec{u}}{\|\vec{u}\|}\right\| = \left\|\frac{1}{\|\vec{u}\|}\,\vec{u}\right\| = \left|\frac{1}{\|\vec{u}\|}\right| \|\vec{u}\| = \frac{1}{\|\vec{u}\|}\,\|\vec{u}\| = 1$$

Teorema 7-18 Seja E um espaço euclidiano ou unitário não nulo. Então, existem em E bases ortonormadas. Se E é o espaço nulo, podemos, por convenção, considerar que a sua base, que é o conjunto vazio, é igualmente ortonormada.

Demonstração:
Faremos a demonstração de um modo construtivo, descrevendo um processo através do qual, partindo de uma base qualquer do espaço em causa, se pode construir uma base ortonormada. O método que vamos apresentar é chamado **método de Gram-Schmidt**.

Seja então E um espaço euclidiano ou unitário, de dimensão n, e fixemos uma base qualquer, $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$.

Tomemos o versor de

$$\vec{e}_1 : \ \vec{w}_1 = \frac{\vec{e}_1}{\|\vec{e}_1\|}$$

Consideremos agora o vector auxiliar $\vec{z}_2 = \vec{e}_2 - (\vec{e}_2 \bullet \vec{w}_1)\,\vec{w}_1$.

Este vector $\vec{z}_2$ não é nulo, porque de $\vec{z}_2 = \vec{0}$ resultaria $\vec{e}_2 - (\vec{e}_2 \bullet \vec{w}_1)\,\vec{w}_1 = \vec{0}$, ou seja,

$$\vec{e}_2 = (\vec{e}_2 \bullet \vec{w}_1)\,\vec{w}_1 = \frac{\vec{e}_2 \bullet \vec{w}_1}{\|\vec{e}_1\|}\vec{e}_1$$

o que é impossível, visto que $\vec{e}_1$ e $\vec{e}_2$ são vectores linearmente independentes, por fazerem parte de uma base de E.

Por outro lado, tem-se $\vec{z}_2 \perp \vec{w}_1$, uma vez que, sendo $\vec{w}_1$ um vector unitário, será

$$\vec{z}_2 \bullet \vec{w}_1 = \left(\vec{e}_2 - (\vec{e}_2 \bullet \vec{w}_1)\,\vec{w}_1\right) \bullet \vec{w}_1 = \vec{e}_2 \bullet \vec{w}_1 - (\vec{e}_2 \bullet \vec{w}_1)\,\vec{w}_1 \bullet \vec{w}_1 = \vec{e}_2 \bullet \vec{w}_1 - (\vec{e}_2 \bullet \vec{w}_1)1 = 0$$

Tomando, então o versor $\vec{w}_2 = \vec{z}_2 / \|\vec{z}_2\|$, os vectores $\vec{w}_1, \vec{w}_2$ constituem um sistema ortonormado.

Consideremos, em seguida, o vector auxiliar

$$\vec{z}_3 = \vec{e}_3 - (\vec{e}_3 \bullet \vec{w}_1)\,\vec{w}_1 - (\vec{e}_3 \bullet \vec{w}_2)\,\vec{w}_2$$

Uma vez que $\vec{w}_1$ e $\vec{w}_2$ se escrevem à custa de $\vec{e}_1$ e $\vec{e}_2$, o vector $\vec{z}_3$ não pode ser nulo, porque, sendo

$$\vec{e}_3 - (\vec{e}_3 \bullet \vec{w}_1)\,\vec{w}_1 - (\vec{e}_3 \bullet \vec{w}_2)\,\vec{w}_2 = \vec{0}$$

teríamos

$$\vec{e}_3 = (\vec{e}_3 \bullet \vec{w}_1)\,\vec{w}_1 + (\vec{e}_3 \bullet \vec{w}_2)\,\vec{w}_2$$

ficando, em última análise, $\vec{e}_3$ a ser combinação linear de $\vec{e}_1$ e $\vec{e}_2$, o que é impossível.

Ao mesmo tempo, $\vec{z}_3$ é ortogonal a $\vec{w}_1$ e $\vec{w}_2$, já que, por exemplo,

$$\begin{aligned}\vec{z}_3 \bullet \vec{w}_1 &= \left(\vec{e}_3 - (\vec{e}_3 \bullet \vec{w}_1)\,\vec{w}_1 - (\vec{e}_3 \bullet \vec{w}_2)\,\vec{w}_2\right) \bullet \vec{w}_1 = \\ &= \vec{e}_3 \bullet \vec{w}_1 - (\vec{e}_3 \bullet \vec{w}_1)\,\vec{w}_1 \bullet \vec{w}_1 - (\vec{e}_3 \bullet \vec{w}_2)\,\vec{w}_2 \bullet \vec{w}_1 = \\ &= \vec{e}_3 \bullet \vec{w}_1 - (\vec{e}_3 \bullet \vec{w}_1)1 - (\vec{e}_3 \bullet \vec{w}_2)0 = 0\end{aligned}$$

Consequentemente, se tomarmos o versor $\vec{w}_3 = \vec{z}_3 / \|\vec{z}_3\|$, verificamos que os vectores $\vec{w}_1, \vec{w}_2, \vec{w}_3$ formam um sistema ortonormado.

É fácil mostrar que, considerando

$$\vec{z}_k = \vec{e}_k - \sum_{i=1}^{k-1} (\vec{e}_k \bullet \vec{w}_i)\,\vec{w}_i \quad (k = 2,3,\ldots,n)$$

e os respectivos versores

$$\vec{w}_k = \frac{\vec{z}_k}{\|\vec{z}_k\|} \quad (k = 2,3,\ldots,n)$$

os vectores $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_n$ formam um sistema ortonormado. Como um sistema ortonormado é linearmente independente e a dimensão do espaço E é n, estes vectores $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_n$ formam uma base ortonormada de E.

As bases ortonormadas revestem-se de especial importância pelo facto de a matriz da métrica em relação a uma base ortonormada ser particularmente simples.

Com efeito, se $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_n$ é uma base ortonormada, tem-se

$$\vec{w}_i \bullet \vec{w}_j = \begin{cases} 0 & , \quad \text{se } j \neq i \\ \|\vec{w}_i\|^2 = 1, & \quad \text{se } j = i \end{cases}$$

pelo que a matriz da métrica, em relação a qualquer base ortonormada, é I_n.

Daí resulta que, tendo os vectores $\vec{x}$ e $\vec{y}$ as componentes

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

respectivamente, em relação à base ortonormada considerada, será

$$\vec{x} \bullet \vec{y} = X^T I_n \overline{Y} = X^T \overline{Y} = x_1 \overline{y_1} + x_2 \overline{y_2} + \cdots + x_n \overline{y_n}$$

ou, num espaço real,

$$\vec{x} \bullet \vec{y} = X^T I_n Y = X^T Y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Trata-se, portanto, da expressão do produto interno canónico no espaço vectorial real $\mathbb{C}^n$ ou $\mathbb{R}^n$, respectivamente.

Em particular,

$$\|\vec{x}\| = \sqrt{\vec{x} \bullet \vec{x}} = \sqrt{|x_1|^2 + |x_2|^2 + \cdots |x_n|^2}$$

ou, no caso real,

$$\|\vec{x}\| = \sqrt{\vec{x} \bullet \vec{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

Para exemplificarmos a utilização do método de Gram-Schmidt, consideremos, no espaço vectorial real $\mathbb{R}^3$, o produto interno definido por

$$\vec{x} \bullet \vec{y} = 4x_1y_1 - x_1y_2 + x_1y_3 - x_2y_1 + 3x_2y_2 + 2x_2y_3 + x_3y_1 + 2x_3y_2 + 6x_3y_3$$

Conforme vimos, a matriz da métrica deste produto interno em relação à base canónica $\vec{e}_1, \vec{e}_2, \vec{e}_3$ é

$$G = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 2 \\ 1 & 2 & 6 \end{bmatrix}$$

O facto de esta matriz não ser a matriz identidade revela imediatamente que a base canónica não é ortonormada para este produto interno. Vamos, pois, usar o método de Gram-Schmidt para construir uma base ortonormada a partir da base canónica.

Começamos por observar que

$$\|\vec{e}_1\| = \sqrt{\vec{e}_1 \bullet \vec{e}_1} = \sqrt{g_{11}} = \sqrt{4} = 2$$

Tomamos, portanto, como primeiro vector da base ortonormada a construir, o versor

$$\vec{w}_1 = \frac{\vec{e}_1}{\|\vec{e}_1\|} = \frac{(1,0,0)}{2} = \left(\frac{1}{2}, 0, 0\right)$$

Calculamos, agora, o vector auxiliar

$$\vec{z}_2=\vec{e}_2-(\vec{e}_2\bullet\vec{w}_1)\,\vec{w}_1=(0,1,0)-\begin{bmatrix}0&1&0\end{bmatrix}\begin{bmatrix}4&-1&1\\-1&3&2\\1&2&6\end{bmatrix}\begin{bmatrix}1/2\\0\\0\end{bmatrix}\left(\frac{1}{2},0,0\right)=$$

$$=(0,1,0)-\frac{1}{4}\begin{bmatrix}-1&3&2\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix}(1,0,0)=(0,1,0)+\frac{1}{4}(1,0,0)=\left(\frac{1}{4},1,0\right)$$

Por sua vez,

$$\|\vec{z}_2\|=\sqrt{\begin{bmatrix}\frac{1}{4}&1&0\end{bmatrix}\begin{bmatrix}4&-1&1\\-1&3&2\\1&2&6\end{bmatrix}\begin{bmatrix}1/4\\1\\0\end{bmatrix}}=\frac{\sqrt{11}}{2}$$

pelo que o segundo vector da base ortonormada será

$$\vec{w}_2=\frac{\vec{z}_2}{\|\vec{z}_2\|}=\frac{\left(\frac{1}{4},1,0\right)}{\sqrt{11}/2}=\left(\frac{1}{2\sqrt{11}},\frac{2}{\sqrt{11}},0\right)$$

Para obtermos o terceiro e último vector da base procurada, começamos por construir

$$\vec{z}_3=\vec{e}_3-(\vec{e}_3\bullet\vec{w}_1)\,\vec{w}_1-(\vec{e}_3\bullet\vec{w}_2)\,\vec{w}_2=$$

$$=(0,0,1)-\begin{bmatrix}0&0&1\end{bmatrix}\begin{bmatrix}4&-1&1\\-1&3&2\\1&2&6\end{bmatrix}\begin{bmatrix}1/2\\0\\0\end{bmatrix}\left(\frac{1}{2},0,0\right)-\begin{bmatrix}0&0&1\end{bmatrix}\begin{bmatrix}4&-1&1\\-1&3&2\\1&2&6\end{bmatrix}\begin{bmatrix}1/(2\sqrt{11})\\2/\sqrt{11}\\0\end{bmatrix}\left(\frac{1}{2\sqrt{11}},\frac{2}{\sqrt{11}},0\right)=$$

$$(0,0,1)-\frac{1}{4}\begin{bmatrix}1&2&6\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix}(1,0,0)-\frac{4}{11}\begin{bmatrix}1&2&6\end{bmatrix}\begin{bmatrix}1/4\\1\\0\end{bmatrix}\left(\frac{1}{4},1,0\right)=\left(-\frac{5}{11},\frac{9}{11},1\right)$$

Donde

$$\|\vec{z}_3\|=\sqrt{\begin{bmatrix}-\frac{5}{11}&\frac{9}{11}&1\end{bmatrix}\begin{bmatrix}4&-1&1\\-1&3&2\\1&2&6\end{bmatrix}\begin{bmatrix}-5/11\\9/11\\1\end{bmatrix}}=\frac{17\sqrt{5}}{11}$$

e, portanto,

$$\vec{u}_3 = \frac{\vec{z}_3}{\|\vec{z}_3\|} = \frac{\left(-\frac{5}{11}, \frac{9}{11}, 1\right)}{\frac{17\sqrt{5}}{11}} = \left(-\frac{\sqrt{5}}{17}, \frac{9}{17\sqrt{5}}, \frac{11}{17\sqrt{5}}\right)$$

Na prática, podemos procurar uma base ortonormada de um modo mais simples, por tentativas:

Fixemos, por exemplo, o vector $\vec{e}_1 = (1,0,0)$ e procuremos um vector ortogonal a este. Se pensarmos num vector genérico (a,b,c), ele será ortogonal a $(1,0,0)$ desde que

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 2 \\ 1 & 2 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 0$$

isto é, desde que

$$4a - b + c = 0 \quad \text{ou seja,} \quad b = 4a + c$$

Escolhendo, por exemplo, $a = 0$ e $c = 1$, encontramos o vector $\vec{v}_2 = (0,1,1)$. Os vectores $\vec{e}_1, \vec{v}_2$ formam um sistema ortogonal.

Procuremos, agora, um terceiro vector ortogonal a ambos, para o que basta resolver o sistema de equações

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 2 \\ 1 & 2 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix} \quad \text{ou seja,} \quad \begin{cases} 4a - b + c = 0 \\ 5b + 8c = 0 \end{cases}$$

Este sistema de equações conduz-nos a

$$\begin{cases} a = -\dfrac{13}{20}c \\ b = -\dfrac{8}{5}c \end{cases}$$

pelo que, com $c = -20$, obtemos o vector $\vec{v}_3 = (13, 32, -20)$.

Os vectores $\vec{e}_1, \vec{v}_2, \vec{v}_3$ formam, portanto, uma base ortogonal do espaço. Para se chegar a uma base ortonormada, bastaria dividir cada um pela respectiva norma.

Dos resultados anteriores podemos ainda inferir o seguinte:

Corolário 7-19 Num espaço euclidiano ou unitário, a matriz da métrica em relação a uma base qualquer é uma matriz invertível.

Demonstração:
Seja G a matriz da métrica em relação a uma base qualquer $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$ do espaço vectorial E em causa, e seja $\vec{u}_1, \vec{u}_2, ..., \vec{u}_n$ uma base ortonormada de E. A matriz da métrica em relação à base $\vec{u}_1, \vec{u}_2, ..., \vec{u}_n$ será I_n, e existirá uma matriz de mudança de base Q, tal que $G = Q^T I_n Q = Q^T Q$.

Ora, a matriz de mudança de base é invertível, pelo que também G é invertível, tendo-se $G^{-1} = \left(Q^T Q\right)^{-1} = Q^{-1}\left(Q^{-1}\right)^T$.

7-3 Complemento ortogonal

Dado um espaço vectorial, real ou complexo, E, com um produto interno fixo, e tomando um conjunto qualquer, $X \subseteq E$, não vazio, podemos considerar o conjunto dos vectores ortogonais a todos os vectores de X, ou seja, o conjunto

$$X^{\perp} = \left\{\vec{v} \in E : \vec{v} \perp \vec{x},\ \forall \vec{x} \in X\right\}$$

Provamos então o seguinte teorema:

Teorema 7-20 Sendo E um espaço vectorial real, ou complexo, e $X \subseteq E$, um conjunto não vazio, o conjunto $X^{\perp}$ é um subespaço vectorial de E.

Demonstração:
Sabemos que o vector $\vec{0}$ é ortogonal a qualquer vector do espaço; logo, em particular, a qualquer vector do conjunto X. Então, $\vec{0} \in X^{\perp}$ e $X^{\perp} \neq \varnothing$.

Por outro lado, tomemos vectores $\vec{a}, \vec{b} \in X^{\perp}$ e escalares quaisquer α e β, e vamos ver se se tem $\alpha\vec{a} + \beta\vec{b} \in X^{\perp}$.

Para isso, tomemos um vector arbitrário $\vec{x} \in X$. Vemos que

$$\left(\alpha\vec{a} + \beta\vec{b}\right) \bullet \vec{x} = \alpha\left(\vec{a} \bullet \vec{x}\right) + \beta\left(\vec{b} \bullet \vec{x}\right)$$

Ora, sendo $\vec{a},\vec{b} \in X^{\perp}$ e $\vec{x} \in X$, teremos, por definição, $\vec{a} \bullet \vec{x} = 0$ e $\vec{b} \bullet \vec{x} = 0$, pelo que

$$\left(\alpha\vec{a} + \beta\vec{b}\right) \bullet \vec{x} = \alpha(\vec{a} \bullet \vec{x}) + \beta\left(\vec{b} \bullet \vec{x}\right) = \alpha 0 + \beta 0 = 0$$

Portanto, $\left(\alpha\vec{a} + \beta\vec{b}\right) \perp \vec{x}$, e, dada a arbitrariedade de $\vec{x}$, concluímos que $\alpha\vec{a} + \beta\vec{b} \in X^{\perp}$.

Estes subespaços, assim construídos, gozam das seguintes propriedades:

Teorema 7-21 Seja E um espaço vectorial, real ou complexo, com um produto interno. Para subconjuntos arbitrários (não vazios) de E, tem-se:

a) Se $X \subseteq Y$, então $Y^{\perp} \subseteq X^{\perp}$.

b) $X \subseteq \left(X^{\perp}\right)^{\perp}$.

c) $\langle X \rangle^{\perp} = X^{\perp}$; em particular, $\langle \vec{v}_1, \vec{v}_2, ..., \vec{v}_k \rangle^{\perp} = \{\vec{v}_1, \vec{v}_2, ..., \vec{v}_k\}^{\perp}$.

d) Se $X \cap X^{\perp} \neq \varnothing$, então $X \cap X^{\perp} = \left\{\vec{0}\right\}$.

Demonstração:

a) Suponhamos que se tem $X \subseteq Y$.

Tomando $\vec{u} \in Y^{\perp}$, teremos $\vec{u} \perp \vec{x}$, qualquer que seja o vector $\vec{x} \in Y$. Como $X \subseteq Y$, os vectores de X estarão também em Y, pelo que podemos afirmar que se tem $\vec{u} \perp \vec{x}$, qualquer que seja o vector $\vec{x} \in X$, ou seja, que $\vec{u} \in X^{\perp}$.

b) Resulta trivialmente da definição.

c) Atendendo a que $X \subseteq \langle X \rangle$, resulta da primeira propriedade que se terá $\langle X \rangle^{\perp} \subseteq X^{\perp}$.

Por outro lado, tomemos um vector $\vec{u} \in X^{\perp}$ e um vector qualquer $\vec{z} \in \langle X \rangle$. O vector $\vec{z}$ será combinação linear de alguns vectores de X:

$$\vec{z} = \lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \cdots + \lambda_t\vec{x}_t \text{, com } \vec{x}_1, \vec{x}_2, ..., \vec{x}_t \in X$$

Atendendo a que $\vec{u} \in X^{\perp}$ e $\vec{x}_1, \vec{x}_2, ..., \vec{x}_t \in X$, sabemos que $\vec{x}_i \bullet \vec{u} = 0$ (para $i = 1, 2, ..., t$), pelo que

$$\vec{z} \bullet \vec{u} = (\lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \cdots + \lambda_t\vec{x}_t) \bullet \vec{u} = \lambda_1(\vec{x}_1 \bullet \vec{u}) + \lambda_2(\vec{x}_2 \bullet \vec{u}) + \cdots + \lambda_t(\vec{x}_t \bullet \vec{u}) = 0$$

Assim, $\vec{u}$ é ortogonal a qualquer vector de $\langle X \rangle$, ou seja, $\vec{u} \in \langle X \rangle^{\perp}$.

d) Supondo que seja $X \cap X^{\perp} \neq \varnothing$, tomemos um vector $\vec{w} \in X \cap X^{\perp}$: o vector $\vec{w}$ é ortogonal a qualquer vector de X; logo, em particular, ortogonal a si mesmo, o que obriga a que seja $\vec{w} = \vec{0}$. Portanto, $X \cap X^{\perp} = \left\{\vec{0}\right\}$.

Estas construções têm particular interesse quando aplicadas a subespaços vectoriais de E, muito em especial se este tem dimensão finita. Com efeito, reparemos no seguinte:

Teorema 7-22 Seja E um espaço euclidiano ou unitário, e seja F um subespaço vectorial de E. Tem-se, neste caso:

a) $E = F \oplus F^{\perp}$

b) $\left(F^{\perp}\right)^{\perp} = F$

Demonstração:

a) De acordo com o que se viu no teorema anterior, sabemos que F e $F^{\perp}$ são subespaços de E, pelo que $\vec{0} \in F \cap F^{\perp}$; consequentemente, $F \cap F^{\perp} \neq \varnothing$, pelo que $F \cap F^{\perp} = \left\{\vec{0}\right\}$. Assim, a soma $F + F^{\perp}$ é uma soma directa.

Para vermos que esta soma coincide com E, fixemos em F uma base ortonormada $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k$. Dado um vector arbitrário $\vec{x} \in E$, podemos então escrever:

$$\vec{x} = \sum_{i=1}^{k} \left(\vec{x} \bullet \vec{f}_i\right) \vec{f}_i + \left(\vec{x} - \sum_{i=1}^{k} \left(\vec{x} \bullet \vec{f}_i\right) \vec{f}_i\right)$$

É claro que

$$\sum_{i=1}^{k} \left(\vec{x} \bullet \vec{f}_i\right) \vec{f}_i \in F$$

Se provarmos que

$$\vec{x} - \sum_{i=1}^{k} \left(\vec{x} \bullet \vec{f}_i\right) \vec{f}_i \in F^{\perp}$$

a demonstração ficará concluída. Para isso, atendendo à terceira propriedade enunciada no teorema anterior, basta ver que o vector

$$\vec{x} - \sum_{i=1}^{k} \left(\vec{x} \bullet \vec{f}_i\right) \vec{f}_i$$

é ortogonal aos vectores da base fixa em F. Ora, tomando um desses vectores, $\vec{f}_j$ $(j = 1, 2, \ldots, k)$, temos

$$\left(\vec{x} - \sum_{i=1}^{k} \left(\vec{x} \bullet \vec{f}_i\right) \vec{f}_i\right) \bullet \vec{f}_j = \vec{x} \bullet \vec{f}_j - \sum_{i=1}^{k} \left(\vec{x} \bullet \vec{f}_i\right)\left(\vec{f}_i \bullet \vec{f}_j\right)$$

Recordando que a base considerada é ortonormada, sabemos que

$$\vec{f}_i \bullet \vec{f}_j = \begin{cases} 0, & \text{se} \quad i \neq j \\ 1, & \text{se} \quad i = j \end{cases}$$

pelo que

$$\left(\vec{x} - \sum_{i=1}^{k}\left(\vec{x} \bullet \vec{f}_i\right)\vec{f}_i\right) \bullet \vec{f}_j = \vec{x} \bullet \vec{f}_j - \sum_{i=1}^{k}\left(\vec{x} \bullet \vec{f}_i\right)\left(\vec{f}_i \bullet \vec{f}_j\right) = \vec{x} \bullet \vec{f}_j - \vec{x} \bullet \vec{f}_j = 0$$

b) A segunda propriedade do teorema anterior diz-nos que $F \subseteq \left(F^{\perp}\right)^{\perp}$. Ao mesmo tempo, o que acabamos de provar mostra que, por um lado,

$$E = F \oplus F^{\perp}$$

e, por outro lado,

$$E = F^{\perp} \oplus \left(F^{\perp}\right)^{\perp}$$

Por conseguinte:

$$\begin{aligned} \dim(F) + \dim\left(F^{\perp}\right) &= \dim\left(F \oplus F^{\perp}\right) = \dim(E) = \\ &= \dim\left(F^{\perp} \oplus \left(F^{\perp}\right)^{\perp}\right) = \dim\left(F^{\perp}\right) + \dim\left(\left(F^{\perp}\right)^{\perp}\right) \end{aligned}$$

Daqui resulta que $\dim(F) = \dim\left(\left(F^{\perp}\right)^{\perp}\right)$, pelo que, dada a inclusão entre os dois subespaços, concluímos que $\left(F^{\perp}\right)^{\perp} = F$.

É interessante notar que as propriedades expressas no teorema anterior não são válidas num espaço vectorial de dimensão infinita.

A fim de apresentarmos um contra-exemplo, consideremos o espaço vectorial real E, de dimensão infinita, constituído por todas as funções reais de variável real definidas e contínuas no intervalo $[0,1]$, com as operações usuais, e fixemos em E o produto interno definido, para duas funções quaisquer, por

$$f \bullet g = \int_0^1 f(x)g(x)dx$$

Seja $F = \{f \in F : f(0) = 0\}$. É muito fácil verificar que F é um subespaço vectorial de E.

Vamos procurar $F^{\perp}$:

Seja $h \in E$ uma função não nula. É claro que, se for $h(0)=0$, então $h \in F$, e, sendo $h \neq 0$, por certo $h \notin F^{\perp}$.

Suponhamos, então, que se tem $h(0) \neq 0$. A função $\left(h(x)\right)^2$ é sempre maior que ou igual a zero e não é nula, pelo que

$$\int_0^1 \left(h(x)\right)^2 dx = c > 0$$

Para cada $u \in]0,1]$, consideremos a função contínua h_u, definida, no intervalo $[0,1]$, da seguinte maneira:

$$h_u(x) = \begin{cases} \dfrac{h(u)}{u} x, & \text{se} \quad 0 \leq x \leq u \\ h(x), & \text{se} \quad u \leq x \leq 1 \end{cases}$$

Tem-se $h_u(0)=0$, pelo que $h_u \in F$.

Da definição de h_u, resulta que

$$h_u(x)\, h(x) = \begin{cases} \dfrac{h(u)}{u} x h(x), & \text{se} \quad 0 \leq x \leq u \\ \left(h(x)\right)^2, & \text{se} \quad u \leq x \leq 1 \end{cases}$$

Pondo $M = \max\limits_{x \in [0,1]} |h(x)|$, tem-se

$$\int_0^u \left|h_u(x)\, h(x)\right| dx = \int_0^u \left|\frac{h(u)}{u} x h(x)\right| dx = \frac{|h(u)|}{u} \int_0^u x |h(x)|\, dx \leq \frac{M}{u} \int_0^u x M dx = \frac{M^2}{u} \int_0^u x dx = \frac{1}{2} M^2 u$$

Ora, evidentemente,

$$\lim_{u \to 0} M^2 u = 0, \text{ pelo que } \lim_{u \to 0} \int_0^u h_u(x)\, h(x)\, dx = 0$$

Mas, por outro lado, de

$$\lim_{u \to 0^+} \int_u^1 \left(h(x)\right)^2 dx = \int_0^1 \left(h(x)\right)^2 dx = c$$

resulta que

$$\lim_{u\to 0}\int_0^1 h_u(x)h(x)\,dx = \lim_{u\to 0}\left(\int_0^u h_u(x)h(x)\,dx + \int_u^1 h_u(x)h(x)\,dx\right) =$$

$$= \lim_{u\to 0}\left(\int_0^u h_u(x)h(x)\,dx + \int_u^1 \left(h(x)\right)^2 dx\right) = 0 + c = c$$

Consequentemente, poderemos fixar

$$u \in \left]0,1\right], \text{ tal que } \int_0^1 h_u(x)h(x)\,dx > \frac{c}{2} > 0$$

Assim,

$$\int_0^1 h_u(x)h(x)\,dx \neq 0\text{, ou seja, } h_u \bullet h \neq 0$$

Como $h_u \in F$, concluímos daqui que $h \notin F^\perp$.
Portanto, seja qual for a função $h \in E$, não nula, tem-se $h \notin F^\perp$, pelo que $F^\perp = \{0\}$.
Logo,

$$F \oplus F^\perp = F \oplus \{0\} = F \neq E \quad \text{e} \quad \left(F^\perp\right)^\perp = \{0\}^\perp = E \neq F$$

No caso de lidarmos com um espaço vectorial de dimensão finita, podemos demonstrar o seguinte teorema, válido também num espaço de dimensão infinita, desde que nele sejam válidas as condições $E = F \oplus F^\perp$ e $\left(F^\perp\right)^\perp = F$:

Teorema 7-23

Teorema de Riesz
Seja E um espaço vectorial euclidiano. Seja $\varphi : E \to \mathbb{R}$ uma aplicação linear qualquer. Então, existe um e um só vector $\vec{a} \in E$ tal que, para qualquer vector $\vec{x} \in E$, se tem $\varphi(\vec{x}) = \vec{a} \bullet \vec{x}$.

Demonstração:
O resultado é trivial no caso de φ ser a aplicação nula, bastando, obviamente, tomar $\vec{a} = \vec{0}$. Por isso, consideraremos o caso de φ não ser a aplicação nula.

Sendo φ uma aplicação não nula, será $\varphi(E) \neq \left\{\vec{0}\right\}$, pelo que, como $\varphi(E) \subseteq \mathbb{R}$, será $\dim\left(\varphi(E)\right) = 1$. Ora, sabemos que

$$\dim(E) = \dim\left(\mathrm{Nuc}(\varphi)\right) + \dim\left(\varphi(E)\right)$$

e daí podemos concluir que $\dim\left(\mathrm{Nuc}(\varphi)\right) = n - 1$.

Por sua vez, por ser $E = \mathrm{Nuc}(\varphi) \oplus \left(\mathrm{Nuc}(\varphi)\right)^\perp$, resulta da igualdade anterior então que $\dim\left(\left(\mathrm{Nuc}(\varphi)\right)^\perp\right) = 1$.

Suponhamos que $\left(\mathrm{Nuc}(\varphi)\right)^{\perp} = \langle\vec{u}\rangle$, onde $\vec{u}$ é um vector de norma 1, e tomemos o vector $\vec{a} = \varphi(\vec{u})\vec{u}$.

Dado um vector qualquer $\vec{x} \in E$, podemos escrever $\vec{x} = \vec{x}_0 + \xi\vec{u}$, com $\vec{x}_0 \in \mathrm{Nuc}(\varphi)$. Daqui resulta que

$$\varphi(\vec{x}) = \varphi(\vec{x}_0 + \xi\vec{u}) = \varphi(\vec{x}_0) + \xi\varphi(\vec{u}) = \vec{0} + \xi\varphi(\vec{u})$$

e, por outro lado,

$$\begin{aligned}\vec{a}\bullet\vec{x} &= \left(\varphi(\vec{u})\vec{u}\right)\bullet\vec{x} = \varphi(\vec{u})(\vec{u}\bullet\vec{x}) = \varphi(\vec{u})\left(\vec{u}\bullet(\vec{x}_0 + \xi\vec{u})\right) = \\ &= \varphi(\vec{u})(\vec{u}\bullet\vec{x}_0 + \xi\vec{u}\bullet\vec{u}) = \varphi(\vec{u})(0 + \xi\,1) = \xi\varphi(\vec{u})\end{aligned}$$

pelo que

$$\varphi(\vec{x}) = \vec{a}\bullet\vec{x}$$

Resta apenas demonstrar a unicidade de $\vec{a}$.

Ora, supondo que os vectores $\vec{a}_1$ e $\vec{a}_2$ verificam as igualdades $\varphi(\vec{x}) = \vec{a}_1\bullet\vec{x}$ e $\varphi(\vec{x}) = \vec{a}_2\bullet\vec{x}$, para qualquer $\vec{x} \in E$, temos

$$\vec{a}_1\bullet\vec{x} = \vec{a}_2\bullet\vec{x} \qquad \text{isto é,} \qquad (\vec{a}_1 - \vec{a}_2)\bullet\vec{x} = 0$$

Como $\vec{x}$ é arbitrário, terá de ser $\vec{a}_1 - \vec{a}_2 = 0$, ou seja, $\vec{a}_1 = \vec{a}_2$.

Por outro lado as propriedades encontradas no Teorema 7-22 dão sentido à seguinte definição:

Definição 7-24 Seja E um espaço euclidiano ou unitário, e seja F um seu subespaço vectorial. O subespaço $F^{\perp}$ designa-se por **complemento ortogonal** de F.

Escrevendo um vector $\vec{v} \in E$ na forma $\vec{v} = \vec{f} + \vec{f}'$, com $\vec{f} \in F$ e $\vec{f}' \in F^{\perp}$, o vector $\vec{f}$ é a **projecção ortogonal** de $\vec{v}$ sobre F, e o vector $\vec{f}'$ é a projecção ortogonal de $\vec{v}$ sobre $F^{\perp}$: $\vec{f} = p_F^{\perp}(\vec{v})$, $\vec{f}' = p_{F^{\perp}}^{\perp}(\vec{v})$.

A projecção ortogonal de um vector sobre um subespaço goza de uma propriedade interessante, que vamos passar a expor, introduzindo previamente uma nova definição:

Definição 7-25 Seja E um espaço vectorial, real ou complexo, com um produto interno fixo. Chama-se **distância** de um vector $\vec{a}$ a um vector $\vec{b}$ ao número real não negativo $d\left(\vec{a},\vec{b}\right) = \left\|\vec{a} - \vec{b}\right\|$.

Este conceito de distância, definido num espaço com produto interno, verifica algumas propriedades muito simples, que seria legítimo esperar de uma noção de "distância", a saber:

Teorema 7-26 Nas condições da definição anterior, tem-se, para vectores arbitrários de E:

a) $d\left(\vec{a},\vec{b}\right)=0$ se e só se $\vec{a}=\vec{b}$

b) $d\left(\vec{a},\vec{b}\right)=d\left(\vec{b},\vec{a}\right)$

c) $d\left(\vec{a},\vec{b}\right)\leq d\left(\vec{a},\vec{c}\right)+d\left(\vec{c},\vec{b}\right)$

Demonstração:

a) $d\left(\vec{a},\vec{b}\right)=0 \Leftrightarrow \left\|\vec{a}-\vec{b}\right\|=0 \Leftrightarrow \vec{a}-\vec{b}=\vec{0} \Leftrightarrow \vec{a}=\vec{b}$

b) $d\left(\vec{a},\vec{b}\right)=\left\|\vec{a}-\vec{b}\right\|=\left\|-\left(\vec{b}-\vec{a}\right)\right\|=|-1|\left\|\vec{b}-\vec{a}\right\|=1\left\|\vec{b}-\vec{a}\right\|=d\left(\vec{b},\vec{a}\right)$

c) $d\left(\vec{a},\vec{b}\right)=\left\|\vec{a}-\vec{b}\right\|=\left\|\vec{a}-\vec{c}+\vec{c}-\vec{b}\right\|\leq\left\|\vec{a}-\vec{c}\right\|+\left\|\vec{c}-\vec{b}\right\|=d\left(\vec{a},\vec{c}\right)+d\left(\vec{c},\vec{b}\right)$

Podemos agora apresentar a anunciada propriedade fundamental da projecção ortogonal de um vector sobre um subespaço:

Teorema 7-27 Seja E um espaço euclidiano ou unitário, e seja F um subespaço vectorial de E. Para cada vector $\vec{v}\in E$, a sua projecção ortogonal sobre F é o vector de F mais próximo (isto é, à distância mínima) de $\vec{v}$.

Demonstração:

Ponhamos $\vec{v}=\vec{f}+\vec{f}'$, com $\vec{f}\in F$ e $\vec{f}'\in F^{\perp}$.

Consideremos, por sua vez, um vector arbitrário $\vec{f}_0\in F$. Por definição, tem-se

$$d\left(\vec{v},\vec{f}_0\right)=\left\|\vec{v}-\vec{f}_0\right\|$$

Deste modo:

$$\begin{aligned}d^2\left(\vec{v},\vec{f}_0\right)&=\left(\vec{v}-\vec{f}_0\right)\bullet\left(\vec{v}-\vec{f}_0\right)=\vec{v}\bullet\vec{v}-\vec{v}\bullet\vec{f}_0-\vec{f}_0\bullet\vec{v}+\vec{f}_0\bullet\vec{f}_0=\\&=\left(\vec{f}+\vec{f}'\right)\bullet\left(\vec{f}+\vec{f}'\right)-\left(\vec{f}+\vec{f}'\right)\bullet\vec{f}_0-\vec{f}_0\bullet\left(\vec{f}+\vec{f}'\right)+\vec{f}_0\bullet\vec{f}_0=\\&=\vec{f}\bullet\vec{f}+\vec{f}\bullet\vec{f}'+\vec{f}'\bullet\vec{f}+\vec{f}'\bullet\vec{f}'-\vec{f}\bullet\vec{f}_0-\vec{f}'\bullet\vec{f}_0-\vec{f}_0\bullet\vec{f}-\vec{f}_0\bullet\vec{f}'+\vec{f}_0\bullet\vec{f}_0\end{aligned}$$

Atendendo a que $\vec{f}'\in F^{\perp}$ e $\vec{f}_0\in F$, sabemos que $\vec{f}'\bullet\vec{f}=\vec{f}'\bullet\vec{f}_0=0$, pelo que, das igualdades anteriores, resulta que

$$d^2\left(\vec{v},\vec{f}_0\right)=\vec{f}\bullet\vec{f}+\vec{f}'\bullet\vec{f}'-f\bullet\vec{f}_0-\vec{f}_0\bullet\vec{f}+\vec{f}_0\bullet\vec{f}_0$$

Ora,

$$\vec{f}\bullet\vec{f}-f\bullet\vec{f}_0-\vec{f}_0\bullet\vec{f}+\vec{f}_0\bullet\vec{f}_0=\left(\vec{f}-\vec{f}_0\right)\bullet\left(\vec{f}-\vec{f}_0\right)=\left\|\vec{f}-\vec{f}_0\right\|^2$$

Portanto,

$$d^2\left(\vec{v},\vec{f}_0\right)=\left\|\vec{f}-\vec{f}_0\right\|^2+\vec{f}'\bullet\vec{f}'\geq\vec{f}'\bullet\vec{f}'$$

Finalmente, observamos que

$$d^2\left(\vec{v},\vec{f}\right)=\left\|\vec{v}-\vec{f}\right\|^2=\left\|\vec{f}'\right\|^2=\vec{f}'\bullet\vec{f}'$$

pelo que

$$d^2\left(\vec{v},\vec{f}_0\right)\geq d^2\left(\vec{v},\vec{f}\right)$$

e como as distâncias não são negativas, podemos concluir que

$$d\left(\vec{v},\vec{f}_0\right)\geq d\left(\vec{v},\vec{f}\right)$$

A tese resulta imediatamente da arbitrariedade de $\vec{f}_0\in F$.

7-4 Produto externo

Concentremo-nos num espaço euclidiano E de dimensão 3, que, como se sabe, será isomorfo a $\mathbb{R}^3$.

Fixemos em E dois vectores linearmente independentes, $\vec{u}$ e $\vec{v}$.

Começaremos por provar o seguinte:

Lema 7-28 Nas condições anteriores, existe uma infinidade de vectores $\vec{z}$ tais que: $\vec{z}\perp\vec{u}$, $\vec{z}\perp\vec{v}$.

Demonstração:

Atendendo a que os vectores $\vec{u}$ e $\vec{v}$ são linearmente independentes, será $\dim\left(\langle\vec{u},\vec{v}\rangle\right)=2$, pelo que

$$\dim\left(\langle\vec{u},\vec{v}\rangle^{\perp}\right)=\dim(E)-\dim\left(\langle\vec{u},\vec{v}\rangle\right)=3-2=1$$

Consequentemente, será $\langle\vec{u},\vec{v}\rangle^{\perp}=\langle\vec{z}_0\rangle$, para algum vector $\vec{z}_0\neq\vec{0}$. Por isso, os vectores $\vec{z}$ que verificam a condição $\vec{z}\perp\vec{u}$, $\vec{z}\perp\vec{v}$ são os vectores da forma $\vec{z}=\alpha\vec{z}_0$, com $\alpha\in\mathbb{R}$, havendo, portanto, uma infinidade deles.

Podemos restringir o número de possibilidades se introduzirmos novas condições para a escolha dos vectores que nos interessam. Assim:

Lema 7-29

Nas condições anteriores, seja ξ um número positivo qualquer. Então, existem dois e só dois vectores $\vec{z}$, tais que:

a) $\vec{z} \perp \vec{u}$, $\vec{z} \perp \vec{v}$

b) $\|\vec{z}\| = \xi$

Demonstração:

Usando a notação empregada na demonstração do lema anterior, sabemos que os vectores $\vec{z}$ que satisfazem a condição a) são da forma $\vec{z} = \alpha \vec{z}_0$, para certo vector $\vec{z}_0 \neq \vec{0}$.

Então,

$$\|\vec{z}\| = \|\alpha \vec{z}_0\| = |\alpha| \, \|\vec{z}_0\|$$

pelo que, para ser $\|\vec{z}\| = \xi$, deveremos pôr

$$|\alpha| \, \|\vec{z}_0\| = \xi \quad \text{ou seja,} \quad |\alpha| = \frac{\xi}{\|\vec{z}_0\|}$$

ou, ainda,

$$\alpha = \pm \frac{\xi}{\|\vec{z}_0\|}$$

Temos, portanto, dois e só dois vectores a satisfazer simultaneamente as condições a) e b). São eles os vectores

$$\vec{z}_1 = \frac{\xi}{\|\vec{z}_0\|} \vec{z}_0 \quad \text{e} \quad \vec{z}_2 = -\frac{\xi}{\|\vec{z}_0\|} \vec{z}_0$$

O nosso objectivo será o de introduzir uma nova condição que permita seleccionar um único dos vectores a que se refere o lema anterior. Para isso, começamos por apresentar o seguinte conceito:

Definição 7-30

Seja E um espaço vectorial, de dimensão n, sobre um corpo K. Fixe-se em E uma base $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$. Dados n vectores arbitrários $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$, consideremos a matriz

$$V = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{bmatrix}$$

cujas colunas são as componentes dos vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ em relação à base considerada. O determinante da matriz V toma então o nome de **determinante dos vectores** $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ em relação à base $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$:

$$\det_{(\vec{e}_i)}(\vec{v}_1, \vec{v}_2, ..., \vec{v}_n) = |V|$$

Com este conceito, podemos facilmente enunciar o seguinte teorema:

Teorema 7-31 Seja E um espaço euclidiano de dimensão 3, com uma base fixa $\vec{e}_1, \vec{e}_2, \vec{e}_3$. Seja ξ um número positivo qualquer. Então, existe um e um só vector $\vec{z}$ tal que:

a) $\vec{z} \perp \vec{u}, \vec{z} \perp \vec{v}$

b) $\|\vec{z}\| = \xi$

c) $\det_{(\vec{e}_i)}(\vec{u}, \vec{v}, \vec{z}) > 0$

Demonstração:

Segundo vimos na demonstração do Lema 7-29, há dois e só dois vectores, $\vec{z}_1$ e $\vec{z}_2$, que satisfazem as condições a) e b), tendo-se $\vec{z}_2 = -\vec{z}_1$.

Ora, resulta imediatamente da definição que

$$\det_{(\vec{e}_i)}(\vec{u}, \vec{v}, \vec{z}_2) = \det_{(\vec{e}_i)}(\vec{u}, \vec{v}, -\vec{z}_1) = -\det_{(\vec{e}_i)}(\vec{u}, \vec{v}, \vec{z}_1)$$

pelo que um e um só dos vectores $\vec{z}_1$ e $\vec{z}_2$ satisfaz a condição c).

O que acabámos de ver permite-nos, então, introduzir a definição de "produto externo" de dois vectores, nos seguintes termos:

Definição 7-32 Seja E um espaço euclidiano de dimensão 3, com uma base fixa $\vec{e}_1, \vec{e}_2, \vec{e}_3$. Dados dois vectores quaisquer $\vec{u}, \vec{v} \in E$, chama-se **produto externo** (ou **produto vectorial**) de $\vec{u}$ por $\vec{v}$, e representa-se por $\vec{u} \times \vec{v}$, ao vector assim definido:

a) Se os vectores $\vec{u}, \vec{v}$ são linearmente dependentes, então $\vec{u} \times \vec{v} = \vec{0}$.

b) Se os vectores $\vec{u}, \vec{v}$ são linearmente independentes, então $\vec{u} \times \vec{v}$ é o (único) vector de E que verifica as seguintes condições:

b1) $\vec{u} \times \vec{v} \perp \vec{u}$, $\vec{u} \times \vec{v} \perp \vec{v}$

b2) $\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \; \|\vec{v}\| \; \text{sen} \angle(\vec{u}, \vec{v})$

b3) $\det_{(\vec{e}_i)}(\vec{u}, \vec{v}, \vec{u} \times \vec{v}) > 0$

Deve observar-se que, do ponto de vista meramente intuitivo, quando interpretamos os vectores do espaço como segmentos orientados em $\mathbb{R}^3$, o valor $\|\vec{u}\|\|\vec{v}\| \text{sen} \angle(\vec{u}, \vec{v})$ representa a "área do paralelogramo" construído sobre os segmentos orientados $\vec{u}$ e $\vec{v}$.

A determinação do produto externo de dois vectores linearmente independentes é, em geral, muito laboriosa. No entanto, se conhecermos as componentes dos dois vectores em relação a uma base ortonormada, o cálculo fica muito mais simples, em virtude do seguinte:

Teorema 7-33 Seja E um espaço euclidiano de dimensão 3, com uma base ortonormada fixa $\vec{e}_1, \vec{e}_2, \vec{e}_3$. Dados dois vectores independentes $\vec{u}, \vec{v} \in E$, suponhamos que são conhecidas as suas componentes, em relação à base considerada: $\vec{u} \equiv (u_1, u_2, u_3)$, $\vec{v} \equiv (v_1, v_2, v_3)$. Então, tem-se

$$\vec{u} \times \vec{v} \equiv \left(u_2 v_3 - u_3 v_2,\ u_3 v_1 - u_1 v_3,\ u_1 v_2 - u_2 v_1\right)$$

Demonstração:

Atendendo à unicidade referida no Teorema 7-31, para vermos que

$$\vec{u} \times \vec{v} \equiv \left(u_2 v_3 - u_3 v_2,\ u_3 v_1 - u_1 v_3,\ u_1 v_2 - u_2 v_1\right)$$

bastará verificar que o vector $\vec{z}$, de componentes

$$\left(u_2 v_3 - u_3 v_2,\ u_3 v_1 - u_1 v_3,\ u_1 v_2 - u_2 v_1\right)$$

satisfaz as três condições da definição de produto externo. Uma vez que a base considerada é ortonormada, temos:

a) $$\vec{z} \bullet \vec{u} = \left(u_2 v_3 - u_3 v_2,\ u_3 v_1 - u_1 v_3,\ u_1 v_2 - u_2 v_1\right) \bullet \left(u_1, u_2, u_3\right) =$$
$$= \left(u_2 v_3 - u_3 v_2\right) u_1 + \left(u_3 v_1 - u_1 v_3\right) u_2 + \left(u_1 v_2 - u_2 v_1\right) u_3 =$$
$$= u_1 u_2 v_3 - u_1 u_3 v_2 + u_2 u_3 v_1 - u_1 u_2 v_3 + u_1 u_3 v_2 - u_2 u_3 v_1 = 0$$

Deste modo, $\vec{z} \perp \vec{u}$. Analogamente se verifica que $\vec{z} \perp \vec{v}$.

b) $$\det_{(\vec{e}_i)}\left(\vec{u}, \vec{v}, \vec{z}\right) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ u_2 v_3 - u_3 v_2 & u_3 v_1 - u_1 v_3 & u_1 v_2 - u_2 v_1 \end{vmatrix} =$$
$$= u_1 v_2 \left(u_1 v_2 - u_2 v_1\right) + u_2 v_3 \left(u_2 v_3 - u_3 v_2\right) + u_3 v_1 \left(u_3 v_1 - u_1 v_3\right) -$$
$$- u_3 v_2 \left(u_2 v_3 - u_3 v_2\right) - u_2 v_1 \left(u_1 v_2 - u_2 v_1\right) - u_1 v_3 \left(u_3 v_1 - u_1 v_3\right) =$$
$$= u_1^2 v_2^2 - u_1 u_2 v_1 v_2 + u_2^2 v_3^2 - u_2 u_3 v_2 v_3 + u_3^2 v_1^2 - u_1 u_3 v_1 v_3 -$$
$$- u_2 u_3 v_2 v_3 + u_3^2 v_2^2 - u_1 u_2 v_1 v_2 + u_2^2 v_1^2 - u_1 u_3 v_1 v_3 + u_1^2 v_3^2 =$$
$$= u_1^2 v_2^2 - 2 u_1 u_2 v_1 v_2 + u_2^2 v_3^2 - 2 u_2 u_3 v_2 v_3 + u_3^2 v_1^2 + u_3^2 v_2^2 +$$
$$+ u_2^2 v_1^2 - 2 u_1 u_3 v_1 v_3 + u_1^2 v_3^2 =$$
$$= \left(u_1 v_2 - u_2 v_1\right)^2 + \left(u_2 v_3 - u_3 v_2\right)^2 + \left(u_1 v_3 - u_3 v_1\right)^2 \geq 0$$

Como $\vec{z}$ é ortogonal a $\vec{u}$ e a $\vec{v}$, os três vectores são linearmente independentes, pelo que o determinante não pode ser nulo.

c) $\|\vec{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2}$, $\|\vec{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2}$

$$\operatorname{sen}\angle(\vec{u},\vec{v}) = \sqrt{1-\cos^2(\vec{u},\vec{v})} = \sqrt{1-\left(\frac{\vec{u}\bullet\vec{v}}{\|\vec{u}\|\|\vec{v}\|}\right)^2} =$$

$$= \sqrt{1-\left(\frac{u_1v_1 + u_2v_2 + u_3v_3}{\sqrt{u_1^2+u_2^2+u_3^2}\sqrt{v_1^2+v_2^2+v_3^2}}\right)^2}$$

$$\|\vec{z}\| = \sqrt{(u_2v_3 - u_3v_2)^2 + (u_3v_1 - u_1v_3)^2 + (u_1v_2 - u_2v_1)^2}$$

Por verificação directa, comprovamos que

$$\|\vec{z}\| = \|\vec{u}\|\|\vec{v}\|\operatorname{sen}\angle(\vec{u},\vec{v})$$

7-5 | Endomorfismos adjuntos

Num espaço vectorial onde se encontra definido um produto interno, merecem interesse particular aqueles endomorfismos que, de alguma maneira, respeitam esse mesmo produto interno. Estudaremos, no presente parágrafo, uma dessas situações, começando por introduzir uma definição muito geral:

Definição 7-34 Seja φ um endomorfismo de um espaço vectorial E, onde se encontre definido um produto interno. Diz-se que um outro endomorfismo φ^*, do mesmo espaço E, é **adjunto** de φ quando, para vectores quaisquer do espaço, se tem

$$\varphi(\vec{x})\bullet\vec{y} = \vec{x}\bullet\varphi^*(\vec{y})$$

Podemos facilmente demonstrar as seguintes propriedades:

Teorema 7-35 Seja E um espaço vectorial com produto interno. Sejam φ^* e ψ^* endomorfismos adjuntos de dois dados endomorfismos φ e ψ, do espaço E, respectivamente. Então:

a) $\varphi^* + \psi^*$ é adjunto de $\varphi + \psi$.
b) Para qualquer escalar λ, $\overline{\lambda}\varphi^*$ é adjunto de $\lambda\varphi$.
c) $\psi^* \circ \varphi^*$ é adjunto de $\varphi \circ \psi$.
d) O endomorfismo 1_E é adjunto de si mesmo.
e) Se E tem dimensão finita e φ é um automorfismo, então também φ^* é um automorfismo e $(\varphi^*)^{-1}$ é adjunto de φ^{-1}.

Demonstração:

a) Tem-se

$$(\varphi+\psi)(\vec{x})\bullet\vec{y} = (\varphi(\vec{x})+\psi(\vec{x}))\bullet\vec{y} = \varphi(\vec{x})\bullet\vec{y}+\psi(\vec{x})\bullet\vec{y} =$$
$$= \vec{x}\bullet\varphi^*(\vec{y})+\vec{x}\bullet\psi^*(\vec{y}) = \vec{x}\bullet(\varphi^*(\vec{y})+\psi^*(\vec{y})) = \vec{x}\bullet(\varphi^*+\psi^*)(\vec{y})$$

b) Tem-se

$$(\lambda\varphi)(\vec{x})\bullet\vec{y} = (\lambda\varphi(\vec{x}))\bullet\vec{y} = \lambda(\varphi(\vec{x})\bullet\vec{y}) = \lambda(\vec{x}\bullet\varphi^*(\vec{y})) =$$
$$= \vec{x}\bullet(\overline{\lambda}\varphi^*(\vec{y})) = \vec{x}\bullet(\overline{\lambda}\varphi^*(\vec{y})) = \vec{x}\bullet(\lambda\varphi^*)(\vec{y})$$

c) Tem-se

$$(\varphi\circ\psi)(\vec{x})\bullet\vec{y} = \varphi(\psi(\vec{x}))\bullet\vec{y} = \psi(\vec{x})\bullet\varphi^*(\vec{y}) = \vec{x}\bullet\psi^*(\varphi^*(\vec{y})) = \vec{x}\bullet(\psi^*\circ\varphi^*)(\vec{y})$$

d) Resulta trivialmente da definição.

e) Comecemos por observar que, sendo $\vec{u}\in \mathrm{Nuc}(\varphi^*)$, se tem, para qualquer $\vec{x}\in E$:

$$\varphi(\vec{x})\bullet\vec{u} = \vec{x}\bullet\varphi^*(\vec{u}) = \vec{x}\bullet\vec{0} = 0$$

Como φ é uma aplicação sobrejectiva, $\varphi(\vec{x})$ representa ainda um vector arbitrário de E, pelo que as igualdades anteriores mostram que

$$\vec{u}\in E^{\perp} = \{\vec{0}\}$$

ou seja, que $\vec{u}=\vec{0}$. Daqui se conclui que φ^* é um monomorfismo.

Uma vez que E tem dimensão finita, o facto de φ^* ser um monomorfismo garante que é um automorfismo.

Para vermos, por fim, que $(\varphi^*)^{-1}$ é adjunto de φ^{-1}, tomemos vectores $\vec{x}, \vec{y} \in E$ e ponhamos $\vec{u} = \varphi^{-1}(\vec{x})$, $\vec{v} = (\varphi^*)^{-1}(\vec{y})$. Temos:

$$\varphi^{-1}(\vec{x}) \bullet \vec{y} = \vec{u} \bullet \varphi^*(\vec{v}) = \varphi(\vec{u}) \bullet \vec{v} = \vec{x} \bullet (\varphi^*)^{-1}(\vec{y})$$

o que completa a demonstração.

Se nos concentrarmos agora nos espaços vectoriais de dimensão finita, é muito fácil encontrar o adjunto de um dado endomorfismo, conforme vamos ver:

Teorema 7-36 Num espaço euclidiano ou unitário, qualquer endomorfismo tem um e um só adjunto.

Demonstração:

Fixemos em E uma base $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$, e suponhamos que os endomorfismos φ e φ^*, do espaço vectorial E, são definidos, em relação a essa base, pelas matrizes A e A^*, respectivamente.

A igualdade

$$\varphi(\vec{x}) \bullet \vec{y} = \vec{x} \bullet \varphi^*(\vec{y})$$

pode escrever-se em termos matriciais, desde que representemos os vectores $\vec{x}$ e $\vec{y}$ pelas colunas X e Y, das suas componentes em relação à base considerada. Teremos, então,

$$(AX)^T G\overline{Y} = X^T G\overline{(A^*Y)}$$

onde G designa a matriz da métrica do produto interno fixo no espaço em relação à base considerada.

Esta última igualdade equivale a

$$X^T A^T G\overline{Y} = X^T G\overline{A^*}\,\overline{Y}$$

e, dada a arbitrariedade de X e Y, a

$$A^T G = G\overline{A^*}$$

Daqui resulta, naturalmente, que

$$\overline{A^*} = G^{-1}A^T G \qquad \text{ou seja, que} \qquad A^* = \overline{G}^{-1}\overline{A}^T\overline{G}$$

o que prova que, existindo o endomorfismo adjunto de φ, ele será único.

Mas, supondo fixo o endomorfismo φ, então, em relação à base considerada, a matriz $\overline{G}^{-1}\overline{A}^T\overline{G}$ define um certo endomorfismo φ^*. Pelo que acabamos de ver, este será adjunto de φ.

Da demonstração anterior resulta, portanto, que, sendo φ um endomorfismo de um espaço euclidiano ou unitário E, definido por uma matriz A em relação a determinada base do espaço, o seu endomorfismo adjunto é definido, em relação à mesma base, pela matriz $\overline{G}^{-1}\overline{A}^T\overline{G}$.

É claro que, se a base escolhida for ortonormada, a respectiva matriz da métrica será a identidade, pelo que o endomorfismo adjunto fica definido pela matriz $\overline{I}_n^{-1}\overline{A}^T\overline{I}_n = \overline{A}^T$ (a matriz **transconjugada** de A).

No caso particular de se tratar de um espaço vectorial real, o endomorfismo adjunto do endomorfismo definido pela matriz A é definido pela matriz $G^{-1}A^TG$, ou, caso a base seja ortonormada, pela matriz A^T.

Um caso particularmente importante é aquele em que um endomorfismo de um dado espaço vectorial seja adjunto de si mesmo. É natural dar a seguinte definição:

Definição 7-37 Um endomorfismo que seja adjunto de si mesmo chama-se um **endomorfismo auto-adjunto**.

Assim, um endomorfismo auto-adjunto φ é aquele que verifica, para vectores arbitrários, a igualdade

$$\varphi(\vec{x})\bullet\vec{y} = \vec{x}\bullet\varphi(\vec{y})$$

Se o espaço vectorial onde φ se encontra definido tem dimensão finita, e se φ tem uma certa matriz A em relação a determinada base de E, o endomorfismo φ será auto-adjunto se e só se a sua matriz verificar a igualdade $\overline{G}^{-1}\overline{A}^T\overline{G} = A$, ou, equivalentemente, $A^T = G\overline{A}G^{-1}$. No caso particular de a base escolhida ser ortonormada, a matriz verificará a igualdade $\overline{A}^T = A$. Diz-se, neste caso, que A é uma **matriz hermítica**.

No caso de o espaço E ser um espaço vectorial real, a igualdade anterior transforma-se em $A^T = GAG^{-1}$, ou, sendo a base ortonormada, $A^T = A$ (A **matriz simétrica**).

Os endomorfismos auto-adjuntos verificam propriedades particularmente interessantes, conforme vamos ver:

Teorema 7-38 Um endomorfismo auto-adjunto, num espaço vectorial de dimensão n, admite n valores próprios reais, distintos ou não.

Demonstração:

Seja φ um endomorfismo do espaço vectorial E, definido por uma matriz A em relação a uma base ortonormada do espaço.

Seja agora λ_0 uma raiz complexa da equação $|A - \lambda I_3| = 0$. Existirá uma coluna U, não nula, tal que $AU = \lambda_0 U$.

Vemos que

$$\overline{U}^T AU = \overline{U}^T(AU) = \overline{U}^T(\lambda_0 U) = \lambda_0\overline{U}^T U = \lambda_0\sum_{i=1}^{n} u_i\overline{u_i} = \lambda_0\sum_{i=1}^{n}|u_i|^2$$

pelo que

$$\lambda_0 = \frac{\overline{U}^T A U}{\sum_{i=1}^{n} |u_i|^2}$$

Ora, por sua vez, tem-se, atendendo a que a matriz A é uma matriz hermítica,

$$\overline{\overline{U}^T A U} = \left(\overline{U}^T A U\right)^T = \overline{U^T A^T \overline{U}} = \overline{U}^T \overline{A}^T\ \overline{\overline{U}} = \overline{U}^T A U$$

pelo que

$$\overline{U}^T A U \in \mathbb{R}$$

Concluímos, assim, que λ_0 é um número real, ou seja, que os valores próprios de φ são todos reais (podendo ser distintos ou não).

Teorema 7-39 Se λ_1 e λ_2 são valores próprios distintos de um endomorfismo auto-adjunto, e se H_{λ_1} e H_{λ_2} são os respectivos subespaços próprios, tem-se

$$H_{\lambda_1} \subseteq H_{\lambda_2}^{\perp}$$

Demonstração:
Tomemos $\vec{v}_1 \in H_{\lambda_1}$ e $\vec{v}_2 \in H_{\lambda_2}$.

Pelo facto de o endomorfismo φ ser auto-adjunto, temos

$$\varphi(\vec{v}_1) \bullet \vec{v}_2 = \vec{v}_1 \bullet \varphi(\vec{v}_2)$$

Ora, sendo $\vec{v}_1$ e $\vec{v}_2$ vectores próprios de φ, sabemos que $\varphi(\vec{v}_1) = \lambda_1 \vec{v}_1$ e $\varphi(\vec{v}_2) = \lambda_2 \vec{v}_2$, pelo que a igualdade anterior se transforma em

$$(\lambda_1 \vec{v}_1) \bullet \vec{v}_2 = \vec{v}_1 \bullet (\lambda_2 \vec{v}_2)$$

Daqui podemos tirar, sucessivamente,

$$\lambda_1 (\vec{v}_1 \bullet \vec{v}_2) = \overline{\lambda}_2 (\vec{v}_1 \bullet \vec{v}_2) = \lambda_2 (\vec{v}_1 \bullet \vec{v}_2)$$

$$\lambda_1 (\vec{v}_1 \bullet \vec{v}_2) - \lambda_2 (\vec{v}_1 \bullet \vec{v}_2) = 0$$

$$(\lambda_1 - \lambda_2)(\vec{v}_1 \bullet \vec{v}_2) = 0$$

Ora, sendo λ_1 e λ_2 valores próprios distintos, temos $\lambda_1 - \lambda_2 \neq 0$, pelo que $\vec{v}_1 \bullet \vec{v}_2 = 0$, o que prova o que pretendíamos.

Teorema 7-40 Seja φ um endomorfismo auto-adjunto de um espaço vectorial E, de dimensão finita, e seja F um subespaço vectorial de E. Se F é **invariante** para φ, isto é, se $\varphi(F) \subseteq F$, então também $F^\perp$ é invariante para φ, ou seja, $\varphi(F^\perp) \subseteq F^\perp$.

Demonstração:
Tomemos vectores $\vec{x} \in F$ e $\vec{y} \in F^\perp$.

Por ser $\vec{x} \in F$ e $\varphi(F) \subseteq F$, concluímos que $\varphi(\vec{x}) \in F$.

Consequentemente,

$$\varphi(\vec{y}) \bullet \vec{x} = \vec{y} \bullet \varphi(\vec{x}) = 0$$

uma vez que $\varphi(\vec{x}) \in F$ e $\vec{y} \in F^\perp$.

Dada a arbitrariedade de $\vec{x} \in F$, resulta daqui que $\varphi(\vec{y}) \in F^\perp$. Como, por sua vez, $\vec{y} \in F^\perp$ é igualmente arbitrário, concluímos que $\varphi(F^\perp) \subseteq F^\perp$.

Teorema 7-41 Um endomorfismo auto-adjunto φ, de um espaço euclidiano ou unitário E, é diagonalizável e existe em E uma base ortonormada constituída por vectores próprios de φ.

Demonstração:
Tudo se resume a provar que φ é um endomorfismo diagonalizável. Com efeito, nesse caso, o espaço E será a soma directa dos respectivos subespaços próprios, e juntando bases ortonormadas desses vários subespaços próprios obteremos uma base ortonormada para E, em virtude do Teorema 7-39.

Sejam então $\lambda_1, \lambda_2, \ldots, \lambda_k$ os valores próprios distintos de φ, e sejam H_{λ_1}, $H_{\lambda_2}, \ldots, H_{\lambda_k}$ os respectivos subespaços próprios. Ponhamos ainda $H = H_{\lambda_1} \oplus \oplus H_{\lambda_2} \oplus \cdots \oplus H_{\lambda_k}$.

É claro que H é invariante para φ, pelo que, pelo Teorema 7-40, o mesmo acontece a $H^\perp$.

Sendo assim, a restrição de φ a $H^\perp$ será um endomorfismo auto-adjunto de $H^\perp$, pelo que, sendo $H^\perp \neq \{\vec{0}\}$, essa restrição admitiria um vector próprio $\vec{z}$. Ora, todos os vectores próprios de φ estão, evidentemente, em H, e não pode um vector não nulo (por definição de vector próprio) pertencer a H e a $H^\perp$ simultaneamente.

Logo, terá de ser $H^\perp = \{\vec{0}\}$.

Mas, então, $E = H \oplus H^\perp = H \oplus \{\vec{0}\} = H = H_{\lambda_1} \oplus H_{\lambda_2} \oplus \cdots \oplus H_{\lambda_k}$, pelo que o endomorfismo φ é diagonalizável.

Terminaremos o presente parágrafo com uma propriedade curiosa das matrizes hermíticas:

Teorema 7-42 Sejam A e B matrizes hermíticas. As matrizes A e B são simultaneamente diagonalizáveis, isto é, existe uma matriz invertível P, tal que tanto $P^{-1}AP$ como $P^{-1}BP$ são matrizes diagonais, se e só se as matrizes A e B são permutáveis, ou seja, se e só se $AB = BA$.

Demonstração:

Supondo que $P^{-1}AP = A'$ e $P^{-1}BP = B'$ são matrizes diagonais, tem-se, evidentemente, $A'B' = B'A'$, já que, de modo geral,

$$\begin{bmatrix} u_1 & 0 & \cdots & 0 \\ 0 & u_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 & 0 & \cdots & 0 \\ 0 & v_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} u_1v_1 & 0 & \cdots & 0 \\ 0 & u_2v_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & u_nv_n \end{bmatrix}$$

Portanto,

$$AB = \left(PA'P^{-1}\right)\left(PB'P^{-1}\right) = PA'B'P^{-1} = PB'A'P^{-1} = \left(PB'P^{-1}\right)\left(PA'P^{-1}\right) = BA$$

Reciprocamente, suponhamos que as matrizes A e B são permutáveis.

Sejam $\lambda_1, \lambda_2, \ldots, \lambda_k$ os valores próprios do endomorfismo φ_A, definido pela matriz A em relação a uma base ortonormada fixa num espaço euclidiano ou unitário E de dimensão adequada. Como se sabe, E é a soma directa dos respectivos subespaços próprios: $E = H_{\lambda_1} \oplus H_{\lambda_2} \oplus \cdots \oplus H_{\lambda_k}$.

É fácil verificar que cada subespaço H_{λ_i} é invariante para o endomorfismo φ_B, definido, em relação à mesma base, pela matriz B. Com efeito, dado $\vec{x} \in H_{\lambda_i}$, e designando por X a coluna das suas componentes em relação à base considerada, tem-se

$$\begin{aligned} \varphi_A\left(\varphi_B(\vec{x})\right) &\equiv A(BX) = (AB)X = (BA)X = B(AX) \equiv \\ &\equiv \varphi_B\left(\varphi_A(\vec{x})\right) = \varphi_B(\lambda_i\vec{x}) = \lambda_i\varphi_B(\vec{x}) \end{aligned}$$

pelo que $\varphi_B(\vec{x}) \in H_{\lambda_i}$.

Uma vez que a matriz B é hermítica, o endomorfismo φ_B, definido por essa matriz em relação a uma base ortonormada, é um endomorfismo auto-adjunto, pelo que a restrição de φ_B a cada subespaço H_{λ_i} será ainda um endomorfismo auto-adjunto; logo, diagonalizável.

Haverá, portanto, em cada H_{λ_i} uma base $\vec{u}_{1i}, \vec{u}_{2i}, \ldots, \vec{u}_{n_i i}$, constituída por vectores próprios de φ_B. Esses vectores, por estarem em H_{λ_i}, serão simultaneamente vectores próprios de φ_A. Juntando as bases dos diversos subespaços H_{λ_i} $(i = 1, 2, \ldots, k)$, obteremos uma base de E, formada ao mesmo tempo por vectores próprios de φ_A e de φ_B.

Sendo P a respectiva matriz de mudança de base, as matrizes $P^{-1}AP$ e $P^{-1}BP$ são ambas matrizes diagonais.

8 | Aplicações multilineares

Estudámos já a noção de "aplicação linear" e, por outro lado, a noção de "produto interno". Essas noções podem ser generalizadas: podemos considerar funções de várias variáveis que obedeçam a propriedades semelhantes às das aplicações lineares, sendo o produto interno um caso particular de uma tal situação.

Começaremos por analisar o caso de funções de duas variáveis, que se reveste, por si só, de grande importância. Encontraremos, ao longo do próximo parágrafo, diversas propriedades que generalizam o que encontrámos no estudo do produto interno num espaço real.

8-1 | Aplicações bilineares

Vamos, de imediato, introduzir a noção de aplicação "bilinear", através da seguinte definição:

Definição 8-1 Sejam E, E' e F espaços vectoriais sobre um certo corpo K. Uma aplicação $f: E \times E' \to F$ é uma **aplicação bilinear** quando verifica as seguintes condições, para vectores e escalares arbitrários:

a) $f\left(\vec{x}+\vec{y},\vec{x}'\right)=f\left(\vec{x},\vec{x}'\right)+f\left(\vec{y},\vec{x}'\right)$

b) $f\left(\vec{x},\vec{x}'+\vec{y}'\right)=f\left(\vec{x},\vec{x}'\right)+f\left(\vec{x},\vec{y}'\right)$

c) $f\left(\alpha\vec{x},\vec{x}'\right)=f\left(\vec{x},\alpha\vec{x}'\right)=\alpha f\left(\vec{x},\vec{x}'\right)$

Em particular, quando $E=E'$ e $F=K$, a aplicação bilinear f toma o nome de **forma bilinear**.

Exemplos 8-2

a) Consideremos os espaços vectoriais $\mathrm{M}_{p\times n}(K)$, $\mathrm{M}_{n\times q}(K)$ e $\mathrm{M}_{p\times q}(K)$, formados por matrizes com elementos do corpo K. De acordo com as propriedades conhecidas das operações sobre matrizes, a aplicação

$$f:\ \mathrm{M}_{p\times n}(K)\times \mathrm{M}_{n\times q}(K)\to \mathrm{M}_{p\times q}(K)$$

definida por $f\left(A,A'\right)=AA'$ é uma aplicação bilinear.

b) Qualquer produto interno num espaço real é uma forma bilinear.

c) Seja E um espaço vectorial real, de dimensão 2, com uma base fixa, $\vec{e}_1, \vec{e}_2$. Atendendo às propriedades conhecidas, vemos que a aplicação $d : E \times E \to \mathbb{R}$ definida por

$$d(\vec{u}, \vec{v}) = \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix}$$

onde (u_1, u_2) e (v_1, v_2) são as componentes dos vectores $\vec{u}$ e $\vec{v}$, respectivamente, em relação à base considerada, é uma forma bilinear.

d) Consideremos a aplicação

$$g : \mathbb{R}^2 \times \mathbb{R}^3 \to \mathbb{R}^2$$

definida por

$$g\left((u_1, u_2), (v_1, v_2, v_3)\right) = \left(5u_1v_2 - 7u_2v_2,\ 3u_1v_3 + 9u_2v_1 - u_2v_3\right)$$

É fácil verificar que se trata de uma aplicação bilinear. Por exemplo:

$$\begin{aligned} g\left((u_1, u_2) + (u_1', u_2'), (v_1, v_2, v_3)\right) &= g\left((u_1 + u_1', u_2 + u_2'), (v_1, v_2, v_3)\right) = \\ &= \left(5(u_1 + u_1')v_2 - 7(u_2 + u_2')v_2, 3(u_1 + u_1')v_3 + 9(u_2 + u_2')v_1 - (u_2 + u_2')v_3\right) = \\ &= \left(5u_1v_2 + 5u_1'v_2 - 7u_2v_2 - 7u_2'v_2, 3u_1v_3 + 3u_1'v_3 + 9u_2v_1 + 9u_2'v_1 - u_2v_3 - u_2'v_3\right) = \\ &= \left(5u_1v_2 - 7u_2v_2, 3u_1v_3 + 9u_2v_1 - u_2v_3\right) + \left(5u_1'v_2 - 7u_2'v_2, 3u_1'v_3 + 9u_2'v_1 - u_2'v_3\right) = \\ &= g\left((u_1, u_2), (v_1, v_2, v_3)\right) + g\left((u_1', u_2'), (v_1, v_2, v_3)\right) \end{aligned}$$

$$\begin{aligned} g\left(\alpha(u_1, u_2), (v_1, v_2, v_3)\right) &= g\left((\alpha u_1, \alpha u_2), (v_1, v_2, v_3)\right) = \\ &= \left(5\alpha u_1v_2 - 7\alpha u_2v_2, 3\alpha u_1v_3 + 9\alpha u_2v_1 - \alpha u_2v_3\right) = \\ &= \left(\alpha(5u_1v_2 - 7u_2v_2), \alpha(3u_1v_3 + 9u_2v_1 - u_2v_3)\right) = \\ &= \alpha\left(5u_1v_2 - 7u_2v_2, 3u_1v_3 + 9u_2v_1 - u_2v_3\right) = \alpha g\left((u_1, u_2), (v_1, v_2, v_3)\right) \end{aligned}$$

A verificação das restantes propriedades é inteiramente análoga. ■

A designação de "bilinear", para as aplicações que satisfazem as condições indicadas, é justificada pelas seguintes propriedades:

Teorema 8-3 Seja $f : E \times E' \to F$ é uma aplicação bilinear. Fixemos vectores $\vec{a} \in E$ e $\vec{a}' \in E'$. As aplicações ${}_{\vec{a}}f : E' \to F$ e $f_{\vec{a}'} : E \to F$, definidas por

$${}_{\vec{a}}f(\vec{x}') = f(\vec{a}, \vec{x}') \text{ e } f_{\vec{a}'}(\vec{x}) = f(\vec{x}, \vec{a}'), \text{ são aplicações lineares.}$$

Demonstração:

Resulta trivialmente das definições.

Por sua vez, atendendo às propriedades conhecidas das aplicações lineares, vemos que:

Corolário 8-4

Dada uma aplicação bilinear $f : E \times E' \to F$, tem-se, para quaisquer vectores $\vec{x} \in E$ e $\vec{x}' \in E'$: $f\left(\vec{x},\vec{0}\right) = f\left(\vec{0},\vec{x}'\right) = \vec{0}$.

Demonstração:

Tem-se

$$f\left(\vec{x},\vec{0}\right) = {}_{\vec{x}}f\left(\vec{0}\right) = \vec{0} \quad \text{e} \quad f\left(\vec{0},\vec{x}'\right) = f_{\vec{x}'}\left(\vec{0}\right) = \vec{0}$$

No que se segue, concentrar-nos-emos no estudo das formas bilineares, em espaços vectoriais de dimensão finita, se bem que alguns dos resultados que vamos obter se possam generalizar a aplicações bilineares quaisquer.

Seja, então, E um espaço vectorial de dimensão finita sobre um corpo K, e fixemos em E uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$.

Dada uma forma bilinear $f : E \times E \to K$, tomemos os vectores $\vec{x}, \vec{y} \in E$, de componentes $\vec{x} \equiv (x_1, x_2, \ldots, x_n)$ e $\vec{y} \equiv (y_1, y_2, \ldots, y_n)$ em relação à base considerada. Vamos calcular a imagem $f(\vec{x},\vec{y})$; utilizando as propriedades das aplicações bilineares, obtemos:

$$f(\vec{x},\vec{y}) = f\left(\sum_{i=1}^{n} x_i \vec{e}_i, \vec{y}\right) = \sum_{i=1}^{n} x_i f(\vec{e}_i, \vec{y}) = \sum_{i=1}^{n} x_i f\left(\vec{e}_i, \sum_{j=1}^{n} y_j \vec{e}_j\right) =$$

$$= \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} y_j f\left(\vec{e}_i, \vec{e}_j\right)\right) = \sum_{i,j=1}^{n} f\left(\vec{e}_i, \vec{e}_j\right) x_i y_j$$

Escrevendo, abreviadamente, $f\left(\vec{e}_i, \vec{e}_j\right) = a_{ij}$, para $i, j = 1, 2, \ldots, n$, a expressão anterior assume a forma

$$f(\vec{x},\vec{y}) = \sum_{i,j=1}^{n} a_{ij} x_i y_j$$

Se, por sua vez, considerarmos a matriz $A = [a_{ij}]$, do tipo $n \times n$, bem como as colunas

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ e } Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

formadas pelas componentes dos vectores em causa em relação à base considerada, podemos ainda escrever, como facilmente se verifica,

$$f(\vec{x},\vec{y}) = \sum_{i,j=1}^{n} a_{ij}x_i y_j = X^T A Y$$

Podemos, em face do que fica exposto, enunciar o seguinte teorema:

Teorema 8-5

Seja E um espaço vectorial sobre um corpo K, com uma base fixa $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$. Dada uma forma bilinear $f : E \times E \to K$, seja A a matriz do tipo $n \times n$, cujos elementos são $a_{ij} = f(\vec{e}_i, \vec{e}_j)$, com $i, j = 1, 2, \ldots, n$. A forma bilinear f fica perfeitamente determinada pela matriz A, uma vez que, para $\vec{x}, \vec{y} \in E$ quaisquer, se tem

$$f(\vec{x},\vec{y}) = \sum_{i,j=1}^{n} a_{ij}x_i y_j = X^T A Y$$

onde X e Y são as colunas formadas pelas componentes dos vectores $\vec{x}$ e $\vec{y}$, respectivamente, em relação à base considerada.

Reciprocamente, dada uma matriz B qualquer, do tipo $n \times n$, existe uma e uma só forma bilinear $g : E \times E \to K$ tal que $g(\vec{e}_i, \vec{e}_j) = b_{ij}$, para quaisquer $i, j = 1, 2, \ldots, n$.

Demonstração:

A primeira afirmação resulta imediatamente do que se viu acima.

Quanto à segunda, consideremos a aplicação $g : E \times E \to K$, definida por

$$g(\vec{x},\vec{y}) = X^T B Y$$

É fácil verificar que g é uma forma bilinear. Por exemplo:

$$g(\vec{x}+\vec{x}',\vec{y}) = (X+X')^T BY = (X^T + X'^T)BY =$$
$$= X^T BY + X'^T BY = g(\vec{x},\vec{y}) + g(\vec{x}',\vec{y})$$

$$g(\alpha\vec{x},\vec{y}) = (\alpha X)^T BY = \alpha X^T BY = \alpha g(\vec{x},\vec{y})$$

Por sua vez,

$$g\left(\vec{e}_i,\vec{e}_j\right)=\begin{bmatrix}0 & \cdots & 0 & 1 & 0 & \cdots & 0\end{bmatrix}B\begin{bmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{bmatrix}=\begin{bmatrix}b_{i1} & \cdots & b_{ij} & \cdots & b_{in}\end{bmatrix}\begin{bmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{bmatrix}=b_{ij}$$

Finalmente, a unicidade de g resulta da primeira parte do enunciado, já provada, uma vez que as igualdades $g\left(\vec{e}_i,\vec{e}_j\right)=b_{ij}$ caracterizam inteiramente a forma g.

A matriz A, associada a uma forma bilinear f, nas condições anteriores, tem uma designação natural:

Definição 8-6 Nas condições anteriores, a matriz A, constituída pelos elementos $a_{ij}=f\left(\vec{e}_i,\vec{e}_j\right)$, é a **matriz da forma bilinear** f em relação à base considerada.

Do teorema anterior resulta também imediatamente o seguinte corolário:

Corolário 8-7 Fixando-se num espaço vectorial E, de dimensão finita, uma determinada base, fica estabelecida uma bijecção entre o conjunto $\mathbf{B}(E)$, formado por todas as formas bilineares definidas em E, e o conjunto $\mathbf{M}_{n\times n}(K)$, das matrizes quadradas, de ordem n, com elementos de K.

Demonstração:
Resulta imediatamente do teorema anterior.

Uma vez que a matriz de uma dada forma bilinear depende da base considerada, é natural procurar-se relacionar as matrizes de uma mesma forma bilinear em relação a duas bases distintas. É precisamente o que se faz no próximo teorema:

Teorema 8-8 Sejam A e A' as matrizes de uma certa forma bilinear f num espaço vectorial E, de dimensão n, em relação às bases $\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n$ e $\vec{e}\,'_1,\vec{e}\,'_2,\ldots,\vec{e}\,'_n$, respectivamente. Sendo P a matriz de mudança de base, tem-se $A'=P^TAP$.

Demonstração:
Atendendo à maneira como se definiu a matriz de uma forma bilinear em relação a certa base, basta observar que:

$$a'_{ij} = f\left(\vec{e}'_i, \vec{e}'_j\right) = f\left(\sum_{r=1}^{n} p_{ri}\vec{e}_r, \sum_{s=1}^{n} p_{sj}\vec{e}_j\right) = \sum_{r,s=1}^{n} p_{ri}p_{sj} f\left(\vec{e}_r, \vec{e}_s\right) = \sum_{r,s=1}^{n} p_{ri}a_{rs}p_{sj}$$

8-2 Formas bilineares simétricas

Algumas formas bilineares, que obedecem a propriedades especiais, merecem certo relevo. Assim, por exemplo, numa forma bilinear qualquer, os valores de $f\left(\vec{x},\vec{y}\right)$ e $f\left(\vec{y},\vec{x}\right)$ não têm qualquer relação especial entre si.

Por exemplo, se considerarmos a forma bilinear *f*, definida, no espaço vectorial real $\mathbb{R}^2$, pela matriz

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

em relação à base canónica do espaço, e se tomarmos os vectores $\vec{x} = (1,3)$ e $\vec{y} = (2,-1)$, vemos que

$$f\left(\vec{x},\vec{y}\right) = \begin{bmatrix} 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \end{bmatrix}\begin{bmatrix} 0 \\ 2 \end{bmatrix} = 6 \qquad f\left(\vec{y},\vec{x}\right) = \begin{bmatrix} 2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 & -1 \end{bmatrix}\begin{bmatrix} 7 \\ 15 \end{bmatrix} = -1$$

Salientaremos, contudo, as seguintes situações especiais:

Definição 8-9 Seja *f* uma forma bilinear num espaço vectorial *E*. Diz-se que a forma *f* é **simétrica** (resp.: **anti-simétrica**) quando se tem $f\left(\vec{x},\vec{y}\right) = f\left(\vec{y},\vec{x}\right)$ [resp.: $f\left(\vec{x},\vec{y}\right) = -f\left(\vec{y},\vec{x}\right)$], para vectores arbitrários de *E*.

Deve observar-se que, na sua maioria, as formas bilineares não são simétricas nem anti-simétricas. É o que acontece na forma bilinear em $\mathbb{R}^2$ de que nos servimos, como exemplo, antes da última definição.

No caso de E ser um espaço vectorial de dimensão finita, vimos acima que as formas bilineares em E estão associadas a matrizes. O facto de uma dada forma bilinear ser simétrica ou anti-simétrica pode deduzir-se do exame da respectiva matriz em relação a uma base qualquer, em virtude do seguinte teorema:

Teorema 8-10 Seja f uma forma bilinear num espaço vectorial de dimensão finita, definida por certa matriz A em relação a uma base fixa no espaço. A forma f é simétrica (resp.: anti-simétrica) se e só se a matriz A verifica a igualdade $A^T = A$ (resp.: $A^T = -A$).

Uma matriz tal que $A^T = A$ (resp.: $A^T = -A$) chama-se **matriz simétrica** (resp.: uma **matriz anti-simétrica**).

Demonstração:
Se f é simétrica, então, para quaisquer vectores da base fixa no espaço, tem-se $f(\vec{e}_i, \vec{e}_j) = f(\vec{e}_j, \vec{e}_i)$. Daqui resulta que, na matriz A, teremos $a_{ij} = a_{ji}$, pelo que $A^T = A$.

Reciprocamente, se $A^T = A$, então, para vectores arbitrários, temos

$$f(\vec{x}, \vec{y}) = X^T A Y = \left(X^T A Y\right)^T = Y^T A^T \left(X^T\right)^T = Y^T A X = f(\vec{y}, \vec{x})$$

Deve notar-se que a igualdade $X^T A Y = \left(X^T A Y\right)^T$ resulta trivialmente do facto de $X^T A Y$ ser uma matriz do tipo 1×1.

A demonstração é análoga para o caso das formas anti-simétricas.

Faremos em seguida um estudo pormenorizado das formas bilineares simétricas em espaços vectoriais de dimensão finita. O estudo que vamos fazer, porém, exigirá que se considerem espaços vectoriais sobre corpos que verifiquem uma condição especial.

Com efeito, necessitaremos de que, no corpo K considerado, se verifique o seguinte: $\forall \lambda \in K : \lambda \neq 0 \Rightarrow 2\lambda = \lambda + \lambda \neq 0$.

Esta condição é, obviamente, válida nos corpos mais utilizados, como $\mathbb{R}$ (corpo dos números reais), $\mathbb{C}$ (corpo dos números complexos) ou $\mathbb{Q}$ (corpo dos números racionais). Há, contudo, outros corpos em que esta propriedade não se verifica.

Por exemplo, no corpo formado pelo conjunto $\{0,1\}$, onde as operações de adição e multiplicação são definidas pelas tabelas

+	0	1
0	0	1
1	1	0

×	0	1
0	0	0
1	0	1

vemos que se tem $1 \neq 0$, mas $1 + 1 = 0$.

Para distinguir os corpos que verificam a propriedade em causa dos que a não verificam, introduzimos uma designação, que se insere em considerações mais gerais da Teoria dos Corpos, que não caberia aqui desenvolver:

Definição 8-11 Diz-se que um corpo K tem **característica diferente de 2** quando verifica a propriedade seguinte:

$$\forall \lambda \in K : \lambda \neq 0 \Rightarrow 2\lambda = \lambda + \lambda \neq 0$$

O resultado fundamental que nos propomos demonstrar acerca das formas bilineares simétricas, num espaço de dimensão finita sobre um corpo de característica diferente de 2, consiste na possibilidade de escolhermos uma base do espaço vectorial em relação à qual a expressão da forma bilinear se torne particularmente simples.

Em primeiro lugar, necessitaremos de dois lemas, o primeiro dos quais de carácter eminentemente técnico:

Lema 8-12 Seja E um espaço vectorial qualquer sobre um corpo K arbitrário, e seja θ um endomorfismo de E tal que $\theta^2 = \theta$. Então, tem-se

$$E = \theta(E) \oplus (1_E - \theta)(E)$$

Demonstração:

Dado um vector qualquer $\vec{x} \in E$, podemos escrevê-lo na forma

$$\vec{x} = \theta(\vec{x}) + \vec{x} - \theta(\vec{x}) = \theta(\vec{x}) + (1_E - \theta)(\vec{x}) \in \theta(E) + (1_E - \theta)(E)$$

o que mostra que $E = \theta(E) + (1_E - \theta)(E)$.

Por outro lado, tomemos um vector $\vec{z} \in \theta(E) \cap (1_E - \theta)(E)$. Teremos, então,

$$\vec{z} = \theta(\vec{a}) \text{ e } \vec{z} = (1_E - \theta)(\vec{b}) = \vec{b} - \theta(\vec{b}),$$

para certos vectores $\vec{a}, \vec{b} \in E$.

Daqui resulta que

$$\vec{z} = \theta(\vec{a}) = \theta^2(\vec{a}) = \theta(\theta(\vec{a})) = \theta(\vec{z}) = \theta(\vec{b} - \theta(\vec{b})) =$$
$$= \theta(\vec{b}) - \theta^2(\vec{b}) = \theta(\vec{b}) - \theta(\vec{b}) = \vec{0}$$

pelo que a soma $\theta(E) + (1_E - \theta)(E)$ é uma soma directa.

O segundo lema refere-se já a formas bilineares:

Lema 8-13

Seja f uma forma bilinear simétrica, não nula, num espaço vectorial E, de dimensão finita, sobre um corpo K, de característica diferente de 2. Então, existe um vector $\vec{u} \in E$ tal que $f(\vec{u},\vec{u}) \neq 0$.

Demonstração:
Fixemos em E uma base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$.

Caso exista, entre os vectores da base, um vector $\vec{e}_k$ tal que $f(\vec{e}_k,\vec{e}_k) \neq 0$, bastará, naturalmente, tomar $\vec{u} = \vec{e}_k$.

Suponhamos, então, que se tem $f(\vec{e}_k,\vec{e}_k) = 0$, para qualquer vector da base. Se designarmos por A a matriz da forma f em relação à base considerada, esta matriz não poderá ser nula, visto a forma não ser nula; os elementos da diagonal principal serão todos iguais a zero, mas haverá índices i, j, distintos, tais que $a_{ij} = f(\vec{e}_i,\vec{e}_j) \neq 0$.

Tomemos, neste caso, o vector $\vec{u} = \vec{e}_i + \vec{e}_j$. Tem-se:

$$\begin{aligned} f(\vec{u},\vec{u}) &= f(\vec{e}_i+\vec{e}_j, \vec{e}_i+\vec{e}_j) = \\ &= f(\vec{e}_i,\vec{e}_i) + f(\vec{e}_i,\vec{e}_j) + f(\vec{e}_j,\vec{e}_i) + f(\vec{e}_j,\vec{e}_j) = \\ &= \vec{0} + f(\vec{e}_i,\vec{e}_j) + f(\vec{e}_j,\vec{e}_i) + \vec{0} = \\ &= f(\vec{e}_i,\vec{e}_j) + f(\vec{e}_j,\vec{e}_i) = 2f(\vec{e}_i,\vec{e}_j) \quad \text{(por } f \text{ ser simétrica)} \end{aligned}$$

Atendendo a que $f(\vec{e}_i,\vec{e}_j) \neq 0$ e a que o corpo K tem característica diferente de 2, daqui se pode concluir que $f(\vec{u},\vec{u}) = 2f(\vec{e}_i,\vec{e}_j) \neq 0$.

Estamos agora em condições de enunciar e provar o resultado fundamental a que acima nos referíamos:

Teorema 8-14

Seja f uma forma bilinear simétrica, não nula, num espaço vectorial E, de dimensão finita, sobre um corpo K, de característica diferente de 2. Então, existe uma base $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$, do espaço E, tal que

a) $f(\vec{u}_k,\vec{u}_k) = \xi_k \neq 0$, para $k = 1,2,\ldots,r \leq n$

b) $f(\vec{u}_k,\vec{u}_k) = 0$, para $k = r+1,\ldots,n$

c) $f(\vec{u}_i,\vec{u}_j) = 0$, para $i,j = 1,2,\ldots,n,\ i \neq j$

isto é, tal que a matriz de f em relação a essa base seja do tipo

$$\begin{bmatrix} \xi_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \xi_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \xi_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

com $\xi_1, \xi_2, \ldots, \xi_r \neq 0$.

Demonstração:
Faremos a demonstração por indução na dimensão do espaço. Uma vez que o resultado é trivial quando $n = 1$, admitamos que seja válido para espaços de dimensão $n-1$ e tomemos, então, um espaço vectorial de dimensão n, nas condições do enunciado.

Atendendo às hipóteses e ao Lema 8-13, começamos por fixar em E um vector qualquer $\vec{u}_1$ tal que $f(\vec{u}_1, \vec{u}_1) = \xi_1 \neq 0$.

Definimos, então, uma aplicação $\theta : E \to E$, pondo, para cada vector $\vec{x} \in E$,

$$\theta(\vec{x}) = \frac{f(\vec{x}, \vec{u}_1)}{\xi_1} \vec{u}_1$$

É fácil verificar que θ é uma aplicação linear:

$$\theta(\alpha\vec{x} + \beta\vec{y}) = \frac{f(\alpha\vec{x} + \beta\vec{y}, \vec{u}_1)}{\xi_1} \vec{u}_1 =$$

$$= \frac{\alpha f(\vec{x}, \vec{u}_1) + \beta f(\vec{y}, \vec{u}_1)}{\xi_1} \vec{u}_1 = \left(\alpha \frac{f(\vec{x}, \vec{u}_1)}{\xi_1} + \beta \frac{f(\vec{y}, \vec{u}_1)}{\xi_1} \right) \vec{u}_1 =$$

$$= \alpha \frac{f(\vec{x}, \vec{u}_1)}{\xi_1} \vec{u}_1 + \beta \frac{f(\vec{y}, \vec{u}_1)}{\xi_1} \vec{u}_1 = \alpha\theta(\vec{x}) + \beta\theta(\vec{y})$$

Observando que

$$\theta(\vec{u}_1) = \frac{f(\vec{u}_1, \vec{u}_1)}{\xi_1} \vec{u}_1 = \frac{\xi_1}{\xi_1} \vec{u}_1$$

podemos verificar que

$$\theta^2(\vec{x}) = \theta(\theta(\vec{x})) = \theta\left(\frac{f(\vec{x},\vec{u}_1)}{\xi_1}\vec{u}_1\right) = \frac{f(\vec{x},\vec{u}_1)}{\xi_1}\theta(\vec{u}_1) = \frac{f(\vec{x},\vec{u}_1)}{\xi_1}\vec{u}_1 = \theta(\vec{x})$$

ou seja, que $\theta^2 = \theta$.

De acordo com o Lema 8-12, teremos, neste caso, $E = \theta(E) \oplus (1_E - \theta)(E)$.

Ora, por construção, $\theta(E) \subseteq \langle \vec{u}_1 \rangle$, e, por ser $\vec{0} \neq \vec{u}_1 = \theta(\vec{u}_1) \in \theta(E)$, podemos concluir que $\theta(E) = \langle \vec{u}_1 \rangle$ e que $E = \langle \vec{u}_1 \rangle \oplus (1_E - \theta)(E)$.

O subespaço $(1_E - \theta)(E)$ de E tem a dimensão $n-1$, pelo que, por hipótese de indução, a afirmação que estamos a estudar será válida para a restrição da forma bilinear f a esse subespaço, isto é, existirá em $(1_E - \theta)(E)$ uma base $\vec{u}_2,\ldots,\vec{u}_n$ tal que

a) $f(\vec{u}_k,\vec{u}_k) = \xi_k \neq 0$, para $k = 2,\ldots,r \leq n$

b) $f(\vec{u}_k,\vec{u}_k) = 0$, para $k = r+1,\ldots,n$

c) $f(\vec{u}_i,\vec{u}_j) = 0$, para $i,j = 2,\ldots,n,\ i \neq j$

Os vectores $\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_n$ constituem uma base de f. Para vermos que esta base satisfaz as condições desejadas, falta apenas verificar que, para qualquer índice $j = 2,\ldots,n$, se tem $f(\vec{u}_1,\vec{u}_j) = 0$.

Ora, $\vec{u}_j \in (1_E - \theta)(E)$, pelo que, para certo vector $\vec{w}_j \in E$, se terá

$$\vec{u}_j = (1_E - \theta)(\vec{w}_j) = \vec{w}_j - \theta(\vec{w}_j) = \vec{w}_j - \frac{f(\vec{w}_j,\vec{u}_1)}{\xi_1}\vec{u}_1$$

Consequentemente,

$$f(\vec{u}_1,\vec{u}_j) = f\left(\vec{u}_1, \vec{w}_j - \frac{f(\vec{w}_j,\vec{u}_1)}{\xi_1}\vec{u}_1\right) = f(\vec{u}_1,\vec{w}_j) - \frac{f(\vec{w}_j,\vec{u}_1)}{\xi_1}f(\vec{u}_1,\vec{u}_1) =$$

$$= f(\vec{u}_1,\vec{w}_j) - \frac{f(\vec{w}_j,\vec{u}_1)}{\xi_1}\xi_1 = f(\vec{u}_1,\vec{w}_j) - f(\vec{w}_j,\vec{u}_1) = 0$$

Tendo em vista as aplicações práticas, é muito interessante analisar o processo que permite a obtenção efectiva da base especial cuja existência acabámos de comprovar.

Nas condições do teorema anterior, seja A a matriz da forma bilinear f em relação a uma base arbitrária $\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n$.

Podemos supor que $f(\vec{e}_1,\vec{e}_1) \neq 0$, porque: se for $f(\vec{e}_1,\vec{e}_1) = 0$ mas houver um índice $i > 1$ tal que $f(\vec{e}_i,\vec{e}_i) \neq 0$, uma reordenação dos vectores da base permitirá considerar $\vec{e}_i$ o primeiro vector da base; se $f(\vec{e}_k,\vec{e}_k) = 0$, para qualquer índice k, bastará procurar índices i,j tais que $f(\vec{e}_i,\vec{e}_j) \neq 0$, tomar o vector $\vec{e}_i + \vec{e}_j$ e passar para uma base cujo primeiro vector seja precisamente $\vec{e}_i + \vec{e}_j$.

Admitindo, então, que $a_{11} = f(\vec{e}_1, \vec{e}_1) \neq 0$, tomamos o endomorfismo θ, do espaço E, definido por

$$\theta(\vec{x}) = \frac{f(\vec{x}, \vec{e}_1)}{a_{11}} \vec{e}_1$$

e restringimos a forma bilinear f ao subespaço

$$\begin{aligned}
(1_E - \theta)(E) &= \left\langle (1_E - \theta)(\vec{e}_1), (1_E - \theta)(\vec{e}_2), \dots, (1_E - \theta)(\vec{e}_n) \right\rangle = \\
&= \left\langle \vec{e}_1 - \theta(\vec{e}_1), \vec{e}_2 - \theta(\vec{e}_2), \dots, \vec{e}_n - \theta(\vec{e}_n) \right\rangle = \\
&= \left\langle \vec{e}_1 - \vec{e}_1, \vec{e}_2 - \theta(\vec{e}_2), \dots, \vec{e}_n - \theta(\vec{e}_n) \right\rangle = \\
&= \left\langle \vec{0}, \vec{e}_2 - \theta(\vec{e}_2), \dots, \vec{e}_n - \theta(\vec{e}_n) \right\rangle = \\
&= \left\langle \vec{e}_2 - \theta(\vec{e}_2), \dots, \vec{e}_n - \theta(\vec{e}_n) \right\rangle
\end{aligned}$$

Como o subespaço tem a dimensão $n-1$, os vectores $\vec{e}_2 - \theta(\vec{e}_2), \dots, \vec{e}_n - \theta(\vec{e}_n)$ constituem uma base de $(1_E - \theta)(E)$.

Ora, para cada $j = 2, \dots, n$, tem-se

$$\vec{e}_j - \theta(\vec{e}_j) = \vec{e}_j - \frac{f(\vec{e}_j, \vec{e}_1)}{a_{11}} \vec{e}_1 = \vec{e}_j - \frac{a_{j1}}{a_{11}} \vec{e}_1$$

A matriz de f em relação à base $\vec{e}_1, \vec{e}_2 - \frac{a_{21}}{a_{11}} \vec{e}_1, \dots, \vec{e}_n - \frac{a_{n1}}{a_{11}} \vec{e}_1$ será do tipo

$$A' = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{bmatrix} \quad \text{sendo} \quad B = \begin{bmatrix} b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots \\ b_{n2} & \cdots & b_{nn} \end{bmatrix}$$

a matriz da restrição de f ao subespaço $(1_E - \theta)(E)$ em relação à base $\vec{e}_2 - \frac{a_{21}}{a_{11}} \vec{e}_1, \dots, \vec{e}_n - \frac{a_{n1}}{a_{11}} \vec{e}_1$.

Para $i, j = 2, \dots, n$, tem-se

$$\begin{aligned}
b_{ij} &= f\left(\vec{e}_i - \frac{a_{i1}}{a_{11}} \vec{e}_1, \vec{e}_j - \frac{a_{j1}}{a_{11}} \vec{e}_1 \right) = \\
&= f(\vec{e}_i, \vec{e}_i) - \frac{a_{j1}}{a_{11}} f(\vec{e}_i, \vec{e}_1) - \frac{a_{i1}}{a_{11}} f(\vec{e}_1, \vec{e}_j) + \frac{a_{i1}}{a_{11}} \frac{a_{j1}}{a_{11}} f(\vec{e}_1, \vec{e}_1) = \\
&= a_{ii} - \frac{a_{j1}}{a_{11}} a_{i1} - \frac{a_{i1}}{a_{11}} a_{1j} + \frac{a_{i1}}{a_{11}} \frac{a_{j1}}{a_{11}} a_{11} = \frac{1}{a_{11}} \left(a_{11} a_{ij} - a_{1j} a_{i1} \right)
\end{aligned}$$

O processo é depois repetido até se chegar a uma matriz do tipo procurado.

Por exemplo, consideremos, no espaço vectorial $\mathbb{R}^3$, a forma bilinear definida, em relação à base canónica $\vec{e}_1, \vec{e}_2, \vec{e}_3$, pela matriz

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 3 & 4 \\ 2 & 4 & 1 \end{bmatrix}$$

Tem-se $f(\vec{e}_1, \vec{e}_1) = 1 \neq 0$.

A matriz da restrição de f em relação à base $\vec{e}_2 - \frac{a_{21}}{a_{11}}\vec{e}_1,\ \vec{e}_3 - \frac{a_{31}}{a_{11}}\vec{e}_1$, ou seja,

$$\vec{e}_2 - \frac{1}{1}\vec{e}_1,\quad \vec{e}_3 - \frac{2}{1}\vec{e}_1, \qquad \text{ou ainda} \qquad \vec{e}_2 - \vec{e}_1,\ \vec{e}_3 - 2\vec{e}_1$$

será

$$\frac{1}{1}\begin{bmatrix} 1\times 3 - 1\times 1 & 1\times 4 - 1\times 2 \\ 1\times 4 - 2\times 1 & 1\times 1 - 2\times 2 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & -3 \end{bmatrix}$$

Por sua vez, atendendo a que o primeiro elemento desta matriz não é nulo, consideremos a restrição de f em relação ao subespaço gerado por $(\vec{e}_3 - 2\vec{e}_1) - \frac{2}{2}(\vec{e}_2 - \vec{e}_1) = \vec{e}_3 - \vec{e}_1 - \vec{e}_2$; a sua matriz, em relação a essa base é

$$\frac{1}{2}[2\times(-3) - 2\times 2] = \frac{1}{2}[-10] = [-5]$$

Por conseguinte, a matriz da forma bilinear f em relação à base $\vec{e}_1, \vec{e}_2 - \vec{e}_1, \vec{e}_3 - \vec{e}_1 - \vec{e}_2$ é

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -5 \end{bmatrix}$$

Do Teorema 8-14 resultam, por sua vez, importantes consequências, que vamos passar a expor.

Corolário 8-15 Seja f uma forma bilinear simétrica, não nula, num espaço vectorial real E, de dimensão finita. Então, existe, no espaço E, uma base $\vec{u}_1, \vec{u}_2, ..., \vec{u}_n$, tal que a matriz de f em relação a essa base tenha o aspecto

$$\begin{bmatrix} \delta_1 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \delta_s & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & -\delta_{s+1} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & -\delta_{s+t} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \quad \text{com} \quad \delta_1, \delta_2, ..., \delta_{s+t} > 0$$

Demonstração:
Basta construir uma base nas condições descritas no Teorema 8-14, e, em seguida, reordenar os seus vectores, de modo que os coeficientes positivos precedam os coeficientes negativos.

Corolário 8-16 Nas condições do corolário anterior, existe em E uma base $\vec{u}'_1, \vec{u}'_2, ..., \vec{u}'_n$ em relação à qual a matriz da forma f é

$$\begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & -1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & -1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Demonstração:
Começamos por construir uma base $\vec{u}_1, \vec{u}_2, ..., \vec{u}_n$ nas condições do corolário anterior. Para cada índice $i = 1, 2, ..., n$, construímos o vector

$$\vec{u}'_i = \frac{1}{\sqrt{\delta_i}} \vec{u}_i$$

Os vectores $\vec{u}_1', \vec{u}_2', ..., \vec{u}_n'$, assim construídos, constituem uma nova base. Ora: Para $i, j = 1, 2, ..., n,\ i \neq j$, temos

$$f\left(\vec{u}_i', \vec{u}_j'\right) = f\left(\frac{1}{\sqrt{\delta_i}}\vec{u}_i, \frac{1}{\sqrt{\delta_j}}\vec{u}_j\right) = \frac{1}{\sqrt{\delta_i \delta_j}} f\left(\vec{u}_i, \vec{u}_j\right) = \frac{1}{\sqrt{\delta_i \delta_j}} 0 = 0$$

Para $j = i$, será

$$f\left(\vec{u}_i', \vec{u}_i'\right) = f\left(\frac{1}{\sqrt{\delta_i}}\vec{u}_i, \frac{1}{\sqrt{\delta_i}}\vec{u}_i\right) = \frac{1}{\sqrt{\delta_i^2}} f\left(\vec{u}_i, \vec{u}_i\right) = \frac{1}{\delta_i} f\left(\vec{u}_i, \vec{u}_i\right)$$

Mas

$$f\left(\vec{u}_i, \vec{u}_i\right) = \begin{cases} \delta_i, & \text{se} \quad i = 1, 2, ..., s \\ -\delta_i, & \text{se} \quad i = s+1, ..., s+t \\ 0, & \text{se} \quad i > s+t \end{cases}$$

pelo que

$$f\left(\vec{u}_i, \vec{u}_i\right) = \begin{cases} \frac{1}{\delta_i}\delta_i = 1, & \text{se} \quad i = 1, 2, ..., s \\ \frac{1}{\delta_i}\left(-\delta_i\right) = -1, & \text{se} \quad i = s+1, ..., s+t \\ \frac{1}{\delta_i} 0 = 0, & \text{se} \quad i > s+t \end{cases}$$

Por consequência, a base $\vec{u}_1', \vec{u}_2', ..., \vec{u}_n'$ satisfaz as condições requeridas.

Corolário 8-17 **Lei de Sylvester**

Nas condições do corolário anterior, os números s e t são fixos, para cada forma f, não dependendo da base construída.

Demonstração:

Admitamos que, em relação a bases $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ e $\vec{w}_1, \vec{w}_2, ..., \vec{w}_n$, a forma f tem as expressões

$$f\left(\vec{x}, \vec{y}\right) = x_1 y_1 + \cdots + x_s y_s - x_{s+1} y_{s+1} - \cdots - x_{s+t} y_{s+t}$$

e

$$f\left(\vec{x}, \vec{y}\right) = x_1' y_1' + \cdots + x_p' y_p' - x_{p+1}' y_{p+1}' - \cdots - x_{p+q}' y_{p+q}'$$

respectivamente.

Comecemos por reparar que $s+t$ é a característica da matriz D que representa a forma f em relação à base $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$, enquanto $p+q$ é a característica da matriz D' que representa a forma f em relação à base $\vec{w}_1, \vec{w}_2, ..., \vec{w}_n$. Ora, sendo P a respectiva matriz de mudança de base, tem-se $D' = P^T DP$, e, uma vez que a matriz de mudança de base é invertível,

$$p+q = r\left(D'\right) = r\left(P^T DP\right) = r\left(D\right) = s+t$$

Suponhamos que se tem $t < q$.

Seja F o subespaço vectorial de E, definido, em relação à base $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$, pelas equações cartesianas

$$\begin{cases} x_{s+1} = 0 \\ \vdots \\ x_{s+t} = 0 \end{cases}$$

e seja G o subespaço vectorial de E, definido, em relação à base $\vec{w}_1, \vec{w}_2, ..., \vec{w}_n$, pelas equações cartesianas

$$\begin{cases} x'_1 = 0 \\ \vdots \\ x'_p = 0 \\ x'_{p+q+1} = 0 \\ \vdots \\ x'_n = 0 \end{cases}$$

Tem-se

$$\dim\left(F\right) = n-t \text{ e } \dim\left(G\right) = n - \left(p + \left(n-p-q\right)\right) = q$$

Portanto,

$$\dim\left(F+G\right) = \dim\left(F\right) + \dim\left(G\right) - \dim\left(F \cap G\right) = n-t+q-\dim\left(F \cap G\right)$$

Se fosse $F \cap G = \left\{\vec{0}\right\}$, teríamos $\dim\left(F+G\right) = n + \left(q-t\right) > n$, uma vez que supusemos $t < q$. Como tal é impossível, será $F \cap G \neq \left\{\vec{0}\right\}$.

Tomemos então um vector não nulo $\vec{z} \in F \cap G$.

Mas, sendo $\vec{z} \in F$, então, usando a expressão de f em relação à base $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$, e atendendo ao modo como o subespaço vectorial F foi definido, vemos que $f(\vec{z}, \vec{z}) \geq 0$. Por sua vez, sendo $\vec{z} \in G$, então, usando a expressão de f em relação à base $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_n$, e atendendo ao modo como o subespaço vectorial G foi definido, vemos que $f(\vec{z}, \vec{z}) < 0$, o que é absurdo.

Assim, concluímos que tem de ser $t \geq q$. Analogamente se provaria que $q \geq t$, pelo que, finalmente, $t = q$.

Como verificámos acima que $p + q = s + t$, daí resulta que também $p = s$.

A lei de Sylvester tem uma importante aplicação na classificação das formas bilineares simétricas. Para a estudarmos, começamos por introduzir a seguinte definição:

Definição 8-18 Seja f uma forma bilinear, num espaço vectorial E sobre um corpo K qualquer. Chama-se **forma diagonal** ou **forma quadrática** associada a f à aplicação $Q_f : E \to K$, definida por $Q_f(\vec{x}) = f(\vec{x}, \vec{x})$.

Assim, por exemplo, consideremos a forma bilinear f, no espaço vectorial real $\mathbb{R}^3$, definida por

$$f(\vec{x}, \vec{y}) = 3x_1y_1 - 7x_1y_2 + x_1y_3 + 3x_2y_1 - 4x_2y_2 + 6x_2y_3 + x_3y_1 + 2x_3y_2 - x_3y_3$$

A forma quadrática associada a f é

$$Q_f(\vec{x}) = 3x_1^2 - 4x_1x_2 + 2x_1x_3 - 4x_2^2 + 8x_2x_3 - x_3^2$$

Por sua vez, a forma bilinear g definida por

$$g(\vec{x}, \vec{y}) = 3x_1y_1 - 4x_1y_2 + 3x_1y_3 - 4x_2y_2 + 4x_2y_3 - 2x_3y_1 + 4x_3y_2 - x_3y_3$$

tem precisamente a mesma forma quadrática que f:

$$Q_g(\vec{x}) = Q_f(\vec{x})$$

O mesmo acontece com a forma bilinear p definida por

$$p(\vec{x}, \vec{y}) = 3x_1y_1 - 2x_1y_2 + x_1y_3 - 2x_2y_1 - 4x_2y_2 + 4x_2y_3 + x_3y_1 + 4x_3y_2 - x_3y_3 \; :$$

$$Q_g(\vec{x}) = Q_f(\vec{x}) = Q_p(\vec{x})$$

De modo geral, no caso de termos duas formas bilineares f e g, num espaço vectorial de dimensão finita n, definidas pelas matrizes A e B, em relação a certa base do espaço, as suas formas quadráticas coincidirão se e só se

a) $a_{ii} = b_{ii} \quad (i = 1, 2, \ldots, n)$

b) $a_{ij} + a_{ji} = b_{ij} + b_{ji} \quad (i, j = 1, 2, \ldots, n; \; i \neq j)$

conforme resulta imediatamente das definições dadas.

Em particular, a forma bilinear f^t definida, em relação à mesma base, pela matriz A^T, tem a mesma forma quadrática. De resto, vemos imediatamente que, para qualquer vector, se tem

$$Q_{f^t}(\vec{x}) = X^T A^T X = \left(X^T A^T X\right)^T = X^T A X = Q_f(\vec{x})$$

De entre as formas bilineares que têm uma mesma forma quadrática associada, num espaço de dimensão finita, podemos, muitas vezes, salientar uma:

Teorema 8-19 Dada uma forma bilinear f, num espaço vectorial E, de dimensão n, sobre um corpo K, de característica diferente de 2, existe uma e uma só forma bilinear simétrica p que tem a mesma forma quadrática que f.

Demonstração:

Suponhamos que a forma f tem uma certa matriz A, em relação a uma base fixa no espaço.

Se uma forma bilinear p, definida, em relação à mesma base, por uma matriz B, tem a mesma forma quadrática que f, então, como observámos acima, os elementos das duas matrizes, A e B, verificam as condições

a) $a_{ii} = b_{ii} \quad (i = 1, 2, \ldots, n)$

b) $a_{ij} + a_{ji} = b_{ij} + b_{ji} \quad (i, j = 1, 2, \ldots, n;\ i \neq j)$

Sendo p uma forma bilinear simétrica, a sua matriz terá de ser, como vimos oportunamente, uma matriz simétrica, pelo que as igualdades anteriores se transformam em

a) $a_{ii} = b_{ii} \quad (i = 1, 2, \ldots, n)$

b) $a_{ij} + a_{ji} = 2b_{ij} \quad (i, j = 1, 2, \ldots, n;\ i \neq j)$

Note-se que, num corpo de característica diferente de 2, podemos encarar "2" como representando um elemento do corpo; com efeito, designando por 1 o elemento neutro da multiplicação em K, podemos considerar $2 = 1 + 1 = 2 \times 1 \neq 0$. Para qualquer escalar α, temos $2\alpha = \alpha + \alpha = (1 + 1)\,\alpha = 2\alpha$, igualdades em que o símbolo "2" foi usado em duas acepções distintas!

Nestas condições, a matriz B fica inteiramente definida pelas igualdades

a) $b_{ii} = a_{ii} \quad (i = 1, 2, \ldots, n)$

b) $b_{ij} = \dfrac{a_{ij} + a_{ji}}{2} \quad (i, j = 1, 2, \ldots, n;\ i \neq j)$

Globalmente, podemos resumir estas condições escrevendo

$$B = \frac{1}{2}\left(A + A^T\right) = \frac{A + A^T}{2}$$

Ora, a forma bilinear p, definida, em relação à base considerada, pela matriz

$$\frac{A+A^T}{2}$$

é, de facto, uma forma bilinear simétrica, uma vez que

$$\left(\frac{A+A^T}{2}\right)^T = \frac{1}{2}\left(A+A^T\right)^T = \frac{1}{2}\left(A^T+A\right) = \frac{A+A^T}{2}$$

e a sua forma quadrática é, por construção, igual à de f.

A forma bilinear simétrica que tem a mesma forma quadrática que uma forma bilinear dada receberá uma designação específica:

Definição 8-20 Seja f uma forma bilinear, num espaço vectorial E, de dimensão n, sobre um corpo K, de característica diferente de 2. À (única) forma bilinear simétrica p que tem a mesma forma quadrática que f chama-se **forma polar** de f.

Por sua vez, dos resultados anteriores resulta imediatamente o seguinte:

Corolário 8-21 Seja f uma forma bilinear, num espaço vectorial E, de dimensão n, sobre um corpo K, de característica diferente de 2. Existe uma base em E em relação à qual a expressão da forma quadrática de f é do tipo

$$Q_f(\vec{x}) = x_1 y_1 + \cdots + x_s y_s - x_{s+1} y_{s+1} - \cdots - x_{s+t} y_{s+t}$$

com $s+t \leq n$, sendo os números s e t independentes da base considerada.

Demonstração:
Basta tomar a forma polar de f e aplicar os resultados obtidos nos Corolários 8-16 e 8-17.

Introduzimos, agora, a classificação das formas bilineares, a que já nos referimos:

Definição 8-22 Seja f uma forma bilinear num espaço vectorial real. Diz-se que a forma f é:

a) **definida positiva** quando se tem:

a1) $f(\vec{x},\vec{x}) \geq 0, \forall \vec{x} \in E$

a2) $f(\vec{x},\vec{x}) = 0 \Rightarrow \vec{x} = \vec{0}$

b) **definida negativa** quando se tem:

b1) $f(\vec{x},\vec{x}) \leq 0, \forall \vec{x} \in E$

b2) $f(\vec{x},\vec{x}) = 0 \Rightarrow \vec{x} = \vec{0}$

c) **semidefinida positiva** quando se tem:

c1) $f(\vec{x},\vec{x}) \geq 0, \forall \vec{x} \in E$

c2) $\exists \vec{x} \neq \vec{0} : f(\vec{x},\vec{x}) = 0$

d) **semidefinida negativa** quando se tem:

d1) $f(\vec{x},\vec{x}) \leq 0, \forall \vec{x} \in E$

d2) $\exists \vec{x} \neq \vec{0} : f(\vec{x},\vec{x}) = 0$

e) **indefinida** quando se tem:

$\exists \vec{x},\vec{y} : f(\vec{x},\vec{x}) > 0, f(\vec{y},\vec{y}) < 0$

Para classificarmos uma dada forma bilinear, bastará utilizar as conclusões anteriores, da seguinte maneira:

Teorema 8-23 Seja f uma forma bilinear, num espaço vectorial real E, de dimensão n. Fixe-se em E uma base em relação à qual a expressão da forma quadrática de f seja

$$Q_f(\vec{x}) = x_1 y_1 + \cdots + x_s y_s - x_{s+1} y_{s+1} - \cdots - x_{s+t} y_{s+t}$$

com $s + t \leq n$. Então, a forma bilinear f é

a) definida positiva se e só se $s = n$ (pelo que $t = 0$)

b) definida negativa se e só se $s = 0,\ t = n$

c) semidefinida positiva se e só se $s < n,\ t = 0$

d) semidefinida negativa se e só se $s = 0,\ t < n$

e) indefinida se e só se $s \neq 0,\ t \neq 0$

Demonstração:

Resulta imediatamente das definições e do Corolário 8-21.

8-3 Aplicações multilineares

Muitas das considerações que ficaram feitas a respeito de aplicações e formas bilineares podem, na verdade, estender-se a aplicações com qualquer número de variáveis. Alguns pormenores, é claro, são específicos das formas bilineares, entre eles a representação matricial, que resulta fundamentalmente da circunstância de os coeficientes da expressão de uma forma bilinear dependerem de dois índices (para o caso de os coeficientes dependerem de três índices, por exemplo, seria necessário, para fazer alguma coisa parecida, lançar mão de uma espécie de "matrizes a três dimensões", ou seja, uma espécie de "cubos", em que n^3 escalares se dispusessem em planos paralelos às faces...).

Assim, começamos por dar uma definição muito geral:

Definição 8-24 Sejam E', E'',..., $E^{(k)}$, $E^{(k+1)}$ espaços vectoriais sobre um certo corpo K. Uma aplicação $f : E' \times E'' \times \cdots \times E^{(k)} \to E^{(k+1)}$ é uma **aplicação k-linear**, quando, para vectores quaisquer $\vec{x}_i, \vec{y}_i \in E^{(i)}$, com $i = 1,2,...,k$, e escalares quaisquer, se tem:

a) $f(\vec{x}_1,...,\vec{x}_{i-1},\vec{x}_i + \vec{y}_i,\vec{x}_{i+1},...,\vec{x}_k) =$

$$= f(\vec{x}_1,...,\vec{x}_{i-1},\vec{x}_i,\vec{x}_{i+1},...,\vec{x}_k) + f(\vec{x}_1,...,\vec{x}_{i-1},\vec{y}_i,\vec{x}_{i+1},...,\vec{x}_k)$$

b) $f(\vec{x}_1,...,\vec{x}_{i-1},\alpha\vec{x}_i,\vec{x}_{i+1},...,\vec{x}_k) = \alpha f(\vec{x}_1,...,\vec{x}_{i-1},\vec{x}_i,\vec{x}_{i+1},...,\vec{x}_k)$

Usam-se, de modo geral, os prefixos habituais: uma aplicação 1-linear (resp.: 2-linear, 3-linear, 4-linear, etc.) chama-se linear (resp.: **bilinear**, **trilinear**, **tetralinear**, etc.). Sempre que $k \geq 2$, uma aplicação k-linear diz-se **multilinear**.

Em particular, quando $E' = E'' = \cdots = E^{(k)} = E$ e $E^{(k+1)} = K$, uma aplicação multilinear $f : E \times E \times \cdots \times E = E^k \to K$ diz-se uma **forma multilinear**.

É fácil provar que se verificam, para aplicações e formas multilineares, propriedades que estendem, de modo natural, as que estudámos para o caso particular das aplicações e das formas bilineares. Por exemplo:

Teorema 8-25 Seja $f : E' \times E'' \times \cdots \times E^{(k)} \to E^{(k+1)}$ uma forma k-linear. Fixemos $\vec{x}_i \in E^{(i)}$, com $i = 1,2,...,k$, $i \neq j$, e definamos uma aplicação $f_j : E^{(j)} \to E^{(k+1)}$ pondo, para cada vector

$$\vec{v}_j \in E^{(j)},\ f_j(\vec{v}_j) = f(\vec{x}_1,...,\vec{x}_{j-1},\vec{x}_j,\vec{x}_{j+1},...,\vec{x}_k)$$

As aplicações f_j ($j = 1,2,...,k$) são aplicações lineares.

Demonstração:
Resulta imediatamente das definições.

Corolário 8-26 Uma aplicação multilinear qualquer dá a imagem $\vec{0}$ sempre que alguma das variáveis toma o valor $\vec{0}$.

Demonstração:
Resulta trivialmente do teorema anterior.

Podemos agora demonstrar o seguinte:

Teorema 8-27 Seja E um espaço vectorial de dimensão finita, com uma base fixa $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$. Uma forma k-linear $f : E^k \to K$ fica perfeitamente determinada pelos n^k escalares $f\left(\vec{e}_{i_1}, \vec{e}_{i_2}, ..., \vec{e}_{i_k}\right)$, onde $i_1, i_2, ..., i_k$ tomam todos os valores de 1 a n.

Reciprocamente, dados n^k escalares arbitrários $a_{i_1 i_2 \dots i_k}$ ($i_1, i_2, ..., i_k$ variando de 1 a n), existe uma e uma só forma k-linear em E, tal que

$$f\left(\vec{e}_{i_1}, \vec{e}_{i_2}, ..., \vec{e}_{i_k}\right) = a_{i_1 i_2 \dots i_k}$$

Demonstração:

A demonstração é análoga à das propriedades correspondentes que estudámos para formas bilineares. Em particular, se representarmos cada vector do espaço E pelas suas componentes,

$$\vec{x}_j = \sum_{i_j=1}^{n} x_{i_j j} \vec{e}_{i_j} \quad (j = 1, 2, ..., k)$$

em relação à base considerada, teremos

$$f\left(\vec{x}_1, \vec{x}_2, ..., \vec{x}_k\right) = \sum_{i_1, i_2, ..., i_k=1}^{n} x_{i_1 1} x_{i_2 2}, ..., x_{i_k k} f\left(\vec{e}_{i_1}, \vec{e}_{i_2}, ..., \vec{e}_{i_k}\right) = \sum_{i_1, i_2, ..., i_k=1}^{n} a_{i_1 i_2 \dots i_k} x_{i_1 1} x_{i_2 2}, ..., x_{i_k k}$$

Um outro conceito que se pode estender, de modo natural, a formas multilineares quaisquer, é o de "diagonal":

Definição 8-28 Dada uma forma k-linear f, num espaço vectorial E, chama-se **forma diagonal** associada a f à aplicação $D_f : E \to K$ definida por

$$D_f(\vec{x}) = f(\vec{x}, \vec{x}, ..., \vec{x})$$

Também a classificação que introduzimos para formas bilineares se pode aplicar a formas multilineares:

Definição 8-29 Seja f uma forma multilinear num espaço vectorial real. Diz-se que a forma f é:

a) **definida positiva** quando se tem:

a1) $D_f \geq 0, \ \forall \vec{x} \in E$

a2) $D_f(\vec{x}) = 0 \Rightarrow \vec{x} = \vec{0}$

b) **definida negativa** quando se tem:

b1) $D_f(\vec{x}) \leq 0, \ \forall \vec{x} \in E$

b2) $D_f(\vec{x}) = 0 \Rightarrow \vec{x} = \vec{0}$

c) **semidefinida positiva** quando se tem:

c1) $D_f(\vec{x}) \geq 0, \ \forall \vec{x} \in E$

c2) $\exists \vec{x} \neq \vec{0} : D_f(\vec{x}) = 0$

d) **semidefinida negativa** quando se tem:

d1) $D_f(\vec{x}) \leq 0, \ \forall \vec{x} \in E$

d2) $\exists \vec{x} \neq \vec{0} : D_f(\vec{x}) = 0$

e) **indefinida** quando se tem:

$\exists \vec{x}, \vec{y} : D_f(\vec{x}) > 0, \ D_f(\vec{y}) < 0$

8-4 Formas multilineares alternadas

Quando estudámos as formas bilineares, mereceram-nos especial atenção as formas bilineares simétricas, que eram aquelas cujos valores não dependiam da ordem das variáveis, isto é, aquelas em que, para vectores arbitrários, se tivesse $f(\vec{x},\vec{y}) = f(\vec{y},\vec{x})$.

Este conceito pode, evidentemente, aplicar-se a formas multilineares quaisquer:

Definição 8-30 Uma forma k-linear f diz-se **simétrica** quando se tem

$$f(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_k) = f\left(\vec{x}_{i_1},\vec{x}_{i_2},\ldots,\vec{x}_{i_k}\right)$$

para qualquer permutação $i_1, i_2, \ldots, i_k$ dos índices $1,2,\ldots,k$.

Já a noção de "forma anti-simétrica" não é tão fácil de generalizar. Com efeito, uma forma bilinear anti-simétrica satisfaz $f(\vec{x},\vec{y}) = -f(\vec{y},\vec{x})$, ou seja, ao trocar entre si a posição das variáveis $\vec{x}$ e $\vec{y}$, o valor fornecido pela forma passa ao simétrico. Mas, no caso de termos mais que duas variáveis, não faria sentido pensar em casos em que uma troca na posição de duas variáveis fizesse passar o valor da forma ao simétrico; na verdade, se pensarmos, por exemplo, numa forma trilinear, essa condição conduzir-nos-ia às igualdades

$$f(\vec{x},\vec{y},\vec{z}) = -f(\vec{y},\vec{x},\vec{z}) = -\left(-f(\vec{y},\vec{z},\vec{x})\right) = f(\vec{y},\vec{z},\vec{x})$$

que mostram que a passagem de $(\vec{x},\vec{y},\vec{z})$ para $(\vec{y},\vec{z},\vec{x})$ não modificaria o valor da forma.

Por estas razões, definiremos o conceito de "forma multilinear anti-simétrica" do seguinte modo:

Definição 8-31 Uma forma k-linear f diz-se **anti-simétrica** quando, para qualquer permutação $i_1, i_2, \ldots, i_k$ dos índices $1, 2, \ldots, k$, se tem

$$f\left(\vec{x}_{i_1}, \vec{x}_{i_2}, \ldots, \vec{x}_{i_k}\right) = \begin{cases} f\left(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\right), & \text{se } i_1, i_2, \ldots, i_k \text{ é uma permutação par} \\ -f\left(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\right), & \text{se } i_1, i_2, \ldots, i_k \text{ é uma permutação ímpar} \end{cases}$$

Por sua vez, um tipo de forma multilinear que, conforme veremos, é muito próximo das formas anti-simétricas, é o seguinte:

Definição 8-32 Uma forma k-linear f, num certo espaço vectorial, diz-se **alternada** quando se tem $f(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k) = 0$, sempre que, entre os vectores $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k$, haja dois iguais.

Podemos, desde já, demonstrar o seguinte:

Teorema 8-33 Toda a forma multilinear alternada é anti-simétrica.

Demonstração:

Seja f uma forma k-linear alternada, num certo espaço vectorial E.

Tomando vectores $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k$ quaisquer, teremos, por definição,

$$f\left(\vec{x}_1, \ldots, \vec{x}_{i-1}, \vec{x}_i + \vec{x}_j, \vec{x}_{i+1}, \ldots, \vec{x}_{j-1}, \vec{x}_i + \vec{x}_j, \vec{x}_{j+1}, \ldots, \vec{x}_k\right) = 0$$

quaisquer que sejam os índices distintos i e j.

Atendendo às propriedades das formas multilineares, a igualdade anterior pode apresentar-se sob a forma

$$f\left(\ldots, \vec{x}_i, \ldots, \vec{x}_i, \ldots\right) + f\left(\ldots, \vec{x}_i, \ldots, \vec{x}_j, \ldots\right) + f\left(\ldots, \vec{x}_j, \ldots, \vec{x}_i, \ldots\right) + f\left(\ldots, \vec{x}_j, \ldots, \vec{x}_j, \ldots\right) = 0$$

Mas, por sua vez,

$$f\left(\ldots, \vec{x}_i, \ldots, \vec{x}_i, \ldots\right) = f\left(\ldots, \vec{x}_j, \ldots, \vec{x}_j, \ldots\right) = 0$$

pelo que

$$f\left(\ldots, \vec{x}_i, \ldots, \vec{x}_j, \ldots\right) + f\left(\ldots, \vec{x}_j, \ldots, \vec{x}_i, \ldots\right) = 0$$

ou ainda

$$f\left(\ldots, \vec{x}_i, \ldots, \vec{x}_j, \ldots\right) = -f\left(\ldots, \vec{x}_j, \ldots, \vec{x}_i, \ldots\right)$$

Isto significa que a troca de posição (transposição) entre duas das variáveis faz o valor da forma passar ao simétrico.

Ora, como se sabe, uma permutação ímpar é a composição de um número ímpar de transposições, pelo que podemos imediatamente concluir que a forma considerada é anti-simétrica.

A afirmação recíproca da anterior é também válida, em certas condições:

Teorema 8-34 Toda a forma multilinear anti-simétrica, num espaço vectorial sobre um corpo de característica diferente de 2, é alternada.

Demonstração:
Dada uma forma multilinear anti-simétrica *f*, num espaço vectorial sobre um corpo de característica diferente de 2, consideremos os vectores $\vec{x}_1, \vec{x}_2, ..., \vec{x}_k$ e suponhamos que se tem, por exemplo, $\vec{x}_i = \vec{x}_j$, para certos índices distintos i e j.

Uma vez que os dois vectores são iguais, podemos, evidentemente, escrever

$$f\left(..., \vec{x}_i, ..., \vec{x}_j, ...\right) = f\left(..., \vec{x}_j, ..., \vec{x}_i, ...\right)$$

Mas, sendo a forma em causa anti-simétrica, será

$$f\left(..., \vec{x}_i, ..., \vec{x}_j, ...\right) = -f\left(..., \vec{x}_j, ..., \vec{x}_i, ...\right)$$

Daqui resulta que

$$f\left(..., \vec{x}_i, ..., \vec{x}_j, ...\right) = -f\left(..., \vec{x}_i, ..., \vec{x}_j, ...\right)$$

ou seja,

$$2f\left(..., \vec{x}_i, ..., \vec{x}_j, ...\right) = 0$$

Como o corpo tem característica diferente de 2, resulta daqui que

$$f\left(..., \vec{x}_i, ..., \vec{x}_j, ...\right) = 0$$

pelo que a forma *f* é alternada.

As formas multilineares alternadas (que, no caso de se trabalhar com um corpo de característica diferente de 2, são as mesmas que as formas multilineares anti-simétricas, como acabámos de ver) satisfazem propriedades especiais, como vamos ver, começando por um lema:

Lema 8-35 Seja f uma forma k-linear alternada, num espaço vectorial E. O valor de $f\left(\vec{x}_1, \vec{x}_2, ..., \vec{x}_k\right)$ não se altera se a um dos vectores $\vec{x}_1, \vec{x}_2, ..., \vec{x}_k$ somarmos uma combinação linear dos restantes.

Demonstração:
Dados os vectores $\vec{x}_1, \vec{x}_2, ..., \vec{x}_k$, temos

$$f\left(\vec{x}_1, \vec{x}_2, ..., \vec{x}_i + \sum_{j \neq i} \alpha_j \vec{x}_j, ..., \vec{x}_k\right) = f\left(\vec{x}_1, ..., \vec{x}_i, ..., \vec{x}_k\right) + \sum_{j \neq i} \alpha_j f\left(\vec{x}_1, ..., \vec{x}_j, ..., \vec{x}_k\right)$$

Cada parcela de

$$\sum_{j \neq i} \alpha_j f\left(\vec{x}_1, ..., \vec{x}_j, ..., \vec{x}_k\right)$$

é nula, já que a forma considerada é alternada, e, por construção, o vector $\vec{x}_j$ aparece sempre duas vezes em $f\left(\vec{x}_1, ..., \vec{x}_j, ..., \vec{x}_k\right)$: uma vez na posição número i; outra, na posição número j.

Por consequência,

$$f\left(\vec{x}_1, \vec{x}_2, ..., \vec{x}_i + \sum_{j \neq i} \alpha_j \vec{x}_j, ..., \vec{x}_k\right) = f\left(\vec{x}_1, ..., \vec{x}_i, ..., \vec{x}_k\right)$$

Aplicando o lema anterior, podemos demonstrar o seguinte:

Teorema 8-36 Uma forma k-linear f, num espaço vectorial E, é alternada se e só se

$$f\left(\vec{x}_1, \vec{x}_2, ..., \vec{x}_k\right) = 0$$

sempre que os vectores $\vec{x}_1, \vec{x}_2, ..., \vec{x}_k$ são linearmente dependentes.

Demonstração:
A condição é obviamente suficiente, uma vez que se, entre os vectores $\vec{x}_1, \vec{x}_2, ..., \vec{x}_k$, houver dois iguais, esses vectores $\vec{x}_1, \vec{x}_2, ..., \vec{x}_k$ serão linearmente dependentes.

Para vermos que a condição é necessária, admitamos que a forma multilinear f é alternada e suponhamos que os vectores $\vec{x}_1, \vec{x}_2, ..., \vec{x}_k$ são linearmente dependentes. Neste caso, algum deles será combinação linear dos restantes. Sem perda de generalidade, podemos admitir que é o primeiro deles que se escreve à custa dos restantes [já que uma mudança na ordem dos vectores afectaria apenas o sinal de $f\left(\vec{x}_1, \vec{x}_2, ..., \vec{x}_k\right)$, visto a forma f ser anti-simétrica]:

$$\vec{x}_1 = \sum_{j=2}^{n} \gamma_j \vec{x}_j$$

Tendo em atenção o Lema 8-35, vemos que

$$f(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_k)=f\left(\vec{0}+\sum_{j=2}^{n}\gamma_j\vec{x}_j,\vec{x}_2,\ldots,\vec{x}_k\right)=f\left(\vec{0},\vec{x}_2,\ldots,\vec{x}_k\right)=0$$

No que se segue, analisaremos uma situação ainda mais particular:

Teorema 8-37 Seja f uma forma n-linear alternada, num espaço vectorial E, de dimensão n. Fixando em E uma base $\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n$, a forma f fica perfeitamente determinada pelo valor $f(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n)$, sendo a forma nula se e só se $f(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n)=0$.

Demonstração:
No Teorema 8-27, vimos que a forma n-linear $f:E^n\to K$ fica perfeitamente determinada pelos n^n escalares $a_{i_1i_2,\ldots,i_n}=f\left(\vec{e}_{i_1},\vec{e}_{i_2},\ldots,\vec{e}_{i_n}\right)$, onde $i_1,i_2,\ldots,i_n$ tomam todos os valores de 1 a n, tendo-se

$$f(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n)=\sum_{i_1,i_2,\ldots,i_k=1}^{n}a_{i_1i_2,\ldots,i_k}x_{i_11}x_{i_22},\ldots,x_{i_kk}$$

No caso de a forma ser alternada, teremos $a_{i_1i_2\ldots i_n}=0$ sempre que, entre os índices $i_1,i_2,\ldots,i_n$ haja dois iguais. Por consequência, bastará, neste caso, considerar, na expressão de $f(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n)$, as parcelas em que aqueles índices constituam uma permutação dos números 1,2,…,n:

$$f(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n)=\sum_{\substack{i_1,i_2,\ldots,i_k\\ \text{permutações}\\ \text{de }1,2,\ldots,n}}a_{i_1i_2\ldots i_k}x_{i_11}x_{i_22},\ldots,x_{i_kk}$$

Ora, f é também anti-simétrica, pelo que, para cada permutação $i_1,i_2,\ldots,i_n$ dos números 1,2,…,n, se tem

$$a_{i_1i_2\ldots i_n}=f\left(\vec{e}_{i_1},\vec{e}_{i_2},\ldots,\vec{e}_{i_n}\right)=(-1)^{\sigma}f(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n)$$

com

$$\sigma=\begin{cases}0, & \text{se } i_1,i_2,\ldots,i_n \text{ é uma permutação par}\\ 1, & \text{se } i_1,i_2,\ldots,i_n \text{ é uma permutação ímpar}\end{cases}$$

Portanto,

$$f\left(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n\right)=\sum_{\substack{i_1,i_2,\ldots,i_k\\ \text{permutações}\\ \text{de } 1,2,\ldots,n}}(-1)^{\sigma}\,x_{i_1 1}x_{i_2 2},\ldots,x_{i_k k}\,f\left(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\right)=$$

$$=\sum_{\substack{i_1,i_2,\ldots,i_k\\ \text{permutações}\\ \text{de } 1,2,\ldots,n}}(-1)^{\sigma}\,x_{i_1 1}x_{i_2 2},\ldots,x_{i_k k}\,f\left(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\right)=$$

$$=\left(\sum_{\substack{i_1,i_2,\ldots,i_k\\ \text{permutações}\\ \text{de } 1,2,\ldots,n}}(-1)^{\sigma}\,x_{i_1 1}x_{i_2 2},\ldots,x_{i_k k}\right)f\left(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\right)$$

Do teorema anterior resultam, trivialmente, os seguintes corolários:

Corolário 8-38 Nas condições do Teorema 8-37, e supondo *f* não nula, os vectores $\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_n\in E$ são linearmente independentes se e só se $f\left(\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_n\right)\neq 0$.

Demonstração:
A condição é suficiente, em virtude do Teorema 8-36, e é necessária em face de 8-37.

Corolário 8-39 Nas condições do Teorema 8-37, e tomando a matriz M cujas colunas são as componentes dos vectores $\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n$, em relação à base considerada, tem-se

$$f\left(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n\right)=|M|\,f\left(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\right)$$

Demonstração:
Resulta imediatamente do que se viu na demonstração do Teorema 8-37 e da definição de determinante de uma matriz.

Corolário 8-40 Dada uma base $\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n$, de um certo espaço vectorial E, sobre um corpo K, e dado um escalar qualquer $\eta\in K$, existe uma e uma só forma n-linear em E tal que $f\left(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\right)=\eta$.

Demonstração:
Basta considerar a forma multilinear $f\left(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n\right)=|M|\eta$, onde M é a matriz cujas colunas são as componentes dos vectores $\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n$ em relação à base considerada.

Das propriedades conhecidas dos determinantes resulta imediatamente que *f* é uma forma multilinear alternada e que $f\left(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\right)=\eta$. A unicidade de *f* resulta do Teorema 8-37.

Deve observar-se que poderíamos ter chegado ao Corolário 8-40 sem recorrer ao conceito de determinante de uma matriz. Nesse caso, poderíamos usar estas propriedades para definir o conceito de "determinante": num espaço vectorial de dimensão n, chamaríamos "forma determinante" em relação a certa base fixa $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ à (única) forma n-linear alternada f tal que $f(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) = 1$. Daqui resultaria o conceito de determinante de uma matriz: o determinante de uma matriz quadrada de ordem n seria o valor da forma determinante em $\mathbb{R}^n$ em relação à base canónica, aplicada às n colunas da matriz dada. As propriedades dos determinantes, que estudámos no primeiro capítulo, resultariam imediatamente das propriedades das formas multilineares.

8-5 Menores e formas quadráticas

A classificação de uma forma bilinear (definida, semidefinida ou indefinida) nem sempre é fácil, especialmente em espaços vectoriais de dimensão elevada. Essa classificação tem, no entanto, bastante importância, aplicando-se, por exemplo, ao estudo dos extremos de uma função real de várias variáveis reais. Estudaremos, no presente parágrafo, um processo para chegar a essa classificação.

Definição 8-41 Dada uma matriz A, do tipo $n \times n$, chamam-se **menores principais** de A aos determinantes

$$\Delta_1 = |a_{11}| = a_{11}, \ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},$$

$$\Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \ldots, \ \Delta_n = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = |A|$$

Vamos ver como se podem usar os menores principais de uma matriz para classificar uma forma bilinear:

Teorema 8-42 Seja f uma forma bilinear simétrica, num espaço vectorial real E, de dimensão n. Seja A a matriz de f em relação a certa base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ do espaço. A forma f é definida positiva se e só se os menores principais de A são todos positivos.

Demonstração:
Faremos a demonstração por indução em n.
No caso de ser $n=1$, teremos $A=[a_{11}]$ e $f(\vec{x},\vec{y})=a_{11}x_1y_1$, pelo que

$$Q_f(\vec{x})=a_{11}x_1^2$$

Daqui resulta, trivialmente, que f é definida positiva se e só se $m_1=a_{11}>0$, como pretendíamos.

Admitindo, em seguida, que o teorema seja válido para espaços de dimensão $n-1$, passemos ao estudo do caso de um espaço vectorial de dimensão n.

Dada a forma bilinear f, podemos escrever

$$Q_f(\vec{x})=X^TAX=\sum_{i,j=1}^{n}a_{ij}x_ix_j=\sum_{i,j=1}^{n-1}a_{ij}x_ix_j+2\sum_{i=1}^{n-1}a_{in}x_ix_n+a_{nn}x_n^2$$

Ponhamos

$$g(\vec{x},\vec{y})=\sum_{i,j=1}^{n-1}a_{ij}x_iy_j$$

Definimos, deste modo, uma forma bilinear g, no espaço vectorial

$$F=\langle\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_{n-1}\rangle$$

que tem a dimensão $n-1$.
A matriz de g em relação à base $\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_{n-1}$ é, evidentemente,

$$A'=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} \\ a_{21} & a_{22} & \cdots & a_{2\,n-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n-1\,1} & a_{n-1\,2} & \cdots & a_{n-1\,n-1} \end{bmatrix}$$

e os menores principais de A, até à ordem $n-1$, coincidem com os de A'.

Ora, é claro que, sendo f uma forma definida positiva, também g será definida positiva, pelo que, por hipótese de indução, os menores principais de A' serão todos positivos.

Por outro lado, tendo em atenção o Corolário 8-16, se f é definida positiva, então a sua matriz A é equivalente à matriz identidade, ou seja, existe uma matriz invertível P tal que $A=P^TI_nP=P^TP$, pelo que

$$\Delta_n=|A|=|P^TP|=|P^T|\,|P|=|P|^2>0$$

Reciprocamente, se os menores principais de A são positivos, então, em particular, também os menores principais de A' são positivos, e a forma bilinear g é definida positiva.

Em relação a certa base $\vec{e}_1', \vec{e}_2', ..., \vec{e}_{n-1}'$, do espaço F, a matriz de g será a matriz identidade I_{n-1}. Logo, em relação à base $\vec{e}_1', \vec{e}_2', ..., \vec{e}_{n-1}', \vec{e}_n$, do espaço E, a matriz de f será

$$B = \begin{bmatrix} 1 & 0 & \cdots & 0 & b_{1n} \\ 0 & 1 & \cdots & 0 & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_{n-1\,n} \\ b_{n1} & b_{n2} & \cdots & b_{n\,n-1} & b_{nn} \end{bmatrix}$$

A expressão de $Q_f(\vec{x})$ em relação a esta base é

$$Q_f(\vec{x}) = \sum_{i=1}^{n-1} x_i'^2 + 2\sum_{i=1}^{n-1} b_{in} x_i' x_n' + b_{nn} x_n'^2$$

Ora, $x_i'^2 + 2b_{in} x_i' x_n' = (x_i' + b_{in} x_n')^2 - b_{in}^2 x_n'^2$.

Façamos uma mudança de base tal que, sendo $\vec{x} \equiv (x_1'', x_2'', ..., x_n'')$ em relação à nova base, se tenha

$$\begin{cases} x_i'' = x_i' + b_{in} x_n' \quad (i = 1, 2, ..., n-1) \\ x_n'' = x_n' \end{cases}$$

É fácil ver que estas igualdades definem realmente uma mudança de base. Em relação a essa nova base, a expressão de $Q_f(\vec{x})$ será, então,

$$Q_f(\vec{x}) = \sum_{i=1}^{n-1} x_i''^2 + \left(b_{nn} - \sum_{i=1}^{n-1} b_{in}^2\right) x_n''^2 = \sum_{i=1}^{n-1} x_i''^2 + c x_n''^2 \quad \text{com} \quad c = b_{nn} - \sum_{i=1}^{n-1} b_{in}^2$$

Se for $c > 0$, a forma f será definida positiva, como se pretende.

Ora, a matriz de f na nova base é

$$C = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & c \end{bmatrix}$$

e tem-se $|C| = c$.

Uma vez que os menores principais de A são todos positivos, em particular $\Delta_n = |A| > 0$. Ora, A e C representam a mesma forma bilinear em relação a bases distintas, pelo que existirá uma matriz invertível R tal que $C = R^T AR$. Então,

$$c = |C| = \left|R^T AR\right| = \left|R^T\right| \ |A| \ |R| = |A| \ |R|^2 > 0$$

Assim termina a demonstração.

Por um processo análogo se provaria também o seguinte:

Corolário 8-43 Nas condições do Teorema 8-42, a forma bilinear f será definida negativa se e só se os seus menores principais verificam a condição

$$(-1)^k \Delta_k > 0 \quad (k = 1, 2, \ldots, n)$$

No caso de A ter característica $r < n$, podemos reordenar as suas linhas e as suas colunas (conservando-se os elementos principais de A na diagonal principal da nova matriz) de modo que, na nova matriz, se tenha $\Delta_r \neq 0$. Então, a forma f será semidefinida positiva se e só se $\Delta_k > 0 \ (k = 1, 2, \ldots, r)$ e será semidefinida negativa se e só se $(-1)^k \Delta_k > 0 \ (k = 1, 2, \ldots, r)$.

9 | Geometria analítica

9-1 | Espaços afins

No presente capítulo, aplicaremos os conceitos da Álgebra Linear ao estudo de certas questões de Geometria. Para isso, precisamos, em primeiro lugar, de introduzir um novo tipo de estrutura, que permita relacionar os aspectos vectoriais com as questões geométricas. A estrutura em causa é a de "espaço afim", que passamos a definir:

Definição 9-1

Seja E um espaço vectorial sobre um certo corpo K. Seja $\mathscr{E}$ um conjunto qualquer, não vazio: $\mathscr{E} = \{A, B, C, \ldots, X, Y, \ldots\}$. Dizemos que $\mathscr{E}$ constitui um **espaço afim** associado ao espaço vectorial E quando se define uma aplicação

$$\tau : \mathscr{E} \times \mathscr{E} \to E$$

na qual a imagem de cada par (X, Y) se representa por

$$\tau(X, Y) = \overrightarrow{XY}$$

que satisfaça as duas seguintes condições:

a) $\forall X, Y, Z \in \mathscr{E} : \overrightarrow{XY} + \overrightarrow{YZ} = \overrightarrow{XZ}$

b) $\forall X \in \mathscr{E},\ \forall \vec{v} \in E,\ \exists^1 Y \in \mathscr{E} : \overrightarrow{XY} = \vec{v}$

Os elementos de $\mathscr{E}$ passam, então, a chamar-se **pontos**.

Exemplo 9-2

Dado um corpo qualquer K, consideremos o conjunto $\mathscr{E} = K^n$, onde n é um número natural arbitrário. Vamos introduzir em K^n a estrutura de espaço afim, associado ao espaço vectorial K^n, sobre o corpo K, definindo a aplicação τ da seguinte maneira: para dois pontos $X = (x_1, x_2, \ldots, x_n)$ e $Y = (y_1, y_2, \ldots, y_n)$, pomos

$$\overrightarrow{XY} = (y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n)$$

É muito fácil verificar que esta aplicação satisfaz as duas condições da definição anterior. Na verdade, tomando um terceiro ponto, $Z=(z_1,z_2,\ldots,z_n)$, vemos que

$$\begin{aligned}\overrightarrow{XY}+\overrightarrow{YZ}&=(y_1-x_1,y_2-x_2,\ldots,y_n-x_n)+(z_1-y_1,z_2-y_2,\ldots,z_n-y_n)=\\&=(y_1-x_1+z_1-y_1,y_2-x_2+z_2-y_2,\ldots,y_n-x_n+z_n-y_n)=\\&=(z_1-x_1,z_2-x_2,\ldots,z_n-x_n)=\overrightarrow{XZ}\end{aligned}$$

Por sua vez, tomando um vector qualquer, $\vec{v}=(v_1,v_2,\ldots,v_n)$, para que seja $\overrightarrow{XY}=\vec{v}$, teremos de ter

$$\overrightarrow{XY}=(y_1-x_1,y_2-x_2,\ldots,y_n-x_n)\,\vec{v}=(v_1,v_2,\ldots,v_n)$$

ou seja,

$$\begin{cases}y_1-x_1=v_1\\y_2-x_2=v_2\\\vdots\\y_n-x_n=v_n\end{cases}$$

Ora, este sistema de equações tem uma solução única, a saber

$$\begin{cases}y_1=x_1+v_1\\y_2=x_2+v_2\\\vdots\\y_n=x_n+v_n\end{cases}$$

pelo que o ponto Y procurado será $Y=(x_1+v_1,x_2+v_2,\ldots,x_n+v_n)$.

Naturalmente, no mesmo conjunto K^n, poderíamos introduzir a estrutura de espaço afim de outras maneiras. No entanto, em tudo quanto se segue, sempre que nos referirmos simplesmente ao espaço afim K^n, subentenderemos que a relação entre pontos e vectores é a que foi exposta neste exemplo. ■

Um tanto por influência dos exemplos anteriores (que, como veremos mais adiante, têm, na realidade, um carácter muito geral), é costume utilizar-se a seguinte notação:

Num espaço afim $\mathscr{E}$ qualquer, escrever

$$\tau(X,Y)=\vec{v}\quad\text{ou}\quad\overrightarrow{XY}=\vec{v}$$

é o mesmo que escrever

$$Y=X+\vec{v}\quad\text{ou, ainda,}\quad\vec{v}=Y-X$$

Deve observar-se que se trata aqui de simples notações, não estando definidas quaisquer operações (como a subtracção ou a adição) entre pontos ou entre pontos e vectores.

Podemos desde já deduzir da definição algumas propriedades elementares, válidas num espaço afim qualquer:

Teorema 9-3 Seja $\mathscr{E}$ um espaço afim qualquer, associado a um espaço vectorial E. Para pontos e vectores arbitrários, tem-se:

a) $\overrightarrow{XY} = \vec{0}$ se e só se $X = Y$

b) $\overrightarrow{YX} = -\overrightarrow{XY}$

c) $(X+\vec{u})+\vec{v} = X+(\vec{u}+\vec{v})$

d) $\overrightarrow{(X+\vec{u})(Y+\vec{v})} = \vec{v}-\vec{u}+\overrightarrow{XY}$

e) $\overrightarrow{X(X+\vec{v})} = \vec{v}$

Demonstração:

a) Da propriedade a) da Definição 9-1 de espaço afim, resulta que

$$\overrightarrow{XX}+\overrightarrow{XX} = \overrightarrow{XX} \quad \text{pelo que} \quad \overrightarrow{XX} = \overrightarrow{XX}-\overrightarrow{XX} = \vec{0}$$

Por outro lado, supondo que $\overrightarrow{XY} = \vec{0}$ e atendendo a que também $\overrightarrow{XX} = \vec{0}$, resulta da propriedade b) da Definição 9-1 de espaço afim que $X = Y$.

b) Resulta imediatamente da propriedade anterior, uma vez que

$$\overrightarrow{XY}+\overrightarrow{YX} = \overrightarrow{XX} = \vec{0}$$

c) Pondo $Y = X+\vec{u}$ e $Z = Y+\vec{v}$, temos $\vec{u} = \overrightarrow{XY}$ e $\vec{v} = \overrightarrow{YZ}$, pelo que

$$\vec{u}+\vec{v} = \overrightarrow{XY}+\overrightarrow{YZ} = \overrightarrow{XZ}$$

Resulta então das definições dadas que $Z = X+(\vec{u}+\vec{v})$, pelo que

$$(X+\vec{u})+\vec{v} = Y+\vec{v} = Z = X+(\vec{u}+\vec{v})$$

d) Pondo $X+\vec{u} = W$ e $Y+\vec{v} = Z$, teremos $\vec{u} = \overrightarrow{XW}$ e $\vec{v} = \overrightarrow{YZ}$. Portanto,

$$Z-W = \overrightarrow{WZ} = \overrightarrow{WX}+\overrightarrow{XY}+\overrightarrow{YZ} = -\overrightarrow{XW}+\overrightarrow{XY}+\overrightarrow{YZ} = -\vec{u}+\overrightarrow{XY}+\vec{v}$$

e) Resulta trivialmente da propriedade anterior, tomando $Y = X$ e $\vec{u} = \vec{0}$.

Consideremos agora um espaço afim qualquer $\mathscr{E}$, associado a um espaço vectorial E, e fixemos um ponto $O \in \mathscr{E}$.

Podemos definir uma aplicação

$$\zeta : \mathscr{E} \to E$$

pondo, para cada $X \in \mathscr{E}$,

$$\zeta(X) = \overrightarrow{OX}$$

A propriedade b) da definição de espaço afim garante que esta aplicação ζ é uma bijecção, uma vez que, estando fixo o ponto O, qualquer vector $\vec{v} \in E$ é imagem de um e um só ponto de $\mathscr{E}$.

Nestas condições, podemos, evidentemente, escrever

$$E = \left\{\overrightarrow{OX} : X \in \mathscr{E}\right\} \quad \text{e} \quad \mathscr{E} = \{O + \vec{v} : \vec{v} \in E\}$$

Em particular, admitamos que o espaço E tenha dimensão finita e fixemos em E uma base $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$. Sabemos que o facto de fixarmos uma base permite estabelecer um isomorfismo entre o espaço E e o espaço vectorial K^n, fazendo-se corresponder a cada vector de E a lista das suas componentes em relação à base considerada:

$$\begin{array}{rcccl} \omega & : & E & \to & K^n \\ & & \vec{x} & \mapsto & (x_1, x_2, \dots, x_n) \end{array}$$

A composição de ω com ζ dá-nos uma bijecção

$$\omega \circ \zeta : \mathscr{E} \to K^n$$

na qual teremos, para cada ponto,

$$(\omega \circ \zeta)(X) = \omega\left(\overrightarrow{OX}\right) = (x_1, x_2, \dots, x_n)$$

com $\overrightarrow{OX} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n$.

Perante as construções que acabámos de apresentar, introduzimos as seguintes definições:

Definição 9-4 Seja $\mathscr{E}$ um espaço afim associado a um espaço vectorial E de dimensão finita n; diz-se, neste caso, que a dimensão de $\mathscr{E}$ é também igual a n. O par

$$(O; \vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$$

formado por um ponto fixo em $\mathscr{E}$ e por uma base fixa em E, chama-se um **referencial** de $\mathscr{E}$. O ponto O é a **origem** desse referencial, e, para cada ponto X, o vector $\overrightarrow{OX}$ é o **vector de posição** do ponto X em relação à origem O. As componentes do vector de posição em relação à base do referencial considerado são as **coordenadas** do ponto X, que representaremos por $X \equiv (x_1, x_2, ..., x_n)$, ou, quando for adequado, na forma matricial

$$X \equiv \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

No espaço afim K^n, associado ao espaço vectorial K^n, podemos destacar o chamado **referencial canónico**, $(O; \vec{e}_1, \vec{e}_2, ..., \vec{e}_n)$, no qual a origem é o ponto $O = (0, 0, ..., 0)$ e a base é a base canónica.

Vamos calcular as coordenadas de um ponto arbitrário $A = (a_1, a_2, ..., a_n)$ em relação ao referencial canónico:

Começamos, para o efeito, por calcular o vector de posição

$$\overrightarrow{OA} = A - O = (a_1, a_2, ..., a_n) - (0, 0, ..., 0) = (a_1, a_2, ..., a_n)$$

Seguidamente, procuramos as componentes do vector de posição em relação à base canónica, obtendo-se, como se sabe,

$$\overrightarrow{OA} = (a_1, a_2, ..., a_n) = a_1\vec{e}_1 + a_2\vec{e}_2 + \cdots + a_n\vec{e}_n$$

Por consequência, as componentes do ponto A no referencial canónico serão $A \equiv (a_1, a_2, ..., a_n)$.

Como é natural, as coordenadas de um dado ponto variam consoante o referencial considerado. É lícito perguntar-se como estarão relacionadas entre si as coordenadas de um mesmo ponto em relação a dois referenciais distintos.

Para responder a esta questão, consideremos, num espaço afim $\mathscr{E}$, de dimensão n, os referenciais $(O; \vec{e}_1, \vec{e}_2, ..., \vec{e}_n)$ e $(O'; \vec{e}'_1, \vec{e}'_2, ..., \vec{e}'_n)$.

Suponhamos que um ponto A tem as coordenadas

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

em relação ao primeiro referencial, e as coordenadas

$$X' = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix}$$

em relação ao segundo. Pretendemos relacionar X com X'.

Por definição, teremos

$$\overrightarrow{OA} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n \quad \text{e} \quad \overrightarrow{O'A} = x_1'\vec{e}_1' + x_2'\vec{e}_2' + \cdots + x_n'\vec{e}_n'$$

Por outro lado, sabemos que

$$\overrightarrow{OA} = \overrightarrow{OO'} + \overrightarrow{O'A}$$

O vector $\overrightarrow{OO'}$ é o vector de posição da origem O', do segundo referencial em relação ao primeiro, pelo que as suas componentes em relação à base $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ nos dão as coordenadas de O' no primeiro referencial: ponhamos

$$O' \equiv \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = C$$

Por sua vez, seja P a matriz de mudança de base, isto é, a matriz cujas colunas são constituídas pelas componentes dos vectores da segunda base, $\vec{e}_1', \vec{e}_2', \ldots, \vec{e}_n'$, em relação à primeira base, $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$. Teremos, para cada índice

$$j = 1, 2, \ldots, n, \; \vec{e}_j' = \sum_{i=1}^{n} p_{ij}\vec{e}_i$$

Nestas condições, verificamos que

$$\overrightarrow{OA} = \sum_{i=1}^{n} x_i\vec{e}_i = \overrightarrow{OO} + \overrightarrow{O'O} = \sum_{i=1}^{n} c_i\vec{e}_i + \sum_{j=1}^{n} x_j'\vec{e}_j' = \sum_{i=1}^{n} c_i\vec{e}_i + \sum_{j=1}^{n} x_j'\left(\sum_{i=1}^{n} p_{ij}\vec{e}_i\right) =$$
$$= \sum_{i=1}^{n} c_i\vec{e}_i + \sum_{i=1}^{n}\left(\sum_{j=1}^{n} p_{ij}x_j'\right)\vec{e}_i = \sum_{i=1}^{n}\left(c_i + \sum_{j=1}^{n} p_{ij}x_j'\right)\vec{e}_i$$

Atendendo a que as componentes de um dado vector em relação a certa base são únicas, concluímos que se tem

$$x_i = c_i + \sum_{j=1}^{n} p_{ij} x'_j \qquad (i = 1,2,\dots,n)$$

Ora, estas n igualdades podem ser escritas, na forma matricial, como

$$X = PX' + C$$

Acabamos, portanto, de provar o seguinte:

Teorema 9-5 Sendo X a coluna das coordenadas de um certo ponto de um espaço afim $\mathscr{E}$, de dimensão n, em relação ao referencial $(O;\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n)$, e X' a coluna das coordenadas do mesmo ponto em relação ao referencial $(O';\vec{e}'_1,\vec{e}'_2,\dots,\vec{e}'_n)$; tem-se

$$X = PX' + C$$

onde P é a matriz de mudança de base e C é a coluna das coordenadas da nova origem O' no primeiro referencial.

Deve notar-se que, como a matriz de mudança de base é sempre uma matriz invertível, a igualdade $X = PX' + C$ equivale a

$$X' = P^{-1}(X - C)$$

9-2 Subespaços afins

Associado ao conceito de espaço afim, aparece, de forma natural, a noção de "subespaço afim", que definiremos da seguinte maneira:

Definição 9-6 Seja $\mathscr{E}$ um espaço afim, associado a um espaço vectorial E. Seja F um subespaço vectorial de E e seja $\mathscr{F}$ um subconjunto não vazio de $\mathscr{E}$. Diz-se que $\mathscr{F}$ é um **subespaço afim** de $\mathscr{E}$, associado ao subespaço vectorial F, quando se verificam as duas condições seguintes:

a) $X, Y \in \mathscr{F} \Rightarrow \overrightarrow{XY} \in F$

b) $X \in \mathscr{F},\ \vec{v} \in F \Rightarrow X + \vec{v} \in \mathscr{F}$

Exemplos 9-7

a) Consideremos o espaço afim $\mathbb{R}^3$ associado ao espaço vectorial real $\mathbb{R}^3$. Tomemos o conjunto

$$\mathcal{H} = \{(x,y,z) \in \mathbb{R}^3 : x+y+z=5\}$$

Vamos provar que $\mathcal{H}$ é um subespaço afim de $\mathbb{R}^3$, associado ao subespaço vectorial

$$H = \langle (1,0,-1),(0,1,-1) \rangle$$

Dados dois pontos $A,B \in \mathcal{H}$, teremos $A=(a_1,a_2,a_3)$, com $a_1+a_2+a_3=5$ e $B=(b_1,b_2,b_3)$, com $b_1+b_2+b_3=5$. Ora,

$$\overrightarrow{AB} = B - A = (b_1-a_1, b_2-a_2, b_3-a_3)$$

mas de $a_1+a_2+a_3=b_1+b_2+b_3$ resulta que

$$b_3 - a_3 = a_1 - b_1 + a_2 - b_2$$

pelo que

$$\overrightarrow{AB} = (b_1-a_1, b_2-a_2, a_1-b_1+a_2-b_2) = (b_1-a_1)(1,0,-1)+(b_2-a_2)(0,1,-1) \in H$$

Por outro lado, se tomarmos um ponto $A \in \mathcal{H}$ e um vector $\vec{v} \in H$, teremos $A=(a_1,a_2,a_3)$, com $a_1+a_2+a_3=5$ e

$$\vec{v} = \alpha(1,0,-1)+\beta(0,1,-1) = (\alpha,\beta,-\alpha-\beta) \text{ e } A+\vec{v} = (a_1+\alpha, a_2+\beta, a_3-\alpha-\beta)$$

Ora,

$$a_1+\alpha+a_2+\beta+a_3-\alpha-\beta = a_1+a_2+a_3=5$$

pelo que $A+\vec{v} \in \mathcal{H}$.

b) De modo geral, consideremos uma matriz A, do tipo $p \times n$, com elementos de um certo corpo K. Seja F o subespaço vectorial do espaço vectorial K^n constituído pelas soluções do sistema homogéneo de equações lineares $AX=0$, e $\mathcal{F}$ o conjunto das soluções do sistema de equações lineares $AX=B$, onde B é uma coluna qualquer. É fácil verificar que, desde que não seja vazio, $\mathcal{F}$ é um subespaço afim do espaço afim K^n, associado ao subespaço vectorial F.

Com efeito, se X_1 e X_2 são soluções do sistema de equações $AX=B$, então teremos $AX_1=B$ e $AX_2=B$, pelo que

$$A(X_1-X_2) = AX_1 - AX_2 = B - B = 0$$

e $X_1 - X_2$ é solução do sistema de equações $AX=0$.

Ao mesmo tempo, se X_1 é solução do sistema de equações $AX=B$ e se V é solução do sistema de equações $AX=0$, então $AX_1=B$ e $AV=0$, pelo que

$$A(X_1+V) = AX_1 + AV = B+0 = B$$

e X_1+V é solução do sistema de equações $AX=B$.

c) Um espaço afim $\mathscr{E}$ qualquer é trivialmente um subespaço afim de si mesmo. Por outro lado, dado um ponto qualquer $A \in \mathscr{E}$, o conjunto $\{A\}$ é um subespaço afim de $\mathscr{E}$, associado ao subespaço vectorial $\{\vec{0}\}$ do respectivo espaço vectorial. ■

Deve notar-se que, de acordo com as definições dadas, um subespaço afim $\mathscr{F}$ de um certo espaço afim, associado a um subespaço vectorial F, do respectivo espaço vectorial, é, ele mesmo, um espaço afim.

Daí resulta que, se fixarmos em $\mathscr{F}$ um ponto A qualquer, podemos escrever

$$\mathscr{F} = \{A + \vec{v} : \vec{v} \in F\}$$

Estas observações são completadas pelo seguinte teorema:

Teorema 9-8

Seja $\mathscr{E}$ um espaço afim qualquer, associado ao espaço vectorial E. Se fixarmos um ponto A qualquer em $\mathscr{E}$ e um subespaço vectorial G qualquer de E, existe um e um só subespaço afim de $\mathscr{E}$, associado a G e que passa por A (isto é, tal que o ponto A lhe pertence).

Demonstração:

Nas condições do enunciado, tomemos em $\mathscr{E}$ o conjunto $\mathscr{I} = \{A + \vec{g} : \vec{g} \in G\}$.

Dados $X, Y \in \mathscr{I}$, teremos $X = A + \vec{g}_1$, $Y = A + \vec{g}_2$, com $\vec{g}_1, \vec{g}_2 \in G$, pelo que

$$\overrightarrow{XY} = \overrightarrow{(A + \vec{g}_1)(A + \vec{g}_2)} = \vec{g}_2 + \overrightarrow{AA} - \vec{g}_1 = \vec{g}_2 + \vec{0} - \vec{g}_1 = \vec{g}_2 - \vec{g}_1 \in G$$

Por sua vez, dados $X \in \mathscr{I}$ e $\vec{v} \in G$, será $X = A + \vec{g}_1$, pelo que

$$X + \vec{v} = (A + \vec{g}_1) + \vec{v} = A + (\vec{g}_1 + \vec{v})$$

Ora, $\vec{g}_1 + \vec{v} \in G$, pelo que $X + \vec{v} \in \mathscr{I}$.

Vemos, assim, que $\mathscr{I}$ é um subespaço afim de $\mathscr{E}$ associado ao subespaço vectorial G.

O ponto A pode escrever-se na forma $A = A + \vec{0}$, pelo que $A \in \mathscr{I}$. Por sua vez, a unicidade deste subespaço afim é garantida pelas observações que precedem o enunciado do teorema.

O teorema anterior mostra que um ponto e um subespaço vectorial quaisquer definem um subespaço afim. Tem, por isso, sentido a seguinte designação:

Definição 9-9

Seja $\mathscr{E}$ um espaço afim qualquer associado ao espaço vectorial E. Dados um ponto A qualquer de $\mathscr{E}$ e um subespaço vectorial G qualquer de E, o (único) subespaço afim de $\mathscr{E}$, associado a G e que passa por A chama-se **subespaço afim gerado** por A e G, e representa-se por $\mathscr{I} = \langle A; G \rangle$.

Os espaços (e, portanto, os subespaços) afins de dimensão finita merecer-nos-ão uma atenção especial. Começamos por introduzir a seguinte nomenclatura:

Definição 9-10 Seja $\mathscr{E}$ um espaço afim qualquer, associado ao espaço vectorial E. Um subespaço de dimensão 1 de $\mathscr{E}$ chama-se uma **recta**; um subespaço de dimensão 2 é um **plano**. Se $\mathscr{E}$ tem dimensão n, um subespaço de dimensão $n-1$ chama-se um **hiperplano**. Dois ou mais pontos de $\mathscr{E}$ dizem-se **colineares** (resp.: **complanares**) quando existe uma recta (resp.: um plano) do espaço que os contém a todos. Analogamente, duas rectas serão **complanares** quando estiverem contidas num mesmo plano.

Podemos desde já demonstrar um certo número de propriedades que recuperam factos bem conhecidos da Geometria Euclidiana:

Teorema 9-11 Num espaço afim qualquer:

a) Dois pontos distintos definem uma recta (isto é, dados dois pontos distintos, existe uma e uma só recta do espaço que os contém simultaneamente).
b) Três pontos não colineares definem um plano.
c) Uma recta e um ponto que lhe não pertença definem um plano.

Demonstração:

a) Sejam A e B dois pontos distintos de um espaço afim $\mathscr{E}$. O vector $\overrightarrow{AB}$ não é nulo, pelo que o subespaço vectorial $\langle\overrightarrow{AB}\rangle$ tem dimensão 1 e o subespaço afim $\mathscr{R} = \left\langle A;\left\langle\overrightarrow{AB}\right\rangle\right\rangle$ é uma recta.

Esta recta inclui, por construção, o ponto A, e também o ponto B, visto que $B = A + \overrightarrow{AB}$.

Por outro lado, se $\mathscr{F}$ é uma recta que inclui os pontos A e B, então poderemos escrever $\mathscr{F} = \langle A;F\rangle$, onde F é um subespaço vectorial de dimensão 1. Uma vez que $A, B \in \mathscr{F}$, será $\overrightarrow{AB} \in F$, pelo que, como $\overrightarrow{AB} \neq \vec{0}$ e $\dim(F) = 1$, teremos $F = \langle\overrightarrow{AB}\rangle$ e $\mathscr{F} = \langle A;F\rangle = \left\langle A;\langle\overrightarrow{AB}\rangle\right\rangle = \mathscr{R}$.

b) Sejam A, B e C três pontos não colineares (logo, forçosamente distintos). Se os vectores $\overrightarrow{AB}$ e $\overrightarrow{AC}$ fossem linearmente dependentes, escrevendo $\overrightarrow{AC} = \lambda\overrightarrow{AB}$, veríamos que $C = A + \lambda\overrightarrow{AB} \in \left\langle A;\langle\overrightarrow{AB}\rangle\right\rangle$, e, nesse caso, os três pontos seriam colineares, já que este último subespaço é a recta gerada por A e B. Consequentemente, os vectores considerados são linearmente independentes, e o subespaço vectorial $\left\langle\overrightarrow{AB}, \overrightarrow{AC}\right\rangle$ tem dimensão 2.

O subespaço afim $\mathscr{P} = \left\langle A;\langle\overrightarrow{AB}, \overrightarrow{AC}\rangle\right\rangle$ é um plano que contém, por construção, o ponto A, e também os outros dois pontos, visto que $B = A + \overrightarrow{AB}$ e $C = A + \overrightarrow{AC}$.

Para vermos que este plano é o único nas condições pretendidas, suponhamos que $\mathscr{F}$ é um plano qualquer que contém os pontos A, B e C. Poderemos definir este plano na forma $\mathcal{F} = \langle A; F \rangle$, onde F é um subespaço vectorial de dimensão 2.

Mas os vectores $\overrightarrow{AB}$ e $\overrightarrow{AC}$ pertencem a F, pelo que $F = \langle \overrightarrow{AB}, \overrightarrow{AC} \rangle$. Logo,

$$\mathcal{F} = \langle A; F \rangle = \left\langle A; \langle \overrightarrow{AB}, \overrightarrow{AC} \rangle \right\rangle = \mathcal{P}$$

c) Seja $\mathcal{R} = \langle A; \langle \vec{u} \rangle \rangle$ uma recta, e B, um ponto tal que $B \notin \mathcal{R}$. Os vectores $\overrightarrow{AB}$ e $\vec{u}$ são linearmente independentes, visto que, de $\overrightarrow{AB} = \lambda \vec{u}$ resultaria que $B = A + \lambda \vec{u} \in \mathcal{R}$.

Consequentemente, $\mathcal{P} = \left\langle A; \langle \overrightarrow{AB}, \vec{u} \rangle \right\rangle$ é um plano que contém o ponto $B = A + \overrightarrow{AB}$ e a recta $\mathcal{R}$, cujos pontos são da forma $X = A + \alpha \vec{u}$.

A prova de que este plano é o único nas condições requeridas é feita nos mesmos moldes que nas alíneas anteriores.

Vamos agora estudar a possibilidade de se construírem subespaços afins a partir de outros:

Teorema 9-12 Seja $\mathcal{E}$ um espaço afim, associado ao espaço vectorial E. Se $\mathcal{F}_i$ (com i percorrendo um certo conjunto de índices I) são subespaços afins de $\mathcal{E}$, associados, respectivamente, aos subespaços vectoriais F_i, de E, então

$$\mathcal{F} = \bigcap_{i \in I} Fi$$

ou é vazio ou é um subespaço afim de $\mathcal{E}$, associado ao subespaço vectorial $F = \bigcap_{i \in I} F_i$.

Demonstração:

Supondo que $\mathcal{F} \neq \varnothing$, vamos ver que constitui um subespaço afim.

Sendo $X, Y \in \mathcal{F}$, teremos $X, Y \in \mathcal{F}_i$ para cada $i \in I$, pelo que, como cada $\mathcal{F}_i$ é um subespaço afim, será $\overrightarrow{XY} \in F_i$ e, como i é arbitrário, $\overrightarrow{XY} \in F$.

Por sua vez, sendo $X \in \mathcal{F}$ e $\vec{v} \in F$, teremos, sempre para i arbitrário, $X \in \mathcal{F}_i$ e $\vec{v} \in F_i$, pelo que $X + \vec{v} \in \mathcal{F}_i$, e, assim, pela arbitrariedade do índice i, $X + \vec{v} \in \mathcal{F}$.

Deste teorema resultam, por sua vez, propriedades geométricas bem conhecidas. Começamos por demonstrar um lema:

Lema 9-13 Sejam $\mathcal{F}$ e $\mathcal{G}$ subespaços afins de um espaço afim $\mathcal{E}$, associados aos subespaços vectoriais F e G.

a) Se $\mathcal{F} \subseteq \mathcal{G}$, então $F \subseteq G$.

b) Se $F \subseteq G$ e $\mathcal{F}$ e $\mathcal{G}$ não são disjuntos, então $\mathcal{F} \subseteq \mathcal{G}$.

Demonstração:

a) Fixando um ponto $A \in \mathcal{F}$, tem-se

$$F = \left\{\overrightarrow{AX} : X \in \mathrm{F}\right\} \subseteq \left\{\overrightarrow{AX} : X \in G\right\} \subseteq G$$

b) Sendo A um ponto comum a $\mathcal{F}$ e $\mathcal{I}$, tem-se

$$\mathcal{F} = \{A + \vec{v} : \vec{v} \in F\} \subseteq \{A + \vec{v} : \vec{v} \in G\} = \mathcal{I}$$

Podemos, agora, enunciar o seguinte:

Teorema 9-14 Num espaço afim qualquer:

a) A intersecção de duas rectas distintas ou é vazia ou se reduz a um ponto.

b) A intersecção de uma recta com um plano que a não contenha ou é vazia ou se reduz a um ponto.

c) Se o espaço tem dimensão 3, a intersecção de dois planos distintos ou é vazia ou é uma recta.

Demonstração:

a) Sejam $\mathcal{R}$ e $\mathcal{S}$ duas rectas distintas. Supondo que a intersecção das duas rectas não seja vazia, consideremos um ponto $A \in \mathcal{R} \cap \mathcal{S}$. Podemos, então, escrever $\mathcal{R} = \langle A; F\rangle$ e $\mathcal{S} = \langle A; G\rangle$, sendo F e G subespaços vectoriais de dimensão 1.

De acordo com o teorema anterior, a intersecção das duas rectas será um subespaço afim associado ao subespaço vectorial $F \cap G$. Uma vez que se tem $F \cap G \subseteq F$ e $F \cap G \subseteq G$, sabemos que $\dim(F \cap G) \leq 1$.

Ora, se fosse $\dim(F \cap G) = 1$, teríamos $F = F \cap G = G$ e, então, $\mathcal{R} = \mathcal{S}$, o que vai contra a hipótese. Consequentemente, tem de ser $\dim(F \cap G) = 0$ e $\mathcal{R} \cap \mathcal{S}$ é um ponto.

b) Sejam $\mathcal{R}$ uma recta e $\mathcal{P}$ um plano, e suponhamos que a recta não está contida no plano. Se $\mathcal{R} \cap \mathcal{P} \neq \varnothing$, consideremos um ponto A comum aos dois subespaços. Nestas condições, $\mathcal{R} = \langle A; F\rangle$ e $\mathcal{P} = \langle A; G\rangle$, sendo $\dim(F) = 1$ e $\dim(G) = 2$.

Pelo teorema anterior, $\mathcal{R} \cap \mathcal{P}$ é um subespaço afim associado ao subespaço vectorial $F \cap G$. Por ser $F \cap G \subseteq F$, terá de ser $\dim(F \cap G) \leq 1$. Ora, se fosse $\dim(F \cap G) = 1$, então $F \cap G = F$, ou seja, $F \subseteq G$; nesse caso, de acordo com o lema anterior, teríamos $\mathcal{R} \subseteq \mathcal{P}$, contra a hipótese. Consequentemente, tem de ser $\dim(F \cap G) = 0$ e $\mathcal{R} \cap \mathcal{P}$ é um ponto.

c) Sejam $\mathcal{P}_1$ e $\mathcal{P}_2$ dois planos distintos, num espaço afim de dimensão 3. Supondo que a sua intersecção não é vazia, fixemos um ponto A, comum aos dois planos. Então, $\mathcal{P}_1 = \langle A; F_1\rangle$ e $\mathcal{P}_2 = \langle A; F_2\rangle$, sendo F_1 e F_2 subespaços vectoriais de dimensão 2. A intersecção dos dois planos será um subespaço afim, associado ao subespaço vectorial $F_1 \cap F_2$.

Por ser $F_1 \cap F_2 \subseteq F_1$ (ou $F_1 \cap F_2 \subseteq F_2$), sabemos que $\dim(F_1 \cap F_2) \leq 2$. Se fosse $\dim(F_1 \cap F_2) = 2$, teríamos $F_1 = F_2$ e, portanto, $\mathcal{P}_1 = \mathcal{P}_2$, o que vai contra a hipótese.

Por outro lado, se fosse $\dim(F_1 \cap F_2) = 0$, então, de acordo com o teorema das dimensões,

$$\dim(F_1 + F_2) = \dim(F_1) + \dim(F_2) - \dim(F_1 \cap F_2) = 2 + 2 - 0 = 4$$

o que é impossível num espaço de dimensão 3.

Por conseguinte, será $\dim(F_1 \cap F_2) = 1$ e $\mathcal{P}_1 \cap \mathcal{P}_2$ é uma recta.

Em relação com a primeira alínea do teorema anterior, introduzimos mais uma designação:

Definição 9-15 Duas rectas (resp.: dois planos) cuja intersecção é um ponto (resp.: uma recta) dizem-se **concorrentes**.

Podemos ainda acrescentar o seguinte:

Teorema 9-16 Duas rectas concorrentes de um espaço afim definem um plano.

Demonstração:

Sejam $\mathcal{R}$ e $\mathcal{S}$ duas rectas concorrentes de um certo espaço afim, e seja A o seu (único) ponto comum. Poderemos escrever $\mathcal{R} = \langle A; \langle \vec{u} \rangle \rangle$ e $\mathcal{S} = \langle A; \langle \vec{v} \rangle \rangle$, onde $\vec{u}$ e $\vec{v}$ são vectores não nulos.

Estes dois vectores são linearmente independentes, já que, se fossem dependentes, seria $\langle \vec{u} \rangle = \langle \vec{v} \rangle$ e, portanto, $\mathcal{R} = \mathcal{S}$.

Logo, o subespaço afim $\mathcal{P} = \langle A; \langle \vec{u}, \vec{v} \rangle \rangle$ é um plano, que contém as duas rectas dadas. Facilmente se verifica que este plano é o único nas condições requeridas.

Terminaremos o presente parágrafo com algumas observações quanto à representação de subespaços afins por meio de sistemas de equações lineares.

Vimos já, no Exemplo 9-7 b), que o conjunto de soluções de um sistema de equações lineares, com n incógnitas, desde que não seja vazio, é um subespaço afim do respectivo espaço afim K^n. Veremos agora que a afirmação recíproca é também verdadeira, através do seguinte teorema:

Teorema 9-17 Seja $(O; \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ um referencial fixo num espaço afim $\mathcal{E}$, de dimensão n, associado a um certo espaço vectorial E, sobre um corpo K. Seja $\mathcal{F}$ um subespaço afim de $\mathcal{E}$, de dimensão $n - p$, associado a um subespaço vectorial F. Nestas condições, existem matrizes A e B, dos tipos $p \times n$ e $p \times 1$, respectivamente, com elementos de K, tais que os pontos de $\mathcal{F}$ são precisamente aqueles cujas coordenadas em relação ao referencial fixo satisfazem o sistema de equações lineares $AX = B$.

Demonstração:
Conforme se sabe, o subespaço vectorial F do espaço E pode ser representado através de um sistema de p equações homogéneas, com n incógnitas, isto é, existe uma matriz A, do tipo $p \times n$, tal que os vectores de F são precisamente aqueles cujas componentes em relação à base fixa em E são soluções do sistema de equações $AX = 0$.

Tomemos, então, em $\mathscr{F}$ um ponto X_0 qualquer, e representemos igualmente por X_0 a coluna das suas coordenadas em relação ao referencial considerado (de onde não resulta qualquer confusão). Seja $B = AX_0$.

Como sabemos, os pontos de $\mathscr{F}$ são da forma $X_0 + \vec{v}$, com $\vec{v} \in F$. Exprimindo os pontos pelas suas coordenadas e os vectores pelas suas componentes em relação ao referencial considerado, podemos dizer que os pontos de $\mathscr{F}$ são aqueles cujas coordenadas têm a expressão $X_0 + V$, onde V é solução do sistema de equações $AX = 0$.

Mas, então, temos

$$A(X_0 + V) = AX_0 + AV = B + 0 = B$$

pelo que os pontos de $\mathscr{F}$ satisfazem as equações $AX = B$.

Reciprocamente, se um dado ponto do espaço tem coordenadas Y que verificam as equações $AY = B$, então será

$$AY = AX_0 \quad \text{pelo que} \quad A(Y - X_0) = 0$$

Mas então o vector $\vec{w}$, de coordenadas $Y - X_0$, pertence ao subespaço vectorial F, já que as suas coordenadas verificam o sistema de equações $AX = 0$ e o ponto de coordenadas $Y = X_0 + V$ pertence ao subespaço afim $\mathscr{F}$.

Do teorema anterior resulta, por exemplo, que uma recta de um espaço afim de dimensão n pode ser representada por um sistema de $n-1$ equações lineares com n incógnitas, enquanto um plano necessitará de $n-2$ equações. Por exemplo, no espaço afim $\mathbb{R}^3$, uma recta será definida por duas equações, e um plano, por uma só.

As equações lineares que definem um dado subespaço afim em relação a um referencial fixo são usualmente chamadas as suas **equações cartesianas.**

Na prática, é fácil obter as equações cartesianas de um dado subespaço afim. Por exemplo, consideremos, no espaço afim $\mathbb{R}^3$, o plano

$$\mathscr{P} = \langle (1,2,3); \langle (1,-1,1), (2,-1,3) \rangle \rangle$$

Os pontos de $\mathscr{P}$ (que se podem sempre considerar definidos pelas suas coordenadas em relação ao referencial canónico do espaço) são os que obedecem à seguinte **equação vectorial:**

$$X = (1,2,3) + \alpha(1,-1,1) + \beta(2,-1,3)$$

Tomando o ponto genérico X na forma $X=(x,y,z)$, vemos que esta equação é equivalente às igualdades seguintes, que são as **equações paramétricas** do plano:

$$\begin{cases} x=1+\alpha+2\beta \\ y=2-\alpha-\beta \\ z=3+\alpha+3\beta \end{cases}$$

Eliminando os parâmetros, obtemos

$$\begin{cases} x+y=3+\beta \\ y+z=5+2\beta \end{cases}$$

e, finalmente, a equação cartesiana

$$-2x-y+z=-1$$

Analogamente, consideremos o plano

$$\mathcal{T}=\langle(1,1,1,1);\langle(0,1,-1,2),(1,1,0,-1)\rangle\rangle$$

do espaço afim $\mathbb{R}^4$. A sua equação vectorial é

$$X=(1,1,1,1)+\lambda(0,1,-1,2)+\mu(1,1,0,-1)$$

as equações paramétricas aparecem como

$$\begin{cases} x=1+\mu \\ y=1+\lambda+\mu \\ z=1-\lambda \\ w=1+2\lambda-\mu \end{cases}$$

e as equações cartesianas obtêm-se eliminando os parâmetros, o que nos conduz a

$$\begin{cases} x=1+\mu \\ y+z=2+\mu \\ 2z+w=3-\mu \end{cases} \quad \text{e a} \quad \begin{cases} x+2z+w=4 \\ y+3z+w=5 \end{cases}$$

9-3 Paralelismo

Estudamos, neste parágrafo, um novo conceito geométrico, o de "paralelismo", que definimos da seguinte maneira:

Definição 9-18 Sejam $\mathcal{F}$ e $\mathcal{G}$ dois subespaços afins de um espaço afim $\mathcal{E}$, associados aos subespaços vectoriais F e G, respectivamente, do espaço vectorial E associado a $\mathcal{E}$. Diz-se que $\mathcal{F}$ é **paralelo** a $\mathcal{G}$, o que se representa por $\mathcal{F} /\!/ \mathcal{G}$, quando $F \subseteq G$ ou $G \subseteq F$.

Desta definição resultam imediatamente as seguintes propriedades:

Teorema 9-19 Num espaço afim, são válidas as seguintes propriedades, para subespaços afins $\mathcal{F}$ e $\mathcal{G}$ quaisquer, associados a subespaços vectoriais F e G, respectivamente:

a) $\mathcal{F} /\!/ \mathcal{F}$

b) Se $\mathcal{F} /\!/ \mathcal{G}$, então $\mathcal{G} /\!/ \mathcal{F}$

c) Se $\mathcal{F}$ e $\mathcal{G}$ têm a mesma dimensão finita, então $\mathcal{F} /\!/ \mathcal{G}$ se e só se $F = G$.

Demonstração:
Qualquer destas propriedades resulta trivialmente das definições.

Para que a noção de paralelismo apresentada corresponda à noção intuitiva de "paralelismo", importa realçar a seguinte propriedade:

Teorema 9-20 Num espaço afim, dois subespaços afins paralelos são disjuntos, a menos que um deles esteja contido no outro.

Demonstração:
Sejam $\mathcal{F}$ e $\mathcal{G}$ subespaços afins associados aos subespaços vectoriais F e G, respectivamente. Suponhamos que $\mathcal{F} /\!/ \mathcal{G}$, e seja, por exemplo, $F \subseteq G$. Se os subespaços não fossem disjuntos, então, pelo Lema 9-13, teríamos $\mathcal{F} \subseteq \mathcal{G}$.

Podemos ainda provar outras propriedades, para o que começamos por introduzir o seguinte lema:

Lema 9-21 Sejam $\mathcal{F} = \langle A; F \rangle$ e $\mathcal{G} = \langle B; G \rangle$ dois subespaços afins do espaço afim $\mathcal{E}$, associado ao espaço vectorial E. Se $F + G = E$, então $\mathcal{F} \cap \mathcal{G} \neq \varnothing$.

Demonstração:
Tomemos o vector $\overrightarrow{AB}$. Por ser $F + G = E$, podemos escrever $\overrightarrow{AB} = \vec{f} + \vec{g}$, com $\vec{f} \in F$ e $\vec{g} \in G$.

Consideremos então o ponto $C = A + \vec{f} \in \mathscr{F}$.
De

$$\overrightarrow{BC} = \overrightarrow{BA} + \overrightarrow{AC} = \overrightarrow{AC} - \overrightarrow{AB} = \vec{f} - \left(\vec{f} + \vec{g}\right) = -\vec{g}$$

resulta que $C = B - \vec{g} \in \mathscr{I}$. Logo, $C \in \mathscr{F} \cap \mathscr{I}$ e a intersecção não é vazia.

Podemos agora demonstrar o seguinte:

Teorema 9-22 Consideremos um espaço afim $\mathscr{E}$, associado a um espaço vectorial E.

a) Se $\mathscr{E}$ tem dimensão 2, então duas rectas disjuntas são paralelas.
b) Se $\mathscr{E}$ tem dimensão 3, então uma recta e um plano disjuntos são paralelos.
c) Se $\mathscr{E}$ tem dimensão 3, então dois planos disjuntos são paralelos.
d) De modo geral, num espaço de dimensão n, dois hiperplanos disjuntos são paralelos.

Demonstração:

a) Sejam $\mathscr{R}$ e $\mathscr{S}$ duas rectas disjuntas de um certo espaço afim. Tratando-se de subespaços de dimensão 1, podemos escrever $\mathscr{R} = \langle A; \langle \vec{u} \rangle \rangle$ e $\mathscr{S} = \langle B; \langle \vec{v} \rangle \rangle$, para certos vectores não nulos $\vec{u}$ e $\vec{v}$.

Se os vectores $\vec{u}$ e $\vec{v}$ fossem linearmente independentes, então formariam uma base do espaço (que tem dimensão 2), pelo que $\langle \vec{u} \rangle + \langle \vec{v} \rangle = E$. Mas então, pelo Lema 9-21, as duas rectas não seriam disjuntas, contra a hipótese. Concluímos, portanto, que os vectores $\vec{u}$ e $\vec{v}$ são linearmente dependentes, de onde resulta que $\langle \vec{u} \rangle = \langle \vec{v} \rangle$, e as duas rectas dadas são paralelas.

b) A demonstração é análoga à anterior: se $\mathscr{R}$ é uma recta e $\mathscr{P}$ é um plano, ponhamos $\mathscr{R} = \langle A; \langle \vec{u} \rangle \rangle$ e $\mathscr{P} = \langle B; \langle \vec{v}, \vec{w} \rangle \rangle$ e suponhamos que se trata de subespaços disjuntos.

Se o vector $\vec{u}$ não se pudesse escrever à custa de $\vec{v}$ e $\vec{w}$, os vectores $\vec{u}$, $\vec{v}$ e $\vec{w}$ formariam uma base de E e teríamos $\langle \vec{u} \rangle + \langle \vec{v}, \vec{w} \rangle = E$, contra a hipótese de se tratar de subespaços disjuntos. Logo, o vector $\vec{u}$ é dependente de $\vec{v}$ e $\vec{w}$, pelo que $\langle \vec{u} \rangle \subseteq \langle \vec{v}, \vec{w} \rangle$, e a recta é paralela ao plano.

c) A demonstração é análoga às anteriores. De resto, as afirmações a) e c) são casos particulares da propriedade d):

d) Supondo que $\mathscr{H}_1$ e $\mathscr{H}_2$ são hiperplanos de $\mathscr{E}$, podemos escrever

$$\mathscr{H}_1 = \langle A; \langle \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1} \rangle \rangle \text{ e } \mathscr{H}_2 = \langle A; \langle \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-1} \rangle \rangle$$

sendo os vectores indicados em cada caso linearmente independentes.

Se houvesse em $\langle \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-1} \rangle$ um vector que não se pudesse escrever à custa de $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}$, então teríamos

$$\dim\left(\langle \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1} \rangle + \langle \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-1} \rangle\right) > \dim\left(\langle \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1} \rangle\right) = n - 1$$

Sendo a dimensão do espaço igual a n, daqui resulta que

$$\dim\left(\langle \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}\rangle + \langle \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-1}\rangle\right) = n$$

pelo que

$$\langle \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}\rangle + \langle \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-1}\rangle = E$$

e então, pelo lema anterior, os dois hiperplanos não são disjuntos, o que vai contra a hipótese.

Consequentemente, todos os vectores de $\langle \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-1}\rangle$ se escrevem à custa de $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}$. Mas, neste caso, e atendendo às respectivas dimensões, podemos concluir que se tem $\langle \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-1}\rangle = \langle \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}\rangle$ e que os hiperplanos são paralelos.

Finalmente, podemos ainda observar o seguinte:

Teorema 9-23 Duas rectas paralelas e distintas de um espaço afim definem um plano.

Demonstração:

Sejam $\mathcal{R}$ e $\mathcal{S}$ duas rectas paralelas. Atendendo a que se trata de subespaços afins com a mesma dimensão, estarão associadas ao mesmo subespaço vectorial, pelo que poderemos escrever $\mathcal{R} = \langle A; \langle \vec{u}\rangle\rangle$ e $\mathcal{S} = \langle B; \langle \vec{u}\rangle\rangle$, para certo vector $\vec{u} \neq \vec{0}$.

Se os vectores $\overrightarrow{AB}$ e $\vec{u}$ fossem linearmente dependentes, poderíamos escrever $\overrightarrow{AB} = \xi\vec{u}$. Mas, então, $B = A + \xi\vec{u} \in \mathcal{R}$, pelo que

$$\mathcal{R} = \langle A; \langle \vec{u}\rangle\rangle = \langle B; \langle \vec{u}\rangle\rangle = \mathcal{S}$$

o que vai contra a hipótese.

Consequentemente, os vectores $\overrightarrow{AB}$ e $\vec{u}$ são linearmente independentes, pelo que o subespaço afim $\mathcal{P} = \langle A; \langle \vec{u}, \overrightarrow{AB}\rangle\rangle$ é um plano e inclui as duas rectas dadas.

É também claro que um plano que contenha as duas rectas coincide com este plano $\mathcal{P}$ que acabámos de construir.

9-4 Espaços euclidianos

No que se segue, vamos concentrar-nos no estudo de um espaço afim real, de dimensão finita, em cujo espaço vectorial associado se introduziu um produto interno.

Definição 9-24 Um espaço afim $\mathscr{E}$, de dimensão finita, associado a um espaço vectorial real E, com um produto interno fixo, chama-se **espaço euclidiano**. A um referencial de um espaço euclidiano, cuja base seja ortonormada, chama-se **referencial ortonormado**.

Vimos já que a introdução de um produto interno num espaço vectorial permite a definição de conceitos de natureza geométrica, como os de "norma" ou "comprimento" de um vector, "ângulo de dois vectores", etc. Quando a um tal espaço vectorial está associado um espaço afim, essas noções podem, por sua vez, aplicar-se aos pontos e aos conjuntos de pontos, conforme vamos ver.

Definição 9-25 Sejam X e Y pontos de um espaço euclidiano $\mathscr{E}$. Chama-se **distância** de X a Y ao número real não negativo $d(X,Y)=\left\|\overrightarrow{XY}\right\|$.

Deve observar-se que, se se encontrar fixo no espaço um referencial ortonormado $(O;\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n)$, e se os pontos X e Y tiverem, em relação a esse referencial, as coordenadas $X\equiv(x_1,x_2,\dots,x_n)$ e $Y\equiv(y_1,y_2,\dots,y_n)$, então

$$\overrightarrow{XY}=(y_1-x_1,y_2-x_2,\dots,y_n-x_n)$$

e, por consequência,

$$d(X,Y)=\left\|\overrightarrow{XY}\right\|=\sqrt{(y_1-x_1)^2+(y_2-x_2)^2+\dots+(y_n-x_n)^2}$$

Das definições dadas resultam facilmente as seguintes propriedades:

Teorema 9-26 Num espaço euclidiano, são válidas, para pontos arbitrários, as seguintes afirmações:

a) $d(X,Y)=0$ se e só se $X=Y$

b) $d(X,Y)=d(Y,X)$

c) $d(X,Y)\le d(X,Z)+d(Z,Y)$

Demonstração:
a) De acordo com as definições e as propriedades já conhecidas, vemos que:

$$d(X,Y)=0 \Leftrightarrow \left\|\overrightarrow{XY}\right\|=0 \Leftrightarrow \overrightarrow{XY}=\vec{0} \Leftrightarrow X=Y$$

b) $d(X,Y)=\left\|\overrightarrow{XY}\right\|=\left\|-\overrightarrow{XY}\right\|=\left\|\overrightarrow{XY}\right\|=d(Y,X)$

c) $d(X,Y)=\left\|\overrightarrow{XY}\right\|=\left\|\overrightarrow{XZ}+\overrightarrow{ZY}\right\|\leq\left\|\overrightarrow{XZ}\right\|+\left\|\overrightarrow{ZY}\right\|=d(X,Z)+d(Z,Y)$

A noção de distância, definida entre pontos, pode estender-se a dois conjuntos, da seguinte maneira:

Definição 9-27 Seja $\mathscr{E}$ um espaço euclidiano. Define-se a **distância** entre dois subconjuntos não vazios, $\mathscr{A}$ e $\mathscr{B}$, do espaço $\mathscr{E}$, como sendo

$$d(\mathscr{A},\mathscr{B})=\inf\{d(X,Y): X\in\mathscr{A}, Y\in\mathscr{B}\}$$

A existência do ínfimo do conjunto das distâncias entre os pontos de $\mathscr{A}$ e os pontos de $\mathscr{B}$ é garantida pelo facto de as distâncias serem sempre números reais não negativos: o conjunto $\{d(X,Y) : X\in\mathscr{A}, Y\in\mathscr{B}\}$ é limitado inferiormente, e, por isso, o ínfimo existe.

Em particular, convém observar que se os dois conjuntos não são disjuntos, então a distância entre eles é 0. Com efeito, tomando um ponto Z comum aos dois conjuntos, o valor $0=d(Z,Z)$ pertence ao conjunto $\{d(X,Y): X\in\mathscr{A}, Y\in\mathscr{B}\}$, sendo, automaticamente, o seu mínimo.

Quando os dois conjuntos em causa são subespaços afins, a distância entre eles acaba sempre por ser o mínimo das distâncias entre os pontos de um e os pontos do outro, como veremos a seguir. Para o efeito, começamos por definir um conceito de ortogonalidade. Embora se possa definir, de modo mais geral, o conceito de ortogonalidade entre dois subespaços afins quaisquer, restringir-nos-emos ao caso em que um deles é uma recta:

Definição 9-28 Sejam $\mathscr{E}$ um espaço euclidiano, $\mathscr{R}=\langle A;\langle\vec{u}\rangle\rangle$ uma recta de $\mathscr{E}$ e $\mathscr{F}=\langle B;F\rangle$ um subespaço afim qualquer de $\mathscr{E}$. Diz-se que a recta $\mathscr{R}$ é ortogonal ao subespaço $\mathscr{F}$, o que se representa por $\mathscr{R}\perp\mathscr{F}$, quando $\vec{u}\in F^{\perp}$.

Estamos agora em condições de provar o seguinte:

Teorema 9-29 Seja $\mathscr{E}$ um espaço euclidiano associado ao espaço vectorial E. Seja $\mathscr{F}=\langle Q;F\rangle$ um subespaço afim de $\mathscr{E}$, e P um ponto não pertencente a $\mathscr{F}$. Existe uma e uma só recta que passa por P, é ortogonal a $\mathscr{F}$ e intersecta $\mathscr{F}$.

Demonstração:
Consideremos o subespaço afim $\mathcal{I} = \left\langle P; F^{\perp} \right\rangle$. Pelo facto de ser $F + F^{\perp} = E$, sabemos que se tem $\mathcal{F} \cap \mathcal{I} \neq \varnothing$.

Por consequência, $\mathcal{F} \cap \mathcal{I}$ será um subespaço afim de $\mathcal{E}$ associado ao subespaço vectorial $F \cap F^{\perp}$. Ora, como $F \cap F^{\perp} = \left\{\vec{0}\right\}$, o subespaço afim $\mathcal{F} \cap \mathcal{I}$ reduzir-se-á a um ponto: $\mathcal{F} \cap \mathcal{I} = \left\{P'\right\}$.

Pelo facto de o ponto P não pertencer a $\mathcal{F}$, será distinto do ponto P'. Tomemos, então, a recta $\mathcal{R}$ definida pelos pontos P e P', ou seja, $\mathcal{R} = \left\langle P; \left\langle \overrightarrow{PP'} \right\rangle \right\rangle$.

Os pontos P e P' pertencem ambos a $\mathcal{I}$, pelo que $\overrightarrow{PP'} \in F^{\perp}$ e a recta $\mathcal{R}$ é, por definição, ortogonal a $\mathcal{F}$.

Quanto à unicidade, se $\mathcal{S} = \left\langle P; \left\langle \vec{u} \right\rangle \right\rangle$ é uma recta que passa por P e é ortogonal a $\mathcal{F}$, então $\vec{u} \in F^{\perp}$, pelo que $\mathcal{S} \subseteq \mathcal{I}$. Mas então

$$\mathcal{F} \cap \mathcal{S} \subseteq \mathcal{F} \cap \mathcal{I} = \left\{P'\right\}$$

Portanto, se a intersecção de $\mathcal{F}$ com $\mathcal{S}$ não é vazia, será $\mathcal{F} \cap \mathcal{S} = \left\{P'\right\}$. Ora, se $P, P' \in \mathcal{S}$, então $\mathcal{S}$ é a recta $\mathcal{R}$, definida pelos dois pontos.

O ponto P', que interveio na demonstração do teorema anterior, é único, pelo que lhe podemos atribuir uma designação apropriada:

Definição 9-30 Seja $\mathcal{F}$ um subespaço afim de um espaço euclidiano $\mathcal{E}$, e seja P um ponto não pertencente a $\mathcal{F}$. O ponto P', intersecção com $\mathcal{F}$ da única recta que passa por P, é ortogonal a $\mathcal{F}$ e intersecta $\mathcal{F}$, é o **pé da perpendicular** de P para $\mathcal{F}$.

Podemos agora verificar o seguinte:

Corolário 9-31 Seja $\mathcal{F}$ um subespaço afim de um espaço euclidiano $\mathcal{E}$ e seja P um ponto não pertencente a $\mathcal{F}$. Designemos por P' o pé da perpendicular de P para $\mathcal{F}$. Então, tem-se: $d(P,\mathcal{F}) = d\left(P, P'\right)$.

É claro que, quando P pertence a $\mathcal{F}$, se tem $d(P,\mathcal{F}) = 0$.

Demonstração:
Seja X um ponto qualquer do subespaço afim $\mathcal{F}$. Tem-se:

$$\begin{aligned}\left\|\overrightarrow{PX}\right\|^2 &= \overrightarrow{PX} \bullet \overrightarrow{PX} = \left(\overrightarrow{PP'} + \overrightarrow{P'X}\right) \bullet \left(\overrightarrow{PP'} + \overrightarrow{P'X}\right) = \\ &= \overrightarrow{PP'} \bullet \overrightarrow{PP'} + \overrightarrow{PP'} \bullet \overrightarrow{P'X} + \overrightarrow{P'X} \bullet \overrightarrow{PP'} + \overrightarrow{P'X} \bullet \overrightarrow{P'X}\end{aligned}$$

Ora, por construção, se designarmos por F o subespaço vectorial associado a $\mathcal{F}$, temos $\overrightarrow{P'X} \in F$ e $\overrightarrow{PP'} \in F^{\perp}$, pelo que $\overrightarrow{PP'} \bullet \overrightarrow{P'X} = \overrightarrow{P'X} \bullet \overrightarrow{PP'} = 0$. Por consequência,

$$\left\|\overrightarrow{PX}\right\|^2 = \overrightarrow{PP'} \bullet \overrightarrow{PP'} + \overrightarrow{P'X} \bullet \overrightarrow{P'X} = \left\|\overrightarrow{PP'}\right\|^2 + \left\|\overrightarrow{P'X}\right\|^2 \geq \left\|\overrightarrow{PP'}\right\|^2$$

Como as normas não são negativas, concluímos que

$$\left\|\overrightarrow{PX}\right\| \geq \left\|\overrightarrow{PP'}\right\| \quad \text{ou seja, que} \quad d(P,X) \geq d(P,P')$$

A tese resulta imediatamente do facto de se ter $P' \in \mathcal{F}$ e da arbitrariedade de X neste subespaço afim.

Exemplo 9-32

Consideremos, no espaço vectorial $\mathbb{R}^4$, o produto interno definido, em relação à base canónica, pela matriz da métrica

$$G = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

Vamos calcular a distância do ponto $P = (1,1,-1,2)$ ao plano $\mathcal{P}$ de equações cartesianas

$$\begin{cases} x - y + z + w = 1 \\ x + y + z - 2w = 1 \end{cases}$$

Partindo de

$$\begin{cases} x - y + z + w = 1 \\ x + y + z - 2w = 1 \end{cases} \quad \text{obtemos} \quad \begin{cases} x = 1 - z + \dfrac{1}{2} \\ y = \dfrac{3}{2} w \end{cases}$$

com z e w quaisquer. Por conseguinte, os pontos do plano $\mathcal{P}$ são da forma

$$X = \left(1 - z + \frac{1}{2}w, \frac{3}{2}w, z, w\right) = (1,0,0,0) + z(-1,0,1,0) + w\left(\frac{1}{2}, \frac{3}{2}, 0, 1\right)$$

ou, o que é equivalente,

$$X = (1,0,0,0) + \alpha(-1,0,1,0) + \beta(1,3,0,2) = (1 - \alpha + \beta, 3\beta, \alpha, 2\beta)$$

onde α e β são parâmetros reais.

Consideremos o vector genérico

$$\overrightarrow{PX} = X - P = (1-\alpha+\beta, 3\beta, \alpha, 2\beta) - (1,1,-1,2) = (-\alpha+\beta, 3\beta-1, \alpha+1, 2\beta-2)$$

Para que este vector seja ortogonal ao subespaço vectorial associado ao plano $\mathcal{P}$, o seu produto interno pelos dois geradores desse subespaço deverá ser nulo:

$$\begin{bmatrix}-\alpha+\beta & 3\beta-1 & \alpha+1 & 2\beta-2\end{bmatrix}\begin{bmatrix}1 & 1 & 0 & 0\\ 1 & 2 & 0 & 0\\ 0 & 0 & 2 & 1\\ 0 & 0 & 1 & 2\end{bmatrix}\begin{bmatrix}-1 & 1\\ 0 & 3\\ 1 & 0\\ 0 & 2\end{bmatrix} = \begin{bmatrix}0 & 0\end{bmatrix}$$

Daqui resulta, resolvendo o sistema de equações resultante,

$$\begin{cases}\alpha = -\dfrac{7}{95}\\ \beta = \dfrac{37}{95}\end{cases}$$

pelo que o pé da perpendicular de P para $\mathcal{P}$ será o ponto P' que se obtém substituindo os parâmetros por estes valores na expressão de X:

$$P' = \left(\frac{9}{19}, \frac{111}{95}, -\frac{7}{95}, \frac{74}{95}\right)$$

Quanto à distância de P a $\mathcal{P}$, teremos

$$D(P,\mathcal{P}) = \left\|\overrightarrow{PP'}\right\| = \left\|\left(-\frac{50}{95}, \frac{116}{95}, \frac{88}{95}, -\frac{116}{95}\right)\right\| = \frac{2}{95}\left\|(-25,8,44,-58)\right\| =$$

$$= \frac{2}{95}\sqrt{\begin{bmatrix}-25 & 8 & 44 & -58\end{bmatrix}\begin{bmatrix}1 & 1 & 0 & 0\\ 1 & 2 & 0 & 0\\ 0 & 0 & 2 & 1\\ 0 & 0 & 1 & 2\end{bmatrix}\begin{bmatrix}-25\\ 8\\ 44\\ -58\end{bmatrix}} = \frac{2}{95}\sqrt{5849} \quad \blacksquare$$

Os resultados anteriores podem generalizar-se a subespaços quaisquer, mediante o seguinte teorema:

Teorema 9-33 Se $\mathcal{F}$ e $\mathcal{G}$ são subespaços afins disjuntos de um espaço euclidiano $\mathcal{E}$ associado ao espaço vectorial E, existe pelo menos uma recta ortogonal a ambos e que os encontra. Essa recta será única se e só se os subespaços vectoriais associados a $\mathcal{F}$ e $\mathcal{G}$ estão em soma directa.

Demonstração:
Consideremos $\mathcal{F} = \langle P; F \rangle$ e $\mathcal{G} = \langle Q; G \rangle$.

Pelo facto de os dois subespaços afins serem disjuntos, terá de ser $F + G \neq E$ (pelo Lema 9-21). Teremos, então,

$$(F+G)^{\perp} \neq \{\vec{0}\}$$

Tomemos, então, os subespaços afins auxiliares

$$\mathcal{H}_1 = \left\langle P; F \oplus (F+G)^{\perp} \right\rangle$$

$$\mathcal{H}_2 = \left\langle Q; G \oplus (F+G)^{\perp} \right\rangle$$

$$\mathcal{F}^* = \mathcal{F} \cap \mathcal{H}_1 \cap \mathcal{H}_2$$

$$\mathcal{G}^* = \mathcal{G} \cap \mathcal{H}_1 \cap \mathcal{H}_2$$

Vemos imediatamente que se tem

$$\mathcal{F} \subseteq \mathcal{H}_1 \quad \text{e} \quad \mathcal{G} \subseteq \mathcal{H}_2$$

pelo que

$$\mathcal{F}^* = \mathcal{F} \cap \mathcal{H}_2 \quad \text{e} \quad \mathcal{G}^* = \mathcal{G} \cap \mathcal{H}_1$$

Ora,

$$F + \left(G \oplus (F+G)^{\perp}\right) = (F+G) \oplus (F+G)^{\perp} = E$$

pelo que $\mathcal{F}^* \neq \varnothing$; analogamente se verifica que $\mathcal{G}^* \neq \varnothing$.

O subespaço vectorial associado a $\mathcal{F}^*$ é

$$F \cap \left(G \oplus (F+G)^{\perp}\right) = F \cap G$$

Do mesmo modo, $\mathcal{G}^*$ está associado ao subespaço vectorial $F \cap G$.

Seja agora A um ponto qualquer de $\mathcal{F}^*$, e A', o pé da perpendicular de A para $\mathcal{G}^*$ [note-se que $\mathcal{F}^* \cap \mathcal{G}^* = \mathcal{F} \cap \mathcal{H}_2 \cap \mathcal{G} \cap \mathcal{H}_1 = (\mathcal{F} \cap \mathcal{G}) \cap \mathcal{H}_2 \cap \mathcal{H}_1 = \varnothing \cap \mathcal{H}_2 \cap \mathcal{H}_1 = \varnothing$, pelo que $A \notin \mathcal{G}^*$].

Por construção, $\overrightarrow{AA'} \in (F \cap G)^{\perp} = F^{\perp} + G^{\perp}$.

Mas, pelo facto de ser $A \in \mathcal{F}^* \subseteq \mathcal{H}_1 \cap \mathcal{H}_2$ e $A' \in \mathcal{G}^* \subseteq \mathcal{H}_1 \cap \mathcal{H}_2$, teremos

$$A, A' \in \mathcal{H}_1 \cap \mathcal{H}_2$$

e então

$$\overrightarrow{AA'} \in \left(F \oplus (F+G)^{\perp}\right) \cap \left(G \oplus (F+G)^{\perp}\right) = (F+G)^{\perp} \oplus (F \cap G)$$

Consequentemente, podemos escrever

$$\overrightarrow{AA'} = \vec{f}' + \vec{g}', \quad \text{com} \quad \vec{f}' \in F^{\perp},\ \vec{g}' \in G^{\perp}$$

e também

$$\overrightarrow{AA'} = \vec{u} + \vec{v}, \quad \text{com} \quad \vec{u} \in (F+G)^{\perp},\ \vec{v} \in F \cap G$$

Ora, como $\vec{v} \in F$ e $\vec{v} \in G$, tem-se

$$\overrightarrow{AA'} \bullet \vec{v} = \left(\vec{f}' + \vec{g}'\right) \bullet \vec{v} = \vec{f}' \bullet \vec{v} + \vec{g}' \bullet \vec{v} = 0 + 0 = 0$$

Mas, por outro lado,

$$\overrightarrow{AA'} \bullet \vec{v} = (\vec{u} + \vec{v}) \bullet \vec{v} = \vec{u} \bullet \vec{v} + \vec{v} \bullet \vec{v} = 0 + \vec{v} \bullet \vec{v} = \|\vec{v}\|^2$$

Logo, $\vec{v} = \vec{0}$ e, portanto, $\overrightarrow{AA'} = \vec{u} \in (F+G)^{\perp}$.

Assim, a recta definida pelos pontos A e A' é simultaneamente ortogonal a $\mathcal{F}$ e a $\mathcal{G}$, encontrando $\mathcal{F}$ no ponto A, e $\mathcal{G}$ no ponto A'.

A afirmação respeitante à unicidade resulta imediatamente das construções feitas.

Deste teorema resulta, agora, a seguinte consequência:

Corolário 9-34 Sejam $\mathcal{F}$ e $\mathcal{G}$ dois subespaços afins disjuntos de um espaço euclidiano $\mathcal{E}$. Sejam A e B os pontos de encontro com $\mathcal{F}$ e $\mathcal{G}$, respectivamente, de uma recta simultaneamente ortogonal a ambos os subespaços. Então:

$$d(\mathcal{F},\mathcal{G}) = d(A,B) = \left\|\overrightarrow{AB}\right\|$$

É claro que, se $\mathcal{F}$ e $\mathcal{G}$ não são disjuntos, então $d(\mathcal{F},\mathcal{G}) = 0$.

Demonstração:

Tomemos pontos arbitrários $X \in \mathcal{F}$ e $Y \in \mathcal{G}$. Designando por F e G os subespaços vectoriais associados a $\mathcal{F}$ e a $\mathcal{G}$, respectivamente, teremos $\overrightarrow{XA} \in F$ e $\overrightarrow{BY} \in G$, pelo que $\overrightarrow{XA} \perp \overrightarrow{AB}$ e $\overrightarrow{BY} \perp \overrightarrow{AB}$.

Assim:

$$\begin{aligned}\overrightarrow{AB} \bullet \overrightarrow{XY} &= \overrightarrow{AB} \bullet \left(\overrightarrow{XA} + \overrightarrow{AB} + \overrightarrow{BY}\right) = \overrightarrow{AB} \bullet \overrightarrow{XA} + \overrightarrow{AB} \bullet \overrightarrow{AB} + \overrightarrow{AB} \bullet \overrightarrow{BY} = \\ &= 0 + \overrightarrow{AB} \bullet \overrightarrow{AB} + 0 = \left\|\overrightarrow{AB}\right\|^2\end{aligned}$$

Por outro lado,

$$\overrightarrow{AB} \bullet \overrightarrow{XY} = \left\|\overrightarrow{AB}\right\| \left\|\overrightarrow{XY}\right\| \cos\angle\left(\overrightarrow{AB}, \overrightarrow{XY}\right)$$

Igualando as duas expressões, obtemos

$$\left\|\overrightarrow{AB}\right\|^2 = \left\|\overrightarrow{AB}\right\| \left\|\overrightarrow{XY}\right\| \cos\angle\left(\overrightarrow{AB}, \overrightarrow{XY}\right)$$

ou

$$\left\|\overrightarrow{AB}\right\| = \left\|\overrightarrow{XY}\right\| \cos\angle\left(\overrightarrow{AB}, \overrightarrow{XY}\right)$$

Uma vez que os valores da função cos são sempre não superiores a 1, concluímos que

$$\left\|\overrightarrow{AB}\right\| \leq \left\|\overrightarrow{XY}\right\|$$

Dado que X e Y foram tomados arbitrariamente nos dois subespaços afins, esta desigualdade mostra que

$$d(A,B) = \left\|\overrightarrow{AB}\right\|$$

é, de facto, a distância entre eles.

Um caso particular que merece atenção é aquele em que os subespaços afins considerados são paralelos. Nesse caso, podemos ir um pouco mais longe:

Teorema 9-35 Se F e G são dois subespaços afins paralelos de um espaço afim $\mathscr{E}$, e se $\dim(\mathscr{F}) \leq \dim(\mathscr{G})$, então, dados quaisquer dois pontos $A, B \in \mathscr{F}$, tem-se

$$d(A,\mathscr{G}) = d(B,\mathscr{G})$$

pelo que $d(\mathscr{F},\mathscr{G}) = d(Q,\mathscr{G})$, qualquer que seja o ponto $Q \in \mathscr{G}$.

Demonstração:
Supondo que $\mathscr{F}$ não esteja contido em $\mathscr{G}$, sejam A' e B', respectivamente, os pés das perpendiculares de A e B para $\mathscr{G}$.

Tal como vimos acima,

$$d(A,\mathscr{G}) = \left\|\overrightarrow{AA'}\right\| \quad \text{e} \quad d(B,\mathscr{G}) = \left\|\overrightarrow{BB'}\right\|$$

Ora, $\overrightarrow{AB} = \overrightarrow{AA'} + \overrightarrow{A'B'} + \overrightarrow{BB'}$, onde, designando por F e G os subespaços vectoriais associados aos dois subespaços afins considerados, se tem

$$\overrightarrow{AB} \in F \subseteq G,\ \overrightarrow{A'B'} \in G,\ \overrightarrow{AA'} + \overrightarrow{BB'} \in G^{\perp}$$

Consideremos o vector $\vec{z} = \overrightarrow{AB} - \overrightarrow{A'B'} \in G$.
Da igualdade $\overrightarrow{AB} = \overrightarrow{AA'} + \overrightarrow{A'B'} + \overrightarrow{B'B}$ resulta que $\vec{z} = \overrightarrow{AA'} + \overrightarrow{B'B} \in G^{\perp}$.
Logo,

$$\vec{z} \in G \cap G^{\perp} = \{\vec{0}\}$$

pelo que $\vec{z} = \vec{0}$. Mas, então, $\overrightarrow{AA'} + \overrightarrow{B'B} = \vec{0}$, pelo que

$$\overrightarrow{AA'} = -\overrightarrow{B'B} = \overrightarrow{BB'} \text{ e } d(A,\mathcal{I}) = d(A,A') = \left\|\overrightarrow{AA'}\right\| = \left\|\overrightarrow{BB'}\right\| = d(B,B') = d(B,\mathcal{I})$$

9-5 Problemas métricos em $\mathbb{R}^3$

A importância do espaço afim $\mathbb{R}^3$ leva-nos a considerar, em particular, os problemas de cálculo de distâncias entre os seus subespaços (pontos, rectas ou planos). Na verdade, a aplicação directa dos métodos gerais, descritos no parágrafo anterior, é um tanto fastidiosa, dada a necessidade de se obterem pontos e rectas auxiliares, obedecendo a determinadas condições, a partir dos quais se calculam as distâncias desejadas. O que nos interessa é chegar a fórmulas que, imediatamente a partir dos subespaços dados, nos permitam calcular essas distâncias sem necessidade de recorrer a outras construções.

É necessário observar, também, que as fórmulas que vamos obter se aplicam também a subespaços de $\mathbb{R}^2$, já que este espaço afim se pode, trivialmente, identificar com o plano de $\mathbb{R}^3$ definido pela equação cartesiana $z = 0$.

Consideremos então o espaço $\mathbb{R}^3$, no qual supomos fixo um determinado produto interno. No que se segue, identificaremos os pontos pelas suas coordenadas em relação a um certo referencial, recordando que, se se tratar do referencial canónico, essas coordenadas coincidem com os próprios pontos considerados.

a) **Distância entre dois pontos**
Dados dois pontos, $A \equiv (a_1, a_2, a_3)$ e $B \equiv (b_1, b_2, b_3)$, temos, por definição,

$$d(A,B) = \left\|\overrightarrow{AB}\right\| = \sqrt{\begin{bmatrix} b_1 - a_1 & b_2 - a_2 & b_3 - a_3 \end{bmatrix} G \begin{bmatrix} b_1 - a_1 \\ b_2 - a_2 \\ b_3 - a_3 \end{bmatrix}}$$

onde G designa a matriz da métrica em relação à base do referencial fixo.

No caso de o referencial em causa ser ortonormado, a matriz da métrica será a matriz identidade, e obtemos a fórmula mais simples

$$d(A,B)=\sqrt{(b_1-a_1)^2+(b_2-a_2)^2+(b_3-a_3)^2}$$

b) **Distância de um ponto a um plano**
Seja Q um ponto e $\mathcal{P}=\langle P;F\rangle$, um plano do espaço $\mathbb{R}^3$.

É claro que se $Q\in\mathcal{P}$, se tem $d(Q,\mathcal{P})=0$. Suporemos, por isso, que $Q\notin\mathcal{P}$. Sendo Q' o pé da perpendicular de Q para $\mathcal{P}$, sabemos que

$$d(Q,\mathcal{P})=d(Q,Q')=\left\|\overrightarrow{QQ'}\right\|$$

Ora, por ser $\overrightarrow{QQ'}\perp\overrightarrow{Q'P}$, temos

$$\overrightarrow{QP}\bullet\overrightarrow{QQ'}=\left(\overrightarrow{QQ'}+\overrightarrow{Q'P}\right)\bullet\overrightarrow{QQ'}=\overrightarrow{QQ'}\bullet\overrightarrow{QQ'}+\overrightarrow{Q'P}\bullet\overrightarrow{QQ'}=\left\|\overrightarrow{QQ'}\right\|^2$$

Mas

$$\overrightarrow{QP}\bullet\overrightarrow{QQ'}=\left\|\overrightarrow{QP}\right\|\left\|\overrightarrow{QQ'}\right\|\cos\angle\left(\overrightarrow{QP},\overrightarrow{QQ'}\right)$$

pelo que, igualando as duas expressões obtidas, chegamos a

$$d(Q,\mathcal{P})=\left\|\overrightarrow{QQ'}\right\|=\left\|\overrightarrow{QP}\right\|\cos\angle\left(\overrightarrow{QP},\overrightarrow{QQ'}\right)$$

Esta fórmula, contudo, obrigar-nos-ia ainda a procurar o ponto Q', que não é dado. Podemos, no entanto, evitar um cálculo tão preciso. Com efeito, tomemos um vector qualquer, não nulo, $\vec{w}\in F^\perp$. Uma vez que F tem dimensão 2, o seu complemento ortogonal terá dimensão 1, pelo que poderemos escrever $\vec{w}=\zeta\overrightarrow{QQ'}$, e então

$$\cos\angle\left(\overrightarrow{QP},\overrightarrow{QQ'}\right)=\left|\cos\angle\left(\overrightarrow{QP},\vec{w}\right)\right|=\frac{\left|\overrightarrow{QP}\bullet\vec{w}\right|}{\left\|\overrightarrow{QP}\right\|\left\|\vec{w}\right\|}$$

Substituindo na igualdade anterior, chegamos, finalmente, a

$$d(Q,\mathcal{P})=\frac{\left|\overrightarrow{QP}\bullet\vec{w}\right|}{\left\|\vec{w}\right\|}$$

fórmula esta que, para além dos dados, necessita apenas do vector $\vec{w}\in F^\perp$.

No caso de se ter $F=\langle \vec{f}_1, \vec{f}_2 \rangle$, poderemos, por exemplo, escolher $\vec{w}=\vec{f}_1 \times \vec{f}_2$, uma vez que o produto externo é, por definição, ortogonal a ambos os vectores. Obter-se-á, neste caso,

$$d(Q,\mathcal{P})=\frac{\left|\overrightarrow{QP}\bullet\vec{f}_1\times\vec{f}_2\right|}{\left\|\vec{f}_1\times\vec{f}_2\right\|}$$

Por outro lado, se o referencial fixo é ortonormado e se o plano $\mathcal{P}$ é definido pela equação cartesiana

$$ax+by+cz=d$$

as coordenadas dos pontos do plano têm de satisfazer esta equação.

Se $P\equiv(p_1,p_2,p_3)$ e se $X_0\equiv(x_0,y_0,z_0)$ é um ponto qualquer do plano, teremos

$$\begin{aligned} ax_0+by_0+cz_0&=d\\ ap_1+bp_2+cp_3&=d \end{aligned}$$

pelo que, sucessivamente,

$$\begin{gathered} a(x_0-p_1)+b(y_0-p_2)+c(z_0-p_3)=0\\ (a,b,c)\bullet(x_0-p_1,y_0-p_2,z_0-p_3)=0\\ (a,b,c)\perp(x_0-p_1,y_0-p_2,z_0-p_3) \end{gathered}$$

Dada a arbitrariedade de X_0 no plano $\mathcal{P}$, concluímos que o vector $\vec{w}\equiv(a,b,c)$ é ortogonal a F. Podemos, portanto, usá-lo na fórmula anterior.

Se repararmos que

$$\begin{aligned} \overrightarrow{QP}\bullet\vec{w}&=(q_1-p_1)a+(q_2-p_2)b+(q_3-p_3)c=\\ &=aq_1-ap_1+bq_2-bp_2+cq_3-cp_3=\\ &=aq_1+bq_2+cq_3-(ap_1+bp_2+cp_3)=\\ &=aq_1+bq_2+cq_3-d \end{aligned}$$

obtemos a fórmula

$$d(Q,\mathcal{P})=\frac{|aq_1+bq_2+cq_3-d|}{\sqrt{a^2+b^2+c^2}}$$

c) **Distância de um ponto a uma recta**
Seja Q um ponto e seja $\mathcal{R} = \langle P; \langle \vec{u} \rangle \rangle$ uma recta do espaço $\mathbb{R}^3$. Admitindo que $Q \notin \mathcal{R}$, seja Q' o pé da perpendicular de Q para $\mathcal{R}$.

Tem-se

$$d(Q,\mathcal{R}) = d(Q,Q') = \left\| \overrightarrow{QQ'} \right\|$$

Ora,

$$\vec{u} \times \overrightarrow{QP} = \vec{u} \times \left(\overrightarrow{QQ'} + \overrightarrow{Q'P} \right) = \vec{u} \times \overrightarrow{QQ'} + \vec{u} \times \overrightarrow{Q'P} = \vec{u} \times \overrightarrow{QQ'} + \vec{0} = \vec{u} \times \overrightarrow{QQ'}$$

Daqui resulta que

$$\left\| \vec{u} \times \overrightarrow{QP} \right\| = \left\| \vec{u} \times \overrightarrow{QQ'} \right\| = \left\| \vec{u} \right\| \left\| \overrightarrow{QQ'} \right\| \operatorname{sen} \angle\left(\vec{u}, \overrightarrow{QQ'} \right) = \left\| \vec{u} \right\| \left\| \overrightarrow{QQ'} \right\| \operatorname{sen} \frac{\pi}{2} = \left\| \vec{u} \right\| \left\| \overrightarrow{QQ'} \right\|$$

Então,

$$d(Q,\mathcal{R}) = \left\| \overrightarrow{QQ'} \right\| = \frac{\left\| \vec{u} \times \overrightarrow{QP} \right\|}{\left\| \vec{u} \right\|}$$

fórmula esta que só utiliza os dados do problema.

d) **Distância entre dois planos**
Sejam $\mathcal{P}_1$ e $\mathcal{P}_2$ dois planos de $\mathbb{R}^3$.

Se os planos não são disjuntos, a distância entre eles é 0. Se são disjuntos, então, conforme vimos, são paralelos, pelo que, de acordo com o Teorema 9-35,

$$d(\mathcal{P}_1, \mathcal{P}_2) = d(P, \mathcal{P}_2), \forall P \in \mathcal{P}_1$$

e) **Distância entre uma recta e um plano**
Seja $\mathcal{R}$ uma recta e $\mathcal{P}$ um plano de $\mathbb{R}^3$.

Se a recta encontra o plano, a distância entre eles é 0. Se os dois subespaços são disjuntos, então, conforme vimos, serão paralelos, pelo que

$$d(\mathcal{R}, \mathcal{P}) = d(P, \mathcal{P}), \forall P \in \mathcal{R}$$

f) **Distância entre duas rectas**
Sejam $\mathcal{R}_1$ e $\mathcal{R}_2$ duas rectas de $\mathbb{R}^3$.

Se as duas rectas se encontram, a distância entre elas é 0. Se não se encontram mas são paralelas, teremos, como acima,

$$d(\mathcal{R}_1, \mathcal{R}_2) = d(P, \mathcal{R}_2), \forall P \in \mathcal{R}_1$$

Resta, portanto, estudar o caso em que as duas rectas não se encontram mas também não são paralelas (dizemos, neste caso, que se trata de rectas **enviesadas**).

Seja $\mathcal{S}$ uma recta simultaneamente ortogonal a $\mathcal{R}_1$ e a $\mathcal{R}_2$ e que as encontre, respectivamente, nos pontos R_1 e R_2. Teremos, então,

$$d(\mathcal{R}_1,\mathcal{R}_2) = d(R_1,R_2) = \left\|\overrightarrow{R_1R_2}\right\|$$

Suponhamos que se tem $\mathcal{R}_1 = \langle P_1;\langle\vec{u}_1\rangle\rangle$ e $\mathcal{R}_2 = \langle P_2;\langle\vec{u}_2\rangle\rangle$, e, atendendo a que os vectores $\vec{u}_1$ e $\vec{u}_2$ são independentes visto as duas rectas não serem paralelas, consideremos o plano auxiliar $\mathcal{P} = \langle P_2;\langle\vec{u}_1,\vec{u}_2\rangle\rangle$. Este plano contém a recta $\mathcal{R}_2$ e é paralelo a $\mathcal{R}_1$.

O vector $\overrightarrow{R_1R_2}$ pertence a $\langle\vec{u}_1,\vec{u}_2\rangle^{\perp}$ e $R_2 \in \mathcal{R}_2 \subseteq \mathcal{P}$. Portanto,

$$d(\mathcal{R}_1,\mathcal{R}_2) = d(R_1,R_2) = d(R_1,\mathcal{P}) = d(P_1,\mathcal{P}) = \frac{\left|\vec{u}_1 \times \vec{u}_2 \bullet \overrightarrow{P_1P_2}\right|}{\|\vec{u}_1 \times \vec{u}_2\|}$$

Para além destes problemas relacionados com o cálculo de distâncias, podemos ainda introduzir conceitos de ângulos entre os subespaços de $\mathbb{R}^3$. Começamos por definir o ângulo entre duas rectas:

Definição 9-36

Dadas as rectas $\mathcal{R}_1 = \langle P_1;\langle\vec{u}_1\rangle\rangle$ e $\mathcal{R}_2 = \langle P_2;\langle\vec{u}_2\rangle\rangle$ do espaço euclidiano $\mathbb{R}^3$, chama-se **ângulo entre as duas rectas** ao menor dos ângulos $\angle(\vec{u}_1,\vec{u}_2)$ e $\angle(\vec{u}_1,-\vec{u}_2)$, ou seja,

$$\angle(\mathcal{R}_1,\mathcal{R}_2) = \min\left(\angle(\vec{u}_1,\vec{u}_2),\angle(\vec{u}_1,-\vec{u}_2)\right)$$

Tal como no caso das distâncias, podemos procurar uma fórmula para o cálculo do ângulo entre duas rectas. Com efeito, nas condições da definição anterior, vemos que

$$\cos\angle(\vec{u}_1,-\vec{u}_2) = \frac{\vec{u}_1 \bullet (-\vec{u}_2)}{\|\vec{u}_1\|\|-\vec{u}_2\|} = -\frac{\vec{u}_1 \bullet \vec{u}_2}{\|\vec{u}_1\|\|\vec{u}_2\|} = -\cos\angle(\vec{u}_1,\vec{u}_2)$$

Atendendo a que, no intervalo $[0,\pi]$, a função co-seno é decrescente, ao ângulo menor corresponderá o co-seno maior, ou seja:

$$\angle(\mathcal{R}_1,\mathcal{R}_2) = \left|\cos\angle(\vec{u}_1,\vec{u}_2)\right| = \frac{|\vec{u}_1 \bullet \vec{u}_2|}{\|\vec{u}_1\|\|\vec{u}_2\|}$$

Passemos agora ao estudo do ângulo de uma recta com um plano, dando a respectiva definição:

Definição 9-37 O **ângulo de uma recta com um plano**, do espaço euclidiano $\mathbb{R}^3$, é o ângulo de um gerador do subespaço vectorial associado à recta com a sua projecção ortogonal sobre o subespaço vectorial associado ao plano.

Partindo desta definição, vamos procurar uma fórmula para o cálculo do ângulo de uma recta $\mathcal{R} = \langle P; \langle \vec{u} \rangle \rangle$ com um plano $\mathrm{P} = \langle Q; F \rangle$.

Começaremos por verificar que a definição apresentada não depende do gerador escolhido para o subespaço vectorial associado à recta $\mathcal{R}$. Para isso, fixemos em F uma base ortonormada $\vec{f}_1, \vec{f}_2$; sabe-se que a projecção ortogonal de $\vec{u}$ sobre F e o vector

$$p_F^\perp(\vec{u}) = \vec{u}' = \left(\vec{u} \bullet \vec{f}_1\right)\vec{f}_1 + \left(\vec{u} \bullet \vec{f}_2\right)\vec{f}_2$$

Tem-se

$$\vec{u} \bullet \vec{u}' = \left(\vec{u} \bullet \vec{f}_1\right)^2 + \left(\vec{u} \bullet \vec{f}_2\right)^2 = \|\vec{u}'\|^2 \quad \text{ou seja,} \quad \|\vec{u}\|\,\|\vec{u}'\| \cos\angle\left(\vec{u}, \vec{u}'\right) = \|\vec{u}'\|^2$$

Tomemos agora um vector $\vec{z} = \xi\,\vec{u}$, com $\xi \neq 0$. Se $\vec{u} \in F^\perp$, então também $\vec{z} \in F^\perp$, pelo que poderemos supor que $\vec{u} \notin F^\perp$, e, nesse caso, $\vec{u}' \neq \vec{0}$. Podemos, portanto, concluir, da igualdade anterior, que

$$\cos\angle\left(\vec{u}, \vec{u}'\right) = \frac{\|\vec{u}'\|}{\|\vec{u}\|}$$

Ora, a projecção ortogonal de $\vec{z} = \xi\,\vec{u}$ sobre F é

$$p_F^\perp(\vec{z}) = \vec{z}' = \left(\vec{z} \bullet \vec{f}_1\right)\vec{f}_1 + \left(\vec{z} \bullet \vec{f}_2\right)\vec{f}_2\, p_F^\perp(\vec{u}) = \vec{u}' = \left(\vec{u} \bullet \vec{f}_1\right)\vec{f}_1 + \left(\vec{u} \bullet \vec{f}_2\right)\vec{f}_2 =$$
$$= \left(\xi\,\vec{u} \bullet \vec{f}_1\right)\vec{f}_1 + \left(\xi\,\vec{u} \bullet \vec{f}_2\right)\vec{f}_2 = \xi\left(\left(\vec{u} \bullet \vec{f}_1\right)\vec{f}_1 + \left(\vec{u} \bullet \vec{f}_2\right)\vec{f}_2\right) = \xi\,\vec{u}'$$

Por conseguinte,

$$\cos\angle\left(\vec{z}, \vec{z}'\right) = \frac{\|\vec{z}'\|}{\|\vec{z}\|} = \frac{\|\xi\,\vec{u}'\|}{\|\xi\,\vec{u}\|} = \frac{\|\vec{u}'\|}{\|\vec{u}\|} = \cos\angle\left(\vec{u}, \vec{u}'\right)$$

Para obtermos, então, a fórmula que nos permite calcular facilmente o ângulo de uma recta com um plano, partiremos do seguinte lema:

Lema 9-38 O ângulo de uma recta com um plano é o complementar (isto é, a diferença para $\pi/2$) do ângulo dessa recta com uma recta ortogonal ao plano.

Demonstração:

Usaremos a notação acima.

Por definição, a projecção ortogonal $p_F^{\perp}(\vec{u}) = \vec{u}'$ é um vector de F tal que $\vec{u} = \vec{u}' + \vec{w}$, com $\vec{w} \in F^{\perp}$.

Excluindo os casos triviais em que $\vec{u}$ pertença a F ou ao seu complemento ortogonal, podemos tomar os versores

$$\vec{v}_1 = \frac{\vec{u}'}{\|\vec{u}'\|} \quad \text{e} \quad \vec{v}_2 = \frac{\vec{w}}{\|\vec{w}\|}$$

e, neste caso,

$$\vec{u} = \|\vec{u}'\|\vec{v}_1 + \|\vec{w}\|\vec{v}_2 = (\vec{u}\bullet\vec{v}_1)\vec{v}_1 + (\vec{u}\bullet\vec{v}_2)\vec{v}_2$$

pelo que

$$\|\vec{u}'\| = \vec{u}\bullet\vec{v}_1,\ \|\vec{w}\| = \vec{u}\bullet\vec{v}_2$$

Temos, sucessivamente,

$$\|\vec{u}\|^2 = \|\vec{u}'\|^2 + (\vec{u}\bullet\vec{v}_2)^2$$

$$1 = \frac{\|\vec{u}'\|^2}{\|\vec{u}\|^2} + \frac{(\vec{u}\bullet\vec{v}_2)^2}{\|\vec{u}\|^2}$$

$$1 = \cos^2\angle(\vec{u},\vec{u}') + \frac{\|\vec{u}\|^2\,\|\vec{v}_2\|^2\cos^2\angle(\vec{u},\vec{v}_2)}{\|\vec{u}\|^2}$$

$$1 = \cos^2\angle(\vec{u},\vec{u}') + \cos^2\angle(\vec{u},\vec{w})$$

$$\cos^2\angle(\vec{u},\vec{u}') = 1 - \cos^2\angle(\vec{u},\vec{w}) = \operatorname{sen}^2\angle(\vec{u},\vec{w})$$

Como $\cos\angle(\vec{u},\vec{u}') \geq 0$ e $\operatorname{sen}\angle(\vec{u},\vec{w}) \geq 0$, temos

$$\cos\angle(\vec{u},\vec{u}') = \operatorname{sen}\angle(\vec{u},\vec{w})$$

pelo que

$$\angle(\vec{u},\vec{u}') = \frac{\pi}{2} - \angle(\vec{u},\vec{w})$$

Em face do lema que acabámos de demonstrar e usando a mesma notação que acima, vemos, então, que

$$\operatorname{sen}\angle(\mathcal{R},\mathcal{P}) = \cos\angle(\vec{u},\vec{w})$$

Supondo que $\vec{u} \equiv (u_1, u_2, u_3)$ e que o plano $\mathcal{P}$ tem a equação cartesiana $ax + by + cz = d$ em relação a um referencial ortonormado, teremos ainda

$$\operatorname{sen}\angle(\mathcal{R},\mathcal{P}) = \frac{|au_1 + bu_2 + cu_3|}{\sqrt{a^2 + b^2 + c^2}\sqrt{u_1^2 + u_2^2 + u_3^2}}$$

Finalmente, ocupar-nos-emos do ângulo entre dois planos. Eis a definição:

Definição 9-39 O **ângulo entre dois planos** é o ângulo entre duas rectas ortogonais, respectivamente, a cada um deles.

Usando as fórmulas anteriormente obtidas, vemos imediatamente que, se os planos $\mathcal{P}_1$ e $\mathcal{P}_2$ têm, respectivamente, as equações cartesianas $a_1x + b_1y + c_1z = d_1$ e $a_2x + b_2y + c_2z = d_2$ em relação a um referencial ortonormado, então

$$\cos\angle(\mathcal{P}_1,\mathcal{P}_2) = \frac{|a_1a_2 + b_1b_2 + c_1c_2|}{\sqrt{a_1^2 + b_1^2 + c_1^2}\sqrt{a_2^2 + b_2^2 + c_2^2}}$$

9-6 Cónicas em $\mathbb{R}^2$

Vimos, em parágrafos anteriores, que os subespaços afins de um espaço afim de dimensão finita se podem representar por equações (ou sistemas de equações) lineares. É natural pensar-se nos conjuntos de pontos que possam ser representados por equações polinomiais de segundo grau. Vamos tratar do assunto, restringindo-nos ao espaço vectorial $\mathbb{R}^2$.

Introduzimos, então, a seguinte designação:

Definição 9-40 Chama-se **cónica** ao conjunto dos pontos do espaço afim $\mathbb{R}^2$ que satisfazem uma dada equação polinomial de grau 2, com duas variáveis.

Fixando, no espaço afim $\mathbb{R}^2$, um referencial $(O; \vec{e}_1, \vec{e}_2)$, designemos por Λ a cónica constituída pelos pontos cujas coordenadas, em relação a esse referencial, satisfaçam uma equação polinomial genérica, de segundo grau, que, por comodidade e sem perda de generalidade, apresentaremos sob a forma

$$a_{11}x^2 + a_{22}y^2 + 2a_{12}xy + 2b_1x + 2b_2y + c = 0$$

Nesta equação, algum dos coeficientes a_{ij} $(i, j = 1, 2)$ terá de ser não nulo, visto supormos que se trata de uma equação de segundo grau.

Chamaremos às parcelas $a_{11}x^2, a_{22}y^2$ os **termos quadrados** da equação; à parcela $2a_{12}xy$, o **termo rectangular** da equação; às parcelas $2b_1x, 2b_2y$, os **termos lineares** da equação, e a c, o **termo independente** da equação.

Por sua vez, se considerarmos as matrizes

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \text{ e } X = \begin{bmatrix} x \\ y \end{bmatrix}$$

a equação anterior pode escrever-se, de modo mais abreviado, na forma

$$\Phi(X) = X^T AX + 2B^T X + c = 0$$

Note-se que, por construção, a matriz A é uma matriz simétrica.

É claro que algumas cónicas correspondem a conjuntos facilmente identificáveis.

Assim, por exemplo, a equação

$$x^2 + y^2 + 1 = 0$$

é, obviamente, impossível, pelo que define o conjunto vazio; por sua vez, a equação

$$x^2 + y^2 = 0$$

tem como única solução o ponto $(0,0)$, e a equação

$$x^2 + y^2 + 2xy = 0 \Leftrightarrow (x+y)^2 = 0 \Leftrightarrow x + y = 0$$

define uma recta de $\mathbb{R}^2$.

Pretendemos obter uma classificação exaustiva das cónicas, de modo a poder identificar cada uma delas. Pode acontecer que a equação de Λ em relação ao referencial considerado não seja suficientemente simples para permitir a identificação da cónica, mas que se simplifique, em relação a outros referenciais. Por isso, vamos estudar as alterações que cada mudança de referencial acarreta na equação geral acima considerada, considerando separadamente duas possibilidades: mudança de base e mudança de origem.

9-6-1 Mudança da base

Vejamos agora o que sucede à equação de Λ, quando passarmos do referencial $(O;\vec{e}_1,\vec{e}_2)$ inicial para um referencial $(O;\vec{e}_1',\vec{e}_2')$.

Sendo P a matriz de mudança de base, as coordenadas, X e X', de um mesmo ponto em relação a cada um dos referenciais considerados, relacionam-se pela igualdade

$$X = PX'$$

A equação da cónica transforma-se, portanto, em

$$(PX')^T A(PX') + 2B^T(PX') + c = 0 \quad \text{ou} \quad X'^T(P^TAP)X' + 2(B^TP)X' + c = 0$$

Vemos, pois, que a mudança da base não altera o termo independente da equação da cónica, modificando os demais.

Suponhamos fixo em $\mathbb{R}^2$ um produto interno, e seja G a matriz da métrica em relação à base $\vec{e}_1,\vec{e}_2$, do referencial inicial.

Consideremos a matriz $S = G^{-1}A$.

Teorema 9-41 A matriz S define, em relação à base $\vec{e}_1,\vec{e}_2$, um endomorfismo diagonalizável de $\mathbb{R}^2$, e existe no espaço uma base ortonormada formada por vectores próprios desse endomorfismo.

Demonstração:

Observemos que se tem

$$S^TG = (G^{-1}A)^T G = A^T(G^{-1})^T G = A(G^T)^{-1}G = AG^{-1}G = A = G(G^{-1}A) = GS$$

A igualdade $S^TG = GS$ mostra que o endomorfismo definido por S é um endomorfismo auto-adjunto do espaço vectorial $\mathbb{R}^3$.

A tese resulta então imediatamente do Teorema 7-41.

Em face destes resultados, introduzimos a seguinte definição:

Definição 9-42 Nas condições anteriores, se $\vec{u}_1,\vec{u}_2$ é uma base ortonormada, constituída por vectores próprios do endomorfismo ψ, definido pela matriz S, em relação à base $\vec{e}_1,\vec{e}_2$, dizemos que os vectores $\vec{u}_1,\vec{u}_2$ são **direcções principais** da cónica Λ.

Deve observar-se que, quando a base $\vec{e}_1, \vec{e}_2$, inicialmente considerada, é ortonormada, se tem $S = G^{-1}A = I_2^{-1}A = A$, pelo que, nesse caso, as direcções principais serão vectores próprios ortonormados do endomorfismo definido, em relação à base considerada, pela matriz A.

A partir do que ficou dito, podemos agora analisar o que acontece à equação geral de uma cónica, quando mudamos a base do referencial inicial para uma base formada por direcções principais da cónica.

Sendo P a matriz de mudança de base, sabemos já que a equação geral de Λ assume a forma

$$X'^T\left(P^TAP\right)X' + 2\left(B^TP\right)X' + c = 0$$

Ora, a matriz da métrica em relação à base ortonormada $\vec{u}_1, \vec{u}_2$ é I_3, e sabemos que esta matriz se relaciona com a matriz G através da igualdade $I_2 = P^TGP$, que equivale a $P^T = (GP)^{-1}$.

Por sua vez, a matriz do endomorfismo ψ em relação à base $\vec{u}_1, \vec{u}_2$, formada por vectores próprios de ψ, será

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = P^{-1}SP = P^{-1}G^{-1}AP = (GP)^{-1}AP = P^TAP$$

Consequentemente, a equação da cónica Λ terá a forma

$$X'^T\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}X' + 2\left(B^TP\right)X' + c = 0$$

ou seja, a forma

$$\lambda_1x^2 + \lambda_2y^2 + 2b_1'x + 2b_2'y + c' = 0$$

Esta equação é chamada **equação reduzida às direcções principais** da cónica Λ.

9-6-2 Mudança da origem (translação do referencial)

Suponhamos que passamos do referencial considerado inicialmente para um referencial $\left(O'; \vec{e}_1, \vec{e}_2\right)$.

As coordenadas, X e X', de um mesmo ponto em relação a cada um dos referenciais considerados, relacionam-se, como vimos oportunamente, pela igualdade

$$X = X' + C$$

onde C designa a coluna das coordenadas de O' em relação ao primeiro referencial.

Por isso, a equação de Λ transforma-se em

$$\left(X'+C\right)^T A\left(X'+C\right)+2B^T\left(X'+C\right)+c=0$$

ou

$$X'^T AX'+X'^T AC+C^T AX'+C^T AC+2B^T X'+2B^T C+c=0$$

Observando que a matriz $X'^T AC$ é do tipo 1×1, e recordando que a matriz A é simétrica, vemos que

$$X'^T AC\ \left(X'^T AC\right)^T=C^T A^T\left(X'^T\right)^T=C^T AX'$$

Por consequência, a equação anterior pode escrever-se na forma

$$X'^T AX'+2C^T AX'+C^T AC+2B^T X'+2B^T C+c=0$$

ou seja,

$$X'^T AX'+2\left(AC+B\right)^T X'+C^T AC+2B^T C+c=0$$

ou, finalmente,

$$X'^T AX'+2\left(AC+B\right)^T X'+\Phi\left(C\right)=0$$

Em resumo: a mudança da origem deixa inalterados os coeficientes dos termos quadrados e do termo rectangular, modificando os termos lineares e o termo independente.

Isto revela que a equação se tornará particularmente simples se for possível encontrar um ponto O' tal que $AC+B=0$, isto é, tal que $AC=-B$.

Um tal ponto tem um nome apropriado:

Definição 9-43 | Nas condições anteriores, um ponto O' tal que $AC=-B$ toma o nome de **centro de simetria** da cónica Λ.

Se O' for um centro de simetria de Λ, a equação da cónica em relação ao referencial $\left(O';\vec{e}_1,\vec{e}_2\right)$ terá o aspecto mais simples

$$X'^T AX'+\Phi\left(C\right)=0$$

anulando-se os termos lineares. Esta equação da cónica é chamada **equação reduzida ao centro.**

É necessário notar que o facto de um dado ponto ser ou não ser um centro de simetria de uma dada cónica não depende do referencial em relação ao qual a equação desta se encontra escrita.

Com efeito, suponhamos que a cónica Λ tem a equação

$$\Phi(X) = X^T AX + 2B^T X + c = 0$$

em relação ao referencial $(O;\vec{e}_1,\vec{e}_2)$.

Seja O' um centro de simetria de Λ, ou seja, um ponto cujas coordenadas C em relação ao referencial considerado verificam a igualdade $AC = -B$.

Suponhamos que passamos para um referencial $(D;\vec{e}_1,\vec{e}_2)$. Designaremos também por D (por abuso de linguagem) as coordenadas da nova origem D em relação ao referencial $(O;\vec{e}_1,\vec{e}_2)$.

Conforme se viu, a equação da cónica em relação ao referencial $(D;\vec{e}_1,\vec{e}_2)$ transforma-se em

$$X'^T AX' + 2(AD+B)^T X' + \Phi(D) = 0$$

As coordenadas C', do ponto O', em relação ao referencial $(D;\vec{e}_1,\vec{e}_2)$, relacionam-se com as coordenadas C, do mesmo ponto em relação ao referencial $(O;\vec{e}_1,\vec{e}_2)$, pela igualdade $C = C' + D$, ou seja, $C' = C - D$.

Verificamos, então, que se tem

$$AC' = A(C-D) = AC - AD = -B - AD = -(AD+B)$$

pelo que o ponto O' continuaria a ser considerado um centro de simetria da cónica, quando trabalhamos com o novo referencial.

Analogamente, se passarmos para um referencial $(O;\vec{e}_1,\vec{e}_2)$, a equação da cónica transforma-se, como vimos, em

$$X'^T\left(P^T AP\right)X' + 2\left(B^T P\right)X' + c = 0$$

e as coordenadas do ponto O' são dadas por $C = PC'$, ou $C' = P^{-1}C$.

Daqui resulta que

$$\left(P^T AP\right)C' = P^T APP^{-1}C = P^T AC = P^T(-B) = -\left(P^T B\right)$$

Estamos, agora, em condições de proceder à discussão da equação geral de uma cónica, para o que distinguiremos dois casos: cónicas com centro e cónicas sem centro.

9-6-3 Cónicas com centro

Se Λ é uma cónica com centro, podemos considerar um seu centro de simetria C e as direcções principais $\vec{u}_1, \vec{u}_2$.

Passando para um referencial de origem C, a equação perde os termos lineares, pelo que, ao passar-se, em seguida, para a base $\vec{u}_1, \vec{u}_2$, os termos lineares se mantêm nulos; assim, no referencial $(C; \vec{u}_1, \vec{u}_2)$, obteremos a equação reduzida da cónica, na forma

$$\lambda_1 x^2 + \lambda_2 y^2 + d = 0$$

Consideremos, então, as seguintes possibilidades:

a) Os valores próprios λ_1, λ_2 são ambos diferentes de zero.

b) Um dos valores próprios λ_1, λ_2 é diferente de zero e o outro é zero.

Vejamos, então:

a) Os valores próprios λ_1, λ_2 são ambos diferentes de zero.

a1) λ_1, λ_2 e d têm o mesmo sinal.

Neste caso, a equação

$$\lambda_1 x^2 + \lambda_2 y^2 + d = 0$$

é claramente impossível, pelo que $\Lambda = \varnothing$.

a2) λ_1, λ_2 têm o mesmo sinal e d tem o sinal contrário.

Neste caso, podemos transformar a equação

$$\lambda_1 x^2 + \lambda_2 y^2 + d = 0$$

sucessivamente, em

$$\lambda_1 x^2 + \lambda_2 y^2 = -d$$

$$\frac{\lambda_1 x^2}{-d} + \frac{\lambda_2 y^2}{-d} = 1$$

$$\frac{x^2}{-\frac{d}{\lambda_1}} + \frac{y^2}{-\frac{d}{\lambda_2}} = 1$$

Ora, por hipótese,

$$-\frac{d}{\lambda_1}, -\frac{d}{\lambda_2} > 0$$

pelo que, para certos números reais a e b, podemos pôr

$$-\frac{d}{\lambda_1} = a^2 \quad \text{e} \quad -\frac{d}{\lambda_2} = b^2$$

A equação da cónica fica, então, com a forma

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

Esta cónica toma o nome de **elipse**.

a3) λ_1, λ_2 têm o mesmo sinal e $d=0$.
A equação da cónica é, neste caso,

$$\lambda_1 x^2+\lambda_2 y^2=0 \quad \text{ou} \quad \frac{x^2}{\frac{1}{\lambda_1}}+\frac{y^2}{\frac{1}{\lambda_2}}=0$$

que equivale a

$$\frac{x^2}{\frac{1}{|\lambda_1|}}+\frac{y^2}{\frac{1}{|\lambda_2|}}=0$$

Pondo $\frac{1}{|\lambda_1|}=a^2$ e $\frac{1}{|\lambda_2|}=b^2$, obtemos a equação $\frac{x^2}{a^2}+\frac{y^2}{b^2}=0$

Ora, esta equação equivale, evidentemente, a $x=y=0$, pelo que a cónica se reduz a um só ponto, precisamente a origem C do referencial considerado.

a4) λ_1, λ_2 têm sinais contrários e $d \neq 0$.
Neste caso, podemos transformar a equação

$$\lambda_1 x^2+\lambda_2 y^2+d=0 \quad \text{em} \quad \frac{x^2}{-\frac{d}{\lambda_1}}+\frac{y^2}{-\frac{d}{\lambda_2}}=1$$

O termo d terá o sinal de um dos valores próprios, pelo que a equação da cónica fica, então, com uma das formas

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \quad \text{ou} \quad -\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

Esta cónica toma o nome de **hipérbole**.

a5) λ_1, λ_2 têm sinais contrários e $d=0$.
A equação da cónica toma a forma

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=0 \quad \text{que equivale a} \quad (bx-ay)(bx+ay)=0$$

e a cónica é constituída por duas rectas concorrentes.

b) Um dos valores próprios λ_1, λ_2 é diferente de zero e o outro é zero.

Suponhamos que se tem $\lambda_2 = 0$. A equação da cónica apresenta-se com a forma

$$\lambda_1 x^2 + d = 0$$

e podemos, então, considerar as seguintes três possibilidades:

b1) λ_1 e d têm o mesmo sinal

A equação é impossível, pelo que a cónica é vazia.

b2) λ_1 e d têm sinais contrários

A equação da cónica transforma-se em

$$x^2 = -\frac{d}{\lambda_1} \quad \text{ou seja, em} \quad x^2 = a^2$$

Esta equação equivale a $x = \pm a$, pelo que a cónica é formada por duas rectas paralelas.

b3) $d = 0$

A equação da cónica transforma-se em

$$x^2 = 0$$

pelo que a cónica é uma recta.

9-6-4 Cónicas sem centro

Terminado o estudo das cónicas com centro de simetria, examinemos as cónicas sem centro.

A equação de uma cónica sem centro, reduzida às direcções principais, tem a forma

$$\lambda_1 x^2 + \lambda_2 y^2 + 2mx + 2ny + d = 0$$

e, sendo a matriz

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

dos coeficientes de segundo grau uma matriz não nula, mas tal que o sistema de equações

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = -\begin{bmatrix} m \\ n \end{bmatrix}$$

é impossível (visto não haver centro de simetria), um e um só dos dois valores próprios terá de ser nulo. Supondo que se tem $\lambda_2 = 0$, teremos $n \neq 0$, e a equação simplifica-se um pouco mais:

$$\lambda_1 x^2 + 2mx + 2ny + d = 0$$

Ora, é possível mudar a origem do referencial para um ponto D, conservando a base (formada pelas direcções principais da cónica) de modo que, em relação ao novo referencial, se anulem, na equação de Λ, o termo linear em x e o termo independente.

Com efeito, tal como vimos acima, a mudança da origem para um ponto $D \equiv (d_1, d_2)$ transforma os termos lineares da equação em

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \begin{bmatrix} m \\ n \end{bmatrix}$$

e o termo independente em $\Phi(D)$, pelo que pretendemos encontrar um ponto D tal que

$$\begin{cases} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} 0 \\ n \end{bmatrix} \\ \Phi(D) = 0 \end{cases}$$

Resolvendo este sistema de equações, encontramos o ponto

$$D \equiv \left(-\frac{m}{\lambda_1}, \frac{1}{2n}\left(\frac{m^2}{\lambda_1} - d \right) \right)$$

Em relação ao referencial $(D; \vec{u}_1, \vec{u}_2)$, a equação da cónica reveste o aspecto

$$\lambda_1 x^2 + 2ny = 0$$

que equivale a

$$x^2 = ky \qquad \text{com} \qquad k = -\frac{2n}{\lambda_1}$$

A cónica em questão é uma **parábola**.

Em resumo:

Cónicas:

1.º) Com centro

a) Dois valores próprios não nulos

a1) λ_1, λ_2 e d têm o mesmo sinal — Conjunto vazio

a2) λ_1, λ_2 têm o mesmo sinal e d tem o sinal contrário — Elipse

a3) λ_1, λ_2 têm o mesmo sinal e $d = 0$ — Um ponto

a4) λ_1, λ_2 têm sinais contrários e $d \neq 0$ — Hipérbole

a5) λ_1, λ_2 têm sinais contrários e $d = 0$ — Duas rectas concorrentes

b) Um dos valores próprios λ_1, λ_2 é diferente de zero e o outro é zero

b1) λ_1 e d têm o mesmo sinal — Conjunto vazio

b2) λ_1 e d têm sinais contrários — Duas rectas paralelas

b3) $d = 0$ — Uma recta

2.º) Sem centro — Parábola

9-7 Quádricas em $\mathbb{R}^3$

Assim como se fez o estudo dos conjuntos de pontos do espaço afim $\mathbb{R}^2$ que são definidos por equações polinomiais de segundo grau, pode fazer-se um estudo semelhante em espaços com outras dimensões, merecendo, naturalmente, uma especial atenção o caso do espaço afim $\mathbb{R}^3$.

Introduzimos, para o efeito, a seguinte designação:

Definição 9-44 Chama-se **quádrica** ao conjunto dos pontos do espaço afim $\mathbb{R}^3$ que satisfazem uma dada equação polinomial de grau 2, com três variáveis.

Fixando, no espaço afim $\mathbb{R}^3$, um referencial $(O; \vec{e}_1, \vec{e}_2, \vec{e}_3)$, designemos por Γ a quádrica constituída pelos pontos cujas coordenadas, em relação a esse referencial, satisfaçam uma equação polinomial genérica, de segundo grau, que, por comodidade e sem perda de generalidade, apresentaremos sob a forma

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + 2b_1x + 2b_2y + 2b_3z + c = 0$$

Nesta equação, algum dos coeficientes a_{ij} $(i, j = 1, 2, 3)$ terá de ser não nulo, visto supormos que se trata de uma equação de segundo grau.

Chamaremos às parcelas $a_{11}x^2,\ a_{22}y^2,\ a_{33}z^2$ os **termos quadrados** da equação; às parcelas $2a_{12}xy,\ 2a_{13}xz,\ 2a_{23}yz$, os **termos rectangulares** da equação; às parcelas $2b_1x,\ 2b_2y,\ 2b_3z$, os **termos lineares** da equação, e a c, o **termo independente** da equação.

Por sua vez, se considerarmos as matrizes

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}, \quad B=\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{e} \quad X=\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

a equação anterior pode escrever-se, de modo mais abreviado, na forma

$$\Phi(X) = X^T A X + 2B^T X + c = 0$$

Note-se que, por construção, a matriz A é uma matriz simétrica.

É claro que algumas quádricas correspondem a conjuntos facilmente identificáveis.

Assim, por exemplo, a equação

$$x^2 + y^2 + z^2 + 1 = 0$$

é, obviamente, impossível, pelo que define o conjunto vazio; por sua vez, a equação

$$x^2 + y^2 + z^2 = 0$$

tem como única solução o ponto $(0,0,0)$, e a equação

$$x^2 + y^2 + z^2 + 2xy + 2xz + 2yz = 0$$

que equivale a

$$(x+y+z)^2 = 0 \quad \text{ou a} \quad x+y+z=0$$

define um plano de $\mathbb{R}^3$.

O que nos interessa é chegar a uma classificação exaustiva das quádricas, de modo a poder identificar cada uma delas, tal como ficou feito para as cónicas. Pode acontecer que a equação de Γ em relação ao referencial considerado não seja suficientemente simples para permitir a identificação da quádrica, mas que se simplifique, em relação a outros referenciais. Por isso, vamos estudar as alterações que cada mudança de referencial acarreta na equação geral acima considerada, estudando separadamente duas possibilidades: mudança da base e mudança da origem.

9-7-1 Mudança da base

Vejamos agora o que sucede à equação de Γ, quando passarmos do referencial $(O;\vec{e}_1,\vec{e}_2,\vec{e}_3)$ inicial para um referencial $(O;\vec{e}_1',\vec{e}_2',\vec{e}_3')$.

Sendo P a matriz de mudança de base, as coordenadas, X e X', de um mesmo ponto em relação a cada um dos referenciais considerados, relacionam-se pela igualdade

$$X = PX'$$

A equação da quádrica transforma-se, portanto, em

$$(PX')^T A(PX') + 2B^T(PX') + c = 0 \quad \text{ou} \quad X'^T(P^TAP)X' + 2(B^TP)X' + c = 0$$

Vemos, pois, que a mudança da base não altera o termo independente da equação da quádrica, modificando os demais.

Suponhamos fixo em $\mathbb{R}^3$ um produto interno, e seja G a matriz da métrica em relação à base $\vec{e}_1,\vec{e}_2,\vec{e}_3$ do referencial inicial.

Consideremos a matriz $S = G^{-1}A$. Tal como no caso das cónicas, podemos demonstrar o seguinte:

Teorema 9-45 Nas condições anteriores, a matriz S define, em relação à base $\vec{e}_1,\vec{e}_2,\vec{e}_3$, um endomorfismo diagonalizável de $\mathbb{R}^3$, e existe no espaço uma base ortonormada formada por vectores próprios desse endomorfismo.

Demonstração:
Observemos que se tem

$$S^TG = (G^{-1}A)^T G = A^T(G^{-1})^T G = A(G^T)^{-1} G = AG^{-1}G = A = G(G^{-1}A) = GS$$

A igualdade $S^TG = GS$ mostra que o endomorfismo definido por S é um endomorfismo auto-adjunto do espaço vectorial $\mathbb{R}^3$.
A tese resulta então imediatamente do Teorema 7-41.

Em face destes resultados, introduzimos a seguinte definição:

Definição 9-46 Nas condições anteriores, se $\vec{u}_1,\vec{u}_2,\vec{u}_3$ é uma base ortonormada, constituída por vectores próprios do endomorfismo ψ, definido pela matriz S, em relação à base $\vec{e}_1,\vec{e}_2,\vec{e}_3$, dizemos que os vectores $\vec{u}_1,\vec{u}_2,\vec{u}_3$ são **direcções principais** da quádrica Γ.

Deve observar-se que, quando a base $\vec{e}_1, \vec{e}_2, \vec{e}_3$, inicialmente considerada, é ortonormada, se tem $S = G^{-1}A = I_3^{-1}A = A$, pelo que, nesse caso, as direcções principais serão vectores próprios ortonormados do endomorfismo definido, em relação à base considerada, pela matriz A.

A partir do que ficou dito, podemos agora analisar o que acontece à equação geral de uma quádrica quando mudamos a base do referencial inicial para uma base formada por direcções principais da quádrica.

Sendo P a matriz de mudança de base, sabemos já que a equação geral de Γ assume a forma

$$X'^T\left(P^TAP\right)X' + 2\left(B^TP\right)X' + c = 0$$

Ora, a matriz da métrica em relação à base ortonormada $\vec{u}_1, \vec{u}_2, \vec{u}_3$ é I_3, e sabemos que esta matriz se relaciona com a matriz G através da igualdade $I_3 = P^TGP$, que equivale a $P^T = (GP)^{-1}$.

Por sua vez, a matriz do endomorfismo ψ, em relação à base $\vec{u}_1, \vec{u}_2, \vec{u}_3$, formada por vectores próprios de ψ, será

$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = P^{-1}SP = P^{-1}G^{-1}AP = (GP)^{-1}\,AP = P^TAP$$

Consequentemente, a equação da quádrica Γ terá a forma

$$X'^T\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}X' + 2\left(B^TP\right)X' + c = 0$$

ou seja, a forma

$$\lambda_1x^2 + \lambda_2y^2 + \lambda_3z^2 + 2b_1'x + 2b_2'y + 2b_3'z + c' = 0$$

Esta equação é chamada **equação reduzida às direcções principais**, da quádrica Γ.

9-7-2 | Mudança da origem (translação do referencial)

Suponhamos que passamos do referencial considerado inicialmente para um referencial $\left(O'; \vec{e}_1, \vec{e}_2, \vec{e}_3\right)$.

As coordenadas, X e X', de um mesmo ponto, em relação a cada um dos referenciais considerados, relacionam-se, como vimos oportunamente, pela igualdade

$$X = X' + C$$

onde C designa a coluna das coordenadas de O' em relação ao primeiro referencial.

Por isso, a equação de Γ transforma-se em

$$(X'+C)^T A(X'+C)+2B^T(X'+C)+c=0$$

ou

$$X'^T AX'+X'^T AC+C^T AX'+C^T AC+2B^T X'+2B^T C+c=0$$

Observando que a matriz $X'^T AC$ é do tipo 1×1, e recordando que a matriz A é simétrica, vemos que

$$X'^T AC\left(X'^T AC\right)^T=C^T A^T\left(X'^T\right)^T=C^T AX'$$

Por consequência, a equação anterior pode escrever-se na forma

$$X'^T AX'+2C^T AX'+C^T AC+2B^T X'+2B^T C+c=0$$

ou seja,

$$X'^T AX'+2(AC+B)^T X'+C^T AC+2B^T C+c=0$$

ou, finalmente,

$$X'^T AX'+2(AC+B)^T X'+\Phi(C)=0$$

Em resumo: a mudança da origem deixa inalterados os coeficientes dos termos quadrados e rectangulares, modificando os termos lineares e o termo independente.

O que ficou dito revela que a equação se tornará particularmente simples se for possível encontrar um ponto O' tal que $AC+B=0$, isto é, tal que $AC=-B$.

Um tal ponto tem um nome apropriado:

Definição 9-47 | Nas condições anteriores, um ponto O' tal que $AC=-B$ toma o nome de **centro de simetria** da quádrica Γ.

Se O' for um centro de simetria de Γ, a equação da quádrica em relação ao referencial $\left(O';\vec{e}_1,\vec{e}_2,\vec{e}_3\right)$ terá o aspecto mais simples

$$X'^T AX'+\Phi(C)=0$$

anulando-se os termos lineares. Esta equação da quádrica é chamada **equação reduzida ao centro**.

Tal como no caso das cónicas, é fácil verificar que o facto de um dado ponto ser ou não centro de simetria de uma quádrica é uma propriedade intrínseca desta, não dependendo do referencial considerado.

Distinguiremos, então, dois casos, no estudo de uma quádrica qualquer: quádricas com centro e quádricas sem centro.

9-7-3 Quádricas com centro

Se Γ é uma quádrica com centro, podemos considerar um seu centro de simetria C e as direcções principais $\vec{u}_1, \vec{u}_2, \vec{u}_3$.

Passando para um referencial de origem C, a equação perde os termos lineares, pelo que, ao passar-se, em seguida, para a base $\vec{u}_1, \vec{u}_2, \vec{u}_3$, os termos lineares se mantêm nulos; assim, no referencial $(C; \vec{u}_1, \vec{u}_2, \vec{u}_3)$, obteremos a equação reduzida da quádrica, na forma

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + d = 0$$

Consideremos, então, as seguintes possibilidades, que estudaremos uma a uma:

a) Os valores próprios $\lambda_1, \lambda_2, \lambda_3$ são todos diferentes de zero.

b) Dois dos valores próprios $\lambda_1, \lambda_2, \lambda_3$ são diferentes de zero e o terceiro é zero.

c) Um valores próprios $\lambda_1, \lambda_2, \lambda_3$ é diferente de zero e os outros dois são nulos.

Vejamos, então:

a) Os valores próprios $\lambda_1, \lambda_2, \lambda_3$ são todos diferentes de zero.

a1) $\lambda_1, \lambda_2, \lambda_3$ e d têm o mesmo sinal.

Neste caso, a equação

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + d = 0$$

é claramente impossível, pelo que $\Gamma = \varnothing$.

a2) $\lambda_1, \lambda_2, \lambda_3$ têm o mesmo sinal e d tem o sinal contrário.

Neste caso, podemos transformar a equação

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + d = 0$$

sucessivamente, em

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 = -d$$

$$\frac{\lambda_1 x^2}{-d} + \frac{\lambda_2 y^2}{-d} + \frac{\lambda_3 z^2}{-d} = 1$$

$$\frac{x^2}{-\frac{d}{\lambda_1}} + \frac{y^2}{-\frac{d}{\lambda_2}} + \frac{z^2}{-\frac{d}{\lambda_3}} = 1$$

Ora, por hipótese,

$$-\frac{d}{\lambda_1}, -\frac{d}{\lambda_2}, -\frac{d}{\lambda_3} > 0$$

pelo que, para certos números reais a, b e c, podemos pôr

$$-\frac{d}{\lambda_1}=a^2, \quad -\frac{d}{\lambda_2}=b^2 \quad \text{e} \quad -\frac{d}{\lambda_3}=c^2$$

A equação da quádrica fica, então, com a forma

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

Esta quádrica toma o nome de **elipsóide**.

a3) $\lambda_1, \lambda_2, \lambda_3$ têm o mesmo sinal e $d=0$.

A equação da quádrica é, neste caso,

$$\lambda_1 x^2+\lambda_2 y^2+\lambda_3 z^2=0$$

ou

$$\frac{x^2}{\frac{1}{\lambda_1}}+\frac{y^2}{\frac{1}{\lambda_2}}+\frac{z^2}{\frac{1}{\lambda_3}}=0 \quad \text{ou ainda} \quad \frac{x^2}{\frac{1}{|\lambda_1|}}+\frac{y^2}{\frac{1}{|\lambda_2|}}+\frac{z^2}{\frac{1}{|\lambda_3|}}=0$$

Pondo

$$\frac{1}{|\lambda_1|}=a^2, \quad \frac{1}{|\lambda_2|}=b^2 \quad \text{e} \quad \frac{1}{|\lambda_3|}=c^2$$

obtemos a equação

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=0$$

Ora, esta equação equivale, evidentemente, a $x=y=z=0$, pelo que a quádrica se reduz a um só ponto, precisamente a origem C do referencial considerado.

a4) λ_1, λ_2 têm o mesmo sinal e λ_3 e d têm o sinal contrário.

Por um processo inteiramente semelhante aos anteriores, chegamos à equação

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$

e a quádrica em causa é um **hiperbolóide de uma folha**.

Note-se que se o sinal de d concordar com o de λ_2 (resp.: com o de λ_1), tendo λ_1, λ_3 (resp.: λ_2, λ_3) o sinal contrário, a equação do hiperbolóide de uma folha terá o aspecto

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \text{ (resp.: } -\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1)$$

a5) λ_1, λ_2, d têm o mesmo sinal e λ_3 tem o sinal contrário.

Sempre por processos análogos aos anteriores, chegamos a

$$-\frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

e estamos perante um **hiperbolóide de duas folhas**.

No caso de o sinal de d concordar, respectivamente, com o de λ_2 e λ_3, ou com o de λ_1 e λ_3, a equação do hiperbolóide de duas folhas será

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad \text{ou} \quad -\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

a6) λ_1, λ_2 têm o mesmo sinal, λ_3 tem o sinal contrário e $d = 0$.

A equação da quádrica assume a forma

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$$

Trata-se de um **cone**.

Caso sejam λ_1, λ_3 com o mesmo sinal e λ_2 com sinal contrário, ou λ_2, λ_3 com o mesmo sinal e λ_1 com sinal contrário, a equação ficará com a forma

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0 \quad \text{ou} \quad -\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0$$

respectivamente.

Analisadas todas as possibilidades dentro da alínea a), passemos, então, à situação seguinte.

b) Dois dos valores próprios $\lambda_1, \lambda_2, \lambda_3$ são diferentes de zero e o terceiro é zero.
Para fixar ideias, admitamos que se tem $\lambda_3 = 0$ e $\lambda_1, \lambda_2 \neq 0$. O estudo é análogo caso seja nulo um outro valor próprio.

Neste caso, a equação reduzida da quádrica é ainda mais simples:

$$\lambda_1 x^2 + \lambda_2 y^2 + d = 0$$

b1) λ_1, λ_2, d com o mesmo sinal

A equação é impossível, pelo que $\Gamma = \varnothing$.

b2) λ_1, λ_2 com o mesmo sinal e d com o sinal contrário

Obtemos, sucessivamente,

$$\lambda_1 x^2 + \lambda_2 y^2 = -d$$

$$\frac{\lambda_1 x^2}{-d} + \frac{\lambda_2 y^2}{-d} = 1$$

$$\frac{x^2}{-\frac{d}{\lambda_1}} + \frac{y^2}{-\frac{d}{\lambda_2}} = 1$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

A quádrica em questão é um **cilindro elíptico.**

b3) λ_1, λ_2 com o mesmo sinal e $d = 0$

Por processo análogo ao anterior, obtemos

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$$

Esta equação equivale a

$$\begin{cases} x = 0 \\ y = 0 \end{cases}$$

Conforme sabemos, este sistema de equações lineares define uma recta do espaço $\mathbb{R}^3$.

b4) λ_1, λ_2 com sinais contrários e $d \neq 0$

No que se refere ao sinal, d concordará com um dos valores próprios e discordará do outro, pela que a equação da quádrica se pode transformar em

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad \text{ou} \quad -\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

e estamos perante um **cilindro hiperbólico.**

b5) λ_1, λ_2 com sinais contrários e $d = 0$

A equação da quádrica transforma-se em

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$$

ou em

$$b^2x^2 - a^2y^2 = 0 \quad \text{ou, ainda, em} \quad (bx + ay)(bx - ay) = 0$$

ou, ainda, em

$$bx + ay = 0 \quad \text{ou} \quad bx - ay = 0$$

A quádrica Γ é, neste caso, constituída por dois planos concorrentes.

Finalmente, ainda dentro do primeiro caso, passemos à última alínea:

c) Um dos valores próprios $\lambda_1, \lambda_2, \lambda_3$ é diferente de zero e os outros dois são nulos.

Para fixar ideias, admitamos que se tem $\lambda_2, \lambda_3 = 0$ e $\lambda_1 = 0$. O estudo é análogo caso sejam nulos outros dois valores próprios.

Neste caso, a equação reduzida da quádrica fica com a forma:

$$\lambda_1 x^2 + d = 0$$

c1) λ_1 e d com o mesmo sinal

Mais uma vez, a equação é impossível e $\Gamma = \varnothing$.

c2) λ_1 e d com sinais contrários

Podemos então escrever

$$x^2 = -\frac{d}{\lambda_1} \quad \text{ou também} \quad x^2 = a^2$$

o que equivale a

$$x = a \quad \text{ou} \quad x = -a$$

A quádrica é constituída por dois planos paralelos.

c3) $d = 0$

A equação da quádrica reduz-se a

$$\lambda_1 x^2 = 0 \quad \text{ou seja,} \quad x = 0$$

que é a equação de um plano.

Terminado o estudo das quádricas com centro, podemos passar ao caso seguinte:

9-7-4 Quádricas sem centro

Neste segundo caso, pode fazer-se a redução da equação geral da quádrica apenas às direcções principais, obtendo-se uma equação da forma

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + 2mx + 2ny + 2pz + d = 0$$

Note-se que a matriz dos termos de segundo grau é

$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

e, como a quádrica não tem centro de simetria, o sistema de equações

$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} X = \begin{bmatrix} m \\ n \\ p \end{bmatrix}$$

deverá ser impossível. Nestas circunstâncias, os três valores próprios $\lambda_1, \lambda_2, \lambda_3$ não podem ser todos diferentes de zero, pelo que podemos considerar duas possibilidades distintas:

a) Dois valores próprios não nulos, um igual a zero.
b) Um valor próprio não nulo, dois iguais a zero.

Passemos, então, a examinar cada uma destas possibilidades:

a) Dois valores próprios não nulos, um igual a zero.

Para fixar ideias, admitamos que seja $\lambda_1, \lambda_2 \neq 0$ e $\lambda_3 = 0$, sendo o estudo análogo em qualquer outro caso. A equação da quádrica fica na forma

$$\lambda_1 x^2 + \lambda_2 y^2 + 2mx + 2ny + 2pz + d = 0$$

Para que o sistema de equações

$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} X = \begin{bmatrix} m \\ n \\ p \end{bmatrix}$$

seja impossível, tem de ser $p \neq 0$.

Vamos ver que, embora não seja possível eliminar todos os termos lineares (visto a quádrica não ter centro), é possível passar a origem do referencial para um ponto D, de modo que, no novo referencial, se eliminem os termos em x e em y, bem como o termo independente.

Na verdade, suponhamos que, no referencial $(O;\vec{u}_1,\vec{u}_2,\vec{u}_3)$, o ponto D é definido pelas suas coordenadas: $D\equiv(d_1,d_2,d_3)$.

Atendendo ao que se viu inicialmente acerca da translação do referencial, o que se pretende é encontrar valores d_1,d_2,d_3 tais que

$$\begin{cases}\begin{bmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}d_1\\ d_2\\ d_3\end{bmatrix}+\begin{bmatrix}m\\ n\\ p\end{bmatrix}=\begin{bmatrix}0\\ 0\\ p\end{bmatrix}\\ \lambda_1 d_1^2+\lambda_2 d_2^2+2md_1+2nd_2+2pd_3+d=0\end{cases}$$

isto é, tais que

$$\begin{cases}\lambda_1 d_1+m=0\\ \lambda_2 d_2+n=0\\ p=p\\ \lambda_1 d_1^2+\lambda_2 d_2^2+2md_1+2nd_2+2pd_3+d=0\end{cases}$$

Ora, resolvendo este sistema de equações, encontramos

$$\begin{cases}d_1=-\dfrac{m}{\lambda_1}\\ d_2=-\dfrac{n}{\lambda_2}\\ d_3=\dfrac{1}{2p}\left(\dfrac{m^2}{\lambda_1}+\dfrac{n^2}{\lambda_2}-d\right)\end{cases}$$

Portanto, em relação ao referencial

$$(D;\vec{u}_1,\vec{u}_2,\vec{u}_3)$$

a equação da quádrica toma o aspecto

$$\lambda_1 x^2+\lambda_2 y^2+2pz=0$$

Nestas condições, temos:

a1) λ_1 e λ_2 com o mesmo sinal

A partir da equação anterior, obtemos, sucessivamente,

$$\lambda_1 x^2 + \lambda_2 y^2 = -2pz$$

$$\frac{x^2}{\cancel{1}/\lambda_1} + \frac{y^2}{\cancel{1}/\lambda_2} = -2pz$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = kz$$

Esta quádrica chama-se um **parabolóide elíptico.**

a2) λ_1 e λ_2 com sinais contrários

De modo análogo ao anterior, chegamos à equação

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = kz$$

e a quádrica em causa é um **parabolóide hiperbólico.**

Finalmente, estudemos a segunda possibilidade:

b) Um valor próprio não nulo, dois iguais a zero.

Para fixar ideias, admitamos que seja $\lambda_1 \neq 0$ e $\lambda_2 = \lambda_3 = 0$, sendo o estudo análogo em qualquer outro caso. A equação da quádrica fica na forma

$$\lambda_1 x^2 + 2mx + 2ny + 2pz + d = 0$$

Para que o sistema de equações

$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} X = \begin{bmatrix} m \\ n \\ p \end{bmatrix}$$

seja impossível, tem de ser $n \neq 0$ ou $p \neq 0$.

Vamos ver que é possível passar a origem do referencial para um ponto D, de modo que, no novo referencial, se eliminem o termo em x e o termo independente.

Na verdade, suponhamos que, no referencial $(O; \vec{u}_1, \vec{u}_2, \vec{u}_3)$, o ponto D é definido pelas suas coordenadas: $D \equiv (d_1, d_2, d_3)$.

Atendendo ao que se viu inicialmente acerca da translação do referencial, o que se pretende é encontrar valores d_1, d_2, d_3 tais que

$$\begin{cases} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} + \begin{bmatrix} m \\ n \\ p \end{bmatrix} = \begin{bmatrix} 0 \\ n \\ p \end{bmatrix} \\ \lambda_1 d_1^2 + 2md_1 + 2nd_2 + 2pd_3 + d = 0 \end{cases}$$

isto é, tais que

$$\begin{cases} \lambda_1 d_1 + m = 0 \\ n = n \\ p = p \\ \lambda_1 d_1^2 + 2md_1 + 2nd_2 + 2pd_3 + d = 0 \end{cases}$$

Ora, resolvendo este sistema de equações, encontramos

$$\begin{cases} d_1 = -\dfrac{m}{\lambda_1} \\ 2nd_2 + 2pd_3 = \dfrac{m^2}{\lambda_1} - d \end{cases}$$

Se $n \neq 0$, bastará tomar

$$d_3 = 0 \text{ e } d_2 = \frac{1}{2n}\left(\frac{m^2}{\lambda_1} - d\right)$$

Pelo contrário, se $n = 0$, então $p \neq 0$, e podemos considerar $d_2 = 0$,

$$d_3 = \frac{1}{2p}\left(\frac{m^2}{\lambda_1} - d\right)$$

Portanto, em relação ao referencial $(D; \vec{u}_1, \vec{u}_2, \vec{u}_3)$, a equação da quádrica toma o aspecto

$$\lambda_1 x^2 + 2ny + 2pz = 0$$

No caso de se ter simultaneamente $n \neq 0$ e $p \neq 0$, podemos ainda efectuar uma última simplificação.

Com efeito, consideremos a matriz

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{n}{r} & \frac{p}{r} \\ 0 & \frac{p}{r} & -\frac{n}{r} \end{bmatrix}, \quad \text{com} \quad r = \sqrt{n^2 + p^2}$$

Esta matriz é invertível, uma vez que

$$|M| = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \frac{n}{r} & \frac{p}{r} \\ 0 & \frac{p}{r} & -\frac{n}{r} \end{vmatrix} = -\frac{n^2}{r^2} - \frac{p^2}{r^2} = -\frac{n^2 + p^2}{r^2} = -1 \neq 0$$

Por consequência, a matriz M pode ser tomada como matriz de uma mudança de base. Efectuando essa mudança de base, a matriz da quádrica fica com a forma

$$X^T M^T A M X + 2B^T M X = 0$$

ou seja,

$$X^T \begin{bmatrix} 1 & 0 & 0 \\ 0 & n/r & p/r \\ 0 & p/r & -n/r \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & n/r & p/r \\ 0 & p/r & -n/r \end{bmatrix} X + 2\begin{bmatrix} 0 & n & p \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & n/r & p/r \\ 0 & p/r & -n/r \end{bmatrix} X = 0$$

ou, ainda,

$$X^T \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} X + 2\begin{bmatrix} 0 & r & 0 \end{bmatrix} X = 0$$

ou, finalmente,

$$\lambda_1 x^2 + 2ry = 0$$

Esta última equação pode escrever-se na forma

$$y = kx^2$$

e a quádrica é um **cilindro parabólico**.

Em resumo, temos a seguinte classificação das quádricas:

Quádricas:

Com centro

a) Os valores próprios $\lambda_1, \lambda_2, \lambda_3$ são todos diferentes de zero
- a1) $\lambda_1, \lambda_2, \lambda_3$ e d têm o mesmo sinal — Conjunto vazio
- a2) $\lambda_1, \lambda_2, \lambda_3$ têm o mesmo sinal e d tem o sinal contrário — Elipsóide
- a3) $\lambda_1, \lambda_2, \lambda_3$ têm o mesmo sinal e $d = 0$ — Um ponto
- a4) λ_1, λ_2 têm o mesmo sinal e λ_3 e d têm o sinal contrário — Hiperbolóide de uma folha
- a5) λ_1, λ_2, d têm o mesmo sinal e λ_3 tem o sinal contrário — Hiperbolóide de duas folhas
- a6) λ_1, λ_2 têm o mesmo sinal, λ_3 tem o sinal contrário e $d = 0$ — Cone

b) Dois dos valores próprios $\lambda_1, \lambda_2, \lambda_3$ são diferentes de zero e o terceiro é zero
- b1) λ_1, λ_2, d com o mesmo sinal — Conjunto vazio
- b2) λ_1, λ_2 com o mesmo sinal e d com o sinal contrário — Cilindro elíptico
- b3) λ_1, λ_2 com o mesmo sinal e $d = 0$ — Uma recta
- b4) λ_1, λ_2 com sinais contrários e $d \neq 0$ — Cilindro hiperbólico
- b5) λ_1, λ_2 com sinais contrários e $d = 0$ — Dois planos concorrentes

c) Um valores próprios $\lambda_1, \lambda_2, \lambda_3$ é diferente de zero e os outros dois são nulos
- c1) λ_1 e d com o mesmo sinal — Conjunto vazio
- c2) λ_1 e d com sinais contrários — Dois planos paralelos
- c3) $d = 0$ — Um plano

Sem centro

a) Dois valores próprios não nulos, um igual a zero
- a1) λ_1 e λ_2 com o mesmo sinal — Parabolóide elíptico
- a2) λ_1 e λ_2 com sinais contrários — Parabolóide hiperbólico

b) Um valor próprio não nulo, dois iguais a zero — Cilindro parabólico

Apêndice A

Bases em espaços de dimensão infinita

O conceito de base de um espaço vectorial, tal como foi apresentado no Capítulo 2, aplica-se exclusivamente a espaços finitamente gerados. Sucede, porém, que se podem igualmente definir bases em espaços que não sejam finitamente gerados, como vamos ver.

Começamos por estender a definição de independência linear a conjuntos infinitos, através da seguinte definição:

Definição A-1 Seja X um subconjunto qualquer, não vazio, de um espaço vectorial E, qualquer, sobre um corpo K. Diz-se que os vectores do conjunto X são **linearmente independentes** (ou que o conjunto X é **linearmente independente**) quando qualquer subconjunto finito não vazio de X é linearmente independente, de acordo com a Definição 2-10.

Por exemplo, consideremos, no espaço vectorial real $\mathbb{R}[x]$, o conjunto

$$X = \left\{1, x, x^2, x^3, x^4, \ldots, x^n, \ldots\right\}$$

É fácil ver que o conjunto X é linearmente independente, uma vez que, dado um seu subconjunto finito, digamos $X' = \left\{x^{k_1}, x^{k_2}, \ldots, x^{k_p}\right\}$, da igualdade

$$\alpha_1 x^{k_1} + \alpha_2 x^{k_2} + \cdots + \alpha_p x^{k_p} = 0$$

resulta, em virtude de propriedades bem conhecidas dos polinómios, que se tenha

$$\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$$

Para chegarmos à definição geral de base, necessitamos de introduzir um conceito específico, oriundo da Teoria de Conjuntos:

Definição A-2 Seja A um conjunto arbitrário e seja $\mathcal{A}$ um conjunto qualquer de partes de A. Um conjunto $A_0 \in \mathcal{A}$ diz-se **maximal** em $\mathcal{A}$ quando não existe qualquer conjunto $B \in \mathcal{A}$ tal que $A_0 \subset B$. Isto é, obviamente, o mesmo que dizer que para qualquer $B \in \mathcal{A}$, se $A_0 \subseteq B$, então $A_0 = B$.

Estamos, agora, em condições de dar a seguinte definição de base de um espaço vectorial:

Definição A-3 Seja E um espaço vectorial qualquer, sobre um corpo K. Chama-se **base** de E, um conjunto maximal na família de todos os subconjuntos linearmente independentes de E.

Devemos começar por observar que esta definição abrange o caso dos espaços vectoriais finitamente gerados. Com efeito, seja $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ uma base de um espaço vectorial E, no sentido do Capítulo 2. O conjunto $B = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ é linearmente independente (veja-se o Corolário 2-14), segundo a definição A.1. Por sua vez, dado um vector arbitrário $\vec{v} \in E$, esse vector poderá escrever-se como combinação linear dos vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, pelo que o conjunto $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n, \vec{v}\}$ será já linearmente dependente, o que vale por dizer que não existe em E um conjunto linearmente independente B' tal que $B \subset B'$.

Vimos (Teorema 2-32) que todo o espaço finitamente gerado tem base. Levanta-se, naturalmente, a questão de saber se o mesmo se poderá dizer de um espaço que não seja finitamente gerado. Para vermos que, de facto, todo o espaço vectorial tem base, necessitamos de um instrumento bastante poderoso da Teoria de Conjuntos, vulgarmente conhecido como o princípio (ou lema) de Zorn, e que, na verdade, é logicamente equivalente ao famoso axioma da Escolha (o assunto encontra-se discutido em A. Monteiro e I. Matos, *Álgebra – Um Primeiro Curso*). Começamos por introduzir mais um conceito:

Definição A-4 Seja A um conjunto arbitrário e seja $\mathcal{A}$ um conjunto qualquer de partes de A. Uma parte $\mathcal{C}$, do conjunto $\mathcal{A}$, é uma **cadeia** quando, quaisquer que sejam $X, Y \in \mathcal{C}$, se tem $X \subseteq Y$ ou $Y \subseteq X$.

Estamos agora em condições de apresentar o seguinte princípio (que aceitamos como um axioma):

Axioma

Princípio de Zorn
Seja A um conjunto arbitrário e seja $\mathcal{A}$ um conjunto qualquer de partes de A. Se qualquer cadeia $\mathcal{C}$ de $\mathcal{A}$ tem um majorante em $\mathcal{A}$, isto é, se, dada uma cadeia qualquer $\mathcal{C}$, existe um conjunto $M \in \mathcal{A}$ tal que

$$X \subseteq M, \ \forall X \in \mathcal{C}$$

então $\mathcal{A}$ tem um elemento maximal.

Com o auxílio do princípio de Zorn, podemos, então, demonstrar o seguinte:

Teorema A-5 Qualquer espaço vectorial tem base.

Demonstração:
Podemos, naturalmente, considerar apenas um espaço vectorial não nulo E.

Seja $\mathcal{A}$ o conjunto de todas as partes linearmente independentes de E. Tomemos em $\mathcal{A}$ uma cadeia $\mathcal{C}$ e consideremos $M = \bigcup\{X{:}X{\in}\mathcal{C}\}$. Para vermos que M é um majorante de $\mathcal{C}$, basta, evidentemente, verificar que o conjunto M é linearmente independente.

Tomemos, então, $\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_m\} \subseteq M$: por definição de M, teremos $\vec{u}_1 \in X_1$, $\vec{u}_2 \in X_2, \ldots, \vec{u}_m \in X_m$, para certos conjuntos $X_1, X_2, ..., X_m \in \mathcal{C}$. Ora, como $\mathcal{C}$ é uma cadeia, sabemos que se tem $X_1 \subseteq X_2$ ou $X_2 \subseteq X_1$; supondo que seja, por exemplo, $X_1 \subseteq X_2$, sabemos, depois, que $X_2 \subseteq X_3$ ou $X_3 \subseteq X_2$, e, iterando o raciocínio, verificamos que existirá um índice t tal que $X_i \subseteq X_t$, para $i = 1, 2, ..., m$. Mas, então, teremos $\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_m\} \subseteq X_t$ e, como $X_t \in \mathcal{A}$, o conjunto X_t é linearmente independente, pelo que $\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_m\}$ é independente. Como $\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_m\}$ foi tomado arbitrariamente em M, concluímos que o conjunto M é independente, como se pretendia.

Pelo princípio de Zorn, uma vez que cada cadeia de $\mathcal{A}$ tem majorante, o conjunto $\mathcal{A}$ tem maximal, que, por definição, será uma base de E.

Tal como se provou que duas bases de um espaço finitamente gerado têm o mesmo número de vectores, pode ainda verificar-se que duas bases de um espaço vectorial qualquer são sempre conjuntos equipotentes, o que já não faremos aqui.

Apêndice B

Polinómios interpoladores de Lagrange

Uma aplicação interessante do conceito de base consiste na construção dos chamados "polinómios interpoladores de Lagrange", que vamos examinar a seguir. Para facilidade da exposição, cingir-nos-emos a polinómios de coeficientes reais, embora as mesmas construções se possam aplicar a polinómios com coeficientes noutros corpos infinitos.

Consideremos então o corpo $\mathbb{R}$, dos números reais, e fixemos os números distintos

$$x_0, x_1, \dots, x_n \in \mathbb{R}$$

e, para cada índice $i \in \{1,2,\dots,n\}$, tomemos o polinómio

$$f_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j} = \frac{(x-x_0)(x-x_1)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_0)(x_i-x_1)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)}$$

Os polinómios $f_0(x), f_1(x), f_2(x), \dots, f_n(x)$ tomam o nome de **polinómios interpoladores de Lagrange**.

Deve notar-se que, para cada $i = 0,1,\dots,n$, o polinómio $f_i(x)$ é um polinómio de grau n, ou seja, $f_i(x)$ pertence ao espaço vectorial $\mathbb{R}_n[x]$. Mais do que isso, vamos verificar que os polinómios $f_0(x), f_1(x), f_2(x), \dots, f_n(x)$ formam uma base deste espaço vectorial, para o que bastará verificar que são linearmente independentes, visto sabermos que $\dim(\mathbb{R}_n[x]) = n+1$.

Ora, suponhamos que se tem

$$\lambda_0 f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_n f_n(x) = 0$$

Neste caso, para cada x_j $(j = 0,1,\dots,n)$, será

$$\lambda_0 f_0(x_j) + \lambda_1 f_1(x_j) + \cdots + \lambda_n f_n(x_j) = 0$$

Mas, por construção, vemos imediatamente que, considerando arbitrariamente $i,j \in \{1,2,\ldots,n\}$, se tem

$$f_i(x_j) = \begin{cases} 0, & \text{se} \quad i \neq j \\ 1, & \text{se} \quad i = j \end{cases} \qquad \text{pelo que} \qquad \lambda_0 f_0(x_j) + \lambda_1 f_1(x_j) + \cdots + \lambda_n f_n(x_j) = \lambda_j$$

Concluímos, portanto, que $\lambda_j = 0$, e, como o índice j era arbitrário, que os polinómios $f_0(x), f_1(x), f_2(x), \ldots, f_n(x)$ são linearmente independentes em $\mathbb{R}_n[x]$.

Verificando, então, que os polinómios $f_0(x), f_1(x), f_2(x), \ldots, f_n(x)$ constituem uma base de $\mathbb{R}_n[x]$, podemos afirmar que qualquer polinómio $g(x) \in \mathbb{R}_n[x]$ se poderá escrever como combinação linear de $f_0(x),\ f_1(x),\ f_2(x), \ldots,\ f_n(x)$, isto é, que, dado $g(x)$, existirão coeficientes $\alpha_0, \alpha_1, \ldots, \alpha_n$ tais que

$$g(x) = \alpha_0 f_0(x) + \alpha_1 f_1(x) + \cdots + \alpha_n f_n(x)$$

Na circunstância, é até fácil identificar esses coeficientes, bastando observar que, para cada $j \in \{1,2,\ldots,n\}$, se tem

$$g(x_j) = \alpha_0 f_0(x_j) + \alpha_1 f_1(x_j) + \cdots + \alpha_n f_n(x_j) = \alpha_j$$

Consequentemente, podemos escrever

$$g(x) = g(x_0) f_0(x) + g(x_1) f_1(x) + \cdots + g(x_n) f_n(x) = \sum_{i=0}^{n} g(x_i) f_i(x)$$

Esta expressão é única, pelo que, dados números reais quaisquer, $y_0, y_1, \ldots, y_n$, o polinómio

$$h(x) = b_0 f_0(x) + b_1 f_1(x) + \cdots + b_n f_n(x) = \sum_{i=0}^{n} b_i f_i(x)$$

é o único elemento de $\mathbb{R}_n[x]$ que verifica as igualdades

$$h(x_j) = y_j \qquad (j = 1,2,\ldots,n)$$

Por outras palavras, acabámos de descrever um método para construir o único polinómio de grau não superior a n que, nos pontos distintos dados $x_0, x_1, \ldots, x_n \in \mathbb{R}$, toma os valores também dados $y_0, y_1, \ldots, y_n \in \mathbb{R}$.

Assim, por exemplo, vamos construir um polinómio de grau não superior a 3 que, quando encarado como uma função (polinomial) real de variável real, satisfaça a seguinte tabela de valores

x	y
$x_0 = 0$	2
$x_1 = 1$	0
$x_2 = 2$	½
$x_3 = -1$	2

ou seja, um polinómio $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$ tal que

$$\begin{aligned} p(0) &= 2 \\ p(1) &= 0 \\ p(2) &= \tfrac{1}{2} \\ p(-1) &= 2 \end{aligned}$$

Usando os polinómios interpoladores de Lagrange, temos, então:

$$\begin{aligned} p(x) &= 2\times\frac{(x-1)(x-2)(x+1)}{(0-1)(0-2)(0+1)} + 0\times\frac{(x-0)(x-2)(x+1)}{(1-0)(1-2)(1+1)} + \\ &\quad + \frac{1}{2}\times\frac{(x-0)(x-1)(x+1)}{(2-0)(2-1)(2+1)} + 2\times\frac{(x-0)(x-1)(x-2)}{(-1-0)(-1-1)(-1-2)} = \\ &= (x-1)(x-2)(x+1) + \frac{1}{12}x(x-1)(x+1) - \frac{1}{3}x(x-1)(x-2) = \\ &= 2 - \frac{7}{4}x - x^2 + \frac{3}{4}x^3 \end{aligned}$$

Uma outra consequência importante das construções que acabámos de apresentar é a seguinte: todo o polinómio de coeficientes reais e grau não superior a n, que se anule em mais que n pontos, é necessariamente o polinómio nulo.

Apêndice C

O teorema de Cayley-Hamilton

Vamos apresentar um famoso teorema, relacionado com o polinómio característico de um endomorfismo de um espaço vectorial de dimensão finita, conhecido pelo nome de teorema de Cayley--Hamilton. Começamos por introduzir um importante conceito:

Definição C-1 Seja E um espaço vectorial arbitrário e seja φ um endomorfismo de E. Um subespaço vectorial F, do espaço E, diz-se **invariante** para φ quando se tem $\varphi(F) \subseteq F$.

Por exemplo, seja qual for o endomorfismo φ, os subespaços $\{\vec{0}\}$, E e $\mathrm{Nuc}(\varphi)$ são invariantes para φ. Por sua vez, se λ é um valor próprio de φ, também o respectivo subespaço próprio é invariante para φ.

Um determinado tipo de subespaços invariantes para um dado endomorfismo merecerá especial atenção, e o próximo teorema descreve esse tipo de subespaços:

Teorema C-2 Seja E um espaço vectorial arbitrário e seja φ um endomorfismo de E. Tomando em E um vector não nulo $\vec{v}$, o subespaço $V = \langle \vec{v}, \varphi(\vec{v}), \varphi^2(\vec{v}), \varphi^3(\vec{v}), \ldots \rangle$ é invariante para φ e, se F é um subespaço qualquer, invariante para φ, e tal que $\vec{v} \in F$, então $V \subseteq F$.

Demonstração:
É claro que $\varphi(V) = \langle \varphi(\vec{v}), \varphi^2(\vec{v}), \varphi^3(\vec{v}), \varphi^4(\vec{v}), \ldots \rangle \subseteq V$.

Por outro lado, se F é invariante para φ e $\vec{v} \in F$, então $\varphi(\vec{v}) \in F$, e, sempre que $\varphi^k(\vec{v}) \in F$, para certo natural k, também $\varphi^{k+1}(\vec{v}) = \varphi\left(\varphi^k(\vec{v})\right) \in F$, pelo que

$$\{\vec{v}, \varphi(\vec{v}), \varphi^2(\vec{v}), \varphi^3(\vec{v}), \ldots\} \subseteq F$$

e, portanto, $V \subseteq F$.

Provaremos, em seguida, um resultado eminentemente técnico:

Lema C-3

Seja E um espaço vectorial de dimensão finita n e seja φ um endomorfismo de E. Dado um subespaço vectorial F, do espaço E, invariante para φ, representaremos por φ_F a restrição de φ a F. O polinómio característico de φ_F divide o polinómio característico de φ.

Demonstração:
Seja $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k$ uma base de F. Com o auxílio do teorema de Steinitz, podemos construir para E uma base $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k, \vec{e}_{k+1}, \ldots, \vec{e}_n$.

Seja A a matriz de φ em relação à base $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k, \vec{e}_{k+1}, \ldots, \vec{e}_n$, e A_F, a matriz de φ_F em relação à base $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k$.

Uma vez que se tem

$$\varphi\left(\vec{f}_i\right) \in F \text{ (para } i = 1, 2, \ldots, k)$$

visto F ser invariante para φ, será

$$\varphi\left(\vec{f}_i\right) = a_{1i}\vec{f}_1 + a_{2i}\vec{f}_2 + \cdots + a_{ki}\vec{f}_k + 0\vec{e}_{k+1} + \cdots + 0\vec{e}_n$$

pelo que a matriz A terá a forma

$$A = \begin{bmatrix} A_F & B \\ 0 & C \end{bmatrix}$$

O polinómio característico de φ, que, como se sabe, não depende da base considerada, é

$$p(\lambda) = |A - \lambda I_n| = \begin{vmatrix} A_F - \lambda I_k & B \\ 0 & C - \lambda I_{n-k} \end{vmatrix} = |A_F - \lambda I_k| \cdot |C - \lambda I_{n-k}|$$

sendo $|A_F - \lambda I_k|$ o polinómio característico de φ_F.

Estamos agora em condições de enunciar e demonstrar um teorema fundamental, que é o seguinte:

Teorema C-4 Seja E um espaço vectorial de dimensão finita n e seja $\vec{v}$ um vector não nulo de E. Consideremos o subespaço vectorial $V = \langle \vec{v}, \varphi(\vec{v}), \varphi^2(\vec{v}), \varphi^3(\vec{v}), \ldots \rangle$ e suponhamos que V tem a dimensão m. Então:

a) Os vectores $\vec{v}, \varphi(\vec{v}), \varphi^2(\vec{v}), \ldots, \varphi^{m-1}(\vec{v})$ formam uma base de V.

b) Se $\alpha_0\vec{v} + \alpha_1\varphi(\vec{v}) + \alpha_2\varphi^2(\vec{v}) + \cdots + \alpha_{m-1}\varphi^{m-1}(\vec{v}) + \varphi^m(\vec{v}) = \vec{0}$, então o polinómio característico de φ_V é

$$g(\lambda) = (-1)^m \left(\alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \cdots + \alpha_{m-1}\lambda^{m-1} + \lambda^m\right)$$

Demonstração:

a) O vector $\vec{v} \neq \vec{0}$ é, por si só, linearmente independente. Seja então k o maior número natural tal que os vectores $\vec{v}, \varphi(\vec{v}), \varphi^2(\vec{v}), \ldots, \varphi^{k-1}(\vec{v})$ são linearmente independentes; é claro que k existe, visto que o espaço E tem dimensão finita, tendo-se ainda, forçosamente, $k \leq m$.

Tomemos o subespaço vectorial $W = \langle \vec{v}, \varphi(\vec{v}), \varphi^2(\vec{v}), \ldots, \varphi^{k-1}(\vec{v}) \rangle$.

Os vectores $\vec{v}, \varphi(\vec{v}), \varphi^2(\vec{v}), \ldots, \varphi^{k-1}(\vec{v})$ são linearmente dependentes (por definição de k), pelo que $\varphi^k(\vec{v}) \in W$.

Daqui resulta imediatamente que W é invariante para φ. Consequentemente, de acordo com o lema anterior, tem-se $V \subseteq W$, pelo que $k = m$.

b) Uma vez que, como vimos acima, $\varphi^m(\vec{v}) \in W$, existirão coeficientes tais que $\alpha_0\vec{v} + \alpha_1\varphi(\vec{v}) + \alpha_2\varphi^2(\vec{v}) + \cdots + \alpha_{m-1}\varphi^{m-1}(\vec{v}) + \varphi^m(\vec{v}) = \vec{0}$.

A matriz de φ_V, em relação à base $\vec{v}, \varphi(\vec{v}), \varphi^2(\vec{v}), \ldots, \varphi^{m-1}(\vec{v})$, é, evidentemente,

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & 0 & \cdots & 0 & -\alpha_2 \\ 0 & 0 & 1 & \cdots & 0 & -\alpha_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -\alpha_{m-2} \\ 0 & 0 & 0 & \cdots & 1 & -\alpha_{m-1} \end{bmatrix}$$

cujo polinómio característico é

$$g(\lambda) = (-1)^m \left(\alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \cdots + \alpha_{m-1}\lambda^{m-1} + \lambda^m\right)$$

conforme se pode verificar facilmente, por indução, usando o desenvolvimento de Laplace do determinante, ao longo da primeira linha da matriz.

Assim, por exemplo, consideremos o endomorfismo φ do espaço vectorial real $\mathbb{R}^3$, definido por $\varphi(x_1, x_2, x_3) = (-x_2 + x_3,\ x_1 + x_3,\ 3x_3)$, e tomemos o vector $\vec{v} = (1,0,0)$. Tem-se:

$$\varphi^2(\vec{v}) = (-1,0,0) = -\vec{v}$$

Logo, $V = \langle \vec{v}, \varphi(\vec{v}), \varphi^2(\vec{v}), \ldots \rangle = \langle \vec{v}, \vec{w} \rangle$ tem dimensão 2 e

$$1\vec{v} + 0\varphi(\vec{v}) + \varphi^2(\vec{v}) = \vec{v} + \vec{0} - \vec{v} = \vec{0}$$

pelo que o polinómio característico de φ_V é

$$1 + 0\lambda + \lambda^2 = 1 + \lambda^2$$

Com efeito, a matriz de φ_V em relação à base $\vec{v}, \vec{w}$ é

$$A_V = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

e o seu polinómio característico é

$$|A - \lambda I_2| = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1$$

Podemos, finalmente, enunciar o teorema que nos propúnhamos apresentar:

Teorema C-5

Teorema de Cayley-Hamilton

Seja E um espaço vectorial de dimensão finita n e seja φ um endomorfismo de E.

Seja $f(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n$ o polinómio característico de φ. Então, $f(\varphi) = a_0 1_E + a_1\varphi + a_2\varphi^2 + \cdots + a_n\varphi^n = 0$.

Demonstração:
É claro que $f(\varphi)$ é ainda um endomorfismo de E, pelo que $f(\varphi)(\vec{0}) = \vec{0}$.

Tomemos, então, um vector qualquer $\vec{v} \neq \vec{0}$ e consideremos, mais uma vez, o subespaço $V = \langle \vec{v}, \varphi(\vec{v}), \varphi^2(\vec{v}), \varphi^3(\vec{v}), \ldots \rangle$, de dimensão m.

Ponhamos $\varphi^m(\vec{v}) = -a_0\vec{v} - a_1\varphi(\vec{v}) - \cdots - a_{m-1}\varphi^{m-1}(\vec{v})$.

O polinómio característico de φ_V será, de acordo com o teorema anterior,

$$g(\lambda) = (-1)^m \left(a_0 + a_1\lambda + \cdots + a_{m-1}\lambda^{m-1} + \lambda^m\right)$$

pelo que

$$\begin{aligned} g(\varphi)(\vec{v}) &= (-1)^m \left(a_0 1_V + a_1\varphi + \cdots + a_{m-1}\varphi + \varphi^m\right)(\vec{v}) = \\ &= (-1)^m \left(a_0\vec{v} + a_1\varphi(\vec{v}) + \cdots + a_{m-1}\varphi^{m-1}(\vec{v}) + \varphi^m(\vec{v})\right) = \vec{0} \end{aligned}$$

Ora, pelo lema anterior, $g(\lambda)$ divide $f(\lambda)$, pelo que existe um polinómio $q(\lambda)$ tal que $f(\lambda) = q(\lambda)g(\lambda)$. Assim:

$$f(\varphi)(\vec{v}) = \left(q(\varphi)g(\varphi)\right)(\vec{v}) = q(\varphi)\left(g(\varphi)(\vec{v})\right) = q(\varphi)(\vec{0}) = \vec{0}$$

Enunciado em termos da representação matricial de um endomorfismo, o teorema de Cayley--Hamilton reveste, naturalmente, a seguinte forma:

Corolário C-6 Seja A uma matriz do tipo $n \times n$ e $f(\lambda) = |A - \lambda I_n| = a_0 + a_1\lambda + \cdots + a_n\lambda^n$. Então, $f(A) = a_0 I_n + a_1 A + \cdots + a_n A^n$.

Apêndice D

Factorização de uma matriz

Vimos, no Capítulo 5, que toda a matriz A se pode transformar numa matriz do tipo

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

onde r designa a característica de A, e que essa transformação se pode obter através da multiplicação de A por matrizes invertíveis, do lado esquerdo e do lado direito.

Por outras palavras, sabemos que existem matrizes invertíveis P e Q tais que

$$PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

Na prática, poderemos querer determinar as matrizes P e Q, para o que bastaria, ao aplicar a A o método de condensação de Gauss, ir tomando nota das transformações efectuadas e, bem assim, das matrizes por que se deveria sucessivamente multiplicar A para obter essas transformações.

Supondo que cada transformação efectuada nas linhas da matriz correspondesse a multiplicar A, do lado esquerdo, por uma matriz E e que cada transformação efectuada nas colunas da matriz correspondesse a multiplicar A, do lado direito, por uma matriz D, chegaríamos a uma igualdade da forma

$$E_t \cdots E_2 E_1 A D_1 D_2 \cdots D_s = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

sendo, naturalmente, $P = E_t \cdots E_2 E_1$ e $Q = D_1 D_2 \cdots D_s$.

Na prática, porém, é possível obter as matrizes P e Q de um modo mais expedito, atendendo ao modo como cada matriz E e cada matriz D são construídas.

Com efeito, basta partir de uma disposição gráfica do tipo

$$\begin{bmatrix} I_n & \\ A & I_p \end{bmatrix}$$

e, ao condensar a matriz A, efectuar todas as transformações necessárias sobre as linhas de A nas últimas p linhas completas e todas as transformações necessárias sobre as colunas de A nas primeiras n colunas completas. Quando terminarmos a condensação, chegaremos a uma disposição da forma

$$\begin{bmatrix} Q & \\ A_C & P \end{bmatrix} \quad \text{onde} \quad A_C = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

Assim, por exemplo, tomemos a matriz

$$A = \begin{bmatrix} 1 & -1 & 1 & 2 \\ -1 & 2 & 0 & 1 \\ 0 & -1 & 1 & 2 \end{bmatrix}$$

e comecemos por escrever

$$\begin{bmatrix} 1 & 0 & 0 & 0 & & & \\ 0 & 1 & 0 & 0 & & & \\ 0 & 0 & 1 & 0 & & & \\ 0 & 0 & 0 & 1 & & & \\ 1 & -1 & 1 & 2 & 1 & 0 & 0 \\ -1 & 2 & 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 2 & 0 & 0 & 1 \end{bmatrix}$$

Aplicando então o método de Gauss, obtemos, sucessivamente:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & & & \\ 0 & 1 & 0 & 0 & & & \\ 0 & 0 & 1 & 0 & & & \\ 0 & 0 & 0 & 1 & & & \\ 1 & -1 & 1 & 2 & 1 & 0 & 0 \\ -1 & 2 & 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 2 & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 0 & & & \\ 0 & 1 & 0 & 0 & & & \\ 0 & 0 & 1 & 0 & & & \\ 0 & 0 & 0 & 1 & & & \\ 1 & -1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 3 & 1 & 1 & 0 \\ 0 & -1 & 1 & 2 & 0 & 0 & 1 \end{bmatrix} \to$$

$$\to\begin{bmatrix} 1 & 1 & -1 & -2 & & & \\ 0 & 1 & 0 & 0 & & & \\ 0 & 0 & 1 & 0 & & & \\ 0 & 0 & 0 & 1 & & & \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 3 & 1 & 1 & 0 \\ 0 & -1 & 1 & 2 & 0 & 0 & 1 \end{bmatrix}\to\begin{bmatrix} 1 & 1 & -1 & -2 & & & \\ 0 & 1 & 0 & 0 & & & \\ 0 & 0 & 1 & 0 & & & \\ 0 & 0 & 0 & 1 & & & \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 3 & 1 & 1 & 0 \\ 0 & 0 & 2 & 5 & 1 & 1 & 1 \end{bmatrix}\to$$

$$\to\begin{bmatrix} 1 & 1 & -2 & -5 & & & \\ 0 & 1 & -1 & -3 & & & \\ 0 & 0 & 1 & 0 & & & \\ 0 & 0 & 0 & 1 & & & \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 5 & 1 & 1 & 1 \end{bmatrix}\to\begin{bmatrix} 1 & 1 & -2 & -5 & & & \\ 0 & 1 & -1 & -3 & & & \\ 0 & 0 & 1 & 0 & & & \\ 0 & 0 & 0 & 1 & & & \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 5/2 & 1/2 & 1/2 & 1/2 \end{bmatrix}\to$$

$$\to\begin{bmatrix} 1 & 1 & -2 & 0 & & & \\ 0 & 1 & -1 & -1/2 & & & \\ 0 & 0 & 1 & -5/2 & & & \\ 0 & 0 & 0 & 1 & & & \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1/2 & 1/2 & 1/2 \end{bmatrix}$$

Teremos, portanto,

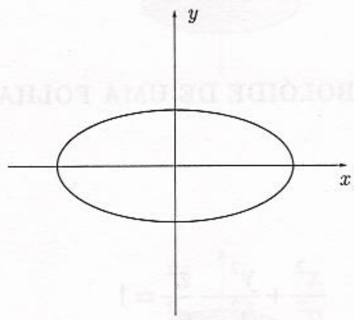
$$P=\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1/2 & 1/2 & 1/2 \end{bmatrix},\quad Q=\begin{bmatrix} 1 & 1 & -2 & 0 \\ 0 & 1 & -1 & -1/2 \\ 0 & 0 & 1 & -5/2 \\ 0 & 0 & 0 & 1 \end{bmatrix}\quad \text{e}\quad PAQ=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Apêndice E

Cónicas

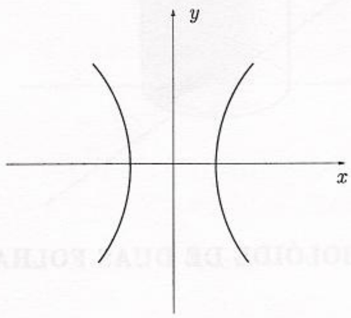
ELIPSE

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$



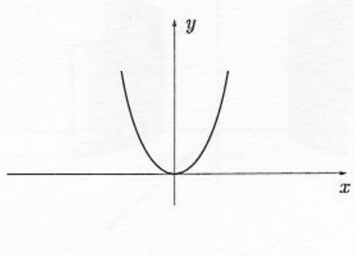
HIPÉRBOLE

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$$



PARÁBOLA

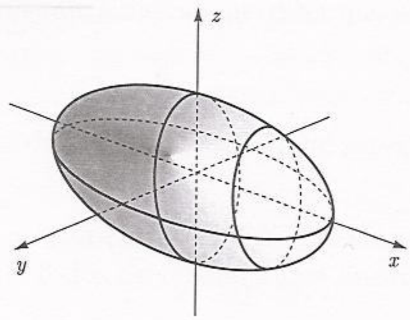
$$x^2 = ay, \quad a > 0$$



Quádricas

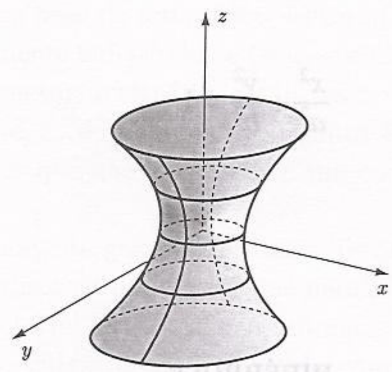
ELIPSÓIDE

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$



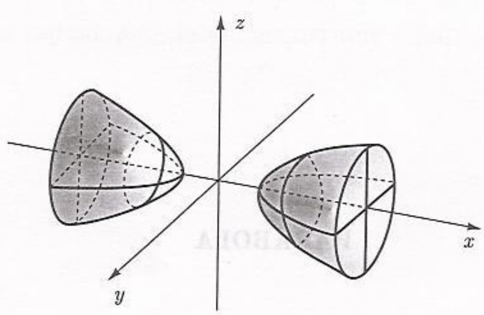
HIPERBOLÓIDE DE UMA FOLHA

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$



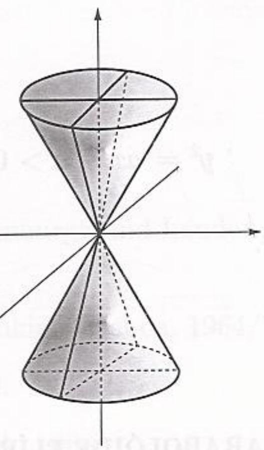
HIPERBOLÓIDE DE DUAS FOLHAS

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$



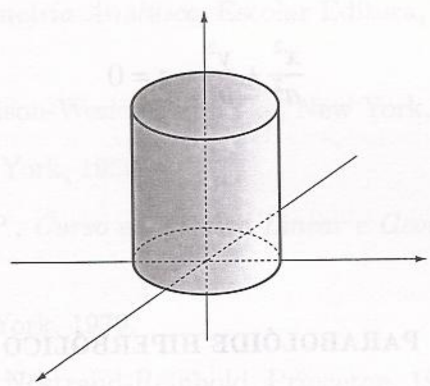
SUPERFÍCIE CÓNICA

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0$$



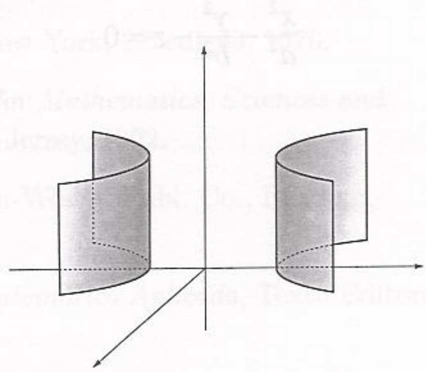
CILINDRO ELÍPTICO

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$



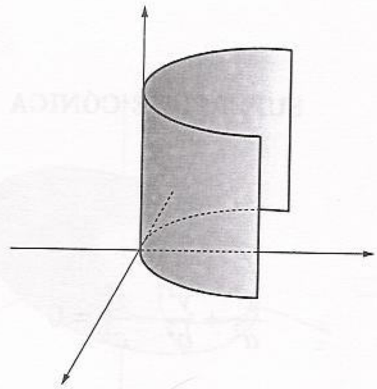
CILINDRO HIPERBÓLICO

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$$



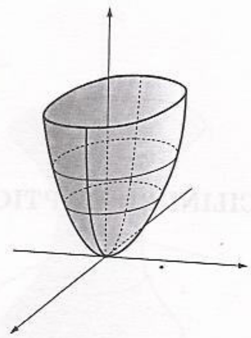
CILINDRO PARABÓLICO

$$y^2 = ax, \quad a > 0$$



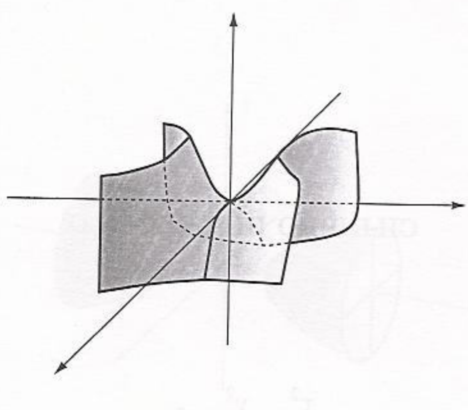
PARABOLÓIDE ELÍPTICO

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - z = 0$$



PARABOLÓIDE HIPERBÓLICO

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - z = 0$$



Bibliografia

Aitken, A. C., *Determinants and Matrices*, Oliver & Boyd, Edinburgh and London, 9.ª edição, 1967.

Almeida Costa, A., *Cours d'Algèbre Générale*, Fund. C. Gulbenkian, Lisboa, 1964/74.

Apostol, T. M., *Calculus*, John Wiley & Sons, New York, 1969.

Birkhoff, G. & MacLane, S., *Algebra*, MacMillan Co., New York, 1968.

Blyth, T. S. & Robertson, E. F., *Basic Linear Algebra*, Springer, London, 2000.

Bourbaki, N., *Théorie des Ensembles*, Hermann, Paris, 1960.

Dias Agudo, F. R., *Introdução à Álgebra Linear e Geometria Analítica*, Escolar Editora, Lisboa, 1983/86.

Farleigh, J. B. & Beauregard, A., *Linear Algebra*, Addison-Wesley Publ. Co., New York, 1990.

Gantmacher, F. R., *Theory of Matrices*, Chelsea, New York, 1959.

Giraldes, E., Fernandes, V. H. & Marques-Smith, M. P., *Curso de Álgebra Linear e Geometria Analítica*, McGraw-Hill, Lisboa, 1995.

Grätzer, G., *Universal Algebra*, Springer-Verlag, New York, 1979.

Halmos, P. R., *Finite-Dimensional Vector Spaces*, Van Nostrand-Reinhold, Princeton, 1958.

Ikrámov, J., *Problemas de Álgebra Lineal*, Editorial Mir, Moscovo, 1990.

Lang, S., *Linear Algebra*, Addison-Wesley Publ. Co., New York, 2.ª edição, 1970.

Laudesman, E. M. & Hestenes, M. R., *Linear Algebra for Mathematics, Sciences and Engineering*, Prentice-Hall International, New Jersey, 1992.

Lay, D. C., *Linear Algebra and its applications*, Addison-Wesley Publ. Co., Reading, Massachusetts, 1994.

Magalhães, L. T., *Álgebra Linear como introdução à Matemática Aplicada*, Texto Editora, Lisboa, 6.ª edição, 1996.

Monteiro, A., *Álgebra Linear e Geometria Analítica*, Associação de Estudantes, Faculdade de Ciências de Lisboa, Lisboa, 1982.

Monteiro, A. & Matos, I. T., *Álgebra — um primeiro curso*, Escolar Editora, Lisboa, 1995.

Monteiro, A., Pinto, G. & Marques, C., *Álgebra Linear e Geometria Analítica — Problemas e Exercícios*, McGraw-Hill, Lisboa, 1997.

Nef, Walter, *Linear Algebra*, Dover Publications Inc., New York, 1988.

Robinson, D. J. S., *A course in Linear Algebra, with applications*, World Scientific, Singapore, 1991.

Santos Guerreiro, J., *Curso de Análise Matemática*, Escolar Editora, Lisboa, 1989.

Índice remissivo

Esta obra apresenta, de forma sistemática, os temas normalmente constantes das disciplinas de Álgebra Linear e Geometria Analítica e que integram os planos curriculares de diversos cursos: Matemática, Física, Engenharia, Informática, Agronomia, Economia, etc.

Trata-se de um livro de texto acessível, que esclarece todas as dúvidas e estimula e desenvolve as capacidades de raciocínio abstracto. A matéria é exposta com todo o pormenor, o que torna a obra facilmente compreensível, não se exigindo pré-requisitos especiais para além de um conhecimento elementar da Teoria de Conjuntos.

O leitor pode melhorar o seu rendimento se acompanhar o estudo com o livro de exercícios *Álgebra Linear e Geometria Analítica Problemas e Exercícios*, do mesmo autor, inserido na colecção Schaum. Neste encontrará uma extensa e cuidada selecção de exercícios de dificuldade variável, que, acompanhando de perto o texto, lhe permitirá pôr à prova os conhecimentos adquiridos.

www.mcgraw-hill.pt

ISBN 972-773-106-6

9 789727 731060

Mc Graw-Hill
A Division of The McGraw-Hill Companies