NOTAS DE GEOMETRIA ANALÍTICA

José Mario Martínez

Departamento de Matemática Aplicada
IMECC-UNICAMP
13 de janeiro de 2004

Prefácio

Estas notas se destinam a um curso semestral de Geometria Analítica para o primeiro ano de estudos universitários.

É impossível traçar uma linha clara separando geometria analítica de álgebra linear. Não vemos, por outra parte, nenhuma utilidade em fazer essa diferenciação, portanto as notas abordam temas de "álgebra linear" sem cerimônia e sem aviso prévio.

A matéria começa com a interpretação e resolução de sistemas lineares de qualquer dimensão. Os motivos para este começo são três. Primeiro, a resolução e interpretação de sistemas de equações estão entre os temas mais aplicáveis e de maior utilidade para engenharia e ciências aplicadas. É desapontadora a freqüência com a qual detectamos que estudantes avançados somente tem consciência da existência de sistemas lineares quadrados. Segundo, os conjuntos de soluções de sistemas lineares fornecem os melhores exemplos para os objetos geométricos tradicionais nesta disciplina: pontos, retas, planos e suas generalizações. Terceiro, a maioria dos problemas geomêtricos de que trata o resto do curso podem ser formulados como sistemas lineares de equações e, de fato, esta costuma ser a forma menos rotineira de resolvê-los.

A prática de, diante de um problema, escrever um sistema de equações que o represente e encarar a resolução desse sistema, é extremamente frutífera na matemática e nas ciências. É um erro pensar que *todo* problema prático ou geomêtrico pode ser resolvido assim. Entretanto, trata-se de um "bom erro". Munido desse "preconceito", o estudante tentará resolver mais problemas dos que efetivamente pode mas, no processo, resolverá de fato muitos problemas e enxergará as ferramentas que são necessárias para resolver outros.

Tradicionalmente, a geometria analítica envolve apenas os espaços de duas e três dimensões. Porém, não vemos escolhos intrasponíveis para falar, em geral, do espaço n-dimensional, como nos textos de álgebra linear. Também é necessária uma dose razoável de álgebra linear para classificar cônicas e quádricas sem que o processo seja excessivamente particularista e entediante.

É importante fazer as listas de exercícios. Como em qualquer área da Matemática e - provavelmente - de qualquer outra ciência, a aprendizagem contemplativa não existe. O estudante aprenderá a matéria na medida exata que faça os problemas das listas.

Embora muitos problemas sejam mecânicos e demandem apenas a aplicação cuidadosa de determinadas regras, outros são "para pensar". Mesmo dentro desta categoria, alguns deles servem para desenvolver o pensamento puramente formal (definir, abstrair, demonstrar) e outros para a interpretação geométrica. Ambos são igualmente importantes.

Por outro lado, nos problemas "mecânicos" deve ser evitada a tentação de aprender uma receita e aplicá-la. Há várias formas de fazer cada exercício e quase sempre há uma que é a mais rápida, mas isso não tem a menor importância neste curso. Para a pergunta "Professor, posso fazer assim?" a única resposta é: "Se você for plenamente consciente do que está fazendo e o porquê, então pode".

Problemas difíceis não devem ser pulados. O tempo "perdido" em cima de um problema desse tipo não é, na verdade, perdido, pois o processo de pensar no problema deixa um sedimento valioso. Vale, a respeito disto, o poema de Drummond:

> Gastei uma hora pensando um verso
> que a pena não quer escrever.
> No entanto ele está cá dentro
> inquieto, vivo.
> Ele está cá dentro
> e não quer sair.
> Mas a poesia deste momento
> inunda minha vida inteira.

Sumário

Capítulo 1

Sistemas lineares

Muitos problemas da realidade podem ser modelados como sistemas de equações lineares. Por exemplo, uma indústria de alimentos deve fabricar uma ração para certo tipo de animais com determinados requisitos em termos nutricionais: cada kilograma de ração deve conter P gramas de proteína, G gramas de gordura e C gramas de carbohidratos. A ração é elaborada usando um conjunto de ingredientes disponíveis no mercado, digamos, arroz, soja, ossos, miúdos de boi, banha de porco e peixe. Chamamos $x_1, x_2, x_3, x_4, x_5, x_6$ à quantidade, em gramas, de cada um dos ingredientes que será usada para fabricar um kilograma de ração. Assim, x_1 será a quantidade de arroz, x_2 a quantidade de soja e assim por diante. Cada ingrediente tem os três nutrientes em diferentes proporções. Para cada $i = 1, \ldots, 6$, chamamos p_i á fração de proteína que contem uma quantidade fixa do ingrediente i, g_i é a fração de gordura e c_i a fração de carbohidratos. Portanto, as equações que representam a necessidade de que as exigências nutricionais sejam cumpridas são:

$$p_1x_1 + p_2x_2 + p_3x_3 + p_4x_4 + p_5x_5 + p_6x_6 = P,$$

$$g_1x_1 + g_2x_2 + g_3x_3 + g_4x_4 + g_5x_5 + g_6x_6 = G,$$

$$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 + c_6x_6 = C.$$

Alem disto:

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 1000,$$

pois a soma das quantidades dos ingredientes deve ser 1000 gramas.

Vemos que este sistema tem quatro equações e seis incógnitas (x_1, x_2, x_3, x_4, x_5 e x_6). Se encontrarmos valores das incógnitas que satisfaçam as quatro equações do sistema e fabricarmos cada kilograma de ração misturando x_1 gramas do ingrediente 1, x_2 gramas do ingrediente 2 etcétera, obteremos um produto aceitável do ponto de vista nutricional. Este sistema não tem uma única solução: os ingredientes podem ser misturados em diferentes proporções para obter os resultados adequados. A escolha da "melhor" solução possível foge ao escopo destas notas.

1.1 Solução geral

Nesta seção veremos o conceito de *solução geral* de um sistema linear de m equações com n incógnitas. Deve-se frisar que os números de equações e de incógnitas são arbitrários. Nada impede, por exemplo que um sistema de equações esteja constituído por única equação e várias incógnitas.

Consideremos o seguinte exemplo:

$$x_1 - x_2 + 3x_3 + 2x_4 = 1 \tag{1.1}$$

$$-x_2 + 5x_3 + 3x_4 = 0. \tag{1.2}$$

Trata-se de um sistema com 2 equações e 4 incógnitas.

Colocando x_2 em evidência na segunda equação, obtemos

$$x_2 = 5x_3 + 3x_4. \tag{1.3}$$

Substituindo na primeira equação, temos:

$$x_1 - 5x_3 - 3x_4 + 3x_3 + 2x_4 = 1,$$

ou seja:

$$x_1 = 2x_3 + x_4 + 1. \tag{1.4}$$

Por (1.3) e (1.4) vemos que, dando valores arbitrários a x_3 e x_4 podemos calcular x_1 e x_2 e, assim, obter diferentes soluções do sistema original.

Por exemplo, se $x_3 = x_4 = 0$, temos $x_1 = 1, x_2 = 0$. Portanto, a "4-upla" $(x_1 = 1,\ x_2 = 0,\ x_3 = 0,\ x_4 = 0)$ é uma solução do sistema original, o que pode ser facilmente verificado substituindo em (1.1) e (1.2).

Se $x_3 = 1, x_4 = 2$ obtemos $x_1 = 5, x_2 = 11$. Com efeito, é fácil ver que a 4-upla $(x_1 = 5,\ x_2 = 11,\ x_3 = 1,\ x_4 = 2)$ é uma solução do sistema original.

Dessa maneira, podemos obter infinitas soluções deste sistema. Como existem dois parâmetros aos quais podemos dar livremente valores para obter diferentes soluções, dizemos que o sistema tem dois *graus de liberdade*.

Uma maneira simples de representar *todas* as soluções do sistema é mediante sua *solução geral*. Por (1.3) e (1.4), a solução geral deste sistema é

$$\begin{pmatrix} 2x_3 + x_4 + 1 \\ 5x_3 + 3x_4 \\ x_3 \\ x_4 \end{pmatrix}.$$

Para expressar esta solução geral de maneira ainda mais clara, vamos usar as regras:

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} a+\alpha \\ b+\beta \\ c+\gamma \\ d+\delta \end{pmatrix} \tag{1.5}$$

e

$$\lambda \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} \lambda a \\ \lambda b \\ \lambda c \\ \lambda d \end{pmatrix}. \tag{1.6}$$

Então, temos:

$$\begin{aligned} &\begin{pmatrix} 2x_3 + x_4 + 1 \\ 5x_3 + 3x_4 \\ x_3 \\ x_4 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 2x_3 \\ 5x_3 \\ x_3 \\ 0 \end{pmatrix} + \begin{pmatrix} x_4 \\ 3x_4 \\ 0 \\ x_4 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ 5 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix}. \end{aligned}$$

Esta é uma segunda forma de expressar a solução geral do sistema.

1.2 Escalonamento

Obter a solução geral do sistema (1.1–1.2) foi bastante fácil. O que fez que fosse assim é que a variável x_1 não aparecia na equação (1.2). Sempre que, em cada equação de um sistema linear, apareça uma variável que não aparece em nenhuma das equações seguintes, o sistema será fácil de resolver. Sistemas com essas características se dizem *escalonados* [1].

Quando um sistema não é naturalmente escalonado a melhor estratégia para resolvê-lo é transformá-lo em um sistema equivalente escalonado. Esse processo se chama *Eliminação Gaussiana.* A ferramenta fundamental para poder fazer isso é o seguinte teorema.

Teorema 1.2. *Suponha que, em um sistema linear, substituímos uma de suas equações por "ela mesma vezes um número diferente de zero mais um múltiplo de outra". Então, o sistema resultante tem as mesmas soluções que o original.*

[1] Sejamos cuidadosos: dizemos que uma variável "aparece" em uma equação quando o coeficiente pela qual está multiplicada é diferente de zero. Por exemplo, a variável x_1 aparece na equação $3x_1 + x_2 - x_3 = 8$ mas "não aparece" na equação $0x_1 + x_2 - x_3 = 9$.

Antes de provar este teorema, veremos como podemos aplicá-lo para conseguir um sistema escalonado a partir de um sistema arbitrário.

Vejamos um exemplo. Consideremos o sistema de 3 equações com 4 incógnitas:

$$3x_1 + 2x_2 - x_3 - x_4 = 4$$

$$x_1 - x_2 + 2x_3 - 3x_4 = 2$$

$$3x_1 + 3x_2 - 3x_3 - 2x_4 = 1.$$

A variável x_1 aparece na primeira equação. Seria desejável, pois, que não aparecesse em nenhuma das equações seguintes. Para que deixe de aparecer na segunda equação podemos substituir esta equação por "3 vezes ela mesma menos a primeira equação". Imediatamente, substituindo a terceira equação por "ela mesma menos a primeira equação" obtemos o sistema equivalente:

$$3x_1 + 2x_2 - x_3 - x_4 = 4$$

$$-5x_2 + 7x_3 - 8x_4 = 2$$

$$x_2 - 2x_3 - x_4 = -3.$$

Neste novo sistema, a variável x_2 aparece na segunda equação. Seria desejável, como antes, que não aparecesse na terceira. Para isso, podemos substituír a terceira equação do novo sistema por "5 vezes ela mesma mais a segunda equação". Assim, obtemos o sistema equivalente:

$$3x_1 + 2x_2 - x_3 - x_4 = 4$$

$$-5x_2 + 7x_3 - 8x_4 = 2$$

$$-3x_3 - 13x_4 = -13.$$

Este sistema já é escalonado. Portanto, pode ser resolvido pelo processo mostrado na Seção 1.1. Mais precisamente, podemos:

(a) Pôr em evidência x_3 na terceira equação e substituír nas duas primeiras equações;

(b) Pôr em evidência x_2 na segunda equação e substituír na primeira equação;

(c) Pôr em evidência x_1 na primeira equação.

Desta maneira, x_1, x_2 e x_3 ficarão em função de x_4.

Às vezes, no processo de escalonamento, pode ser completamente eliminada uma equação ou pode ser detectado que o sistema não tem nenhuma solução. Vejamos exemplos de ambas situações.

Consideremos o sistema

$$5x_1 + 2x_2 - x_3 = 1$$

$$3x_1 - x_2 - 6x_3 = 2$$
$$8x_1 + x_2 - 7x_3 = 3.$$

Substituindo a segunda equação por "ela mesma vezes 5 menos a primeira vezes 3" e a terceira equação por "ela mesma vezes cinco menos a primeira vezes 8" obtemos:

$$5x_1 + 2x_2 - x_3 = 1$$
$$-11x_2 - 27x_3 = 7$$
$$-11x_2 - 27x_3 = 7.$$

Agora, para eliminar x_2 da terceira equação substituimos esta equação por "ela mesma menos a segunda". Assim, obtemos:

$$5x_1 + 2x_2 - x_3 = 1$$
$$-11x_2 - 27x_3 = 7$$
$$0 = 0.$$

A terceira equação ficou sem incógnitas e, certamente, pode ser eliminada, pois a proposição $0 = 0$ é uma tautologia que não oferece nenhuma informação. Portanto, o sistema sem essa equação já ficou escalonado.

Entretanto, consideremos o sistema

$$5x_1 + 2x_2 - x_3 = 1$$
$$3x_1 - x_2 - 6x_3 = 2$$
$$8x_1 + x_2 - 7x_3 = 9.$$

Substituindo, como no exemplo anterior, a segunda equação por "ela mesma vezes 5 menos a primeira vezes 3" e a terceira equação por "ela mesma vezes cinco menos a primeira vezes 8" obtemos:

$$5x_1 + 2x_2 - x_3 = 1$$
$$-11x_2 - 27x_3 = 7$$
$$-11x_2 - 27x_3 = 37.$$

Finalmente, para eliminar x_2 da terceira equação substituimos esta equação por "ela mesma menos a segunda". Assim, obtemos:

$$5x_1 + 2x_2 - x_3 = 1$$

$$-11x_2 - 27x_3 = 7$$

$$0 = 30.$$

Este sistema é equivalente ao colocado inicialmente e, obviamente, não tem nenhuma solução pois a "equação" $0 = 30$ não se cumpre nunca.

O processo de escalonamento pode ser descrito, em geral, pelo seguinte *algoritmo*, ou sequência de instruções.

Algoritmo 1.2
Suponha que o sistema está formado por m equações.
Para $i = 1, \ldots, m-1$, execute o Passo 1.

Passo 1.

1. Se em nenhuma das equações $i, \ldots, m$ aparece nenhuma variável, termine a execução do Algoritmo 1.2. Nesse caso, o escalonamento do sistema está finalizado.

2. Se na equação i não aparece nenhuma variável mas em uma equação posterior aparece alguma, permute a equação i com esta equação (que agora passará a ser a nova equação i).

3. Escolha uma das variáveis que aparecem na (talvez nova) equação i. Chamemos $\mathbf{x}$ a essa variável. Para cada $j = i+1, \ldots, m$, substitua a equação j por um múltiplo diferente de zero dela mesma mais um múltiplo da equação i, de maneira a eliminar a variável $\mathbf{x}$ da equação j.

Ao terminar a execução do Algoritmo 1.2 o sistema terá ficado escalonado. Se alguma equação do sistema ficou da forma $0 = c$ mas $c \neq 0$ o sistema original não tem solução. Se isto não ocorreu, eliminando as equações da forma $0 = 0$, poderemos encontrar a solução geral do sistema seguindo o procedimento explicado na Seção 1.1.

Prova do Teorema 1.2. Suponhamos que o sistema original contém as equações

$$a_1x_1 + \ldots + a_nx_n = b \tag{1.7}$$

e

$$c_1x_1 + \ldots + c_nx_n = d \tag{1.8}$$

e que, no novo sistema, a equação (1.8) foi substituída por

$$(\alpha c_1 + \beta a_1)x_1 + \ldots + (\alpha c_n + \beta a_n)x_n = \alpha d + \beta b, \tag{1.9}$$

onde $\alpha \neq 0$.

A prova do teorema consiste em mostrar que o conjunto de soluções das equações (1.7, 1.8) é o mesmo que o conjunto de soluções das equações (1.7, 1.9).

Suponhamos que $x_1 = y_1, \ldots, x_n = y_n$ é uma solução de (1.7, 1.8). Portanto,

$$a_1 y_1 + \ldots + a_n y_n = b \tag{1.10}$$

e

$$c_1 y_1 + \ldots + c_n y_n = d \tag{1.11}$$

Multiplicando ambos membros de (1.11) por α e ambos membros de (1.10) por β, obtemos:

$$\alpha c_1 y_1 + \ldots + \alpha c_n y_n = \alpha d$$

e

$$\beta a_1 y_1 + \ldots + \beta a_n y_n = \beta b.$$

Somando membro a membro estas duas igualdades, obtemos:

$$\alpha c_1 y_1 + \ldots + \alpha c_n y_n + \beta a_1 y_1 + \ldots + \beta a_n y_n = \alpha d + \beta b. \tag{1.12}$$

Como $x_1 = y_1, \ldots, x_n = y_n$ é solução de (1.10) e de (1.12) resulta que, como esperávamos, toda solução do sistema original é solução do sistema modificado.

Agora suponhamos que $x_1 = z_1, \ldots, x_n = z_n$ é solução do sistema modificado. Portanto

$$a_1 z_1 + \ldots + a_n z_n = b \tag{1.13}$$

e

$$\alpha c_1 z_1 + \ldots + \alpha c_n z_n + \beta a_1 z_1 + \ldots + \beta a_n z_n = \alpha d + \beta b. \tag{1.14}$$

Multiplicando ambos membros de (1.13) por β, obtemos:

$$\beta a_1 z_1 + \ldots + \beta a_n z_n = \beta b. \tag{1.15}$$

Subtraindo, membro a membro, (1.15) de (1.14), deduzimos:

$$\alpha c_1 z_1 + \ldots + \alpha c_n z_n = \alpha d.$$

Como $\alpha \neq 0$, podemos dividir ambos membros desta equação por α, portanto:

$$c_1 z_1 + \ldots + c_n z_n = d. \tag{1.16}$$

Por (1.13) e (1.16), vemos que $x_1 = z_1, \ldots, x_n = z_n$ é solução do sistema original.

Portanto, provamos que os conjuntos de soluções do sistema original e do sistema modificado são os mesmos. Isto significa que os dois sistemas são equivalentes, com o qual o teorema fica demonstrado.

1.3 Exercícios

1. Encontre o conjunto de soluções dos seguintes sistemas lineares escalonados. Escreva a "solução geral" de cada um destes sistemas. Interprete geometricamente. Identifique parâmetros da solução geral e graus de liberdade. Em cada caso onde haja infinitas soluções calcule explicitamente pelo menos duas. Desenhe tudo o que for possível.

   (a)

   $3x_1 + 2x_2 = 4$

   $x_2 = 6$

   (b)

   $3x_1 - x_2 = 1$

   (c)

   $x_1 + x_2 = 8$

   $x_1 = 3$

   (d)

   $x_1 = 5$

   (e)

   $x + y = 2$

   $y = 9$

   (f)

   $x_1 + x_2 - x_3 = 3$

   $x_2 + x_3 = 5$

   $x_3 = 8$

   (g)

   $x_1 - x_2 - x_3 = 100$

   $x_2 = 5$

$x_3 = 6$

(h)

$x_1 - x_2 - x_3 = 7$

$x_1 + x_2 = 987$

$x_1 = 11$

(i)

$3x_1 + x_2 - x_3 = 4$

$2x_2 + 5x_3 = 10$

(j)

$x_1 - x_2 - x_3 = 2$

(k)

$x + y + z + w = 4$

$y + z + w = 3$

$z + w = 2$

$w = 1$

(l)

$3x_1 + x_2 - x_3 - x_4 = 2$

$x_2 - x_3 + x_4 = 1$

$x_3 + 5x_4 = 6$

(m)

$2x - y + z - w = 1$

$x + y + z = 4$

(n)

$4x_1 - x_2 - x_3 - x_4 = 6$

(o)

$2x_1 - x_2 - x_3 - x_4 - x_5 - x_6 = 8$

$x_3 - 3x_4 + 4x_6 = 2001$

2. Explique a razão pela qual cada um dos sistemas do exercício anterior são escalonados. Explique a razão pela qual todo sistema escalonado é "fácil de resolver".

3. Sistematize em forma de algoritmo o procedimento para encontrar a solução geral de um sistema escalonado.

4. Transforme os sistemas abaixo em sistemas escalonados. Depois, encontre o conjunto de soluções. Escreva a "solução geral" de cada um destes sistemas. Interprete geometricamente. Identifique parâmetros da solução geral e graus de liberdade. Em cada caso onde haja infinitas soluções calcule explicitamente pelo menos duas. Desenhe tudo o que for possível.

   (a)

   $3x_1 + 2x_2 = 4$

   $3x_1 + 3x_2 = 10$

   (b)

   $3x_1 - x_2 = 1$

   (c)

   $x_1 + x_2 = 8$

   $2x_1 + x_2 = 11$

   (d)

   $x_1 = 5$

   (e)

   $x + y = 2$

   $x - y = -7$

   (f)

$x_1 + x_2 - x_3 = 3$

$x_1 - 2x_3 = -2$

$2x_1 + x_2 - 2x_3 = 9$

(g)

$x_1 - x_2 - x_3 = 100$

$x_1 - x_3 = 105$

$2x_1 - x_2 - 2x_3 = 211$

(h)

$x_1 - x_2 - x_3 = 7$

$2x_1 - x_3 = 994$

$4x_1 - x_2 - 2x_3 = 1112$

(i)

$3x_1 + x_2 - x_3 = 4$

$3x_1 + 3x_2 + 4x_3 = 14$

(j)

$x_1 - x_2 - x_3 = 2$

(k)

$x + y + z + w = 4$

$x + 2y + 2z + 2w = 7$

$2x + 3y + 4z + 4w = 13$

$4x + 6y + 7z + 8w = 25$

(l)

$3x_1 + x_2 - x_3 - x_4 = 2$

$3x_1 + 2x_2 - 2x_3 = 3$

$6x_1 + 3x_2 - 2x_3 + 4x_4 = 11$

(m)

$2x - y + z - w = 1$

$x + y + z + w = 4$

(n)

$4x_1 - x_2 - x_3 - x_4 = 6$

(o)

$2x_1 - x_2 - x_3 - x_4 - x_5 - x_6 = 8$

$x_1 + x_2 + x_3 - 3x_4 + x_5 + 4x_6 = 38$

(p)

$2x_1 + x_2 = 3$

$x_1 + 2x_2 = 3$

$x_1 + x_2 = 4$

(q)

$2x_1 + x_2 = 3$

$x_1 + 2x_2 = 3$

$x_1 + x_2 = 2$

5. Um sistema linear

$$a_{i,1}x_1 + \ldots + a_{i,n}x_n = b_i, \; i = 1, \ldots, n$$

se diz *homogêneo* se $b_i = 0$ para todo $i = 1, \ldots, n$. Que aspecto deve ter a solução geral de um sistema homogêneo?

6. O exame de sangue do paciente A estabelece que ele tem 37 mg/dl de HDL ("colesterol bom"), 86.2 mg/dl de LDL ("colesterol ruim"), 144 mg/dl de triglicerídeos e 152 mg/dl de Colesterol Total. Jà o exame do paciente B diz que ele tem 53 mg/dl de HDL ("colesterol bom"), 132 mg/dl de LDL ("colesterol ruim"), 82 mg/dl de triglicerídeos e 201 mg/dl de Colesterol Total. Por último, para o paciente C, o relatório do laboratório dá por resultado 50 mg/dl de HDL ("colesterol bom"), 100 mg/dl de LDL ("colesterol ruim"), 150 mg/dl de triglicerídeos e 180 mg/dl de Colesterol Total.

O colesterol total, o colesterol bom e os triglicerídeos são efetivamente medidos no laboratório, mas o colesterol ruim é estimado usando uma fórmula (chamada "fórmula de Friedwald"). Qual você acha que é essa fórmula?

7. Neste capítulo começamos mostrando que o problema de fabricação de uma ração animal envolve a resolução de um sistema linear de equações. Entretanto, esse problema prático não acaba aí, pois muitas soluções desse sistema não servem (aquelas que têm componentes negativas) e algumas soluções do sistema linear são melhores que outras porque os ingredientes podem ser adquiridos com um custo menor. Usando estas observações, formule o problema da ração de uma maneira mais realista. Para o caso de dois ingredientes e um nutriente, interprete geometricamente a nova formulação. Faça o mesmo para os casos de um nutriente e três ingredientes e para o caso de dois nutrientes e três ingredientes.

Capítulo 2

O espaço $I\!R^n$

O grande achado de René Descartes (1596-1650) e Pierre de Fermat (1601-1655) foi a identificação do plano físico com os pares de números reais e do espaço físico tridimensional com as 3-uplas de números reais. Isso permitiu obter todas as propriedades da geometria euclidiana dos antigos gregos a partir de cálculos puramente algêbricos. Mais ainda, a interpretação das n-uplas permitiu dar um sentido preciso e estudar as propriedades de espaços de dimensão superior a 3. Espaços com dimensão maior que 3 não são um mero artifício matemático. Cada vez que um objeto pode ser caracterizado por, digamos, 6 parâmetros podemos pensar que esse objeto é um ponto em um espaço de dimensão 6. A partir daí a distância entre objetos pode ser medida, eles podem ser agrupados de acordo a proximidade ou outras características e tais abstrações podem conduzir ao aprofundamento de conhecimentos ou à tomada de decisões nas mais diversas áreas.

2.1 Soma em $I\!R^n$ e produto por escalares

Chamamos $I\!R^n$ ao conjunto das n-uplas de números reais. Por motivos que ficarão claros mais adiante, estas n-uplas serão colocadas em forma de coluna. Por exemplo,

$$\begin{pmatrix} 3 \\ -8.7 \end{pmatrix} \in I\!R^2, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \in I\!R^4, \quad \begin{pmatrix} -1 \\ -1 \\ 0.3 \end{pmatrix} \in I\!R^3.$$

A n-upla $\begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$ será chamada *origem* e denotada com a letra O.

Dado um "plano físico", desenhando duas retas perpendiculares e estabelecendo a unidade de distância, todo elemento de $I\!R^2$ pode ser interpretado como um ponto do plano. Veja Figura 2.1.

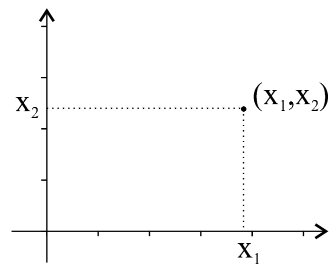

Figura 2.1. **Ponto no plano com coordenadas cartesianas.**

Analogamente, todo elemento de $I\!R^3$ pode ser interpretado como um ponto do "espaço físico", depois de traçados três eixos perpendiculares. Veja Figura 2.2.

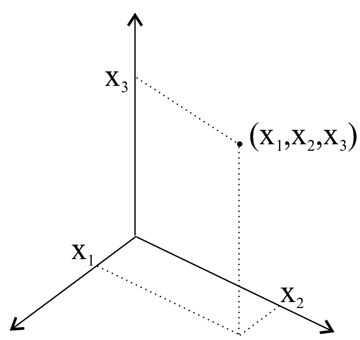

Figura 2.2. **Ponto no espaço com coordenadas cartesianas.**

Muitas vezes é conveniente interpretar os elementos de $I\!R^n$ como *vetores*. O vetor x é "a flecha" com "ponto de aplicação" na origem e "ponta" no ponto x.

Elementos de $I\!R^n$ se somam de acordo com a regra que generaliza a fórmula (1.5), vista no capítulo anterior:

$$\begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n + y_n \end{pmatrix}.$$

Geometricamente, se os vetores x e y são dois lados de um paralelogramo, então $x + y$ é a diagonal que passa pela origem. Veja Figura 2.3.

O conjunto $I\!R^n$ é um bom modelo para representar "vetores livres". O "vetor livre" x tem seu ponto de aplicação em qualquer ponto $P \in I\!R^n$ e sua "ponta" no

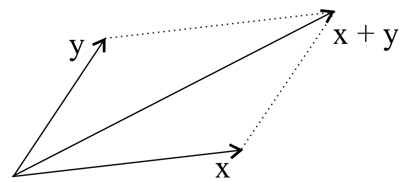

Figura 2.3. **Soma de vetores em $I\!R^n$.**

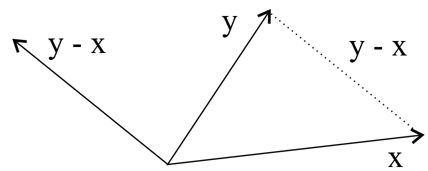

Figura 2.4. **Diferença entre vetores de $I\!R^n$.**

ponto $P + x$. A diferença $y - x$ pode ser interpretada como o vetor livre que tem ponto de aplicação em x e "ponta" em y. Veja Figura 2.4.

Como em (1.6), o produto de elementos de $I\!R^n$ por nũmeros reais é definido assim:

$$\lambda \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \cdot \\ \cdot \\ \cdot \\ \lambda x_n \end{pmatrix} \quad \forall\ \lambda \in I\!R.$$

Veja Figura 2.5.

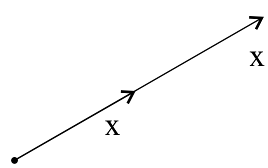

Figura 2.5. **Produto de um vetor de $I\!R^n$ por um escalar.**

2.2 Distâncias e ângulos

A *norma* (ou comprimento, ou módulo) de $x \in I\!R^n$ se define assim:

$$\|x\| = \sqrt{x_1^2 + \ldots + x_n^2}.$$

A quantidade $\|x\|$ é uma medida do "tamanho" do vetor x. Pelo Teorema de

Pitágoras, esta medida é a distância euclidiana entre o ponto x e a origem. Veja Figura 2.6.

Uma propriedade muito fácil de verificar é que, para qualquer $x \in I\!R^n$ e $\lambda \in I\!R$,

$$\|\lambda x\| = |\lambda| \|x\|.$$

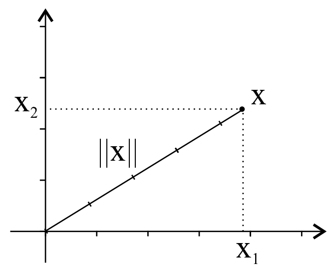

Figura 2.6. **Norma de vetores.**

A *distância* entre dois pontos $x, y \in I\!R^n$ se define como a norma da diferença $y - x$. Veja Figura 2.7.

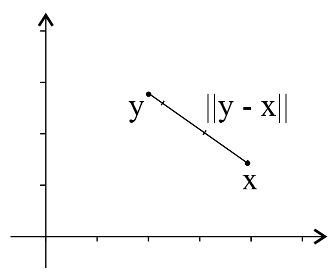

Figura 2.7. **Distância entre pontos de $I\!R^n$.**

Dados $x, y \in I\!R^n$ definimos seu *produto escalar*:

$$\langle x, y \rangle = x_1 y_1 + \ldots + x_n y_n.$$

Uma lista de propriedades bastante simples, cuja prova deixamos por conta do leitor, são enunciadas na seguinte proposição.

Proposição 2.2. *Para quaisquer $x, y, z \in I\!R^n$, $\alpha, \beta \in I\!R$, tem-se:*

$$\langle x, y \rangle = \langle y, x \rangle,$$

$$\langle \alpha x, y \rangle = \langle x, \alpha y \rangle = \alpha \langle x, y \rangle,$$

$$\langle x + y, z \rangle = \langle z, x + y \rangle = \langle x, z \rangle + \langle y, z \rangle.$$

A *Desigualdade Cauchy-Schwarz* é uma propriedade importante que relaciona produto escalar e normas. Veremos sua demonstação no seguinte teorema.

Teorema 2.2. *Para quaisquer $x, y \in I\!R^n$, tem-se que*

$$|\langle x, y\rangle| \leq \|x\|\|y\|. \tag{2.1}$$

Prova. Vejamos a prova desta propriedade no caso $n = 3$. A prova geral é totalmente análoga, mas neste caso particular é mais fácil de visualizar.

As seguinte desigualdade é óbvia:

$$(x_1y_2 - x_2y_1)^2 + (x_1y_3 - x_3y_1)^2 + (x_2y_3 - x_3y_2)^2 \geq 0.$$

Portanto, desenvolvendo os quadrados:

$$(x_1y_2)^2 + (x_2y_1)^2 + (x_1y_3)^2 + (x_3y_1)^2 + (x_2y_3)^2 + (x_3y_2)^2$$

$$\geq 2x_1y_1x_2y_2 + 2x_1y_1x_3y_3 + 2x_2y_2x_3y_3.$$

Somando $(x_1y_1)^2 + (x_2y_2)^2 + (x_3y_3)^2$ a ambos membros da desigualdade acima, temos:

$$(x_1y_1)^2 + (x_2y_2)^2 + (x_3y_3)^2 + (x_1y_2)^2 + (x_2y_1)^2 + (x_1y_3)^2 + (x_3y_1)^2 + (x_2y_3)^2 + (x_3y_2)^2$$

$$\geq (x_1y_1)^2 + (x_2y_2)^2 + (x_3y_3)^2 + 2x_1y_1x_2y_2 + 2x_1y_1x_3y_3 + 2x_2y_2x_3y_3.$$

Portanto:

$$(x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) \geq (x_1y_1 + x_2y_2 + x_3y_3)^2.$$

Logo,

$$\|x\|^2\|y\|^2 \geq \langle x, y\rangle^2.$$

Isto implica a desigualdade (2.1).

Corolário 2.2. *Se $x, y \in I\!R^n$ são não-nulos, então:*

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

O Corolário 2.2 sugere que a quantidade $\frac{\langle x,y\rangle}{\|x\|\|y\|}$ mede alguma propriedade significativa em relação à posição relativa dos vetores x e y. De fato, veremos agora que essa propriedade é o *cosseno* do ângulo entre os vetores não-nulos x e y.

Para ver isto, desenhe os vetores não-nulos x e y de maneira que formem um ângulo agudo. Agora, considere todos os pontos da forma tx com $t \geq 0$. Esses pontos formam a semirreta com origem em O e que contêm o vetor x. Desenhe, nesta semirreta, o ponto que fica mais perto de y. Chamemos P a esse ponto. Claramente, o segmento que une y com P deve ser perpendicular à semirreta, portanto, os pontos y, P e a origem O formam um triângulo retângulo. A hipotenusa mede $\|y\|$ e o "cateto adjacente" mede $\|P\|$. Logo, o que entendemos por *cosseno* do ângulo entre x e y (da trigonometria plana elementar) é o quociente $\|P\|/\|y\|$. (Como o ângulo é agudo, o sinal é positivo.)

Agora vamos procurar uma fórmula analítica para P. Este ponto deve ser aquele elemento da forma tx, com $t \geq 0$, cuja distância a y é a menor possível. Portanto, se trata de encontrar $t \geq 0$ tal que a expressão

$$\|y - tx\|$$

seja mínima. Desenvolvendo, temos:

$$\|y - tx\|^2 = \|y\|^2 + t^2\|x\|^2 - 2t\langle x, y\rangle.$$

Logo, encontrar t tal que minimize $\|y - tx\|$ equivale a encontrar t minimizando a quantidade

$$t^2\|x\|^2 - 2t\langle x, y\rangle.$$

Agora,

$$\begin{aligned} t^2\|x\|^2 - 2t\langle x, y\rangle &= t^2\|x\|^2 - 2t\langle x, y\rangle + \langle x, y\rangle^2/\|x\|^2 - \langle x, y\rangle^2/\|x\|^2 \\ &= (t\|x\| - \langle x, y\rangle/\|x\|)^2 - \langle x, y\rangle^2/\|x\|^2. \end{aligned}$$

O valor de t que minimiza esta expressão é aquele que faz com que o termo elevado ao quadrado se anule, ou seja:

$$t = \frac{\langle x, y\rangle}{\|x\|^2}.$$

Como nosso ângulo é agudo, este t é maior que zero, portanto $\langle x, y\rangle > 0$. Assim, temos:

$$P = \frac{\langle x, y\rangle}{\|x\|^2}x,$$

portanto

$$\|P\| = \frac{\langle x, y\rangle}{\|x\|}.$$

Mas o cosseno trigonométrico do ângulo é $\|P\|/\|y\|$, portanto esse cosseno é igual a

$$\frac{\langle x, y\rangle}{\|x\|\|y\|}$$

como suspeitávamos.

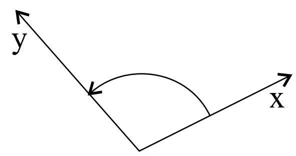

Figura 2.8. **Ângulo entre vetores.**

Fica a cargo do leitor mostrar que o mesmo raciocínio geométrico, com pequenas variações, pode ser reproduzido se o ângulo entre x e y é obtusso (parta da Figura 2.8) e também se o ângulo é reto. Portanto, de agora em diante, a quantidade

$$\frac{\langle x, y\rangle}{\|x\|\|y\|}$$

será adotada como definição de cosseno do ângulo entre os vetores não-nulos x e y.

Assim, os vetores serão *ortogonais* se seu produto escalar é nulo.

Uma conseqüência importantíssima da desigualdade Cauchy-Schwarz é a *Desigualdade triangular.* Esta nos diz que dados $x, y \in I\!R^n$ tem-se que

$$\|x+y\| \leq \|x\| + \|y\|. \tag{2.2}$$

Daqui se deduz que

$$\|x-y\| \geq |\|x\| - \|y\|| \tag{2.3}$$

e que, para todos $x, y, z \in I\!R^n$,

$$\|x-z\| \leq \|x-y\| + \|y-z\|. \tag{2.4}$$

Veja Figura 2.9.

Também é fácil provar, usando as definições de norma e produto escalar, o teorema que diz que se x e y são ortogonais,

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2. \tag{2.5}$$

Esta é a versão analítica do Teorema de Pitágoras.

2.3 Retas

Dados $P, v \in I\!R^n$, $v \neq 0$, definimos a *reta que passa por* P e tem v como *vetor diretor* como o conjunto de pontos x de $I\!R^n$ tais que $x = P + tv$ para algum $t \in I\!R$. Veja Figura 2.10. Esta costuma ser chamada a forma *paramêtrica* de definir a reta. Consequentemente, o escalar t, que determina cada ponto da reta, se denomina "parâmetro".

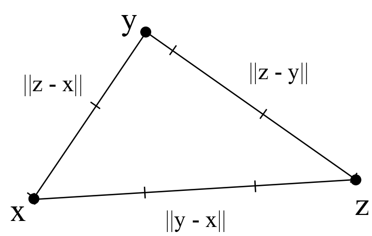

Figura 2.9. **Desigualdade triangular.**

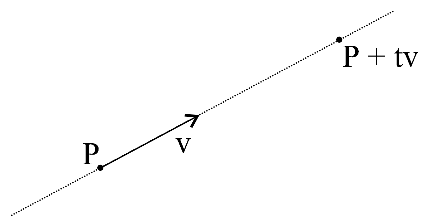

Figura 2.10. **Reta: pontos da forma $P + tv$.**

Observe que, se A e B são dois pontos diferentes de uma reta, então necessariamente essa reta é a reta que passa por A e tem $B - A$ como vetor diretor.

Também é fácil ver que, se v é vetor diretor de uma reta, então μv também é vetor diretor para qualquer $\mu \neq 0$.

Suponhamos que o conjunto de soluções de um sistema linear tem um grau de liberdade. Por exemplo, suponhamos que a solução geral é

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} + t \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}.$$

Então, pela definição de reta, vemos que este conjunto de soluções representa uma reta em $I\!R^n$ ($n = 4$ no exemplo).

2.4 Distância entre ponto e reta

Consideremos o problema de encontrar a distância entre um ponto Q e uma reta L. Veja Figura 2.11.

Essa distância se define como a distância entre o ponto Q e aquele ponto P da reta que fica mais perto do ponto dado. O ponto P é chamado *projeção* de Q na reta L. Como vemos, P deve satisfazer duas condições:

(a) P deve pertencer à reta L;

(b) $Q - P$ deve ser ortogonal ao vetor diretor da reta.

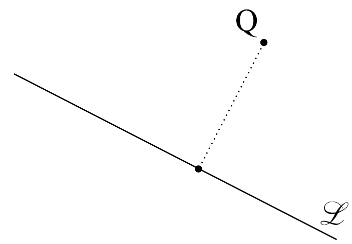

Figura 2.11. **Distância entre ponto e reta.**

A resolução do problema passa por colocar essas duas condições em forma de equações e por resolver o sistema correspondente.

Por exemplo, consideremos o problema de encontrar a distância entre $Q = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} \in I\!R^3$ e a reta que passa por $\begin{pmatrix} 2 \\ -1 \\ 6 \end{pmatrix}$ e $\begin{pmatrix} -1 \\ 9 \\ 8 \end{pmatrix}$.

Portanto, $v = \begin{pmatrix} -1 \\ 9 \\ 8 \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \\ 6 \end{pmatrix} = \begin{pmatrix} -3 \\ 10 \\ 2 \end{pmatrix}$ é vetor diretor da reta dada.

A projeção P deve pertencer à reta, portanto deve satisfazer as seguinte equação:

$$P = \begin{pmatrix} 2 \\ -1 \\ 6 \end{pmatrix} + t \begin{pmatrix} -3 \\ 10 \\ 2 \end{pmatrix} \tag{2.6}$$

para algum $t \in I\!R$. Observe que em (2.6) temos 3 equações e 4 incógnitas (as coordenadas de P e o parâmetro t).

Por outro lado, $Q - P$ deve ser ortogonal ao vetor diretor, ou seja,

$$\left\langle \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} - P, \begin{pmatrix} -3 \\ 10 \\ 2 \end{pmatrix} \right\rangle = 0. \tag{2.7}$$

Em (2.7) temos uma equação com três incógnitas (as coordenadas de P). Subtituindo (2.6) em (2.7) temos:

$$\left\langle \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \\ 6 \end{pmatrix} - t \begin{pmatrix} -3 \\ 10 \\ 2 \end{pmatrix}, \begin{pmatrix} -3 \\ 10 \\ 2 \end{pmatrix} \right\rangle = 0,$$

ou seja,

$$\left\langle \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix} - t \begin{pmatrix} -3 \\ 10 \\ 2 \end{pmatrix}, \begin{pmatrix} -3 \\ 10 \\ 2 \end{pmatrix} \right\rangle = 0.$$

Isto equivale a

$$\left\langle \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ 10 \\ 2 \end{pmatrix} \right\rangle - t \left\langle \begin{pmatrix} -3 \\ 10 \\ 2 \end{pmatrix}, \begin{pmatrix} -3 \\ 10 \\ 2 \end{pmatrix} \right\rangle = 0,$$

ou seja,

$$29 - 113t = 0.$$

Portanto, $t = 29/113$.

Agora, por (2.6),

$$P = \begin{pmatrix} 2 \\ -1 \\ 6 \end{pmatrix} + t \begin{pmatrix} -3 \\ 10 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 6 \end{pmatrix} + \frac{29}{113} \begin{pmatrix} -3 \\ 10 \\ 2 \end{pmatrix}$$
$$= \begin{pmatrix} 2 - 3 \times 29/113 \\ -1 + 290/113 \\ 6 + 2 \times 29/113 \end{pmatrix}.$$

Portanto, a distância entre o ponto Q e a reta dada é:

$$\|P - Q\| = \sqrt{(2.77 - 3)^2 + (-3.566 - 2)^2 + (5.447 - 5)^2} = 5.589.$$

Evidentemente, é possível deduzir uma fórmula para a distância entre ponto e reta, para a projeção e assim por diante. Se decorarmos a fórmula isso certamente aumentará a velocidade de nossos cálculos. Entretanto, a esta altura de nossa aprendizagem, esse procedimento não é recomendável.

2.5 Distância entre duas retas

A distância entre duas retas $\mathcal{L}_1$ e $\mathcal{L}_2$ de $I\!R^n$ se define como a menor distância possível entre pontos de uma e outra. Veja Figura 2.12. Essa menor distância é realizada por um ponto $P \in \mathcal{L}_1$ e um ponto $Q \in \mathcal{L}_2$. O vetor $P - Q$ deve ser ortogonal aos vetores diretores de $\mathcal{L}_1$ e $\mathcal{L}_2$. Colocando as equações adequadamente, isto nos dá todos as informações necessárias para encontrar a distância entre as duas retas.

Suponhamos que $\mathcal{L}_1 \subset I\!R^4$ seja a reta que passa por $\begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix}$ e tem $\begin{pmatrix} 2 \\ 1 \\ 3 \\ -1 \end{pmatrix}$ como vetor diretor e que $\mathcal{L}_2$ seja a reta que passa por $\begin{pmatrix} 2 \\ -1 \\ -1 \\ -1 \end{pmatrix}$ e tem $\begin{pmatrix} 1 \\ 4 \\ 5 \\ -6 \end{pmatrix}$ como vetor diretor.

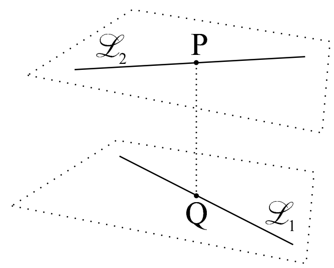

Figura 2.12. **Distância entre retas.**

Por tanto, como P deve pertencer a $\mathcal{L}_1$, existe $s \in I\!R$ tal que

$$P = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix} + s \begin{pmatrix} 2 \\ 1 \\ 3 \\ -1 \end{pmatrix}. \tag{2.8}$$

Como Q deve pertencer a $\mathcal{L}_2$, existe $t \in I\!R$ tal que:

$$Q = \begin{pmatrix} 2 \\ -1 \\ -1 \\ -1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 4 \\ 5 \\ -6 \end{pmatrix}. \tag{2.9}$$

Logo

$$P - Q = \begin{pmatrix} -1 + 2s - t \\ 3 + s - 4t \\ 3s - 5t \\ 2 - s - 6t \end{pmatrix}. \tag{2.10}$$

Este vetor deve ser ortogonal aos vetores diretores das duas retas, portanto:

$$\left\langle \begin{pmatrix} 2 \\ 1 \\ 3 \\ -1 \end{pmatrix}, P - Q \right\rangle = 0, \quad \left\langle \begin{pmatrix} 1 \\ 4 \\ 5 \\ -6 \end{pmatrix}, P - Q \right\rangle = 0.$$

Aqui temos um sistema de duas equações com duas incógnitas (s e t). Resolvendo esse sistema teremos a solução do nosso problema.

2.6 Planos

Sejam $P, v, w \in I\!R^n$. Vamos supor que os vetores v e w são *linearmente independentes*, o que significa que nenhum deles é múltiplo do outro, em particular, nenhum deles é zero. (Mais tarde, este conceito será definido com mais generalidade.) Definimos o *plano que passa por* P e tem v e w como *vetores diretores* como o conjunto de pontos x de $I\!R^n$ tais que $x = P + sv + tw$ para algum par de escalares $s, t \in I\!R$. Veja Figura 2.13.

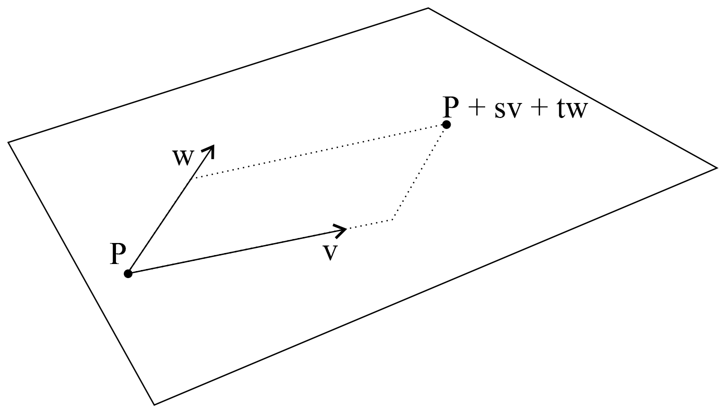

Figura 2.13. **Planos.**

Notemos que, se A, B e C são três pontos diferentes de um plano, tais que $B - A$ e $C - A$ são independentes, então necessariamente esse plano é o plano que passa por A e tem $B - A$ e $C - A$ como vetores diretores.

Também é fácil ver que, se v e w são vetores diretores de um plano, então μv e σw também são vetores diretores para quaisquer $\mu, \sigma \neq 0$.

Suponhamos que o conjunto de soluções de um sistema linear tem dois graus de liberdade. Por exemplo, suponhamos que a solução geral é

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} + s \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix} + t \begin{pmatrix} -1 \\ 3 \\ 4 \\ 4 \end{pmatrix}.$$

Pela definição de plano, vemos que este conjunto de soluções representa um plano em $I\!R^n$ ($n = 4$ no exemplo).

2.7 Distância entre ponto e plano

Consideremos o problema de encontrar a distância entre um ponto Q e um plano $\mathcal{P}$. Veja Figura 2.14.

Essa distância se define como a distância entre o ponto Q e aquele ponto P do plano que fica mais perto de Q. O ponto P é chamado *projeção* de Q no plano $\mathcal{P}$. Como vemos, P deve satisfazer duas condições:

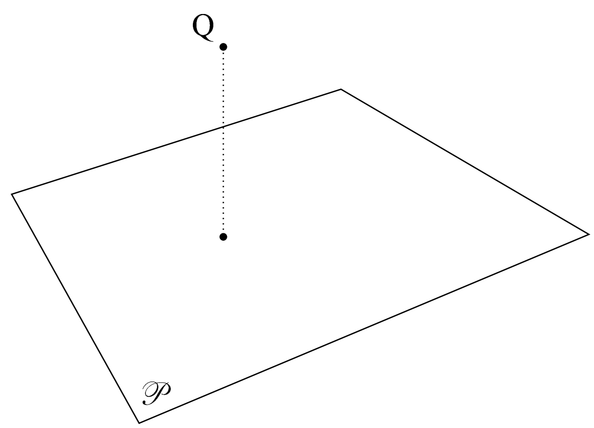

Figura 2.14. **Distância entre ponto e plano.**

(a) P deve pertencer ao plano $\mathcal{P}$;

(b) $Q - P$ deve ser ortogonal aos vetores diretores do plano.

A resolução do problema passa por colocar essas duas condições em forma de equações e por resolver esse sistema.

Por exemplo, consideremos o problema de encontrar a distância entre $Q = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} \in I\!R^3$ e o plano que passa por $\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$, $\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$ e $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Portanto, $v = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} -3 \\ 2 \\ -2 \end{pmatrix}$ e $w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix}$ são vetores diretores do plano.

A projeção P deve pertencer ao plano, portanto deve satisfazer as seguinte equação:

$$P = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + s \begin{pmatrix} -3 \\ 2 \\ -2 \end{pmatrix} + t \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix} \tag{2.11}$$

para algum par de valores $s, t \in I\!R$. Observe que, na realidade, temos aqui 3 equações e 5 incógnitas (as coordenadas de P e os parâmetros s e t).

Por outro lado, $Q - P$ deve ser ortogonal aos vetores diretores, ou seja,

$$\left\langle \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} - P, \begin{pmatrix} -3 \\ 2 \\ -2 \end{pmatrix} \right\rangle = 0. \tag{2.12}$$

e

$$\left\langle \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} - P, \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix} \right\rangle = 0. \tag{2.13}$$

Em (2.12-2.13) temos duas equação com três incógnitas (as coordenadas de P). Substituindo (2.11) em (2.12) e (2.13) obtemos um sistema de duas equações com duas incógnitas. Resolvendo este sistema obtemos s e t. Usando os valores obtidos de s e t em (2.11), a projeção P é calculada. A distância entre o ponto e o plano é $\|P - Q\|$.

De novo, é possível deduzir uma fórmula para a distância entre ponto e reta, para a projeção e assim por diante mas tal coisa seria contraproduzente a esta altura dos acontecimentos.

2.8 Combinação linear e independência linear

Dizemos que o vetor $v \in I\!R^n$ é combinação linear de $v_1, \ldots, v_m \in I\!R^n$ se existem escalares $t_1, \ldots, t_m$ tais que

$$v = t_1 v_1 + \ldots + t_m v_m.$$

Dado um conjunto $\{v_1, \ldots, v_m\} \subset I\!R^n$, se algum deles é combinação linear dos outros ou se o único elemento do conjunto é o vetor nulo, dizemos que eles são *linearmente dependentes*. Em caso contrário, dizemos que $v_1, \ldots, v_m$ são *linearmente independentes*. Veja Figura 2.15.

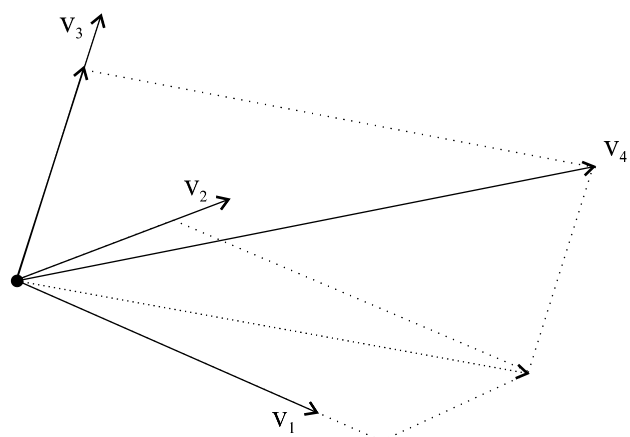

Figura 2.15. **Dependência linear.**

A caracterização mais usual de vetores linearmente independentes vem dada no seguinte teorema.

Teorema 2.8 *As seguintes afirmações são equivalentes:*

(a) $\{v_1, \ldots, v_m\}$ *são linearmente dependentes.*

(b) Existem escalares $t_1, \ldots, t_m$, não todos nulos, tais que

$$t_1 v_1 + \ldots + t_m v_m = 0.$$

Corolário 2.8 *As seguintes afirmações são equivalentes:*
(a) $\{v_1, \ldots, v_m\}$ são linearmente independentes.
(b) Se $t_1, \ldots, t_m$ são tais que

$$t_1 v_1 + \ldots + t_m v_m = 0,$$

então, necessariamente, $t_1 = \ldots = t_m = 0$.

As provas do Teorema 2.8 e o Corolário 2.8 ficam como exercício para o leitor.

2.9 Bases e coordenadas

Vejamos primeiro que, em $I\!R^n$, não pode haver um conjunto linearmente independente com mais de n elementos.

Lema 2.9. *Se $\{v_1, \ldots, v_p\} \subset I\!R^n$ e $p > n$ então os vetores $v_1, \ldots, v_p$ são linearmente dependentes.*

Prova. Observe que

$$v_1 x_1 + \ldots + v_p x_p = 0 \tag{2.14}$$

é um sistema linear com n equações (porque $v_j \in I\!R^n$ para todo j) e p incógnitas $(x_1, \ldots, x_p)$. Submetamos este sistema ao processo de escalonamento.

Se, depois de escalonar o sistema, resulta que todas as equações são $0 = 0$ então valores arbitrários de $x_1, \ldots, x_p$ não-nulos são soluções de (2.14). Logo, $v_1, \ldots, v_p$ são linearmente dependentes neste caso.

Suponhamos, eliminando as equações $0 = 0$, que o sistema escalonado ficou com q equações. Certamente, $1 \leq q \leq n < p$.

Para cada $i = 1, \ldots, q-1$ seja z_i uma incógnita que aparece na equação i e não aparece em nenhuma das seguintes e seja z_q uma incógnita que aparece na equação q e não está no conjunto de incógnitas $z_1, \ldots, z_{q-1}$. (Tal z_q existe porque, em caso contrário a equação q seria $0 = 0$.) Como $q < p$ o conjunto de incógnitas não se esgota em $z_1, \ldots, z_q$. Seja então z uma incógnita que não está em $\{z_1, \ldots, z_q\}$. Dando um valor arbitrário náo-nulo a esta incógnita e valores arbitrários a todas as outras incógnitas diferentes de $z_1, \ldots, z_q$ podemos colocar em evidência, sucessivamente, $z_{q-1}, \ldots, z_1$ obtendo assim uma solução não nula do sistema escalonado e, portanto, de (2.14). Isto implica que $v_1, \ldots, v_p$ são linearmente dependentes.

Teorema 2.9. *As seguintes proposições são equivalentes:*

(a) $\{v_1, \ldots, v_n\} \subset \mathbb{R}^n$ *é um conjunto linearmente independente.*

(b) Todo vetor $v \in \mathbb{R}^n$ *pode ser escrito de maneira única como combinação linear de* $v_1, \ldots, v_n$.

Prova. Vejamos primeiro que (a) implica (b).

Seja $v \in \mathbb{R}^n$. Consideremos o conjunto $\{v_1, \ldots, v_n, v\}$. Pelo Lema 2.9, este conjunto é linearmente dependente. Portanto, existem $t_1, \ldots, t_n, t \in \mathbb{R}$, não todos nulos, tais que

$$t_1 v_1 + \ldots + t_n v_n + tv = 0. \tag{2.15}$$

Agora, t deve ser diferente de zero, pois se fosse zero teriamos

$$t_1 v_1 + \ldots + t_n v_n = 0$$

com algum t_i diferente de zero, o que é impossível pela independência linear de $\{v_1, \ldots, v_n\}$. Logo, por (2.15),

$$v = -(t_1/t)v_1 - \ldots - (t_n/t)v_n.$$

Portanto, v pode ser escrito como combinação linear de $v_1, \ldots, v_n$. Vejamos que v não pode ser combinação linear de $v_1, \ldots, v_n$ de duas maneiras diferentes. Com efeito, se

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n = \mu_1 v_1 + \ldots + \mu_n v_n,$$

deduzimos que

$$(\lambda_1 - \mu_1)v_1 + \ldots + (\lambda_n - \mu_n)v_n = 0.$$

Logo, pela independência linear de $v_1, \ldots, v_n$,

$$\lambda_1 = \mu_1, \ldots, \lambda_n = \mu_n.$$

Com isto, fica provado que (a) implica (b).

Agora provemos que (b) implica (a). Se todo vetor $v \in \mathbb{R}^n$ pode ser escrito de maneira única como combinação linear de $v_1, \ldots, v_n$ então, em particular, isso se aplica a $v = 0$. Logo, se

$$0 = t_1 v_1 + \ldots + t_n v_n.$$

os coeficientes t_i devem ser todos nulos. Portanto, $v_1, \ldots, v_n$ são linearmente independentes.

Um conjunto de n vetores linearmente independentes de $\mathbb{R}^n$ se denomina uma *base*. Vimos que, se $v \in \mathbb{R}^n$ e $\{v_1, \ldots, v_n\}$ é uma base, então v se escreve como $t_1 v_1 + \ldots + t_n v_n$ de maneira única. Esses coeficientes $t_1, \ldots, t_n$ se chamam *coordenadas* de v na base $v_1, \ldots, v_n\}$. .

2.10 Variedades afins

Retas e planos são casos particulares de objetos mais gerais chamados *variedades afins*. De fato, as retas são "variedades afins de dimensão 1" e os planos são "variedades afins de dimensão 2". Dado um ponto $P \in I\!R^n$ e m vetores $v_1, \ldots, v_m \in I\!R^n$, a variedade afim que passa por P e é gerada por $v_1, \ldots, v_m$ é o conjunto de pontos de $I\!R^m$ da forma

$$P + t_1 v_1 + \ldots + t_m v_m \tag{2.16}$$

com $t_1, \ldots, t_m \in I\!R$. Veja Figura 2.16.

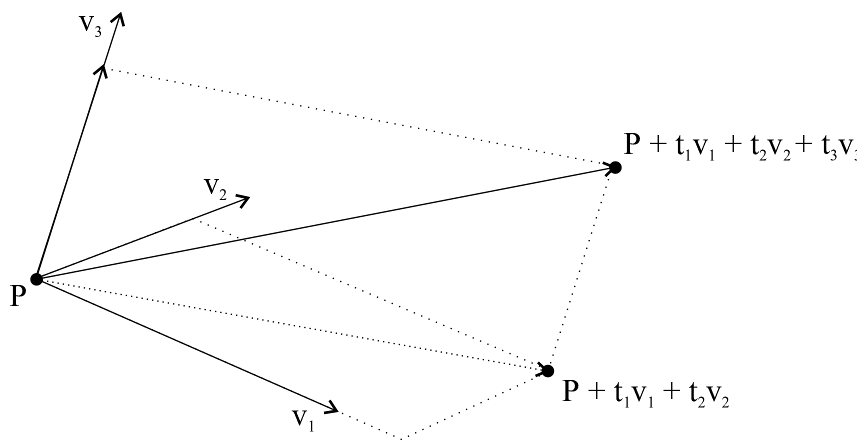

Figura 2.16. **Variedade afim.**

Se P é a origem a variedade afim se diz um *subespaço*.

Dada a variedade afim definida por (2.16) chamamos *subespaço associado* ao conjunto de pontos de $I\!R^n$ da forma $t_1 v_1 + \ldots + t_m v_m$, com $t_1, \ldots, t_m \in I\!R$.

Se $v_1, \ldots, v_m$ são linearmente independentes, dizemos que a variedade afim tem dimensão m. Nesse caso, dizemos que $v_1, \ldots, v_m$ são vetores diretores da variedade afim.

Se $A_0, \ldots, A_m$ são $m+1$ pontos diferentes de uma variedade afim de dimensão m tais que $A_1 - A_0, \ldots, A_m - A_0$ são linearmente independentes, então necessariamente essa variedade afim é a variedade afim que passa por A_0 e tem $A_1 - A_0, \ldots, A_m - A_0$ como vetores diretores.

Também é fácil ver que, se $v_1, \ldots, v_m$ são vetores diretores de uma variedade afim, e $\mu_1, \ldots, \mu_m$ são escalares não nulos, então $\mu_1 v_1, \ldots, \mu_m v_m$ também são vetores diretores.

Suponhamos que o conjunto de soluções de um sistema linear tem m graus de liberdade. Por exemplo, suponhamos que a solução geral é

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} + s \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \\ 2 \end{pmatrix} + t \begin{pmatrix} -1 \\ 3 \\ 4 \\ 4 \\ 21 \end{pmatrix} + r \begin{pmatrix} 100 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Então, pela definição de variedade afim, vemos que este conjunto de soluções representa uma variedade afim em $I\!R^n$ ($n = 5$ e $m = 3$ no exemplo).

2.11 Vetores normais

Se um vetor $v \in I\!R^n$ é ortogonal a todos os vetores diretores de uma variedade afim, dizemos que v é ortogonal ou *normal* à variedade afim. Isto é equivalente a dizer que v é ortogonal a todos os elementos do subespaço associado. Veja Figura 2.17.

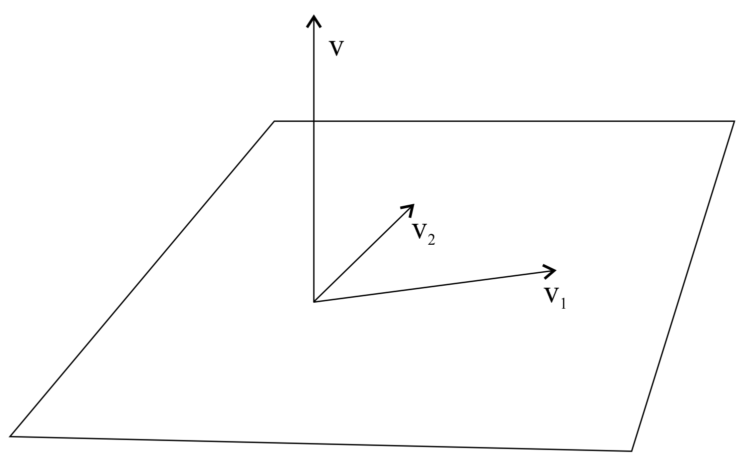

Figura 2.17. **Vetor normal a uma variedade afim.**

Portanto, se $v_1, \ldots, v_m$ são os vetores diretores de uma variedade afim, então o conjunto de todos os vetores normais é o conjunto de soluções do seguinte sistema linear:

$$\langle v_1, x\rangle = 0, \ldots, \langle v_m, x\rangle = 0.$$

Já vimos que o conjunto de soluções de um sistema linear é uma variedade afim. Quando a origem é uma solução esta variedade afim é um subespaço. Portanto, o conjunto de vetores normais a uma variedade afim formam um subespaço. Veja Figura 2.18.

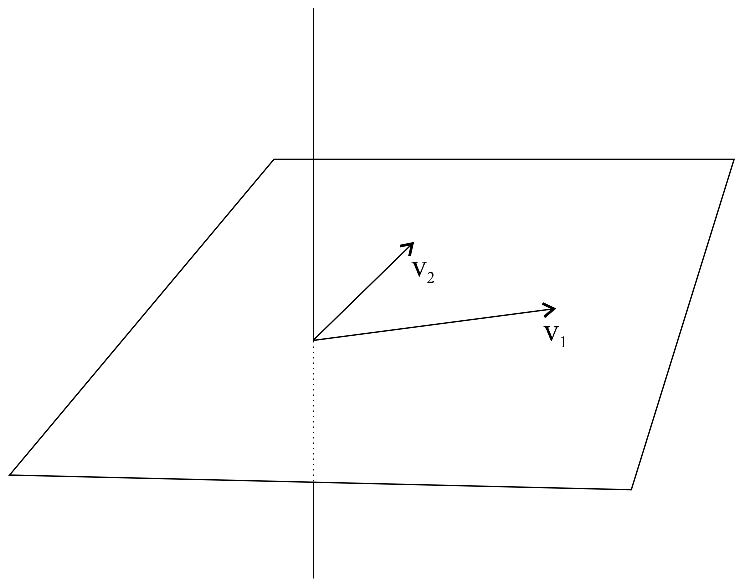

Figura 2.18. **Recta ortogonal a uma variedade afim.**

Por exemplo, se $\mathcal{V}$ é a variedade afim que passa por $\begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \end{pmatrix}$ e tem como diretores os vetores $\begin{pmatrix} 3 \\ 3 \\ 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}$ e $\begin{pmatrix} 0 \\ 0 \\ 2 \\ 2 \end{pmatrix}$ então os vetores normais são os vetores $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ tais que

$$3x_1 + 3x_2 + x_3 + x_4 = 0$$

$$x_2 + 2x_3 + 3x_4 = 0$$

$$2x_3 + 2x_4 = 0.$$

Neste exemplo, o conjunto dos vetores normais é uma reta em $I\!R^4$. Com efeito, se resolvermos este sistema linear e obtermos sua solução geral, veremos que há somente um grau de liberdade e o conjunto de soluções pode ser expressado como uma reta.

Frequentemente nos depararemos com o problema de encontrar um vetor normal a um plano em $I\!R^3$. Por exemplo, suponhamos que os vetores diretores do plano sejam $v = \begin{pmatrix} 1 \\ 6 \\ 10 \end{pmatrix}$ e $w = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$. Portanto, os vetores normais serão as soluções das equações

$$\left\langle x, \begin{pmatrix} 1 \\ 6 \\ 10 \end{pmatrix} \right\rangle = x_1 + 6x_2 + 10x_3 = 0,$$

$$\left\langle x, \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} \right\rangle = 3x_1 - x_2 + 2x_3 = 0.$$

Substituindo a segunda equação por ela mesma menos 3 vezes a primeira, temos o sistema:

$$x_1 + 6x_2 + 10x_3 = 0,$$

$$-19x_2 - 28x_3 = 0.$$

A solução geral deste sistema é:

$$\begin{pmatrix} -(22/19)x_3 \\ -(28/19)x_3 \\ x_3 \end{pmatrix}.$$

Portanto, os vetores normais são os múltiplos do vetor

$$\begin{pmatrix} -22/19 \\ -28/19 \\ 1 \end{pmatrix}.$$

2.12 Produto vetorial

Uma forma detestável de chegar ao mesmo resultado é usar o chamado *produto vetorial*. O produto vetorial entre dos vetores v e w de $I\!R^3$ se define assim:

$$v \times w = \begin{pmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{pmatrix}.$$

No exemplo que estamos considerando,

$$v \times w = \begin{pmatrix} 22 \\ 28 \\ -19 \end{pmatrix},$$

que, naturalmente, é um múltiplo do vetor normal que já havia sido encontrado.

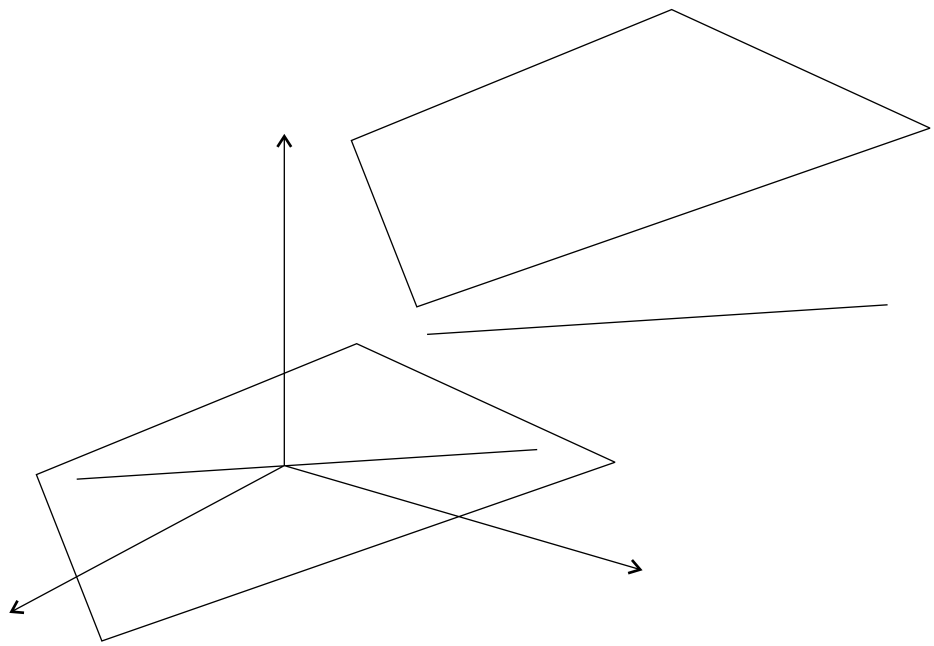

Figura 2.19. **Variedades afins paralelas.**

2.13 Paralelismo

Duas variedades afins se dizem *paralelas* se o subespaço associado de uma delas está contido no subespaço associado da outra. Veja Figura 2.19.

2.14 Exercícios

1. Prove, para todo $x, y, z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$,

$$\langle \lambda x, y \rangle = \langle x, \lambda y \rangle = \lambda \langle x, y \rangle$$

e

$$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle.$$

2. Prove a desigualdade Cauchy-Schwarz para $n = 4$ da seguinte maneira: escreva, secretamente, a desigualdade que você quer provar e desenvolva essa desigualdade até chegar a algo que você sabe ser verdade. Depois, verifique que os passos dados "na direção contrária" são corretos do ponto de vista lógico. (Isto nem sempre é assim.) Feito isto, saia da clandestinidade e escreva a prova na ordem correta.

3. Prove (2.2), (2.3), (2.4) e (2.5).

4. Calcule o ângulo entre os vetores $\begin{pmatrix}1\\2\end{pmatrix}$ e $\begin{pmatrix}3\\-1\end{pmatrix}$. Desenhe e confirme sua intuição geométrica.

5. Mostre que a definição dada de cosseno corresponde exatamente ao conceito de "cateto adjacente dividido pela hipotenusa" da Trigonometria.

6. Escreva as soluções dos sistemas lineares dados na Lista 1 como subconjuntos de $I\!R^n$.

7. Considere um sistema linear homogêneo. Prove que qualquer *combinação linear* de soluções também é solução.

8. Seja v uma solução de um sistema linear e x uma solução do sistema homogêneo correspondente. Prove que $v+w$ é uma solução do sistema linear original.

9. Justifique a definição de perpendicularidade desenvolvendo a igualdade
$$\|v + w\| = \|v - w\|.$$

10. Observe que se o ângulo entre v e w é agudo, então $\|v + w\| > \|v - w\|$. Usando isto, interprete geomêtricamente a positividade ou negatividade do produto escalar. Defina rigorosamente o cosseno do ângulo entre dois elementos de $I\!R^n$.

11. Justifique o fato de que, se P é o ponto da reta $\mathcal{L}$ más próximo de Q, então $P - Q$ deve ser ortogonal ao vetor diretor da reta. (Use Pitágoras.)

12. Calcule a distância entre o ponto $P \in I\!R^n$ e a reta $\mathcal{L}$, assim como o elemento de $\mathcal{L}$ que está mais próximo de P nos casos abaixo. Desenhe sempre.

    (a) $P = \begin{pmatrix}3\\5\end{pmatrix}$; $\mathcal{L}$ é a reta que passa por $\begin{pmatrix}1\\1\end{pmatrix}$ cujo vetor diretor é $\begin{pmatrix}3\\3\end{pmatrix}$.

    (b) $P = \begin{pmatrix}3\\5\end{pmatrix}$; $\mathcal{L}$ é a reta que passa por $\begin{pmatrix}1\\1\end{pmatrix}$ cujo *vetor normal* é $\begin{pmatrix}-1\\2\end{pmatrix}$.

    (c) $P = \begin{pmatrix}2\\10\end{pmatrix}$; $\mathcal{L}$ é o conjunto de soluções daqueles sistemas da Lista 1 cujo conjunto de soluções seja uma reta em $I\!R^2$.

(d) $P = \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix}$; $\mathcal{L}$ a reta que passa por $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ cujo vetor diretor é $\begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix}$.

(e) $P = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$; $\mathcal{L}$ é o conjunto de soluções daqueles sistemas da Lista 1 cujo conjunto de soluções seja uma reta em $\mathbb{R}^3$.

13. Seja $\mathcal{L}_1$ é a reta em $\mathbb{R}^3$ que passa por $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ e cujo vetor diretor é $\begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix}$.
Seja $\mathcal{L}_2$ a reta em $\mathbb{R}^3$ que passa por $\begin{pmatrix} 10 \\ -1 \\ 3 \end{pmatrix}$ e cujo vetor diretor é $\begin{pmatrix} 1 \\ 3 \\ 7 \end{pmatrix}$.
Calcule a distância entre essas retas e os pontos que *realizam* essa distância. Interprete geometricamente e desenhe.

14. Identifique todos os casos em que o conjunto de soluções de um sistema da Lista 1 é uma reta em $\mathbb{R}^3$. Para cada par desses casos, seja $\mathcal{L}_1$ o conjunto de soluções de um sistema e $\mathcal{L}_2$ o conjunto de soluções do outro. Calcule a distância entre essas retas e os pontos que *realizam* essa distância. Interprete geometricamente e desenhe.

15. Justifique o fato de que, se $P \in \mathcal{L}_1$ e $Q \in \mathcal{L}_2$ são os pontos que realizam a distância entre as retas $\mathcal{L}_1$ e $\mathcal{L}_2$, então o vetor $P - Q$ deve ser ortogonal aos vetores diretores de $\mathcal{L}_1$ e $\mathcal{L}_2$.

16. Invente um exercício do tipo do Exercício 8 onde haja infinitos pontos que realizam a distância entre as duas retas. Interprete geometricamente e desenhe.

17. Seja $P = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ e seja $\mathcal{P}$ o plano que passa por $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ e cujos vetores diretores são $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ e $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$. Calcule a distância entre P e $\mathcal{P}$ e o ponto do plano que realiza essa distância.

18. Justifique o fato de que, se P é o ponto do plano $\mathcal{P}$ más próximo de Q, então $P - Q$ deve ser ortogonal aos vetores diretores do plano. (Use Pitágoras.)

19. Seja $P = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ e seja $\mathcal{P}$ o plano que representa o conjunto de soluções de cada um dos sistemas da Lista 1 cujo conjunto de soluções é um plano em $\mathbb{R}^3$. Calcule a distância entre P e $\mathcal{P}$ e o ponto do plano que realiza essa distância.

20. Seja $P = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ e seja $\mathcal{P}$ o plano que passa por $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ e cujo *vetor normal* é $\begin{pmatrix} 2 \\ 4 \\ 7 \end{pmatrix}$. Calcule a distância entre P e $\mathcal{P}$ e o ponto do plano que realiza essa distância.

21. Seja $\mathcal{L}$ a reta que passa por $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ e cujo vetor diretor é $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$. Calcule a reta paralela a $\mathcal{L}$ que passa por $\begin{pmatrix} 6 \\ -2 \end{pmatrix}$. Declare qual é o vetor diretor dessa reta e qual é o vetor normal. Desenhe tudo.

22. Seja $\mathcal{L}$ a reta que passa por $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ e cujo vetor diretor é $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$. Calcule a reta perpendicular a $\mathcal{L}$ que passa por $\begin{pmatrix} 6 \\ -2 \end{pmatrix}$. Declare qual é o vetor diretor dessa reta e qual é o vetor normal. Desenhe tudo.

23. Seja $\mathcal{L}$ a reta que passa por $\begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}$ e cujo vetor diretor é $\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}$. Calcule uma reta paralela a $\mathcal{L}$ que passa por $\begin{pmatrix} 6 \\ -2 \\ 1 \end{pmatrix}$. Declare qual é o vetor diretor dessa reta e defina dois vetores normais independentes (um deles não deve ser múltiplo do outro). Desenhe tudo.

24. Seja $\mathcal{L}$ a reta que passa por $\begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}$ e cujo vetor diretor é $\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$. Calcule a reta perpendicular a $\mathcal{L}$ que passa por $\begin{pmatrix} 6 \\ -2 \\ 2 \end{pmatrix}$. Declare qual é o vetor diretor dessa reta e defina dois vetores normais independentes. Desenhe tudo.

25. Seja $\mathcal{L}$ a reta que passa por $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ e cujo vetor normal é $\begin{pmatrix} 3 \\ -4 \end{pmatrix}$. Calcule a reta paralela a $\mathcal{L}$ que passa por $\begin{pmatrix} 6 \\ -2 \end{pmatrix}$. Declare qual é o vetor diretor dessa reta e qual é o vetor normal. Desenhe tudo.

26. Seja $\mathcal{L}$ a reta que passa por $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ e cujo vetor normal é $\begin{pmatrix} -2 \\ -1 \end{pmatrix}$. Calcule a reta perpendicular a $\mathcal{L}$ que passa por $\begin{pmatrix} 6 \\ -2 \end{pmatrix}$. Declare qual é o vetor diretor dessa reta e qual é o vetor normal. Desenhe tudo.

27. Seja $\mathcal{L}$ a reta que passa por $\begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}$ e cujos vetores normais independentes são $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ e $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. Calcule uma reta paralela a $\mathcal{L}$ que passa por $\begin{pmatrix} 6 \\ -2 \\ 1 \end{pmatrix}$. Declare qual é o vetor diretor dessa reta e defina dois vetores normais independentes (um deles não deve ser múltiplo do outro). Desenhe tudo.

28. Seja $\mathcal{L}$ a reta que passa por $\begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}$ e cujos vetores normais independentes são $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ e $\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$. Calcule a reta perpendicular a $\mathcal{L}$ que passa por $\begin{pmatrix} 6 \\ -2 \\ 2 \end{pmatrix}$. Declare qual é o vetor diretor dessa reta e defina dois vetores normais independentes. Desenhe tudo.

29. Seja $\mathcal{P}$ o plano que passa por $\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$ e cujos vetores diretores independentes

são $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ e $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$. Calcule a reta perpendicular a este plano que passa por $\begin{pmatrix} 2 \\ 1 \\ 10 \end{pmatrix}$. Desenhe.

30. Seja $\mathcal{P}$ o plano que passa por $\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$ e cujo vetor normal é $\begin{pmatrix} 1 \\ 10 \\ -1 \end{pmatrix}$. Calcule a reta perpendicular a este plano que passa por $\begin{pmatrix} 2 \\ 1 \\ 10 \end{pmatrix}$. Desenhe.

31. Justifique o fato de que, se P é o ponto da variedade afim $\mathcal{V}$ más próximo de Q, então $P - Q$ deve ser ortogonal aos vetores diretores da variedade afim. (Use Pitágoras.)

32. Interprete geometricamente os seguintes fatos:

    (a) Um sistema linear de 2×2 tem uma única solução.

    (b) Um sistema linear de 2×2 tem infinitas soluções.

    (c) Um sistema linear de 2×2 não tem nenhuma solução.

    (d) Um sistema linear de 3×3 tem uma única solução.

    (e) As duas primeiras equações de um sistema linear de 3×3 não tem nenhuma solução, mas tanto a primeira com a terceira como a segunda com a terceira tem infinitas soluções.

33. Prove o Teorema 2.8 e o Corolário 2.8.

34. Chamamos *semiespaço* de $\mathbb{R}^n$ a um subconjunto do tipo

$$\{x \in \mathbb{R}^n \mid \langle v, x\rangle \leq c\}.$$

    Interprete geometricamente esta definição. Chamamos *poliedro* a uma intersecção de um número finito de subespacos. Invente um algoritmo para encontrar o ponto de um poliedro más próximo de um ponto dado de $\mathbb{R}^3$. Seu algoritmo deve servir, também, para detectar se o poliedro dado é vazio.

Capítulo 3

Matrizes

Este é o capítulo mais genuinamente "de álgebra linear" destas notas. Do ponto de vista lógico, seu objetivo é chegar de maneira razoavelmente rápida, porém refletida, fatoração QDQ^T de matrizes simêtricas. A decomposição em valores singulares, que vem depois, está certamente fora de qualquer curso básico de Geometria Analítica, mas, ficando tão perto e sendo tão útil, achamos razoável incluí-la no texto.

A decomposição de matrizes simêtricas será usada como ferramenta fundamental para a classificação de equações quadráticas. Cônicas e quádricas podem ser classificadas sem o auxílio dessa técnica, mas essa opção costuma ser excessivamente particular (restrita a 2 ou 3 dimensões) e tediosa.

3.1 Definição e operações básicas

Uma matriz de m linhas e n colunas (simplificadamente, de $m \times n$) é um arranjo de mn números dispostos em forma de "retângulo", da seguinte maneira:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \tag{3.1}$$

Freqüentemente, quando uma matriz A venha dada por (3.1), diremos que a_{ij} é o elemento genêrico de A.

O conjunto das matrizes de m linhas e n colunas é denotado $\mathbb{R}^{m\times n}$. Se $A \in \mathbb{R}^{m\times n}$ vem dada por (3.1) chamamos *Transposta de A* (em símbolos A^T) à matriz de n linhas e m colunas dada por:

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}.$$

Por exemplo, se

$$A = \begin{pmatrix} 1 & 3 & 5 & -1 \\ 0 & 3 & 6 & -2 \end{pmatrix}$$

temos:

$$A^T = \begin{pmatrix} 1 & 0 \\ 3 & 3 \\ 5 & 6 \\ -1 & -2 \end{pmatrix}.$$

Uma matriz de $n \times n$ se diz *simétrica* se $A = A^T$.

A soma de matrizes da mesma ordem se define componente a componente:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}.$$

Também definimos o produto de uma matriz por um número real da seguinte forma:

$$t \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} ta_{11} & \cdots & ta_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ ta_{m1} & \cdots & ta_{mn} \end{pmatrix}.$$

Se $A \in \mathbb{R}^{m \times n}$ e $x \in \mathbb{R}^n$ o produto Ax se define como sendo o seguinte vetor de $\mathbb{R}^n$:

$$Ax = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \cdot \\ \cdot \\ \cdot \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix}. \tag{3.2}$$

É fácil verificar que, para toda matriz $A \in I\!R^{m\times n}$, $x, y \in I\!R^n$, $\alpha, \beta \in I\!R$,

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay. \tag{3.3}$$

O produto AB entre uma matriz $A \in I\!R^{m\times n}$ e uma matriz $B \in I\!R^{n\times p}$ se define como sendo a matriz de $m \times p$ dada por:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1p} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ b_{n1} & \cdots & b_{np} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11} + \ldots + a_{1n}b_{n1} & \cdots & a_{11}b_{1p} + \ldots + a_{1n}b_{np} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ a_{m1}b_{11} + \ldots + a_{mn}b_{n1} & \cdots & a_{m1}b_{1p} + \ldots + a_{mn}b_{np} \end{pmatrix}.$$

Em outras palavras, se a_{ij} é o elemento genêrico de $A \in I\!R^{m\times n}$, b_{ij} é o elemento genêrico de $B \in I\!R^{n\times p}$, e c_{ij} é o elemento genêrico do produto $C = AB \in I\!R^{m\times p}$, temos:

$$c_{ij} = a_{i1}b_{1j} + \ldots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

para todo $i = 1, \ldots, m$, $j = 1, \ldots, p$. É importante destacar que o produto de matrizes AB está definido somente se o número de colunas de A é igual ao número de linhas de B.

Observe que, se $A \in I\!R^{m\times n}$, $x \in I\!R^{n\times 1}$, e $b = Ax$ temos que $b \in I\!R^{m\times 1}$ e:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_{11} \\ . \\ . \\ . \\ x_{n1} \end{pmatrix} = \begin{pmatrix} a_{11}x_{11} + \ldots + a_{1n}x_{n1} \\ . \\ . \\ . \\ a_{m1}x_{11} + \ldots + a_{mn}x_{n1} \end{pmatrix}.$$

Devido a esta definição, e à definição do produto entre uma matriz e um elemento de $I\!R^n$, para todos os efeitos práticos o conjunto das matrizes com n linhas e uma coluna ($I\!R^{n\times 1}$) pode ser identificado com $I\!R^n$. (Daí a utilidade de escrever os elementos de $I\!R^n$ como colunas.)

3.2 Operações em blocos

Consideremos o seguinte exemplo. Seja $A \in I\!R^{7\times 4}$,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \\ a_{61} & a_{62} & a_{63} & a_{64} \\ a_{71} & a_{72} & a_{73} & a_{74} \end{pmatrix}.$$

Definindo

$$A_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \in I\!R^{3\times 2},$$

$$A_{12} = \begin{pmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \\ a_{33} & a_{34} \end{pmatrix} \in I\!R^{3\times 2},$$

$$A_{21} = \begin{pmatrix} a_{41} & a_{42} \\ a_{51} & a_{52} \end{pmatrix} \in I\!R^{2\times 2},$$

$$A_{22} = \begin{pmatrix} a_{43} & a_{44} \\ a_{53} & a_{54} \end{pmatrix} \in I\!R^{2\times 2},$$

$$A_{31} = \begin{pmatrix} a_{61} & a_{62} \\ a_{71} & a_{72} \end{pmatrix} \in I\!R^{2\times 2},$$

$$A_{32} = \begin{pmatrix} a_{63} & a_{64} \\ a_{76} & a_{74} \end{pmatrix} \in I\!R^{2\times 2},$$

podemos escrever:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix}. \tag{3.4}$$

Este tipo de notação pode ser generalizada para matrizes de qualquer tamanho. O mais interessante é que as operações entre matrizes formuladas em termos de blocos, como em (3.4) são, formalmente, às operações definidas na Seção 3.1. Con efeito, suponhamos que

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1N} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ A_{M1} & \cdots & A_{MN} \end{pmatrix}, \tag{3.5}$$

onde cada entrada A_{ij} é uma matriz, o número de linhas de todas as matrizes $A_{ij}, j = 1, \ldots, N$ (deixando fixo i) é o mesmo e o número de colunas das matrizes $A_{ij}, i = 1, \ldots, M$ (deixando fixo j) é o mesmo.

Pode-se verificar, então, que

$$A^T = \begin{pmatrix} A_{11}^T & \cdots & A_{1M}^T \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ A_{N1}^T & \cdots & A_{NM}^T \end{pmatrix}.$$

Ainda mais, se

$$B = \begin{pmatrix} B_{11} & \cdots & B_{1N} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ B_{M1} & \cdots & B_{MN} \end{pmatrix},$$

onde cada matriz B_{ij} tem a mesma ordem que A_{ij}, é fácil ver que:

$$A + B = \begin{pmatrix} A_{11} & \cdots & A_{1N} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ A_{M1} & \cdots & A_{MN} \end{pmatrix} + \begin{pmatrix} B_{11} & \cdots & B_{1N} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ B_{M1} & \cdots & B_{MN} \end{pmatrix}$$

$$= \begin{pmatrix} A_{11} + B_{11} & \cdots & A_{1N} + B_{1N} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ A_{M1} + b_{M1} & \cdots & A_{MN} + B_{MN} \end{pmatrix}.$$

Também é trivial verificar que, para todo $t \in I\!R$,

$$tA = t\begin{pmatrix} A_{11} & \cdots & A_{1N} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ A_{M1} & \cdots & A_{MN} \end{pmatrix} = \begin{pmatrix} tA_{11} & \cdots & tA_{1N} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ tA_{M1} & \cdots & tA_{MN} \end{pmatrix}.$$

O mais interessante é analisar o produto entre uma matriz $A \in I\!R^{m\times n}$ e uma matriz $B \in I\!R^{n\times p}$ que, escrita em blocos, é:

$$B = \begin{pmatrix} B_{11} & \cdots & B_{1P} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ B_{N1} & \cdots & B_{NP} \end{pmatrix}$$

Suponhamos que para todo $i = 1, \ldots, M$, para todo $k = 1, \ldots, N$ e para todo $j = 1, \ldots, P$, o número de colunas de A_{ik} é igual ao número de linhas de B_{kj}. Nesse caso, os produtos $A_{ik}B_{kj}$ podem ser efetuados e:

$$AB = \begin{pmatrix} A_{11}B_{11} + \ldots + A_{1N}B_{N1} & \cdots & A_{11}B_{1P} + \ldots + A_{1n}B_{nP} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ A_{M1}B_{11} + \ldots + A_{MN}B_{N1} & \cdots & A_{M1}B_{1P} + \ldots + A_{MN}B_{NP} \end{pmatrix}. \tag{3.6}$$

Esta propriedade pode ser rigorosamente demonstrada, o que não é difícil, mas é certamente tedioso. Entretanto, sua verificação em casos particulares é instrutivo e recomendável.

3.3 Matrizes e sistemas lineares

Todo sistema linear de equações pode ser escrito na forma $Ax = b$. Com efeito, observe que

$$a_{i1}x_1 + \ldots + a_{in}x_n = b_i \; \forall \; i = 1, \ldots, m$$

é equivalente a:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{pmatrix}.$$

Por sua vez, isto é o mesmo que:

$$\begin{pmatrix} a_{11} \\ \cdot \\ \cdot \\ \cdot \\ a_{m1} \end{pmatrix} x_1 + \ldots + \begin{pmatrix} a_{1n} \\ \cdot \\ \cdot \\ \cdot \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{pmatrix}.$$

Esta expressão nos diz que o sistema tem solução se, e somente se, o termo independente b é combinação linear das colunas de A. Por exemplo, o sistema

$$4x_1 + 5x_2 + 2x_3 = 11,$$

$$3x_1 - x_2 - x_3 - x_4 = 0,$$

$$x_1 + 5x_2 + 6x_3 + x_4 = 13,$$

$$3x_1 - x_2 - x_3 - 2x_4 = -1,$$

$$9x_1 + 8x_2 - 10x_3 - 11x_4 = -4;$$

pode ser escrito

$$\begin{pmatrix} 4 & 5 & 2 & 0 \\ 3 & -1 & -1 & -1 \\ 1 & 5 & 6 & 1 \\ 3 & -1 & -1 & -2 \\ 9 & 8 & -10 & -11 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 11 \\ 0 \\ 13 \\ -1 \\ -4 \end{pmatrix},$$

que equivale a:

$$\begin{pmatrix} 4 \\ 3 \\ 1 \\ 3 \\ 9 \end{pmatrix} x_1 + \begin{pmatrix} 5 \\ -1 \\ 5 \\ -1 \\ 8 \end{pmatrix} x_2 + \begin{pmatrix} 2 \\ -1 \\ 6 \\ -1 \\ -10 \end{pmatrix} x_3 + \begin{pmatrix} 0 \\ -1 \\ 1 \\ -2 \\ -11 \end{pmatrix} x_4 = \begin{pmatrix} 11 \\ 0 \\ 13 \\ -1 \\ -4 \end{pmatrix}.$$

Este sistema tem solução porque $\begin{pmatrix} 11 \\ 0 \\ 13 \\ -1 \\ -4 \end{pmatrix}$ é combinação linear de

$$\begin{pmatrix} 4 \\ 3 \\ 1 \\ 3 \\ 9 \end{pmatrix}, \begin{pmatrix} 5 \\ -1 \\ 5 \\ -1 \\ 8 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 6 \\ -1 \\ -10 \end{pmatrix} \text{ e } \begin{pmatrix} 0 \\ -1 \\ 1 \\ -2 \\ -11 \end{pmatrix}.$$

De fato, observe que

$$\begin{pmatrix} 11 \\ 0 \\ 13 \\ -1 \\ -4 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 1 \\ 3 \\ 9 \end{pmatrix} + \begin{pmatrix} 5 \\ -1 \\ 5 \\ -1 \\ 8 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \\ 6 \\ -1 \\ -10 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \\ 1 \\ -2 \\ -11 \end{pmatrix},$$

Ou seja, $x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ é uma solução do sistema.

3.4 Posto

O *Posto* de uma matriz é o número máximo de colunas linearmente independentes que ela possui. Por exemplo,

(a) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 3 \\ -1 & 1 \end{pmatrix}$ têm posto 2 porque suas duas colunas são linearmente independentes (nehuma delas é combinação linear da outra).

(b) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ têm posto 1 porque $\begin{pmatrix} 2 \\ 4 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Entretanto, o conjunto formado por somente uma coluna (qualquer uma das duas) é linearmente independente.

(b) $\begin{pmatrix} 2 & 0 \\ 3 & 0 \end{pmatrix}$ têm posto 1 porque $\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0\begin{pmatrix} 2 \\ 3 \end{pmatrix}$. Entretanto, o conjunto formado por somente pela segunda coluna é linearmente independente.

(c) $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ têm posto 0 porque qualquer subconjunto de colunas contém a coluna nula.

Se

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ . & \cdots & . \\ . & \cdots & . \\ . & \cdots & . \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

é a matriz do sistema $Ax = b$, chamamos *matriz ampliada* à matriz de m linhas e $n+1$ colunas dada por:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ . & \cdots & . & . \\ . & \cdots & . & . \\ . & \cdots & . & . \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}$$

Vimos que o sistema $Ax = b$ tem solução se, e somente se, o termo independente é combinação linear das colunas da matriz A. Isto é equivalente a dizer que o posto da matriz A é igual ao posto da matriz ampliada.

Vejamos agora um teorema que, entre outras coisas, nos permite calcular o posto de uma matriz por operações de escalonamento e nos permite deduzir que o posto de A é o posto de A^T.

Teorema 3.4a. *Seja $A \in \mathbb{R}^{m\times n}$. As seguintes operações não alteram o posto de A:*

1. *Permutar linhas de A;*

2. *Permutar colunas de A;*

3. *Substituír uma linha por "ela mesma mais um múltiplo de outra";*

4. *Substituír uma coluna por "ela mesma mais um múltiplo de outra".*

Prova. O fato de que nem permutar linhas nem permutar colunas altera o posto é trivial.

Vejamos que a terceira operação mencionada na tese não altera o posto de A. Seja A_1 a matriz que resulta de substituír uma linha de A por ela mesma mais um múltiplo de outra. Pelo Teorema 1.2, os sistemas lineares

$$Ax = 0 \quad \text{e} \quad A_1 x = 0$$

tem as mesmas soluções. Isto implica que um subconjunto de colunas de A é linearmente dependente se, e somente se, o subconjunto de colunas correspondente de A_1 é linearmente dependente. Portanto, A_1 tem mesmo posto que A.

Passemos à quarta operação. Seja A_2 a matriz que resulta de substituír uma coluna de A por ela mesma mais um múltiplo de outra. Seja p o posto de A, cujas colunas são $v_1, \ldots, v_n$, ou seja:

$$A = (v_1, \ldots, v_n).$$

Sem perda de generalidade, suponhamos que $v_1, \ldots, v_p$ são linearmente independentes. Consideramos os seguintes casos:

(a) A_2 resulta de substituír v_j, uma das primeiras p colunas, por $v_j + \beta v_k$, onde v_k também é uma das primeiras p colunas. (Se $p = n$ este é o único caso possível.)

(b) A_2 resulta de substituír v_j, uma das primeiras p colunas, por $v_j + \beta v_k$, onde v_k é uma das últimas $n - p$ colunas.

(c) A_2 resulta de substituír v_j, uma das últimas $n - p$ colunas, por $v_j + \beta v_k$.

Provaremos que, em todos estes casos, o posto de A_2 é maior ou igual que o posto de A.

Consideremos o Caso (a). Sem perda de generalidade, para facilitar a notação, suponhamos que $j = 1, k = 2$. Logo, a primeira coluna de A_2 é $w_1 = v_1 + \beta v_2$. Vejamos que $w_1, v_2, \ldots, v_p$ são linearmente independentes. Com efeito, se

$$t_1 w_1 + t_2 v_2 + \ldots + t_p v_p = 0,$$

temos:

$$t_1(v_1 + \beta v_2) + t_2 v_2 + \ldots + t_p v_p = 0.$$

Logo,

$$t_1 v_1 + (t_1 \beta + t_2) v_2 + \ldots + t_p v_p = 0.$$

Como $v_1, \ldots, v_p$ são linearmente independentes, isto implica que

$$t_1 = 0,\ t_1\beta + t_2 = 0, \ldots, t_p = 0.$$

Portanto, $t_1 = t_2 = \ldots = t_p = 0$. Logo, $w_1, v_2, \ldots, v_p$ são linearmente independentes. Portanto, no Caso (a), deduzimos que o posto de A_2 é maior ou igual que o posto de A_1.

Consideremos o Caso (b). Sem perda de generalidade suponhamos que $j = 1$ e $k = n$. Portanto, v_n é uma combinação linear de $v_1, \ldots, v_p$. Ou seja,

$$v_n = s_1 v_1 + \ldots + s_p v_p. \tag{3.7}$$

Devemos considerar dois sub-casos:
(i) $s_1 \neq 0$;
(ii) $s_1 = 0$.

Vejamos que, se $s_1 \neq 0$, os vetores $v_n, v_2, \ldots, v_p$ são linearmente independentes.

Com efeito, se

$$t_n v_n + t_2 v_2 + \ldots + t_p v_p = 0,$$

temos:

$$t_n(s_1 v_1 + \ldots + s_p v_p) + t_2 v_2 + \ldots + t_p v_p = 0,$$

logo

$$t_n s_1 v_1 + (t_n s_2 + t_2) v_2 + \ldots + (t_n s_p + t_p) v_p = 0.$$

Pela independência linear de $v_1, \ldots, v_p$ e o fato de que $s_1 \neq 0$, deduzimos que $t_n = t_2 = \ldots = t_p = 0$. Portanto, os vetores $v_n, v_2, \ldots, v_p$ são linearmente independentes. Mas estes vetores continuam sendo colunas de A_2. Portanto, no sub-caso (i) também temos que o posto de $\bar{A}$ é maior ou igual que o posto de A.

Consideremos agora o sub-caso (ii). A primeira coluna de A_2 é $w_1 = v_1 + \beta v_n$. Vejamos que $w_1, v_2, \ldots, v_p$ são linearmente independentes.

Com efeito, se

$$t_1 w_1 + t_2 v_2 + \ldots + t_p v_p = 0,$$

temos:

$$t_1(v_1 + \beta v_n) + t_2 v_2 + \ldots + t_p v_p = 0.$$

Logo, por (3.7), desde que $s_1 = 0$, obtemos:

$$t_1[v_1 + \beta(s_2 v_2 + \ldots + s_p v_p)] + t_2 v_2 + \ldots + t_p v_p = 0.$$

Portanto,

$$t_1 v_1 + (t_1\beta s_2 + t_2) v_2 + \ldots + (t_1 \beta s_p + t_p) v_p) = 0.$$

Pela independência linear de $v_1, \ldots, v_p$ isto implica que $t_1 = \ldots = t_p = 0$. Logo, $w_1, v_2, \ldots, v_p$ são linearmente independentes e, mais uma vez, o posto de A_2 é maior ou igual que o posto de A.

Finalmente, consideremos o Caso (c). Neste caso, as colunas independentes $v_1, \ldots, v_p$ continuam sendo colunas de A_2, portanto o posto de A_2 é maior ou igual que o posto de A.

Assim, deduzimos que o posto de A_2 sempre é maior ou igual que o posto de A. Agora, dizer que a coluna v_j de A é substituída pela coluna $v_j + \beta v_k$ para obter a coluna j de A_2 é a mesma coisa que dizer que a coluna w_j de A_2 é substituída pelacoluna $w_j - \beta w_k$ para obter a coluna j de A. Portanto, pelos mesmos argumentos que o posto de A_2 é maior ou igual que o posto de A, temos que o posto de A é maior ou igual que o posto de A_2. Ou seja, o posto de A é igual ao posto de A_2, como queriamos provar.

Teorema 3.4b. *Seja $A \in I\!R^{m \times n}$. Realizando uma seqüência de operações do tipo das indicadas no Teorema 3.4a, a matriz A pode ser transformada em uma matriz diagonal.*

Prova. Este resultado é obviamente verdadeiro para matrizes de 1×1. A prova consiste em mostrar que, se for verdadeiro para matrizes de $m \times n$ também deve ser verdade para matrizes de $(m+1) \times n$ e matrizes de $m \times (n+1)$. Para fixar idéias, vamos provar isto para $m = 5, n = 3$. A prova geral fica como exercício para o leitor.

Consideremos uma matriz A de $(m+1) \times n = 6 \times 3$.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \\ a_{61} & a_{62} & a_{63} \end{pmatrix}.$$

Supondo que o resultado é verdade para matrizes de 5×3, temos que, aplicando uma seqüência de operações como as indicadas, a matriz A pode ser transformada em:

$$A' = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \bar{a}_{61} & \bar{a}_{62} & \bar{a}_{63} \end{pmatrix}.$$

Consideremos os seguintes casos:

1. $d_1 \neq 0, d_2 \neq 0, d_3 \neq 0$;

2. $d_1 \neq 0, d_2 \neq 0, d_3 = 0$;

3. $d_1 \neq 0, d_2 = 0, d_3 = 0$;

4. $d_1 = 0, d_2 = 0, d_3 = 0$.

Qualquer outro caso se reduz a um destes três por permutação de linhas ou colunas.

No primeiro caso, substituindo sucessivamente:

$$\text{Linha } 6 \leftarrow \ \text{Linha } 6 - \frac{\bar{a}_{61}}{d_1} \text{ Linha } 1, \tag{3.8}$$

$$\text{Linha } 6 \leftarrow \ \text{Linha } 6 - \frac{\bar{a}_{62}}{d_2} \text{ Linha } 2, \tag{3.9}$$

e

$$\text{Linha } 6 \leftarrow \ \text{Linha } 6 - \frac{\bar{a}_{63}}{d_3} \text{ Linha } 3, \tag{3.10}$$

a matriz fica transformada em

$$A'' = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

que é uma matriz diagonal.

No segundo caso, fazendo as operações (3.8) e (3.9) e permutando a linha 3 com a linha 6, também chegamos a uma matriz diagonal.

No terceiro caso, fazendo (3.8) obtemos uma matriz da forma:

$$A''' = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & a'_{62} & a'_{63} \end{pmatrix}.$$

Se $a'_{62} \neq 0$ substituimos:

$$\text{Coluna } 3 \leftarrow \ \text{Coluna } 3 - \frac{a'_{63}}{a'_{62}} \text{ Coluna } 2 \tag{3.11}$$

e permutamos a linha 6 com a linha 2.

Se $a'_{63} \neq 0$ mas $a'_{62} = 0$ permutamos a coluna 3 com a coluna 2 e fazemos a substituição (3.11).

Se $a'_{62} = a'_{63} = 0$, A''' já é diagonal e nada mais precisa ser feito.

No quarto caso temos uma matriz da forma

$$A'''' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ a_{61} & a_{62} & a_{63} \end{pmatrix}.$$

Se $a_{61} = a_{62} = a_{63} = 0$, A'''' é diagonal. Em caso contrário, permutamos as colunas e as linhas de maneira que o elemento da primeira linha e primeira coluna seja um escalar diferente de zero. Assim, obtemos:

$$\begin{pmatrix} a'_{11} \neq 0 & a'_{12} & a'_{13} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Finalmente, obtemos uma matriz diagonal mediante as seguintes substituições sucessivas.

$$\text{Coluna } 2 \leftarrow \text{ Coluna } 2 - \frac{a'_{12}}{a'_{11}} \text{ Coluna } 1$$

e

$$\text{Coluna } 3 \leftarrow \text{ Coluna } 3 - \frac{a'_{13}}{a'_{11}} \text{ Coluna } 1.$$

Agora, partindo de que a tese é verdade para matrizes de 5×3, a prova de que também o é para matrizes de 5×4 segue procedimentos análogos ao caso colocado acima. Generalizando este processo, e de forma recursiva, demostramos o teorema para qualquer matriz de $m \times n$.

Corolário 3.4. *O posto de uma matriz é igual ao posto de sua transposta.*

Prova. Usando determinada seqüência de operações 3.4a, transformamos A em uma matriz diagonal. Claramente, o posto dessa matriz é o número de seus elementos diferentes de zero. Agora, fazendo as mesmas operações com A^T (mas trocando "linha" por "coluna" e "coluna" por "linha" em cada passo) a matriz A^T fica transformada em D^T, que também é diagonal e cujo posto é o posto de D. Com isto, o corolário fica provado.

3.5 Matrizes e sistemas quadrados

Uma matriz com o mesmo número de linhas e colunas se diz *quadrada*. Os sistemas "quadrados" de equações (n equações e n incógnitas) cumprem um papel muito importante nas aplicações.

Se a matriz $A \in \mathbb{R}^{n\times n}$ tem posto n suas colunas formam uma base de $\mathbb{R}^n$, então, pelo visto na Seção 2.9, o sistema $Ax = b$ tem solução única.

Matrizes de $n \times n$ com posto igual a n se dizem *não-singulares*. Quando o posto é menor que n as chamamos *singulares*.

3.6 Matriz inversa

A matriz *Identidade* de $n \times n$, é definida por

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & & 0 \\ 0 & 1 & \cdots & & 0 \\ \cdot & & \cdots & & \cdot \\ 0 & & \cdots & 1 & 0 \\ 0 & & \cdots & 0 & 1 \end{pmatrix}.$$

Quando não há lugar a confusão denotamos $I = I_n$.

Observe que, para toda matriz $A \in \mathbb{R}^{n\times n}$, temos

$$AI = IA = A.$$

Seja $A \in \mathbb{R}^{n\times n}$. Dizemos que B é *Inversa* de A se

$$BA = I. \tag{3.12}$$

Suponhamos que B é inversa de A. Então, se

$$Ax = 0, \tag{3.13}$$

multiplicando ambos membros por B, obtemos $x = 0$. Logo $0 \in \mathbb{R}^n$ é a única solução de (3.13). Isto implica que as colunas de A são linearmente independentes. Por tanto, definindo os *vetores canônicos*

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{pmatrix}, \tag{3.14}$$

temos que os sistemas lineares

$$Ax = e_1, \ldots, Ax = e_n$$

tem solução única. Chamando v_1 à solução do primeiro sistema, v_2 à solução do segundo, e assim por diante, e formando a matriz C que tem por colunas os vetores $v_1, v_2, \ldots, v_n$, temos que

$$AC = I.$$

Pre-multiplicando esta igualdade por B e usando (3.12) obtemos que $C = B$, portanto

$$AB = I.$$

Se agora partimos da existência de uma matriz B tal que $AB = I$ e fazemos, com B, o mesmo raciocínio que fizemos acima com A chegaremos a conclusão de que $BA = I$.

Isto significa que $AB = I$ é equivalente a $BA = I$ e que ambas coisas são equivalentes a dizer que A é não singular. A inversa de A será denotada A^{-1}.

Veremos agora que, se uma matriz arbitrária é pre-multiplicada ou pós-multiplicada por uma matriz não-singular, seu posto permanece inalterado.

Teorema 3.6. *Seja $A \in \mathbb{R}^{m\times n}$ tal que Posto$(A) = p$. Sejam $B \in \mathbb{R}^{m\times m}$ e $C \in \mathbb{R}^{n\times n}$ não-singulares. Então:*

$$\textit{Posto}\ (BA) = \textit{Posto}\ (AC) = p.$$

Prova. Provemos primeiro que Posto$(BA) = p$. Suponhamos que $\{v_1, \ldots, v_q\}$ é um subconjunto de colunas de A. Então,

$$t_1 Bv_1 + \ldots + t_q Bv_q = 0$$

se, e somente se:

$$B(t_1 v_1 + \ldots + t_p v_q) = 0.$$

Pre-multiplicando esta igualdade por B^{-1} vemos que ela é equivalente a

$$t_1 v_1 + \ldots + t_q v_q = 0.$$

Portanto, as colunas $v_1, \ldots, v_q$ são linearmente independentes se, e somente se, os vetores $Bv_1, \ldots, Bv_q$ são linearmente independentes. Mas as colunas de BA são exatamente os vetores da forma Bv_j, com v_j coluna de A. Logo, o posto de BA deve ser igual ao posto de A.

O fato de que Posto $(AC) = p$ resulta da primeira parte do teorema usando o Corolário 3.4 e o fato de que C^T é não singular.

3.7 Determinantes

Uma permutação ν de $\{1,\ldots,n\}$ é uma n-upla de números inteiros (que denotaremos $(\nu(1),\ldots,\nu(n))$) tal que $\nu(j)\in\{1,\ldots,n\}$ para todo $j=1,\ldots,n$ e $\nu(i)\neq\nu(j)$ sempre que $i\neq j$. Por exemplo, $(3,1,4,2)$ é uma permutação de $\{1,2,3,4\}$, $(2,1)$ é uma permutação de $\{1,2\}$ e o próprio $(1,2,3)$ é uma permutação de $\{1,2,3\}$. Pela análise combinatória elementar sabemos que existem $n! = 1\times 2\times\ldots\times n$ permutações de $\{1,\ldots,n\}$.

É fácil ver que sempre se pode "chegar" de uma permutação a outra por uma seqüência finita de "trocas elementares" que, como seu nome indica, consiste em trocar os lugares de dois elementos da permutação entre si. Por exemplo:

$$(1,2,3,4)\to(3,2,1,4)\to(3,1,2,4)\to(3,1,4,2). \tag{3.15}$$

O *sinal* de uma permutação ν (denotado por $s(\nu)$) se define pelas seguintes regras:

(a) $s(1,\ldots,n)=1$;

(b) Se ν' é a permutação que se obtém a partir da permutação ν por uma troca elementar, então $s(\nu')=-s(\nu)$.

Como conseqüência da regra (b), se ν' é a permutação que se obtém a partir da permutação ν por um número ímpar de trocas elementares, então $s(\nu')=-s(\nu)$. Mas se ν' se obtém a partir da permutação ν por um número par de trocas elementares, então $s(\nu')=s(\nu)$.

Por exemplo, $s(3,1,4,2)=-1$, mas $s(3,1,2,4)=1$.

A combinação das duas regras permite deduzir, ao mesmo tempo, que o sinal de uma permutação sempre é 1 ou -1.

De agora em diante, denotamos $\mathcal{U}$ o conjunto das permutações de $\{1,\ldots,n\}$. Por exemplo, se $n=3$,

$$\mathcal{U}=\{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\}.$$

Agora podemos definir o *determinante* de uma matriz $A\in\mathbb{R}^{n\times n}$, cujo elemento genérico é $a_{i,j}$. Definimos

$$\det\ (A)\sum_{\nu\in\mathcal{U}} s(\nu)a_{\nu(1),1}\ldots a_{\nu(n),n}. \tag{3.16}$$

Ou seja, o determinante de A é uma soma de produtos precedidos de um sinal $+$ ou um sinal $-$; cada um desses produtos tem como fatores um elemento de cada coluna de A e as linhas correspondentes são determinadas por uma permutação de $\{1,\ldots,n\}$. Na somatória, todas as permutações são consideradas e o sinal é positivo se o sinal da permutação é 1 e negativo se o sinal é -1. Assim, o determinante tem $n!$ somandos e cada somando é o produto de n números (com sinal $+$ ou $-$).

Por exemplo, se $A \in I\!R^{3\times 3}$, os somandos são:
o correspondente à permutação $(1,2,3)$ com sinal $+$:

$$a_{11}a_{22}a_{33};$$

o correspondente à permutação $(1,3,2)$ com sinal $-$:

$$-a_{11}a_{32}a_{23};$$

o correspondente à permutação $(2,1,3)$ com sinal $-$:

$$-a_{21}a_{12}a_{33};$$

o correspondente à permutação $(2,3,1)$ com sinal $+$:

$$a_{21}a_{32}a_{13};$$

o correspondente à permutação $(3,1,2)$ com sinal $+$:

$$a_{31}a_{12}a_{23};$$

e o correspondente à permutação $(3,2,1)$ com sinal $-$:

$$-a_{31}a_{22}a_{13}.$$

Podemos verificar, fazendo os cálculos que, para matrizes de 2×2:

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Para matrizes de 3×3,

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

Mais ainda, os determinantes de matrizes de $n\times n$ se definem seguindo a mesma regra recursiva usada para definir os determinantes de 3×3 usando os de 2×2.

Existem muitas propriedades interessantes de determinantes. A que nos interessa aqui é que o determinante de uma matriz quadrada é zero se, e somente se, a matriz é singular.

As seguintes propriedades preliminares preparam o resultado principal:

Propriedade 3.7a. *Se a matriz A' é a matriz A com duas colunas trocadas, então*

$$det\ (A') = -\ det\ (A).$$

Prova. Surge da definição geral de determinante e da definição de sinal das permutações.

Propriedade 3.7b. *Se A tem duas colunas iguais, seu determinante é zero.*

Prova. Pela Propriedade 3.7a, $\det\ (A) = -\ \det\ (A)$. Logo $\det\ (A) = 0$.

Propriedade 3.7c. *Se A' resulta de multiplicar uma coluna de A por um escalar α, então*

$$det\ (A') = \alpha\ det\ (A).$$

Prova. Surge, diretamente, da definição de determinante.

Propriedade 3.7d. *Se A e B são duas matrizes iguais exceto a coluna j, e C é a mesma matriz exceto que a coluna j resulta de somar a coluna j de A com a coluna j de B, então:*

$$det\ (C) =\ det\ (A) +\ det\ (B).$$

Prova. De novo, se deduz imediatamente da definição de determinante.

Propriedade 3.7e. *Se coluna de A é uma combinação linear de outras, o determinante de A é zero.*

Prova. Aplicamos, sucessivamente, as propriedades 3.7d, 3.7c e 3.7b.

Pela Propriedade 3.7e, já sabemos que toda matriz singular tem determinante nulo. Precisamos provar, ainda, que se uma matriz tem determinante nulo, então é singular. Para isso, veremos ainda a seguinte propriedade simples:

Propriedade 3.7f. *Se uma coluna de A é substituída por "ela mesma mais uma combinação linear de outras" o determinante permanece o mesmo.*

Prova. Usamos, sucessivamente, as propriedades 3.7d e 3.7e.

Diremos que uma matriz cujo elemento genérico é a_{ij} é *triangular superior* se $a_{ij} = 0$ sempre que $i > j$. Por exemplo,

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}, \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}$$

são matrizes triangulares superiores.

O seguinte lema sobre matrizes triangulares superiores será útil. Sua prova usa argumentos largamente usados quando tratamos do posto de matrizes arbitrárias.

Lema 3.7. *Uma matriz triangular superior de $n \times n$ é singular se, e somente se, algum dos elementos da diagonal é zero.*

Mediante operações de *permutar colunas* e *substituír uma coluna por "ela mesma mais um múltiplo de outra"* conseguimos "transformar" qualquer matriz quadrada em uma matriz triangular superior. A prova deste fato é similar a provas anteriores feitas nestas notas e fica por conta do leitor. Portanto, pelas propriedades 3.7f e 3.7a, em cada uma destas operações o módulo do determinante permanece o mesmo. Assim, a matriz triangular superior que resulta depois destas operações tem um determinante com o mesmo módulo que a matriz original. A prova da seguinte proposição, que diz qual é o determinante de uma matriz triangular superior, resulta diretamente da definição de determinante.

Propriedade 3.7g. *Seja A uma matriz triangular superior. Então, seu determinante é o produto dos elementos de sua diagonal.*

Teorema 3.7. *Uma matriz $A \in I\!R^{n\times n}$ é singular se, e somente se, seu determinante é zero.*

Prova. O fato de que, se a matriz é singular seu determinante é zero resulta diretamente da Propriedade 3.7e.

Suponhamos, reciprocamente, que o determinante de A é zero. Seja U a matriz triangular superior obtida por sucessivas aplicações das operações "permutar colunas"e "substituír uma coluna por ela mesma mais um múltiplo de outra". Pelas propriedades 3.7a e 3.7f o determinante de U é zero. Logo, um dos elementos da diagonal de U, digamos u_{jj}, é zero. Se $j = 1$, isto implica que a coluna 1 de U é nula, portanto o posto de U é menor que n. Se $j > 1$, a coluna j de U é combinacao linear das $j - 1$ primeiras, de maneira que, neste caso, também o posto de U é menor que n. Mas as operacoes que conduzem de A até U não alteram o posto, portanto o posto de A é menor que n. Logo, A é singular.

3.8 Matrizes unitárias

Uma matriz $Q \in I\!R^{n\times n}$ se diz *unitária* ou *ortonormal* se suas colunas, consideradas como vetores de $I\!R^n$, são ortogonais e têm norma igual a 1. A matriz *Identidade* de $n \times n$ é a matriz unitária mais popular.

Pela definição de matriz unitária, temos que:

$$Q^T Q = I. \tag{3.17}$$

Portanto, $QQ^T = I$, ou seja Q é não singular e $Q^{-1} = Q^T$. Observe que a igualdade $QQ^T = I$ implica que as linhas de Q também são ortogonais e têm norma igual a 1.

As matrizes unitárias são não-singulares, portanto suas n colunas são linearmente independentes e formam uma base de $\mathbb{R}^n$. O seguinte teorema, embora de demonstração muito simples, nos será de grande utilidade no próximo capítulo.

Teorema 3.8. *Seja $x \in \mathbb{R}^n$ e seja $Q = (v_1, \dots, v_n)$ uma matriz unitária. Então, chamando $y_1, \dots, y_n$ às coordenadas de x na base $\{v_1, \dots, v_n\}$ e escrevendo*

$$y = \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix},$$

temos que

$$y = Q^T x.$$

Prova. Observe que, como $QQ^T = I$, escrevendo $z = Q^T x$, temos:

$$x = QQ^T x = Qz = z_1 v_1 + \ldots + z_n v_n.$$

É fácil ver que as matrizes unitárias tem a seguinte propriedade:

$$\langle Qv, Qw \rangle = \langle v, w \rangle \tag{3.18}$$

para todo $v, w \in \mathbb{R}^n$.

Daqui se deduz que, para todo $v \in \mathbb{R}^n$,

$$\|Qv\| = \|v\| \tag{3.19}$$

e que o ângulo entre Qv e Qw é o mesmo que o ângulo entre v e w.

3.9 Ortogonalização

Suponhamos que $\{v_1, \dots, v_m\} \subset \mathbb{R}^n$ é um conjunto de vetores linearmente independentes. Chamemos S_m ao subespaço de todas as combinações lineares de $\{v_1, \dots, v_m\}$, S_{m-1} às combinações lineares de $\{v_1, \dots, v_{m-1}\}$, e assim por diante.

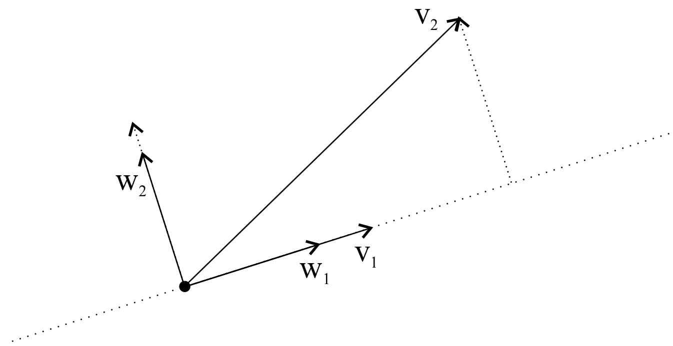

Figura 3.1. **Ortogonalização.**

Em geral, S_j é o subespaço das combinações lineares de $\{v_1, \ldots, v_j\}$, $j = 1, \ldots, m$. Veja Figura 3.1.

Em particular, S_1 é a reta que tem v_1 como vetor diretor e passa pela origem.

Definimos $w_1 = v_1/\|v_1\|$. Claramente, S_1 é, também, a reta que passa pela origem e tem w_1 como vetor diretor.

Agora definimos w_2' como a projeção de v_2 em S_1. Podemos verificar que $w_2' = \langle v_2, w_1 \rangle w_1$. Definimos

$$w_2 = \frac{v_2 - w_2'}{\|v_2 - w_2'\|}.$$

É fácil ver que S_2 é o subespaço das combinações lineares de $\{w_1, w_2\}$ e que w_1, w_2 são ortogonais e têm norma 1.

Repetindo este processo para todo j até m, w_j' será a projeção de v_j no subespaço das combinações lineares de $\{w_1, \ldots, w_j\}$.

$$w_j' = \langle v_j, w_1 \rangle w_1 + \ldots + \langle v_j, w_{j-1} \rangle w_{j-1}$$

e

$$w_j = \frac{v_j - w_j'}{\|v_j - w_j'\|}.$$

No fim, $\{w_1, \ldots, w_m\}$ será um conjunto de vetores ortogonais e de norma unitária tal que S_j é o subespaço das combinações lineares de $\{w_1, \ldots, w_j\}$ para todo $j = 1, \ldots, m$.

O processo definido, que permite obter $\{w_1, \ldots, w_m\}$ a partir de $\{v_1, \ldots, v_m$ se chama *ortogonalização de Gram-Schmidt.*

3.10 Autovalores e autovetores

Dizemos que o escalar λ é um *autovalor* de $A \in I\!R^{n\times n}$ se a matriz $A-\lambda I$ é singular. Em outras palavras, existe $v \neq 0$ tal que

$$Av = \lambda v. \tag{3.20}$$

O vetor $v \neq 0$ que verifica (3.20) se chama *autovetor* associado ao autovalor λ.
Pelo Teorema 3.7, λ será um autovalor de A se, e somente se,

$$\det(A - \lambda I) = 0. \tag{3.21}$$

A expressão $\det(A - \lambda I)$, quando desenvolvida, se revela um polinômio de grau n. Portanto, a equação (3.21), chamada *equação característica* tem a forma:

$$c_n\lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0 = 0.$$

Equações deste tipo podem ter raízes reais ou complexas. O Teorema Fundamental da Álgebra diz que o número total de raízes, reais ou complexas, contando suas multiplicidades, é igual a n.

Felizmente, quando a matriz A é simétrica, a equação característica somente tem raízes reais. Para demonstrar isto, precisamos usar algum conhecimento de propriedades de números complexos. Suponhamos que o número complexo λ e o "vetor complexo" v satisfazem (3.20). Denotamos

$$\bar{v} = \begin{pmatrix} \bar{v}_1 \\ \cdot \\ \cdot \\ \cdot \\ \bar{v}_n \end{pmatrix}$$

onde $\bar{z}$ é o conjugado do complexo z. (Portanto, $z\bar{z} = \bar{z}z = |z|^2 \in I\!R$.)
Pre-multiplicando (3.20) por $\bar{v}^T$, obtemos:

$$\bar{v}^T Av = \lambda(|v_1|^2 + \ldots + |v_n|^2). \tag{3.22}$$

Agora, usando a propriedade $(AB)^T = B^T A^T$ e "transpondo" ambos membros da igualdade acima, temos, graças à simetria de A:

$$v^T A\bar{v} = \lambda(|v_1|^2 + \ldots + |v_n|^2). \tag{3.23}$$

Por último, usando as propriedades $\overline{zw} = \bar{z}\bar{w}$ e $\overline{z + w} = \bar{z} + \bar{w}$ e "conjugando" ambos membros da igualdade (3.23), obtemos:

$$\bar{v}^T Av = \bar{\lambda}(|v_1|^2 + \ldots + |v_n|^2). \tag{3.24}$$

Como v não é o vetor nulo, (3.22) e (3.24) implicam que $\lambda = \bar{\lambda}$. Portanto λ é um número real.

A outra propriedade fundamental de autovalores e autovetores de matrizes simétricas é que autovetores correspondentes a autovalores diferentes são ortogonais. Para provar isto, suponhamos que λ é autovalor com autovetor v, μ é autovalor com autovetor w e $\lambda \neq \mu$. Logo,

$$Av = \lambda v \text{ e } Aw = \mu w.$$

Pre-multiplicando a primeira igualdade por w^T e a segunda por v^T e usando que $v^T w = w^T v = \langle v, w \rangle$, obtemos:

$$w^T A v = \lambda \langle v, w \rangle \text{ e } v^T A w = \mu \langle v, w \rangle.$$

Mas $(w^T A v)^T = v^T A w \in I\!R$, portanto $v^T A w = w^T A v$. Logo,

$$\lambda \langle v, w \rangle = \mu \langle v, w \rangle.$$

Como $\lambda \neq \mu$ isto implica que $\langle v, w \rangle = 0$.

3.11 Localização de autovalores

O cálculo dos autovalores de uma matriz é uma tarefa complicada. Por isso, resultam úteis os resultados que permitem estimar sua localização sem necessidade de saber exatamente o valor.

Lema 3.11. *Seja λ autovalor da matriz de $n \times n$ A, e seja $v = \begin{pmatrix} v_1 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{pmatrix}$ um autovetor associado. Suponhamos que*

$$|v_i| = \max\{|v_1|, \ldots, |v_n|\}.$$

Então, se a_{ij} é o elemento genérico de A,

$$|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$$

Prova. Como $Av = \lambda v$, temos que, em particular,

$$\sum_{j=1}^{n} a_{ij} v_j = \lambda v_i.$$

Portanto,

$$a_{ii}v_i + \sum_{j\neq i} a_{ij}v_j = \lambda v_i.$$

Como $v \neq 0$, v_i é diferente de zero. Divindindo ambos membros da igualdade acima por v_i, temos:

$$a_{ii} + \sum_{j\neq i} a_{ij}v_j/v_i = \lambda.$$

Logo,

$$|\lambda - a_{ii}| = |\sum_{j\neq i} a_{ij}v_j/v_i|.$$

Portanto, como $|v_j/v_i| \leq 1$ para todo j,

$$|\lambda - a_{ii}| \leq \sum_{j\neq i} |a_{ij}|,$$

como queriamos provar.

O Teorema dos *Discos de Gerschgorin*, enunciado em continuação, nos diz que todos os autovalores de uma matriz quadrada A estão na união de n "discos" centrados nos elementos da diagonal de A e cujos "rádios" são as somas dos módulos dos elementos de fora da diagonal de cada linha. Falamos de "discos" e não meramente de "intervalos" porque, embora não sendo central para estas notas, este resultado vale para autovalores complexos.

Teorema 3.11. *Seja λ autovalor da matriz de $n \times n$ A, cujo elemento genêrico é a_{ij}. Então, existe $i \in \{1, \dots, n\}$ tal que*

$$|\lambda - a_{ii}| \leq \sum_{j\neq i} |a_{ij}|,$$

Prova. Imediata, pelo Lema 3.11.

Vejamos um exemplo. Seja $A = \begin{pmatrix} 2 & -1 & 3 & 0 \\ -7 & 5 & -1 & 3 \\ 2 & 2 & 9 & 1 \\ -2 & -1 & 1 & 1 \end{pmatrix}$, e seja λ um autovalor de A. O Teorema dos Discos de Gerschgorin nos diz que um das seguintes desigualdades necessariamente se cumpre:

$$|\lambda - 2| \leq 4,$$

$$|\lambda - 5| \leq 11,$$

$$|\lambda - 9| \leq 5$$

ou

$$|\lambda - 1| \leq 4.$$

Portanto, procurar autovalores de A fora dessas quatro regiões é, neste caso, inútil.

3.12 Diagonalização de matrizes simétricas

Uma matriz $D \in I\!R^{m\times n}$, com elementos denotados por d_{ij}, se diz *diagonal* se $d_{ij} = 0$ sempre que $i \neq j$. Por exemplo,

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 0 & 5 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix}$$

são matrizes diagonais.

As matrizes diagonais de $n \times n$ têm a seguinte forma:

$$D = \begin{pmatrix} d_{11} & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & d_{22} & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & 0 & d_{nn} \end{pmatrix}.$$

O seguinte teorema tem múltiplas aplicações em diversas áreas da matemática. Em particular, é a ferramenta mais eficaz para classificar cônicas, quádricas e equações quadráticas em geral.

Teorema 3.12 *Toda matriz simétrica $A \in I\!R^{n\times n}$ pode ser decomposta na forma*

$$A = QDQ^T \tag{3.25}$$

onde D é uma matriz diagonal e Q é unitária.

Uma prova parcial deste teorema, para o caso em que A tem n autovalores diferentes, resulta dos resultados "autovalores reais" e "autovetores ortogonais" provados na Seção 3.10. Esta prova parcial será deixada como exercício para o leitor.

Antes de provar o Teorema 3.12, vamos analisar algumas de suas conseqüências. Para isso, escrevemos

$$Q = (v_1, \ldots, v_n),$$

onde $v_j \in I\!R^n$ para todo $j = 1, \ldots, n$. Os vetores v_j são as colunas de Q.

Usando que $Q^TQ = I$ e multiplicando a direita ambos membros de (3.25) por Q, temos:

$$AQ = QD.$$

Ou seja:

$$A(v_1,\dots,v_n) = (v_1,\dots,v_n)D.$$

Usando que D é diagonal e desenvolvendo a última igualdade, obtemos:

$$(Av_1,\dots,Av_n) = (d_{11}v_1,\dots,d_{nn}v_n).$$

Portanto,

$$Av_j = d_{jj}v_j$$

para todo $j = 1,\dots,n$.

Portanto, cada elemento d_{jj} da diagonal de D é um autovalor de A e v_j (a coluna j de Q) é um autovetor associado a esse autovalor.

Prova do Teorema 3.12.

Se $n = 1$ o teorema é obviamente verdadeiro. Neste caso, $A \in I\!R^{1\times 1}$ é um número e podemos definir $Q = (1)$. Assim, Q é uma matriz unitária de 1×1. A matriz diagonal D pode ser definida como a própria A.

Provaremos que, se o teorema é verdadeiro para matrizes de $(n-1)\times(n-1)$ então necessariamente deve ser verdadeiro para matrizes de $n\times n$. Se isto for verdade, do fato do teorema estar certo para 1×1, se deduz que está certo para 2×2, deste fato se deduz o teorema para matrizes de 3×3 e assim por diante. Este processo se denomina *indução matemática.*

Suponhamos, então, que o teorema é verdade para todas as matrizes simétricas de $(n-1)\times(n-1)$.

Pelo Teorema Fundamental da Álgebra, a equação característica tem pelo menos uma solução λ, e esta solução, pelo visto na seção anterior é um autovalor real de A. Seja $v_1 \in I\!R^n$ um autovetor associado a λ. Sejam $v_2,\dots,v_n$ tais que $\{v_1,v_2,\dots,v_n\}$ são linearmente independentes. Aplicando ortogonalização de Gram-Schmidt a $\{v_1,v_2,\dots,v_n\}$ obtemos um conjunto de vetores ortogonais e unitários $\{w_1,w_2,\dots,w_n\}$. Definimos $Q_0 = (w_1,w_2,\dots,w_n)$. Claramente, Q_0 é unitária.

Agora,

$$Q_0^T A Q_0 = (w_1,w_2,\dots,w_n)^T A(w_1,\dots,w_n) = \begin{pmatrix} w_1^T \\ w_2^T \\ \cdot \\ \cdot \\ \cdot \\ w_n^T \end{pmatrix} A(w_1,w_2,\dots,w_n)$$

$$= \begin{pmatrix} w_1^T \\ w_2^T \\ \cdot \\ \cdot \\ \cdot \\ w_n^T \end{pmatrix} (Aw_1, Aw_2, \ldots, Aw_n) = \begin{pmatrix} w_1^T \\ w_2^T \\ \cdot \\ \cdot \\ \cdot \\ w_n^T \end{pmatrix} (\lambda w_1, Aw_2, \ldots, Aw_n)$$

$$= \begin{pmatrix} w_1^T Aw_1 & w_1^T Aw_2 & \cdots & w_1^T Aw_n \\ w_2^T Aw_1 & w_2^T Aw_2 & \cdots & w_2^T Aw_n \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ w_n^T Aw_1 & w_n^T Aw_2 & \cdots & w_n^T Aw_n \end{pmatrix} = \begin{pmatrix} \lambda & w_2^T Aw_1 & \cdots & w_n^T Aw_1 \\ \lambda w_2^T w_1 & w_2^T Aw_2 & \cdots & w_2^T Aw_n \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \lambda w_n^T w_1 & w_n^T Aw_2 & \cdots & w_n^T Aw_n \end{pmatrix}$$

$$= \begin{pmatrix} \lambda & \lambda w_2^T w_1 & \cdots & \lambda w_n^T w_1 \\ 0 & w_2^T Aw_2 & \cdots & w_2^T Aw_n \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & w_n^T Aw_2 & \cdots & w_n^T Aw_n \end{pmatrix} = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & w_2^T Aw_2 & \cdots & w_2^T Aw_n \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & w_n^T Aw_2 & \cdots & w_n^T Aw_n \end{pmatrix}.$$

Agora, excluindo a primeira linha e a primeira coluna desta matriz, temos uma matriz de $(n-1)\times(n-1)$. Portanto, podemos supor que existem $Q_1 \in I\!R^{(n-1)\times(n-1)}$ unitária $D_1 \in I\!R^{(n-1)\times(n-1)}$ diagonal tais que

$$\begin{pmatrix} w_2^T Aw_2 & \cdots & w_2^T Aw_n \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ w_n^T Aw_2 & \cdots & w_n^T Aw_n \end{pmatrix} = Q_1 D_1 Q_1^T.$$

Portanto,

$$\begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & w_2^T Aw_2 & \cdots & w_2^T Aw_n \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & w_n^T Aw_2 & \cdots & w_n^T Aw_n \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & D_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q_1^T \end{pmatrix}.$$

É fácil ver que

$$Q_2 = \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}$$

é unitária e que

$$D = \begin{pmatrix} \lambda & 0 \\ 0 & D_1 \end{pmatrix}$$

é diagonal.

Portanto,

$$Q_0^T A Q_0 = Q_2 D Q_2^T.$$

Portanto,

$$A = Q_0 Q_2 D Q_2^T Q_0^T = (Q_0 Q_2) D (Q_0 Q_2)^T.$$

Como $Q_0 Q_2$ é unitária, o teorema está provado.

Corolário 3.12a. *Seja $A \in I\!R^{n\times n}$ uma matriz simêtrica. Então, as seguintes afirmações são equivalentes:*

(a) Todos os autovalores de A são não-negativos;

(b) Para todo $v \in I\!R^n$, $v^T A v \geq 0$.

Prova. Vejamos primeiro que (a)⇒(b). Pelo Teorema 3.12 temos que as colunas de Q formam uma base ortonormal de $I\!R^n$. Cada elemento de esta base é um autovetor v_i associado ao autovalor $\lambda_i \geq 0$. Portanto,

$$v_i^T A v_i = v_i^T \lambda_i v_i = \lambda \|v_i\|^2 = \lambda \geq 0.$$

Seja $v \in I\!R^n$, arbitrário. Como $\{v_1, \ldots, v_n\}$ é uma base de $I\!R^n$, existem $t_1, \ldots, t_n$ tais que

$$v = t_1 v_1 + \ldots + t_n v_n.$$

Portanto, usando que $v_i^T v_i = 1$ e $v_i^T v_j = 0$ para todo $i \neq j$, temos:

$$\begin{aligned} v^T A v &= (t_1 v_1 + \ldots + t_n v_n)^T A (t_1 v_1 + \ldots + t_n v_n) \\ &= (t_1 v_1 + \ldots + t_n v_n)^T (t_1 A v_1 + \ldots + t_n A v_n) \\ &= (t_1 v_1 + \ldots + t_n v_n)^T (t_1 \lambda_1 v_1 + \ldots + t_n \lambda_n v_n) \\ &= t_1^2 \lambda_1 + \ldots + t_n^2 \lambda_n \geq 0. \end{aligned}$$

Logo, $v^T A v \geq 0$, como queriamos provar.

O fato de que (b)⇒(a) é mais simples. Suponhamos que λ_i é um autovalor menor que zero de A e que v_i é um autovetor unitário associado. Então:

$$v_i^T A v_i = \lambda_i v_i^T v_i = \lambda_i < 0.$$

Logo, a negação de (a) implica a negação de (b), o que completa nossa prova.

A prova dos seguintes corolários fica como exercício para o leitor.

Corolário 3.12b. *Seja $A \in I\!R^{n \times n}$ uma matriz simêtrica. Então, as seguintes afirmações são equivalentes:*

(a) Todos os autovalores de A são menores ou iguais que zero;

(b) Para todo $v \in I\!R^n$, $v^T A v \leq 0$.

Corolário 3.12c. *Seja $A \in I\!R^{n \times n}$ uma matriz simêtrica. Então, as seguintes afirmações são equivalentes:*

(a) Todos os autovalores de A são maiores que zero;

(b) Para todo $v \in I\!R^n$, tal que $v \neq 0$, tem-se $v^T A v > 0$.

Corolário 3.12d. *Seja $A \in I\!R^{n \times n}$ uma matriz simêtrica. Então, as seguintes afirmações são equivalentes:*

(a) Todos os autovalores de A são menores que zero;

(b) Para todo $v \in I\!R^n$, tal que $v \neq 0$, tem-se $v^T A v < 0$.

Por último, mencionaremos um corolário simples mas de grande utilidade. Ele diz que a simples inspecção da diagonal de uma matriz simêtrica fornece informações sobre seus autovalores.

Corolário 3.12e. *Seja $A \in I\!R^{n \times n}$ uma matriz simêtrica cujo elemento genêrico é a_{ij}. Então:*

(a) Se todos os autovalores de A são maiores ou iguais que zero, então todos os elementos da diagonal de A são maiores ou iguais que zero. (Logo, a presença de um elemento negativo na diagonal implica a existência de um autovalor menor que zero.)

(b) Se todos os autovalores de A são menores ou iguais que zero, então todos os elementos da diagonal de A são menores ou iguais que zero. (Analogamente, a presença de um elemento positivo na diagonal implica a existência de um autovalor maior que zero.)

(c) Se todos os autovalores de A são maiores que zero, então todos os elementos da diagonal de A são maiores que zero. (A presença de um elemento negativo ou nulo na diagonal de A implica a existência de um autovalor menor ou igual que zero.)

(d) Se todos os autovalores de A são menores que zero, então todos os elementos da diagonal de A são menores que zero. (Logo, a presença de um elemento positivo ou nulo na diagonal implica a existência de um autovalor maior ou igual que zero.)

Prova. Sejam $e_1, \ldots, e_n$ definidos por (3.14). A prova deste corolário se segue dos anteriores usando que

$$a_{ii} = e_i^T A e_i$$

para todo $i = 1, \ldots, n$.

3.13 Decomposição em valores singulares

Nesta seção veremos uma ferramenta importante para a compreensão de quase todas as propriedades de cálculo matricial: toda matriz $A \in I\!R^{m\times n}$ pode ser decomposta como $A = UDV^T$, onde U e V são unitárias e $D \in I\!R^{m\times n}$ é diagonal não-negativa. Os escalares positivos que aparecerão na diagonal de D se chamam *valores singulares* de A. As conseqüências teóricas e práticas deste fato são imensas. No livro de Golub e Van Loan [1], cuja leitura supõe certa maturidade em Álgebra Linear, a decomposição em valores singulares é, essencialmente, o primeiro resultado demonstrado e a maior parte de definições e propriedades da teoria de matrizes são baseadas nele.

Teorema 3.13 *Seja $A \in I\!R^{m\times n}$. Então existem matrizes unitárias $U \in I\!R^{m\times m}$, $V \in I\!R^{n\times n}$ e uma matriz diagonal $D \in I\!R^{m\times n}$, onde $d_{ii} > 0 \;\; \forall \;\; i \leq p$ e $d_{ii} = 0$ se $i > p$, tais que*

$$A = UDV^T.$$

A prova do Teorema 3.13 está baseada no seguinte lema:

Lema 3.13 *Seja $A \in I\!R^{m\times n}$. Seja u autovetor de AA^T, $\|u\| = 1$, com autovalor λ. Então:*

$$A^T A(A^T u) = \lambda A^T u \tag{3.26}$$

e

$$\|A^T u\| = \sqrt{\lambda}. \tag{3.27}$$

Prova. Como u é autovetor de AA^T com autovalor λ, temos:

$$AA^T u = \lambda u.$$

Pre-multiplicando ambos membros desta equação por A^T, obtemos (3.26). Agora:

$$\|A^T u\|^2 = \langle A^T u, A^T u\rangle = (A^T u)^T A^T u = u^T AA^T u = \lambda u^T u = \lambda \|u\|^2 = \lambda.$$

Como $\lambda \geq 0$, segue-se (3.27).

Pelo Lema 3.13, se u é um autovetor de AA^T associado a um autovalor não-nulo, então A^Tu é um autovetor de A^TA associado ao mesmo autovalor. Por outro lado, se u é um autovetor associado a um autovalor nulo de AA^T, então $A^Tu = 0$.

Como conseqüência, os autovalores não-nulos de AA^T e A^TA são os mesmos, mas uma destas matrizes pode ter um autovalor nulo, enquanto a outra não tem. Por exemplo, se

$$A = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in I\!\!R^{2\times 1},$$

temos

$$AA^T = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in I\!\!R^{2\times 2} \text{ e } A^TA = (\,2\,) \in I\!\!R^{1\times 1}.$$

A matriz AA^T tem dois autovalores: 0 e 2. Porém, 2 é o único autovalor da matriz A^TA.

Prova do Teorema 3.13. Pelo Teorema 3.12 existe uma matriz unitária $U \in I\!\!R^{m\times m}$ e uma matriz diagonal $D \in I\!\!Rm \times m$ com p elementos da diagonal não nulos, tais que:

$$AA^T = UDU^T.$$

Sem perda de generalidade podemos supor que os elementos não nulos da diagonal de D são os p primeiros $(d_{11}, \ldots, d_{pp})$. (Se não forem, basta trocar a ordem das colunas de U.)

O resto da prova exige algumas manipulações de operações matriciais com as quais o leitor pode não estar acostumado. Para que possa ser melhor acompanhada, vamos fazer esta prova para um caso particular: $m = 5, n = 3, p = 2$. O procedimento geral usa as mesmas idéias que as deste caso e é sugerido ao leitor como exercício.

Escrevemos

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{pmatrix},$$

$$U = (\, u_1 \quad u_2 \quad u_3 \quad u_4 \quad u_5 \,),$$

$$u_j \in I\!\!R^5, j = 1, \ldots, 5.$$

Pelo Teorema 3.12, u_1, u_2 são autovetores de AA^T correspondentes a autovalores diferentes de zero e u_3, u_4, u_5 são autovetores de AA^T correspondentes ao autovalor zero.

Transpondo, temos:

$$U^T = \begin{pmatrix} u_1^T \\ u_2^T \\ u_3^T \\ u_4^T \\ u_5^T \end{pmatrix}.$$

Como U é unitária, temos:

$$A = UU^T A = U \begin{pmatrix} u_1^T A \\ u_2^T A \\ u_3^T A \\ u_4^T A \\ u_5^T A \end{pmatrix}.$$

Pelo Lema 3.13, $A^T u_3 = A^T u_4 = A^T u_5 = 0$, portanto:

$$A = U \begin{pmatrix} u_1^T A \\ u_2^T A \\ 0\ 0\ 0 \\ 0\ 0\ 0 \\ 0\ 0\ 0 \end{pmatrix}$$

$$= U \begin{pmatrix} \|A^T u_1\| & 0 & 0 & 0 & 0 \\ 0 & \|A^T u_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} (A^T u_1)^T / \|A^T u_1\| \\ (A^T u_2)^T / \|A^T u_2\| \\ 0\ 0\ 0 \\ 0\ 0\ 0 \\ 0\ 0\ 0 \end{pmatrix}$$

$$= U \begin{pmatrix} \|A^T u_1\| & 0 & 0 \\ 0 & \|A^T u_2\| & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0; & 0 & 0 \end{pmatrix} \begin{pmatrix} (A^T u_1)^T / \|A^T u_1\| \\ (A^T u_2)^T / \|A^T u_2\| \\ a \quad b \quad c \end{pmatrix}$$

para quaisquer $a, b, c \in I\!R$.

Agora, como d_{22} é o autovalor de AA^T associado ao autovetor u_2,

$$\left\langle A^T u_1, A^T u_2 \right\rangle = u_1^T A A^T u_2 = d_{22} \langle u_1, u_2 \rangle = 0.$$

Portanto, se $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in I\!R^3$ é tal que

$$\left\langle \begin{pmatrix} a \\ b \\ c \end{pmatrix}, A^T u_1 \right\rangle = \left\langle \begin{pmatrix} a \\ b \\ c \end{pmatrix}, A^T u_2 \right\rangle = 0$$

e

$$\left\| \begin{pmatrix} a \\ b \\ c \end{pmatrix} \right\| = 1$$

temos que a matriz $\begin{pmatrix} (A^T u_1)^T / \|A^T u_1\| \\ (A^T u_2)^T / \|A^T u_2\| \\ a \quad b \quad c \end{pmatrix}$ é unitária.

Logo, definindo

$$D = \begin{pmatrix} \|A^T u_1\| & 0 & 0 \\ 0 & \|A^T u_2\| & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

e

$$V^T = \begin{pmatrix} (A^T u_1)^T / \|A^T u_1\| \\ (A^T u_2)^T / \|A^T u_2\| \\ a \quad b \quad c \end{pmatrix},$$

o teorema está provado.

3.14 Exercícios

1. Prove que $(AB)^T = B^T A^T$.

2. Prove (3.3).

3. Seja $A \in I\!R^{m \times n}$, $B \in I\!R^{n \times p}$ tal que

$$B = (v_1, \ldots, v_p).$$

(Logo, $v_1, \ldots, v_p$, elementos de $I\!R^n$, são as colunas de B.) Prove que

$$AB = (Av_1, \ldots, Av_p).$$

4. Mostre que todo sistema linear de $m \times n$ pode ser escrito na forma $Ax = b$ onde $A \in I\!R^{m \times n}$. Para cada sistema da Lista 1, escreva a *matriz do sistema* A, o termo independente b e a *matriz ampliada*. Mostre como o algoritmo de escalonamento pode ser bastante simplificado usando matrizes.

5. Dizemos que um conjunto de vetores é *linearmente independente* se nenhum deles é combinação linear dos outros. Interprete geometricamente esta definição. Prove:

(a) Se um conjunto de vetores é linearmente independente e outro conjunto está contido nele, então o segundo é linearmente independente.

(b) Se um conjunto de vetores é *linearmente dependente* e está contido em outro, então o segundo é linearmente dependente.

6. Chamamos *posto* de uma matriz ao máximo número de colunas linearmente independentes. Prove:

   (a) Um sistema linear tem solução se, e somente se, o termo independente é combinação linear das colunas da matriz do sistema.

   (b) Um sistema linear tem solução se, e somente se, o posto da matriz do sistema é igual ao posto da matriz ampliada.

7. Para cada uma das matrizes dos sistemas da Lista 1, calcule o posto.

8. Prove que o posto de uma matriz diagonal é o número de suas entradas diferentes de zero.

9. Suponha que a matriz $A \in \mathbb{R}^{m\times n}$ tem posto n e que o sistema $Ax = b$ tem solução. Prove que esta solução deve ser única. Pode ser que o sistema não tenha solução? Forneça exemplos.

10. Suponha que a matriz $A \in \mathbb{R}^{m\times n}$ tem posto menor que n e que o sistema $Ax = b$ tem solução. Prove que o sistema tem infinitas soluções.

11. Seja $A \in \mathbb{R}^{2\times 2}$. Prove que, diretamente, que A tem posto 2 se, e somente se,
$$\det(A) \neq 0.$$

12. Prove que $(AB)^{-1} = B^{-1}A^{-1}$.

13. Prove que A é não singular se, e somente se, A^T é não-singular e que
$$(A^T)^{-1} = (A^{-1})^T.$$

14. Seja $x \in \mathbb{R}^n$ e seja $A = (v_1 \quad \cdots \quad v_n)$ uma matriz não-singular. Então, chamando $y_1, \ldots, y_n$ às coordenadas de x na base $\{v_1, \ldots, v_n\}$ e escrevendo $y = \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix}$, temos que
$$y = A^{-1}x.$$

15. Prove (3.18) e (3.19).

16. Seja $v \in I\!R^n$, $v \neq 0$. Mostre vetores $\{v_2, \ldots, v_n\}$ tais que $\{v, v_2, \ldots, v_n\}$ são linearmente independentes. (Use vetores canônicos!.)

17. Prove que, de fato, no processo de Gram-Schmidt o subespaço das combinações lineares de $w_1, \ldots, w_j$ é o subespaço das combinações lineares de $v_1, \ldots, v_j$ para todo $j = 1, \ldots, m$.

18. Aplique o processo de Gram-Schmidt aos vetores

    (a) $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ e $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

    (b) $\begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ 3 \end{pmatrix}, \begin{pmatrix} 10 \\ 1 \\ 0 \end{pmatrix}$.

19. Prove que o produto de matrizes unitárias é uma matriz unitária.

20. Prove que, se x é autovetor associado a λ e $c \neq 0$, então cx é autovetor associado a λ. Prove que combinaçãos lineares não-nulas de autovetores associados a λ são autovetores associados a λ.

21. Seja $A \in I\!R^{m \times n}$. Prove que AA^T e $A^T A$ são matrizes simétricas e que seus autovalores são não-negativos.

22. Usando os resultados da Seção 3.10 prove o Teorema 3.25 para o caso em que A tem n autovalores diferentes.

23. Prove os corolários 3.12a, 3.12b, 3.12c, 3.12d e 3.12e.

24. Peça para seu vizinho inventar matrizes simêtricas de 5×5 e diga o que pode afirmar sobre os autovalores das mesmas usando o Corolario 3.12e e o Teorema 3.11.

25. Para as seguintes matrizes simétricas $A \in I\!R^{2 \times 2}$, calcule autovalores e autovetores e calcule uma matriz *unitária* Q e uma matriz diagonal D tais que $A = QDQ^T$.

    (a) $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

    (b) $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

    (c) $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

(d) $A = \begin{pmatrix} -5 & 0 \\ 0 & 3 \end{pmatrix}$

(e) $A = \begin{pmatrix} 0 & 0 \\ 0 & 1976 \end{pmatrix}$

(f) $A = \begin{pmatrix} 10 & 1 \\ 1 & 8 \end{pmatrix}$

(g) $A = \begin{pmatrix} 3 & -3 \\ -3 & 3 \end{pmatrix}$

(h) $A = \begin{pmatrix} 5 & 4 \\ 4 & -6 \end{pmatrix}$

(i) $A = \begin{pmatrix} -8 & 7 \\ 7 & -11 \end{pmatrix}$.

26. Escreva a demonstração do Teorema 3.13 em toda sua generalidade.

27. Mostre como resolver o problema de *quadrados mínimos*

$$\text{Minimizar } \|Ax - b\|$$

usando a decomposição em valores singulares.

28. Prove que toda matriz não-nula $A \in I\!R^{m\times n}$ pode ser escrita na forma

$$A = \sigma_1 u_1 v_1^T + \ldots + \sigma_p u_p v_p^T \tag{3.28}$$

onde $\sigma_1 > \ldots > \sigma_p > 0$ são os valores singulares de A, os vetores $u_1, \ldots, u_p \in I\!R^m$ são ortogonais e tem norma igual a 1 e os vetores $v_1, \ldots, v_p \in I\!R^n$ são ortogonais e tem norma igual a 1.

29. A tela de um computador tem 1024×768 pixels. Atribuindo números às diferentes cores, uma imagem pode identificada com uma matriz de 1024×768. Portanto, para armazenar totalmente uma imagem precisamos de aproximadamente 770000 números. Suponhamos que calculamos a decomposição em valores singulares da imagem (da matriz) e constatamos que apenas 14 valores singulares são significativamente diferentes de zero. Quanto espaço necessitaremos para armazenar a imagem, usando (3.28)? Quanto tempo precisaremos para transmití-la?

Capítulo 4
Cônicas e Quádricas

As equações mais simples são as lineares, para as quais vimos formas de resolução e interpretações geométricas nos capítulos anteriores. As soluções de uma equação linear de duas variáveis é, em geral, uma reta, as de uma equação linear de três variáveis é um plano e assim por diante. As equações quadráticas tem um grau de complexidade imediatamente superior ao das equações lineares. Nelas, as variáveis podem aparecer elevadas ao quadrado ou multiplicadas entre si. O estudo geométrico das soluções de equações quadráticas de duas e três variáveis neste capítulo nos levará aos conceitos de *cônicas* e *quádricas* respectivamente.

4.1 Equações quadráticas

Uma equação quadrática de n variáveis é uma igualdade da forma

$$\langle c, x\rangle + \langle x, Bx\rangle = a \tag{4.1}$$

onde $a \in I\!R$, $c \in I\!R^n$ e B, uma matriz simétrica de $n \times n$, são dados e $x \in I\!R^n$ é um vetor de incógnitas ou variáveis. É usual escrever a fórmula das equações quadráticas usando a notação puramente matricial:

$$c^T x + x^T Bx = a \tag{4.2}$$

Seja $c = \begin{pmatrix} c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{pmatrix}$, $B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ b_{1n} & \cdots & b_{nn} \end{pmatrix}$. Então, podemos verificar que:

$$c^T x + x^T Bx = c_1x_1 + \ldots + c_nx_n + (b_{11}x_1^2 + \ldots + b_{nn}x_n^2) + 2\sum_{i<j} b_{ij}x_ix_j.$$

Por exemplo, se $n = 2$,

$$c^T x + x^T B x = c_1 x_1 + c_2 x_2 + b_{11} x_1^2 + b_{22} x_2^2 + 2 b_{12} x_1 x_2.$$

Se $n = 3$,

$$c^T x + x^T B x = c_1 x_1 + c_2 x_2 + c_3 x_3 + b_{11} x_1^2 + b_{22} x_2^2 + b_{33} x_3^2 + 2(b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3).$$

E, mesmo para $n = 1$,

$$c^T x + x^T B x = c_1 x + b_{11} x^2.$$

Uma equação quadrática se diz *linear* se a matriz B é nula.

4.2 Cilindros

Quando na equação (4.38) ou (4.39) não aparecem todas as variáveis dizemos que o conjunto de soluções é um *cilindro*. A razão de esta denominação vem da consideração do seguinte caso particular: considere o conjunto definido por

$$A = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1 \right\}.$$

Claramente, A é a circunferência unitária de $\mathbb{R}^2$. Entretanto,

$$A' = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \mid x_1^2 + x_2^2 = 1 \right\}$$

é o "cilindro infinito" da Figura 4.1.

Outro exemplo é o conjunto de soluções de $x_1^2 = 1$. Se olhamos este conjunto como um subconjunto de $\mathbb{R}^1$ ele está formado por dois pontos (1 e -1). Entretanto, como subconjunto de $\mathbb{R}^2$ está constituído por duas retas paralelas, como subconjunto de $\mathbb{R}^3$ consta de dois planos paralelos e assim por diante.

Naturalmente, a compreensão geomêtrica de conjuntos cilíndricos depende somente da compreensão de sua forma no espaço $\mathbb{R}^n$ com a menor dimensão possível onde a equação pode ser interpretada. Por exemplo, em princípio não há nenhuma necessidade de interpretar a equação $x_1^2 + x_2^2 = 1$ em $\mathbb{R}^3$. Entretanto, essa necessidade apareceria se estivéssemos interessados no conjunto de soluções do sistema

$$x_1^2 + x_2^2 = 1, \quad x_1 + x_2 + x_3 = 0.$$

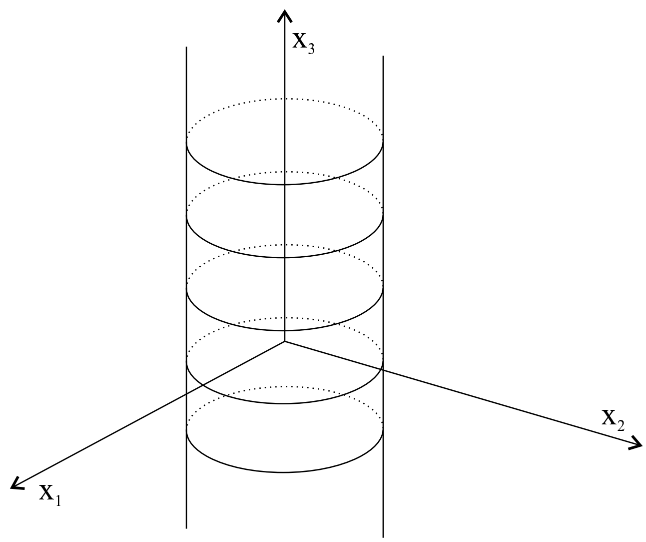

Figura 4.1. **Cilindro de base circular.**

4.3 Eliminação dos termos cruzados

Uma das dificuldades para visualizar equações quadráticas é a presença de "termos cruzados", isto é, termos da forma $b_{ij}x_ix_j$ com $b_{ij} \neq 0$. Nesta seção veremos que, com uma mudança de variáveis, esses termos podem ser eliminados. Em outras palavras, que formulando a equação (4.1) em outro sistema de coordenadas ortogonais os termos cruzados desaparecem. Para entender como isto é possível vejamos, primeiro, um exemplo.

Consideremos a equação

$$x_1 + 6x_2 + 2x_1^2 - x_1x_2 + 2x_2^2 = 10. \tag{4.3}$$

Definamos as novas variáveis (mais tarde veremos com que critério):

$$y_1 = \frac{\sqrt{2}}{2}x_1 + \frac{\sqrt{2}}{2}x_2, \tag{4.4}$$

$$y_2 = -\frac{\sqrt{2}}{2}x_1 + \frac{\sqrt{2}}{2}x_2. \tag{4.5}$$

Escrevendo as "velhas" variáveis em função das novas, temos:

$$x_1 = \frac{\sqrt{2}}{2}y_1 - \frac{\sqrt{2}}{2}y_2, \tag{4.6}$$

$$x_2 = \frac{\sqrt{2}}{2}y_1 + \frac{\sqrt{2}}{2}y_2. \tag{4.7}$$

Substituindo (4.6-4.7) em (4.3), temos que a equação nas novas variáveis fica:

$$\frac{\sqrt{2}}{2}y_1 - \frac{\sqrt{2}}{2}y_2 + 6\Big(\frac{\sqrt{2}}{2}y_1 + \frac{\sqrt{2}}{2}y_2\Big) + 2\Big(\frac{\sqrt{2}}{2}y_1$$
$$-\frac{\sqrt{2}}{2}y_2\Big)^2 - \Big(\frac{\sqrt{2}}{2}y_1 - \frac{\sqrt{2}}{2}y_2\Big)\Big(\frac{\sqrt{2}}{2}y_1 + \frac{\sqrt{2}}{2}y_2\Big) + 2\Big(\frac{\sqrt{2}}{2}y_1 + \frac{\sqrt{2}}{2}y_2\Big)^2 = 10.$$

Ou seja,

$$\frac{7\sqrt{2}}{2}y_1 + \frac{5\sqrt{2}}{2}y_2 + 2\Big(\frac{1}{2}y_1^2 + \frac{1}{2}y_2^2 - y_1y_2\Big) - \frac{1}{2}y_1^2 + \frac{1}{2}y_2^2 + 2\Big(\frac{1}{2}y_1^2 + \frac{1}{2}y_2^2 + y_1y_2\Big) = 10.$$

Portanto, a equação nas novas variáveis é:

$$\frac{7\sqrt{2}}{2}y_1 + \frac{5\sqrt{2}}{2}y_2 + \frac{3}{2}y_1^2 + \frac{5}{2}y_2^2 = 10,$$

onde o coeficiente do termo cruzado y_1y_2 é nulo.

Vamos tentar entender o significado geométrico da mudança de variáveis efetuada.

Observe que a matriz

$$Q = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

é unitária. Portanto, como vimos no Capítulo 3, as coordenadas de um ponto $x \in I\!\!R^2$ com respeito à base formada por suas duas colunas vêm dadas por Q^Tx. Mas Q^Tx é exatamente a fórmula escolhida para y em (4.4-4.5). Portanto, o sistema de coordenadas no qual a equação (4.3) toma uma forma sem termos cruzados é aquele que corresponde à base de vetores ortonormais

$$v_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix} \text{ e } v_2 = \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}.$$

Veja Figura 4.2.

Depois deste exemplo, o caso geral nos resultará mais familiar. Consideremos a equação quadratica (4.1).

Como B é simétrica, existem $D \in I\!\!R^{n\times n}$ diagonal e $Q \in I\!\!R^n$ unitária tais que

$$B = QDQ^T.$$

(Sabemos também que os elementos diagonais de D são os autovalores de B e que as colunas correspondentes de Q são autovetores associados.) Portanto,

$$c^Tx + x^TBx = c^Tx + x^TQDQ^Tx = c^Tx + (Qx)^TDQ^Tx.$$

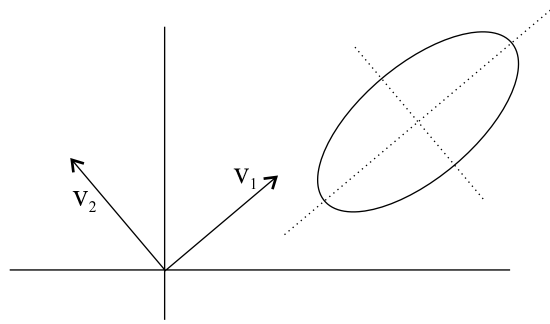

Figura 4.2. **Efeito da eliminação de termos cruzados.**

Assim, definindo $y = Q^T x$, a equação (4.2) nas novas variáveis é:

$$(Q^T c)^T y + y^T D y = a,$$

equação na qual, evidentemente, não aparecem termos cruzados. Portanto, a mudança de coordenadas que deve ser feita em uma equação quadrática para fazer desaparecer os termos cruzados é a que resulta de escolher, como nova base, um conjunto de n vetores linearmente independentes de B.

4.4 Redução à forma diagonal

No processo de entender a geometria das equações quadráticas já nos livramos da complicação mais séria: a presença de termos cruzados. Portanto, nesta seção, vamos supor que na matriz B da equação quadrática (4.1) temos sempre $b_{ij} = 0$ se $i \neq j$.

Diremos que uma equação quadrática tem *forma (ou estrutura) diagonal* se não tem termos cruzados e cada variável aparece como máximo uma vez com coeficiente diferente de zero. Por exemplo, as equações quadráticas seguintes tem estrutura diagonal:

$$x_1^2 - 3x_3^2 + 9x_5 - x_7 = 3,$$
$$-2x_1 + 8x_2^2 - 3x_3 - 5x_4 = 11,$$
$$7x_1^2 - 9x_2^2 + 8x_3^2 = -32.$$

Veremos que a técnica fundamental para simplificar estas quadráticas e levá-las a estrutura diagonal consiste em "completar quadrados".

Se $n = 2$, uma equação quadrática sem termos cruzados tem a forma:

$$c_1 x_1 + c_2 x_2 + b_{11} x_1^2 + b_{22} x_2^2 = a. \tag{4.8}$$

E, para n geral:

$$c_1 x_1 + \ldots + c_n x_n + b_{11} x_1^2 + \ldots + b_{nn} x_n^2 = a.$$

Para quadráticas de duas variáveis sem termos cruzados, vamos analisar três casos:

(a) $b_{11} \neq 0,\ b_{22} \neq 0$;

(b) $b_{11} \neq 0,\ b_{22} = 0$;

(c) $b_{11} = 0,\ b_{22} = 0$.

No caso (c), trata-se de uma equação linear. Seu conjunto de soluções (se $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$) é a reta

$$c_1 x_1 + c_2 x_2 = a. \tag{4.9}$$

Como já vimos, para cada constante diferente a, o conjunto de $I\!R^2$ definido por (4.9) será uma reta ortogonal a $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$.

Mais interessantes são os casos (a) e (b). A simplificação de (4.8) em cada um destes três casos passa pela técnica de *completar quadrados.* No Caso (a) observamos que

$$b_{11}x_1^2 + c_1x_1 = b_{11}\left(x_1^2 + \frac{c_1}{b_{11}}x_1\right)$$

$$= b_{11}\left(x_1^2 + \frac{c_1}{b_{11}}x_1 + \left(\frac{c_1}{2b_{11}}\right)^2 - \left(\frac{c_1}{2b_{11}}\right)^2\right)$$

$$= b_{11}\left(x_1 + \frac{c_1}{2b_{11}}\right)^2 - \frac{c_1^2}{4b_{11}}.$$

Analogamente,

$$b_{22}x_2^2 + c_2x_2 = b_{22}\left(x_2 + \frac{c_2}{2b_{22}}\right)^2 - \frac{c_2^2}{4b_{22}}.$$

Portanto, no Caso (a) temos:

$$c^T x + x^T Bx = b_{11}\left(x_1 + \frac{c_1}{2b_{11}}\right)^2 + b_{22}\left(x_2 + \frac{c_1}{2b_{22}}\right)^2 - \frac{c_1^2}{4b_{11}} - \frac{c_2^2}{4b_{22}}.$$

Definindo as "novas variáveis":

$$y_1 = x_1 + \frac{c_1}{2b_{11}},$$

$$y_2 = x_2 + \frac{c_1}{2b_{22}}$$

e a "nova constante"

$$\beta = -\frac{c_1^2}{4b_{11}} - \frac{c_2^2}{4b_{22}}$$

temos

$$c^T x + x^T Bx = (b_{11}y_1^2 + b_{22}y_2^2) + \beta. \tag{4.10}$$

Esta já é uma forma bastante simples, que analisaremos mais adiante.

No Caso (b), temos

$$c^T x + x^T Bx = c_1 x_1 + c_2 x_2 + b_{11} x_1^2.$$

Como no Caso (a), temos que

$$b_{11} x_1^2 + c_1 x_1 = b_{11}\left(x_1 + \frac{c_1}{2b_{11}}\right)^2 - \frac{c_1^2}{4b_{11}}$$

mas é impossível completar quadrados em relação à variável x_2. Portanto,

$$c^T x + x^T Bx = b_{11}\left(x_1 + \frac{c_1}{2b_{11}}\right)^2 + c_2 x_2 - \frac{c_1^2}{4b_{11}}.$$

Definindo a nova variável

$$y_1 = x_1 + \frac{c_1}{2b_{11}}$$

e a nova constante

$$\beta = -\frac{c_1^2}{4b_{11}}$$

obtemos

$$c^T x + x^T Bx = b_{11} y_1^2 + c_2 x_2 + \beta.$$

Definindo, por razões estéticas, $y_2 = x_2$, temos:

$$c^T x + x^T Bx = b_{11} y_1^2 + c_2 y_2 + \beta. \tag{4.11}$$

Esta expressão já é suficientemente simples para uma análise geomêtrica.

Generalizando o procedimento anterior para quadráticas de n variáveis sem termos cruzados vemos que sempre podemos encontrar novas variáveis $y_1, \ldots, y_n$ tais que, nas novas variáveis, a equação quadrática tem a forma:

$$\sum_{i=1}^{q} \alpha_i y_i^2 - \sum_{i=1}^{r} \alpha_{q+i} y_{q+i}^2 + \sum_{i=1}^{s} c_i y_{q+r+i} = a, \tag{4.12}$$

onde $\alpha_1, \ldots, \alpha_{q+r} > 0$.

Ou seja, q variáveis aparecem elevadas ao quadrado com coeficiente positivo, r variáveis aparecem elevadas ao quadrado com coeficiente negativo e s variáveis aparecem linearmente. Em (4.12) consideramos que as variáveis que aparecem ao quadrado e com coeficiente positivo são as q primeiras, e assim por diante, apenas para simplificar a notação.

Um esclarecimento relativo à notação de somatória merece ser feito. Se, digamos, não há nenhuma variável que apareça elevada ao quadrado com coeficiente negativo, então $r = 0$ e a segunda somatória em (4.12) iria desde $i = 1$ até $i = 0$. Neste caso, será sub-entendido que tal somatória não aparece. De fato, essa é a convenção adotada sobre o símbolo de somatória em toda a literatura matemática.

A mudança de variáveis que permite transformar uma quadrática geral sem termos cruzados na forma (4.12) é uma mera mudança de origem, do tipo:

$$y_i = x_i - a_i,\ i = 1, \dots, q + r,\ \ y_j = x_j,\ j = q + r + 1, \dots, n. \tag{4.13}$$

4.5 Cônicas com estrutura diagonal

Chamaremos *cônicas* aos conjuntos de soluções de equações quadráticas com duas variáveis. A razão para este nome é que a forma geométrica de todas as cônicas resulta de cortar o cone mostrado na Figura 4.14 com diferentes planos.

De acordo com (4.12), as cônicas com estrutura diagonal possíveis são as seguintes:

$$\alpha_1 x_1^2 + \alpha_2 x_2^2 = a, \tag{4.14}$$

$$\alpha_1 x_1^2 - \alpha_2 x_2^2 = a, \tag{4.15}$$

$$\alpha_1 x_1^2 + c_2 x_2 = a, \tag{4.16}$$

A equação

$$-\alpha_1 x_1^2 + c_2 x_2 = a$$

se reduz a (4.16) multiplicando por -1. Portanto, não a consideraremos como um caso separado. O mesmo acontece com a equação

$$-\alpha_1 x_1^2 - \alpha_2 x_2^2 = a$$

em relação a (4.14). Também não consideraremos casos separados àqueles que podem ser levados a um dos casos mencionados trocando os nomes das variáveis.

4.5.1 Elipses

Consideremos a equação

$$\alpha_1 x_1^2 + \alpha_2 x_2^2 = a, \tag{4.17}$$

onde $\alpha_1 > 0, \alpha_2 > 0$. Como o membro da esquerda de (4.17) sempre é maior ou igual que zero, o conjunto é vazio se $a < 0$. Esse é o caso, por exemplo, da equação $x_1^2 + x_2^2 = -1$. Mais ainda, se $a = 0$ o conjunto de soluções de (4.17) é um único ponto ($x_1 = x_2 = 0$).

Portanto, o caso geomêtricamente interessante aparece quando $a > 0$. Nesse caso, dividindo membro a membro por a, a equação pode ser formulada assim:

$$\frac{\alpha_1}{a}x_1^2 + \frac{\alpha_2}{a}x_2^2 = 1. \tag{4.18}$$

Daqui resultam as seguintes quatro soluções evidentes:

$$\begin{pmatrix} \sqrt{\frac{a}{\alpha_1}} \\ 0 \end{pmatrix}, \begin{pmatrix} -\sqrt{\frac{a}{\alpha_1}} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \sqrt{\frac{a}{\alpha_2}} \end{pmatrix}, \begin{pmatrix} 0 \\ -\sqrt{\frac{a}{\alpha_2}} \end{pmatrix}. \tag{4.19}$$

Entretanto, o conjunto de soluções de (4.18) não se esgota em (4.19). Com efeito, podemos reformular (4.19) como

$$\left\| \begin{pmatrix} \sqrt{\frac{\alpha_1}{a}}x_1 \\ \sqrt{\frac{\alpha_2}{a}}x_2 \end{pmatrix} \right\| = 1. \tag{4.20}$$

Ou seja, o conjunto de pontos que satisfazem a equação (4.18) é o conjunto de pontos $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ tal que a distância de $\begin{pmatrix} \sqrt{\frac{\alpha_1}{a}}x_1 \\ \sqrt{\frac{\alpha_2}{a}}x_2 \end{pmatrix}$ a $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ é igual a 1. Não é difícil se convencer de que estes são os pontos da elipse cujos "pontos extremos" são os dados em (4.19). Veja Figura 4.3.

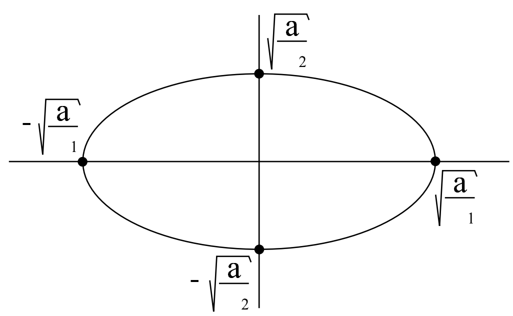

Figura 4.3. **Elipse.**

Uma elipse pode ser definida como o conjunto de pontos do plano tal que a soma das distâncias a dois pontos dados, chamados *focos*, é constante. O leitor interessado nesta caracterização pode consultar o Capítulo 16 de [2].

4.5.2 Hipérboles

Consideremos agora

$$\alpha_1 x_1^2 - \alpha_2 x_2^2 = a, \tag{4.21}$$

com $\alpha_1 > 0$ e $\alpha_2 > 0$.

É interessante analisar o conjunto (4.21) quando $a = 0$. Neste caso, a equação (4.21) é equivalente a

$$(\sqrt{\alpha_1}x_1 + \sqrt{\alpha_2}x_2)(\sqrt{\alpha_1}x_1 - \sqrt{\alpha_2}x_2) = 0. \tag{4.22}$$

Claramente, o conjunto de soluções de (4.22) é a união das retas definidas por

$$\sqrt{\alpha_1}x_1 + \sqrt{\alpha_2}x_2 = 0 \quad \text{e} \quad \sqrt{\alpha_1}x_1 - \sqrt{\alpha_2}x_2 = 0. \tag{4.23}$$

Veja Figura 4.4. Ou seja, se $a = 0$ as soluções de (4.21) são os pontos das das retas que passam pela origem e são ortogonais, uma delas a $\begin{pmatrix}\sqrt{\alpha_1}\\ \sqrt{\alpha_2}\end{pmatrix}$ e a outra a $\begin{pmatrix}\sqrt{\alpha_1}\\ -\sqrt{\alpha_2}\end{pmatrix}$. Estas retas serão chamadas *assíntotas*.

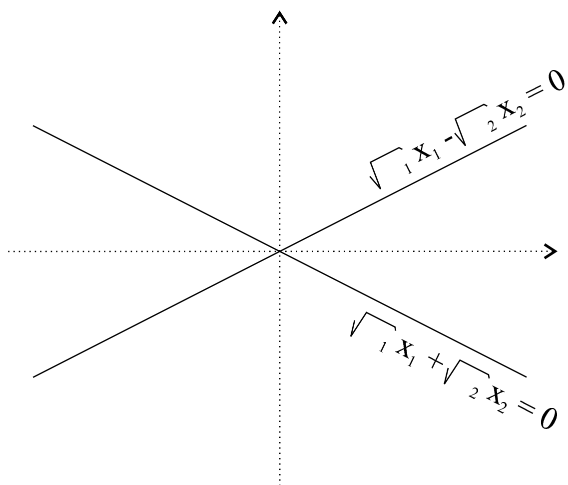

Figura 4.4. **Assíntotas.**

As assíntotas dividem o plano em duas regiões. Em uma delas $\sqrt{\alpha_1}x_1 + \sqrt{\alpha_2}x_2$ e $\sqrt{\alpha_1}x_1 - \sqrt{\alpha_2}x_2$ têm o mesmo sinal e na outra essas expressões têm sinais opostos. Veja Figura 4.5. Na primeira região,

$$\alpha_1 x_1^2 - \alpha_2 x_2^2 \geq 0$$

enquanto, na segunda,

$$\alpha_1 x_1^2 - \alpha_2 x_2^2 \leq 0.$$

Portanto, as soluções das equações (4.21) para $a > 0$ estão totalmente contidos na primeira região e as soluções de (4.21) para $a < 0$ estão totalmente contidos na

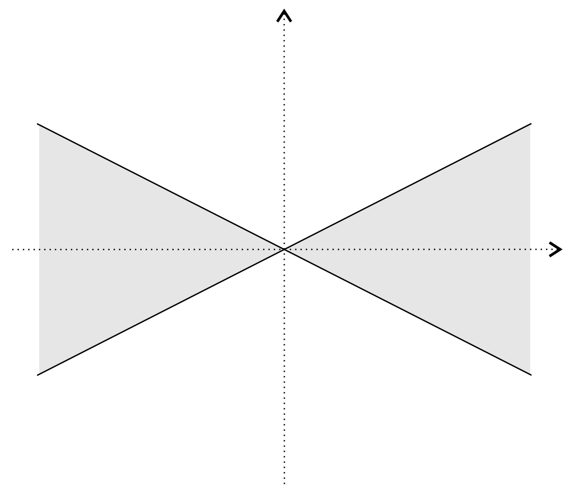

Figura 4.5. **Duas regiões limitadas pelas assíntotas.**

segunda região. Já vimos que, se $a = 0$, o conjunto de soluções está formado pelas duas assíntotas.

Suponhamos que $a > 0$ e tentemos visualizar o conjunto de soluções (4.21) nesse caso. É útil verificar que os pontos

$$\begin{pmatrix} \sqrt{a/\alpha_1} \\ 0 \end{pmatrix} \text{ e } \begin{pmatrix} -\sqrt{a/\alpha_1} \\ 0 \end{pmatrix} \tag{4.24}$$

pertencem ao conjunto de soluções.

Agora, colocando x_2 em evidência em (4.21), obtemos (se $\alpha_1 x_1^2 - a \geq 0!$):

$$x_2 = \sqrt{\frac{\alpha_1 x_1^2 - a}{\alpha_2}} \text{ ou } x_2 = -\sqrt{\frac{\alpha_1 x_1^2 - a}{\alpha_2}}, \tag{4.25}$$

Quando $\alpha_1 x_1^2 - a < 0$ (ou seja, quando $-\sqrt{a/\alpha_1} < x_1 < \sqrt{a/\alpha_1}$) não é possível pôr em evidência x_2 em função de x_1, isto é, não existe nenhuma solução da equação nessa faixa do plano. Veja Figura 4.6. Ao mesmo tempo, para $x_1 < -\sqrt{a/\alpha_1}$ e para $x_1 > \sqrt{a/\alpha_1}$ existem dois valores de x_2 tais que $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ são soluções. Finalmente, quando x_1 toma os valores $-\sqrt{a/\alpha_1}$ ou $\sqrt{a/\alpha_1}$ o único valor possível para x_2 é 0.

Para valores grandes e positivos de x_1, a expressão $x_2 = \sqrt{\frac{\alpha_1 x_1^2 - a}{\alpha_2}}$ toma a forma:

$$x_2 = x_1 \sqrt{\frac{\alpha_1 - a/x_1^2}{\alpha_2}}.$$

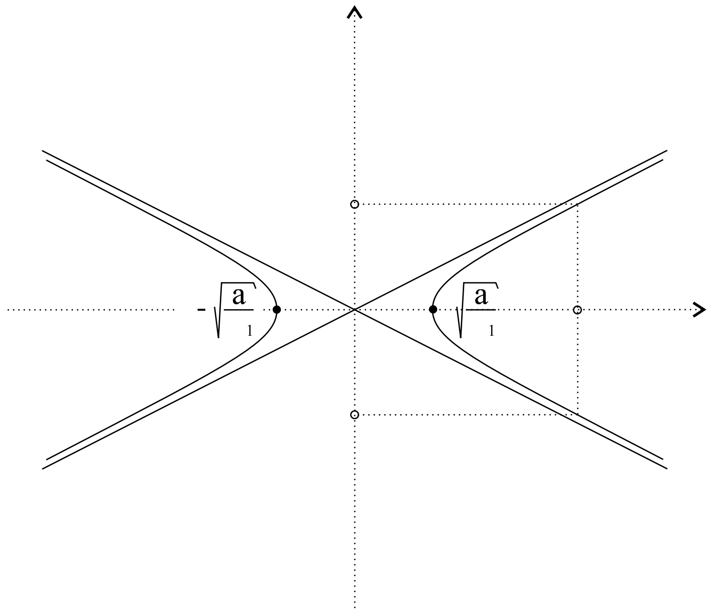

Figura 4.6. **Hipérbole.**

Portanto, quanto maior for x_1, mais "se parece" esta equação com a assíntota

$$x_2 = \sqrt{\frac{\alpha_1}{\alpha_2}}x_1.$$

Analogamente, para x_1 grande,

$$-\sqrt{\frac{\alpha_1 x_1^2 - a}{\alpha_2}} \approx -\sqrt{\frac{\alpha_1}{\alpha_2}}x_1.$$

Raciocinando da mesma maneira, para valores grandes de $-x_1$, a expressão $x_2 = \sqrt{\frac{\alpha_1 x_1^2 - a}{\alpha_2}}$ se aproxima de $-\sqrt{\frac{\alpha_1}{\alpha_2}}x_1$ e a expressão $x_2 = -\sqrt{\frac{\alpha_1 x_1^2 - a}{\alpha_2}}$ se aproxima de $\sqrt{\frac{\alpha_1}{\alpha_2}}x_1$.

Em outras palavras, quanto maior for $|x_1|$, mais parecida será a curva de soluções com a assíntota "correspondente". Resumindo, para $a > 0$ o conjunto de soluções é o par de curvas indicado na Figura 4.6. Curvas como estas se denominam *hipérboles*.

Por último, podemos analisar as soluções de (4.21) no caso $a < 0$. Evidentemente, o conjunto de pontos que verificam $\alpha_1 x_1^2 - \alpha_2 x_2^2 = a$ para $a < 0$ é o mesmo que o conjunto de pontos que verificam

$$-\alpha_1 x_1^2 + \alpha_2 x_2^2 = -a.$$

Como $-a > 0$ podemos repetir os argumentos acima trocando as variáveis x_1 e x_2. Na Figura 4.7 mostramos, juntas, as hipérboles $\alpha_1 x_1^2 - \alpha_2 x_2^2 = a$ e $-\alpha_1 x_1^2 + \alpha_2 x_2^2 = a$ para $a > 0$.

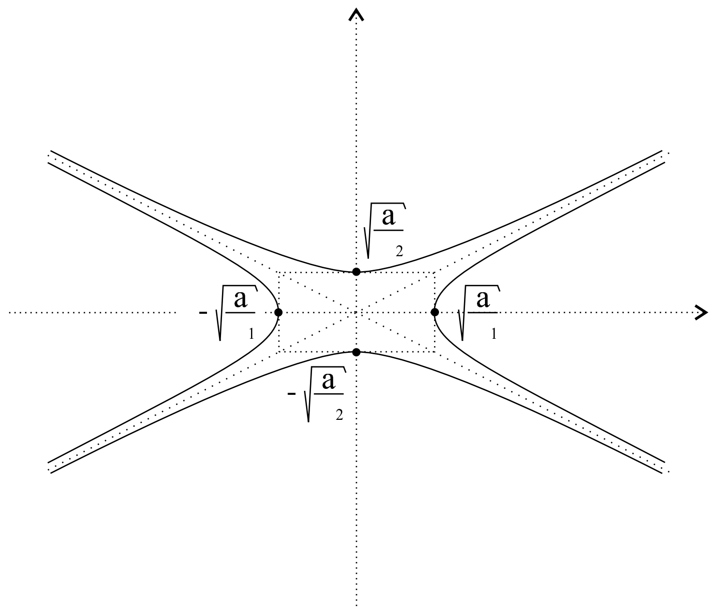

Figura 4.7. **Duas hipérboles.**

Uma hipérbole pode ser definida como o conjunto de pontos do plano tal que a diferença das distâncias a dois pontos dados, chamados *focos*, é constante. Como no caso das elipses, o leitor interessado nesta caracterização pode consultar o Capítulo 17 de [2].

4.5.3 Parábolas

Consideremos a forma (4.16).

Nossa tarefa consiste em analisar as soluções da equação

$$\alpha_1 x_1^2 + c_2 x_2 = a. \tag{4.26}$$

Devemos distinguir dois casos:

(a) $c_2 = 0$;

(b) $c_2 \neq 0$.

No caso (a), a equação (4.26) toma a forma

$$x_1^2 = \frac{a}{\alpha_1}. \tag{4.27}$$

O conjunto de pontos que verifica esta equação é vazio se $a < 0$ (lembremos que $\alpha_1 > 0$). Se $a = 0$, este conjunto é a reta $x_1 = 0$. Finalmente, se $a > 0$, o conjunto de pontos que verifica (4.27) está formado pelas duas retas:

$$x_1 = \sqrt{\frac{a}{\alpha_1}} \text{ e } x_1 = -\sqrt{\frac{a}{\alpha_1}}.$$

Veja Figura 4.8. Tanto neste caso como no caso da reta $x_1 = 0$, o conjunto de soluções é um cilindro originado na extensão de dois pontos e de um ponto em $I\!R^1$ respectivamente.

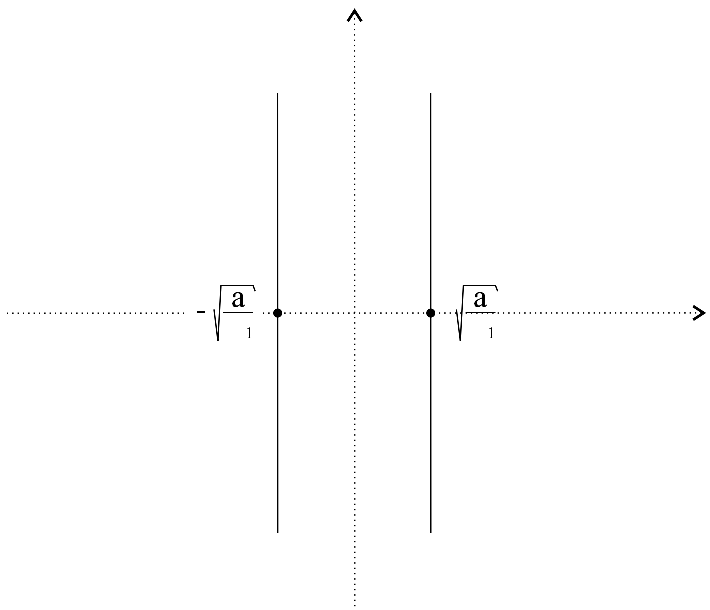

Figura 4.8. **Conjunto de soluções: duas retas paralelas.**

No caso (b), a equação (4.26) pode ser escrita:

$$x_2 = -\frac{\alpha_1}{c_2}x_1^2 + \frac{a}{c_2}. \tag{4.28}$$

A expressão (4.28) é a de uma parábola com vértice em $\begin{pmatrix} 0 \\ a/c_2 \end{pmatrix}$. Se $-\frac{\alpha_1}{c_2} > 0$ a parábola é *convexa* e se $-\frac{\alpha_1}{c_2} < 0$ a parábola é *côncava*. Veja Figura 4.9 (onde $a = 0$).

A parábola pode ser caracterizada como o conjunto dos pontos do plano que são equidistantes de um ponto e uma reta dada. De novo, o leitor é remitido ao Capítulo 18 de [2] para o desenvolvimento desta idéia.

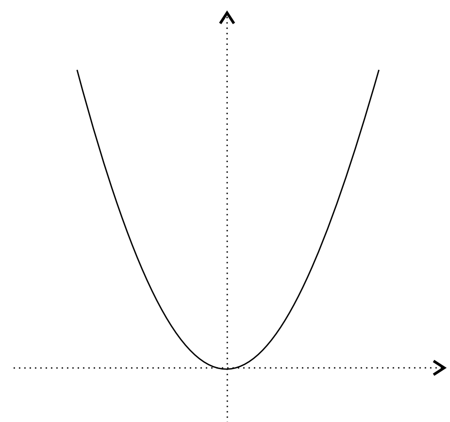

Figura 4.9. **Parábola convexa.**

4.5.4 Classificação rápida de cônicas

Às vezes, uma breve olhada na matriz B que define uma equação quadrática de duas variáveis nos permite - se não reconhecer uma cônica - ao menos descartar algumas possibilidades.

Já conhecemos os seguintes princípios:

(a) Se os autovalores de B são maiores que zero, o conjunto de soluções pode ser vazio, pode ser um ponto ou pode ser uma elipse. Mas, certamente, não pode ser uma hipérbole, nem uma parábola, nem duas retas, nem uma reta. O mesmo acontece se os autovalores de B são menores que zero, pois este caso se reduz ao anterior.

(b) Se um autovalor é maior que zero e o outro é menor que zero, o conjunto de soluções pode ser uma hipérbole ou duas retas que se cortam (assíntotas). Mas elipses, parábolas, o vazio, o ponto único, a reta e as duas retas paralelas estão descartadas.

O Teorema 3.11 nos permite reconhecer rapidamente o Caso (a) em algumas circunstâncias. Por exemplo, a matriz $\begin{pmatrix} 3 & -2 \\ -2 & 4 \end{pmatrix}$ têm autovalores positivos porque, de acordo com o resultado mencionado, todos eles devem satisfazer

$$|\lambda - 3| \leq 2 \text{ ou } |\lambda - 4| \leq 2,$$

o que exclui totalmente valores negativos de λ.

A outra ferramenta importante para a identificação rápida é o Corolário 3.12e. Por exemplo, considere a matriz $\begin{pmatrix} 2 & 3 \\ 3 & 1 \end{pmatrix}$. Apesar de que os elementos de sua diagonal são positivos, o Teorema 3.11 não garante que seus (um ou dois) autovalores são não-negativos. Entretanto, o Corolário 3.12e garante que a matriz tem, pelo menos, um autovalor positivo.

Mais interessante é o caso da matriz $\begin{pmatrix} 1 & 5 \\ 5 & -1 \end{pmatrix}$. De novo, o Teorema 3.11 não fornece nenhuma informação segura. Entretanto, o Corolário 3.12e garante que a matriz tem um autovalor positivo e um autovalor negativo. Como conseqüência, o conjunto de soluções da equação quadrática correspondente somente pode ser uma hipérbole ou duas retas que se cortam.

4.5.5 Um exemplo detalhado

Consideremos a equação quadrática

$$3x_1 - 2x_2 + 2x_1^2 - 6x_1x_2 + 4x_2^2 = 0.$$

Nos termos de (4.2), temos:

$$c = \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \; B = \begin{pmatrix} 2 & -3 \\ -3 & 4 \end{pmatrix}, \; a = 0.$$

A inspecção preliminar da matriz B nos mostra que deve haver algum autovalor positivo, pois os termos da são positivos, mas o Teorema 3.11 não chega a garantir que os dois autovalores sejam positivos.

Para calcular os autovalores, formulamos a equação característica

$$\det\,(B - \lambda I) = (b_{11} - \lambda)(b_{22} - \lambda) - b_{12}^2 = 0.$$

Esta equação nos dá as seguintes duas soluções:

$$\lambda_1 = 6.162278 \text{ e } \lambda_2 = -0.1622777.$$

Estes são os dois autovalores da matriz B. O fato de que um autovalor é positivo e o outro negativo já nos antecipa que o conjunto de solucoes será uma hipérbole ou duas retas que se cortam em um ponto.

O autovetor v associado a λ_1 deve satisfazer

$$(B - \lambda_1 I)v = 0,$$

ou seja,

$$-4.162278v_1 - 3v_2 = 0$$

e

$$-3v_1 - 2.162278v_2 = 0.$$

Estas duas equacoes, descartando o erro de arredondamento infiltrado, devem ser equivalentes (em caso contrário a única solução de $(B - \lambda_1 I)v = 0$ seria $v = 0$) portanto uma delas pode ser descartada. Ficando apenas com a primeira, obtendo uma solução e dividindo por sua norma, para que o autovetor tenha norma unitária, chegamos ao autovetor

$$v = \begin{pmatrix} 0.5847103 \\ -0.8112422 \end{pmatrix}.$$

Procedendo da mesma maneira com λ_2, obtemos o autovetor

$$w = \begin{pmatrix} 0.8.112422 \\ 0.5.847103 \end{pmatrix}.$$

Como antecipado pela teoria, estes dois autovetores são ortogonais.

Observe que teriamos liberdade para escolher $-v$ em vez de v e $-w$ em vez de w. Por comodidade, sempre usaremos esta liberdade para fazer que o segundo autovetor seja produto de uma rotação de 90 graus do primeiro, no sentido antihorário. A razão para esta conveniência é que essa é a relação entre os vetores canônicos $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ e $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, o que facilitará nossa visualização geométrica.

Os vetores v e w constituirão a nova base e, nas coordenadas relativas a esta base, a equação quadrática terá estrutura diagonal.

Escrevemos

$$Q = (v, w) = \begin{pmatrix} 0.5847103 & 0.8.112422 \\ -0.8112422 & 0.5.847103 \end{pmatrix}.$$

Portanto,

$$Q^T = (v, w)^T = \begin{pmatrix} v^T \\ w^T \end{pmatrix} = \begin{pmatrix} 0.5847103 & -0.8.112422 \\ 0.8112422 & 0.5.847103 \end{pmatrix}.$$

Para eliminar o termo cruzado a mudança de variáveis recomendada é

$$y = Q^T x,$$

ou seja,

$$y_1 = 0.5847103x_1 - 0.8.112422x_2,$$

$$y_2 = 0.8112422x_1 + 0.5.847103x_2.$$

Dessa maneira, $x = Qy$ (ou seja

$$x_1 = 0.5847103y_1 + 0.8.112422y_2,$$

$$x_2 = -0.8112422y_1 + 0.5.847103y_2)$$

e

$$c^T x + x^T Bx = c^T x + x^T Q^T \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} Qx = c^T Qy + y^T \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} y$$

$$= c_{novo}^T y + y^T \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} y,$$

onde

$$c_{novo} = Q^T c = \begin{pmatrix} 0.5847103 & -0.8.112422 \\ 0.8112422 & 0.5.847103 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 3.376615 \\ 1.264306 \end{pmatrix}.$$

Portanto a equacao nas variáveis y_1, y_2 é:

$$3.376615y_1 + 1.264306y_2 + 6.162278y_1^2 - 0.1622777y_2^2 = 0.$$

Completando quadrados de maneira que cada variável apareça uma única vez na equacao, obtemos:

$$6.162278(y_1 + 0.2739746)^2 - 0.1622777(y_2 - 3.895502)^2 = -2.$$

Portanto, se definimos as novas variáveis

$$z_1 = y_1 + 0.2739746 = 0.5847103x_1 - 0.8.112422x_2 + 0.2739746, \tag{4.29}$$

$$z_2 = y_2 - 3.895502 = 0.8112422x_1 + 0.5.847103x_2 - 3.89502, \tag{4.30}$$

a equacao toma a forma

$$6.162278z_1^2 - 0.1622777z_2^2 = -2.$$

Observe que, enquanto a primeira mudança de variáveis envolveu uma rotação dos eixos coordenados, a segunda foi uma mudança de origem.

Claramente, a última equacao é a equacao de uma hipérbole, cujas assíntotas são

$$2.48239z_1 - 0.402837z_2 = 0$$

e

$$2.48239z_1 + 0.402837z_2 = 0.$$

Para obter as equacoes destas assíntotas nas variáveis originais x_1, x_2 é suficiente substituír z_1 e z_2 pelas fórmulas (4.29) e (4.30). Observe que sua interseccao acontece no ponto $x_1 = 3, x_2 = 2.5$.

O leitor é convidado a acompanhar todos estes cálculos com desenhos.

4.6 Redução a uma única variável linear

Os conceitos mostrados até agora são suficientes para analisar equações quadráticas em duas variáveis mas, para entender equações com mais de duas variáveis, algumas reduções são ainda necessárias.

O processo que conduziu a forma (4.12) nos autoriza a concentrar-nos em equações da forma

$$\sum_{i=1}^{q} \alpha_i x_i^2 - \sum_{i=1}^{r} \alpha_{q+i} x_{q+i}^2 + \sum_{i=1}^{s} c_i x_{q+r+i}. \tag{4.31}$$

onde $\alpha_1, \ldots, \alpha_{q+r} > 0$. Com efeito, pelas duas seções anteriores, depois de uma mudança de base ortonormal seguida de uma mudança de origem, toda equação quadrática pode ser vista desta forma. A presença de várias variáveis que aparecem linearmente (ou seja, com $c_i \neq 0$) dificultará a visualização. Portanto, nesta seção veremos como reduzir o caso em que há mais de uma variável que aparece linearmente a um caso em que somente uma variável aparece dessa maneira.

Se $s = 0$, de acordo com a convenção sobre a somatória já mencionada, nenhuma variável aparece linearmente. Claramente, o caso em que todos los coeficientes c_i são iguais a zero se reduz trivialmente ao caso $s = 0$ e, portanto, não nos interessa. Portanto, vamos focar a atenção na situação em que $s > 0$ e algum dos coeficientes $c_1, \ldots, c_s$ é diferente de zero. Chamemos

$$\tilde{c} = \begin{pmatrix} c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_s \end{pmatrix} \in I\!R^s.$$

Logo, pela suposição acima, $c \neq 0$. Definamos, também,

$$\tilde{x} = \begin{pmatrix} x_{q+r+1} \\ \cdot \\ \cdot \\ \cdot \\ x_{q+r+s} \end{pmatrix} \in I\!R^s.$$

Então, a equação (4.31) é

$$\alpha_1 x_1^2 + \ldots + \alpha_q x_q^2 - \alpha_{q+1} x_{q+1}^2 - \ldots - \alpha_{q+r} x_{q+r}^2 + \tilde{c}^T \tilde{x} = a \tag{4.32}$$

Seja $Q \in I\!R^{s \times s}$ unitária tal que sua primeira coluna é $\tilde{c}/\|\tilde{c}\|$. Como $QQ^T = I_s$, a equação (4.32) é:

$$\alpha_1 x_1^2 + \ldots + \alpha_q x_q^2 - \alpha_{q+1} x_{q+1}^2 - \ldots - \alpha_{q+r} x_{q+r}^2 + (Q^T \tilde{c})^T Q^T \tilde{x} = a. \tag{4.33}$$

Seja $\tilde{y} = Q^T \tilde{x}$. Pela definição de Q, a única coordenada diferente de zero de $Q^T \tilde{c}$ é a primeira. Chamando α_{q+r+1} a esta coordenada, vemos que (4.33) toma a forma

$$\alpha_1 x_1^2 + \ldots + \alpha_q x_q^2 - \alpha_{q+1} x_{q+1}^2 - \ldots - \alpha_{q+r} x_{q+r}^2 + \alpha_{q+r+1} \tilde{y}_1 = a. \tag{4.34}$$

Para resumir: partindo da equação (4.31) e fazendo a mudança de variáveis definida por

$$y_1 = x_1, \ldots, y_{q+r} = x_{q+r},$$

$$\begin{pmatrix} y_{q+r+1} \\ \cdot \\ \cdot \\ \cdot \\ y_{q+r+s} \end{pmatrix} = Q^T \begin{pmatrix} x_{q+r+1} \\ \cdot \\ \cdot \\ \cdot \\ x_{q+r+s} \end{pmatrix},$$

e

$$y_{q+r+s+1} = x_{q+r+s+1}, \ldots, y_n = x_n$$

a equação (4.12) toma a forma

$$\alpha_1 y_1^2 + \ldots + \alpha_q y_q^2 - \alpha_{q+1} y_{q+1}^2 - \ldots - \alpha_{q+r} y_{q+r}^2 + \alpha_{q+r+1} y_{q+r+1} = a. \tag{4.35}$$

As novas coordenadas de x são suas coordenadas com respeito à base dada pelas colunas da matriz definida, por blocos, da seguinte maneira:

$$Q_1 = \begin{pmatrix} I_{q+r} & 0 & 0 \\ 0 & Q & 0 \\ 0 & 0 & I_{n-(q+r+s)} \end{pmatrix},$$

a qual, como pode ser facilmente verificado, é uma matriz unitária.

4.7 Redução a uma forma canônica

Vimos que com uma mudança de variáveis a equação quadrática fica sem termos cruzados. Depois, com uma mudança de origem, cada variável aparece uma única vez na equação. Por último uma mudança restrita a algumas variáveis pode ser necessária para que haja uma única variável que aparece linearmente.

As mudanças de variáveis de que falamos acima foram mudanças “ortonormais” no sentido de que as novas bases sempre foram bases com vetores unitários e ortogonais entre si.

Assim, ficamos autorizados a nos concentrarmos em equações da forma

$$\alpha_1 x_1^2 + \ldots + \alpha_q x_q^2 - \alpha_{q+1} x_{q+1}^2 - \ldots - \alpha_{q+r} x_{q+r}^2 + \alpha_{q+r+1} x_{q+r+1} = a, \tag{4.36}$$

onde $\alpha_1, \ldots, \alpha_{q+r} > 0$ e α_{q+r+1} pode ser zero ou não. (De fato, se $q + r = n$ este último termo pode não existir.)

Para simplificar ainda mais nossa equação e facilitar sua visualização, vamos fazer uma nova mudança de variáveis. Mas, a diferença das anteriores, esta mudança não conservará intactas as distâncias e representará, de fato, uma mera mudança de escalas. A base que corresponde às novas coordenadas estará composta por vetores ortogonais mas não necessariamente unitários.

As novas variáveis $y_1, \ldots, y_{q+r}$ serão definidas assim:

$$y_1 = \sqrt{\alpha_1} x_1, \ldots, y_{q+r} = \sqrt{\alpha_{q+r}} x_{q+r}. \tag{4.37}$$

Se o termo $\alpha_{q+r+1}x_{q+r+1}$ aparece na equação com $\alpha_{q+r+1} \neq 0$, a mudança de variáveis inclui

$$y_{q+r+1} = |\alpha_{q+r+1}|x_{q+r+1}.$$

Finalmente, se $q + r + 1 < n$, incluimos:

$$y_j = x_j \text{ para todo } j > q + r + 1.$$

Nas novas variáveis, se a variável linear não aparece, a equação (4.36) toma a forma

$$y_1^2 + \ldots + y_q^2 - y_{q+1}^2 - \ldots - y_{q+r}^2 = a. \tag{4.38}$$

Entretanto, se a variável linear aparece, a forma canônica da equação é:

$$y_1^2 + \ldots + y_q^2 - y_{q+1}^2 - \ldots - y_{q+r}^2 + y_{q+r+1} = a. \tag{4.39}$$

ou

$$y_1^2 + \ldots + y_q^2 - y_{q+1}^2 - \ldots - y_{q+r}^2 - y_{q+r+1} = a. \tag{4.40}$$

4.8 Quádricas canônicas

Seguindo o roteiro da Seção 4, analisaremos aqui as formas canônicas das quadráticas de 3 variáveis. Os conjuntos de soluções de equações de esse tipo se denominam *quádricas*. Veremos que, em geral, são superfícies de $I\!R^3$.

Vimos que, com uma mudança de origem e através do procedimento de completar quadrados todas as quadráticas sem termos cruzados podem ser reduzidas a uma forma na qual cada variável aparece, como máximo, uma única vez, elevada ao quadrado ou em forma linear. Depois de outra mudança ortonormal de coordenadas, não mais de uma variável aparece linearmente. Finalmente, com uma mudança de escala, chegamos às formas canônicas (4.38) e (4.39).

Eliminando as formas cuja análise se reduz obviamente a outras, as que devem ser analisadas no caso de 3 variáveis são:

$$x_1^2 + x_2^2 + x_3^2 = a. \tag{4.41}$$

$$x_1^2 + x_2^2 - x_3^2 = a. \tag{4.42}$$

$$x_1^2 + x_2^2 - x_3 = a \text{ ou } x_1^2 + x_2^2 + x_3 = a \tag{4.43}$$

$$x_1^2 - x_2^2 - x_3 = a \text{ ou } x_1^2 - x_2^2 + x_3 = a. \tag{4.44}$$

Para ser justos, também devemos mencionar

$$x_1^2 + x_2^2 = a, \tag{4.45}$$

$$x_1^2 - x_2^2 = a, \tag{4.46}$$

$$x_1^2 + x_2 = a \text{ ou } x_1^2 - x_2 = a \tag{4.47}$$

e, até,

$$x_1^2 = a \tag{4.48}$$

e

$$x_1 = a. \tag{4.49}$$

Mas todos estes casos são cilindros que podem ser compreendidos em $I\!R^2$ e estendidos a $I\!R^3$ como indicado na Seção 4.2. O conjunto de soluções de (4.45) para $a > 0$ se denomina *cilindro elítico* e os conjuntos de soluções de (4.46) e (4.47) se chamam *cilindro hiperbólico* e *cilindro parabólico* respectivamente. O cilindro parabólico está representado na Figura 4.10.

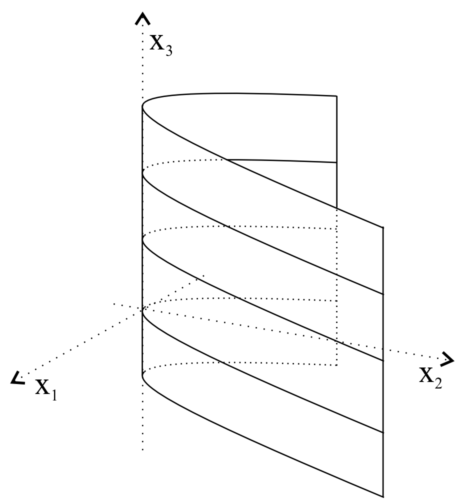

Figura 4.10. **Cilindro parabólico**

4.8.1 Elipsoides

Consideremos a forma canônica (4.41).

Como no caso das elipses em duas variáveis, o conjunto de soluções é vazio se $a < 0$ e se reduz a um único ponto (a origem) se $a = 0$.

O caso interessante é quando $a > 0$. Neste caso, a equação (4.41) pode ser escrita:

$$\|x\| = \sqrt{a},$$

portanto o conjunto de pontos que a satisfazem é a esfera de com centro na origem e rádio $\sqrt{a}$.

Não esqueça, entretanto, que este conjunto resultou uma esfera nas variáveis "escaladas" definidas em (4.37). Se não tivéssemos realizado essa mudança de variáveis (ou seja, nas variáveis "velhas") os conjunto seria um elipsoide. Veja Figura (4.11).

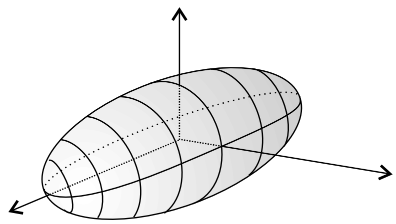

Figura 4.11. **Elipsoide.**

4.8.2 Cones e Hiperboloides

Analisemos as soluções de

$$x_1^2 + x_2^2 - x_3^2 = a. \tag{4.50}$$

Como no caso das hipérboles, o caso $a = 0$ é interessante e instrutivo. Com efeito, se $a = 0$ a equação (4.50) toma a forma

$$x_3^2 = x_1^2 + x_2^2. \tag{4.51}$$

O conjunto de pontos $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ que cumprem (4.51) é a união daqueles que satisfazem

$$x_3 = \sqrt{x_1^2 + x_2^2} \tag{4.52}$$

com os que satisfazem

$$x_3 = -\sqrt{x_1^2 + x_2^2}. \tag{4.53}$$

Por outra parte,

$$\sqrt{x_1^2 + x_2^2} = \left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\|,$$

portanto, em (4.52) o valor de x_3 é sempre a distância entre $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ e $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ e em (4.53) x_3 é essa mesma distância com sinal trocado.

Juntando estas informações, vemos que o conjunto (4.52) é o *cone* em forma de *sorvete* da Figura 4.12 e que o conjunto (4.53) é o sorvete invertido da Figura 4.13. Portanto, o conjunto de soluções de (4.51) é o *sorvete duplo* mostrado na Figura 4.14.

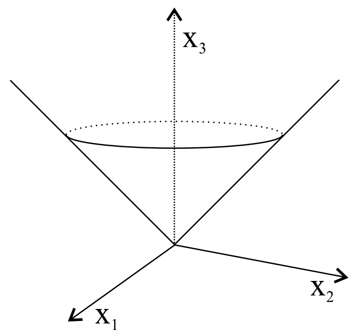

Figura 4.12. **Sorvete** $x_3 = \sqrt{x_1^2 + x_2^2}$**.**

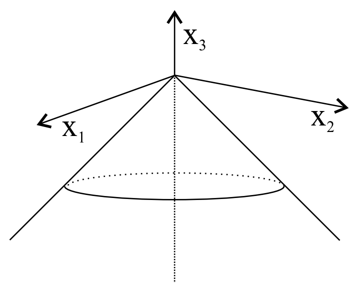

Figura 4.13. **Sorvete invertido** $x_3 = -\sqrt{x_1^2 + x_2^2}$**.**

Na parte de dentro deste sorvete duplo temos que

$$x_3^2 > x_1^2 + x_2^2,$$

enquanto, na parte de fora,

$$x_3^2 < x_1^2 + x_2^2.$$

Portanto, as soluções de (4.50) quando $a < 0$ estão na parte de dentro do sorvete duplo e os conjuntos de soluções com $a > 0$ estão na parte de fora.

Analisemos o conjunto de soluções de (4.50) para $a < 0$. Colocando em evidência x_3 temos que os pontos que verificam (4.50) é a união daqueles que verificam

$$x_3 = \sqrt{x_1^2 + x_2^2 - a} \tag{4.54}$$

com os que satisfazem

$$x_3 = -\sqrt{x_1^2 + x_2^2 - a}. \tag{4.55}$$

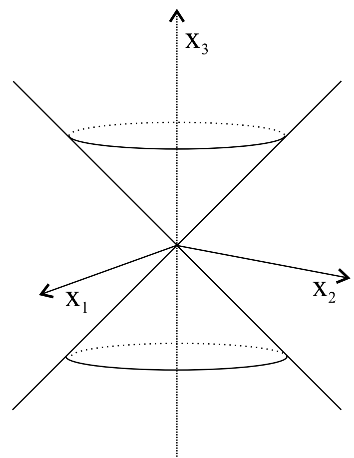

Figura 4.14. **Cone.**

Ou seja, para cada $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ existem dois valores de x_3 que fazem que $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ esteja no conjunto de soluções. Em particular, para $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ temos que

$$x_3 = \sqrt{-a} \text{ ou } x_3 = -\sqrt{-a}.$$

Veja Figura 4.15. Mais ainda, por argumentos similares aos usados no caso das hipérboles, quando x_1 e x_2 são muito grandes,

$$\sqrt{x_1^2 + x_2^2 - a} \approx \sqrt{x_1^2 + x_2^2},$$

o que significa que, nesse caso, o ponto $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ que está no conjunto de soluções se aproxima do ponto que tem as coordenadas x_1, x_2 no sorvete. Isto completa nossa intuição geomêtrica sobre as soluções de (4.50) para $c < 0$. Vemos que esta superfície tem duas partes (a superior e a inferior) disjuntas, pelo qual é denominada *hiperboloide de duas folhas.*

Se $a > 0$, também temos que as soluções de (4.50) são os pontos que satisfazem

$$x_3 = \sqrt{x_1^2 + x_2^2 - a} \tag{4.56}$$

junto com os que satisfazem

$$x_3 = -\sqrt{x_1^2 + x_2^2 - a}, \tag{4.57}$$

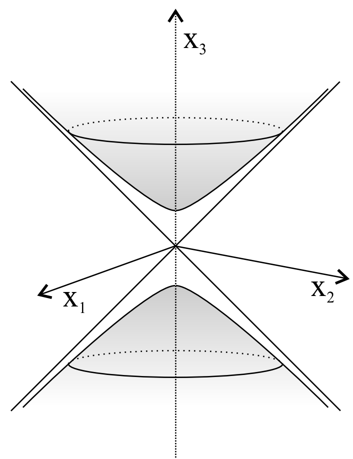

Figura 4.15. **Hiperbolóide de duas folhas.**

mas (4.56) e (4.57) somente têm sentido quando $x_1^2 + x_2^2 > a$. Quando $x_1^2 + x_2^2 = a$ temos que, necessariamente, $x_3 = 0$ enquanto que para $x_1^2 + x_2^2 > a$ (4.56) e (4.57) dão dois valores de x_3 com sinais trocados.

Fixando um $x_3 > 0$, os valores de x_1, x_2 que satisfacem (4.50) com $a > 0$ vêm dados correspondem aos pontos da circunferência $x_1^2 + x_2^2 = x_3^2 + a$ e a mesma circunferência é obtida "cortando" o conjunto (4.57) por um plano do tipo $x_3 =$ constante < 0. Finalmente, como no caso $a < 0$, quando $x_1^2 + x_2^2$ é muito grande, x_3 se aproxima a $\sqrt{x_1^2 + x_2^2}$ no caso de (4.56) e a $-\sqrt{x_1^2 + x_2^2}$ em (4.57).

Colocando estas informações juntas, o aspecto do conjunto de soluções de (4.50) quando $a > 0$ é o *hiperboloide de uma folha* mostrado na Figura 4.16.

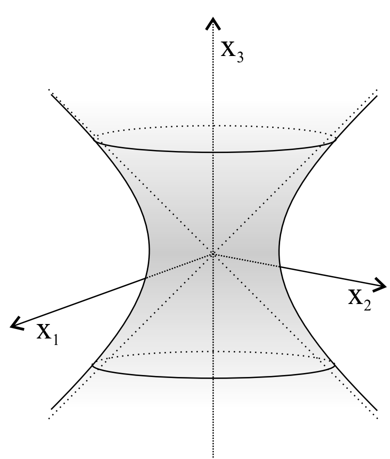

Figura 4.16. **Hiperbolóide de uma folha.**

4.8.3 Paraboloides elípticos

O paraboloide elíptico é a quádrica originada na equação (4.43). Ela pode ser escrita como

$$x_3 = x_1^2 + x_2^2 + a \text{ ou } x_3 = -(x_1^2 + x_2^2) + a. \tag{4.58}$$

Não é difícil se convencer de que os pontos que satisfazem (4.58) são resultado de fazer girar a parábola $x_3 = x_1^2$ (ou $x_3 = -x_1^2$) ao redor do eixo x_3.

4.8.4 Paraboloides hiperbólicos

Consideraremos somente a forma

$$x_1^2 - x_2^2 - x_3 = a \tag{4.59}$$

da equação (4.44). A análise de $x_1^2 - x_2^2 + x_3 = a$ se reduz ao de (4.59) trocando as variáveis x_1 e x_2 e o sinal de a.

Colocando em evidência x_3, a equação (4.59) se reduz a

$$x_3 = x_1^2 - x_2^2 - a. \tag{4.60}$$

Portanto, a intersecção desta superfície com o plano $x_2 = 0$ dá, no plano x_1x_3, a parábola convexa $x_3 = x_1^2 - a$. Mas a intersecção com o plano $x_1 = 0$ dá, no plano x_2x_3, a parábola côncava $x_3 = -x_2^2 - a$. Veja Figura 4.17 (na qual $a = 0$).

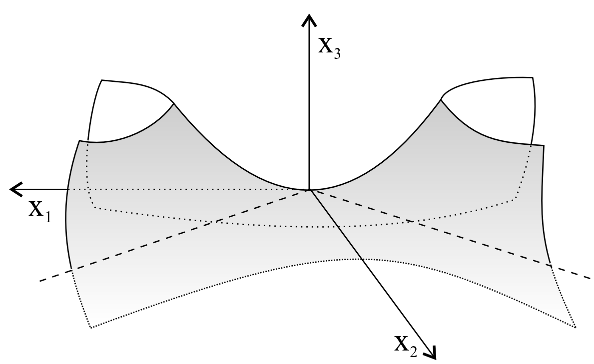

Figura 4.17. **Parabolóide hiperbólico.**

4.9 Exercícios

1. Considere a função quadrática dada por

$$f(x) = c^T x + x^T B x.$$

Definimos, para todo $z \in I\!R^n$,

$$\nabla f(z) = 2Bz + c.$$

Prove que, para todo $z \in I\!R^n, x \in I\!R^n$,

$$f(x) = f(z) + \nabla f(z)^T(x - z) + (x - z)^T B(x - z).$$

2. Classifique e desenhe os conjuntos de $I\!R^2$ definidos pelas equações abaixo. Em todos os casos, procure usar o Teorema 3.11 e o Corolário 3.12e para obter informação sobre o conjunto de soluções antes de começar a fazer contas pesadas.

   (a) $x^2 + y^2 = 1$

   (b)$x^2 + y^2 = 0$

   (c)$x^2 + y^2 - 5 = 0$

   (d)$x^2/9 + y^2/16 = 1$

   (e)$3x^2 + 18y^2 = 7$

   (f)$5x^2 + 9y^2 = -2$

   (g)$x^2 - y^2 = 1$

   (h)$x^2 - y^2 = 0$

   (i)$x^2 - y^2 - 7 = 0$

   (j)$y^2 - x^2 = 1$

   (k)$y^2 - x^2 = 0$

   (l)$y^2 - x^2 - 78 = 0$

   (m)$y^2 - x^2 + 14 = 0$

   (n)$y = x^2$

   (o)$x = y^2$

   (p)$y - x^2 - 4 = 0$

   (q)$x - y^2 + 7 = 0$

3. Classifique e desenhe os conjuntos de $I\!R^2$ definidos pelas equações abaixo:

   (a)$x^2 + y^2 + 6x + 8y - 5 = 0$

   (b)$x^2 + 2x + y^2 + 2y + 2 = 0$

   (c)$3x^2 + 5y^2 + 4x - 2y - 10 = 0$

   (d)$x^2 - y^2 - 4x + 2y + 2 = 0$

(e)$x^2 - 2x + y^2 - 4y + 5 = 0$

(f)$x^2 - y^2 - 7x - 5y - 8 = 0$

(g)$y^2 - x^2 - 2x - 9y - 8 = 0$

(h)$x^2 + 5x + y - 9 = 0$

(i)$y^2 + x - y + 16 = 0$

4. Escreva as seguintes expressões na forma

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

(a) $\alpha x^2 + \beta xy + \gamma y^2$

(b) $3x^2 + 5xy + 6y^2$

(c) $-3x^2 - xy$

5. Classifique e desenhe os conjuntos de $I\!R^2$ definidos pelas seguintes equações:

(a)$x^2 + y^2 + \frac{1}{3}xy + 6x + 8y - 5 = 0$

(b)$x^2 + 2x + y^2 + 2y + 2 + \frac{1}{5}xy = 0$

(c)$3x^2 + 5y^2 + 4x - 2y - 10 + xy = 0$

(d)$4x^2 - 4y^2 - 16x + 8y + 8 - xy = 0$

(e)$5x^2 - 10x + 5y^2 - 20y + 25 + xy = 0$

(f)$10x^2 - 10y^2 - 70x - 50y - 80 + xy = 0$

(g)$3y^2 - 3x^2 - 6x - 27y - 24 + xy = 0$

(h)$x^2 + 2xy + y^2 + 5x + y - 9 = 0$

(i)$3y^2 + 6xy + 3y^2 + 3x - 3y + 48 = 0$

6. Escreva a generalização da Seção 4.5.4 destas notas para o caso de quadráticas com 3 variáveis.

7. Classifique e desenhe os conjuntos de $I\!R^3$ definidos pelas seguintes equações. Em todos os casos, procure usar o Teorema 3.11 e o Corolário 3.12e para obter informação sobre o conjunto de soluções antes de começar a fazer contas pesadas.

(a) $x^2 + y^2 + z^2 = 1$

(b)$x^2 + y^2 + z^2 = 0$

(c)$x^2 + y^2 + z^2 - 5 = 0$

(d)$x^2/9 + y^2/16 + z^2/25 = 1$

(e)$3x^2 + 18y^2 + 4z^2 = 7$

(f)$5x^2 + 9y^2 + 4z^2 = -2$

(g)$x^2 - y^2 + z^2 = 1$

(h)$x^2 - y^2 + z^2 = 0$

(i)$x^2 - y^2 + z^2 - 7 = 0$

(j)$y^2 - x^2 + z^2 = 1$

(k)$y^2 - x^2 + z^2 = 0$

(l)$y^2 - x^2 + z^2 - 78 = 0$

(m)$y^2 - x^2 + z^2 + 14 = 0$

(n)$x^2 - y^2 - z^2 = 1$

(o)$x^2 - y^2 - z^2 = 0$

(p)$x^2 - y^2 - z^2 - 7 = 0$

(q)$y^2 - x^2 - z^2 = 1$

(r)$y^2 - x^2 - z^2 = 0$

(s)$y^2 - x^2 - z^2 - 78 = 0$

(t)$y^2 - x^2 - z^2 + 14 = 0$

(n)$y = x^2 + z^2$

(o)$x = y^2 + z^2$

(p)$y - x^2 - 4 + z = 0$

(q)$x - y^2 + 7 + z = 0$

8. Classifique e desenhe os conjuntos de $I\!R^3$ definidos pelas seguintes equações:

(a)$x^2 + y^2 + z^2 + 6x + 8y - 2z - 5 = 0$

(b)$x^2 + 2x + y^2 + z^2 + 3z + 2y + 2 = 0$

(c)$3x^2 + 5y^2 + 4z^2 + 4x - 2y - z - 10 = 0$

(d)$x^2 - y^2 + z^2 - 4x + 2y - z + 2 = 0$

(e)$x^2 - 2x + y^2 - 4y + 5 + z^2 + z = 0$

(f)$x^2 - y^2 - 7x - 5y - 8 - z^2 - z = 0$

(g)$y^2 - x^2 - 2x - 9y - 8 - 2z^2 - 3z = 0$

(h)$x^2 + 5x + y - 9 + z^2 - 4z = 0$

(i)$y^2 + x - y + 16 - 2z^2 - 2z = 0$

9. Escreva as seguintes expressões na forma

$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

   (a) $\alpha x^2 + \beta xy + \delta yz + \varepsilon xz + \gamma y^2$

   (b) $3x^2 + 5xy + 9xz + 11yz + 6y^2$

   (c) $-3x^2 - xy - xz$

10. Classifique e desenhe os conjuntos de $I\!R^3$ definidos pelas seguintes equações:

   (a)$x^2 + y^2 + z^2 + xz + yz + \frac{1}{3}xy + 6x + 8y - 5 = 0$

   (b)$x^2 + 2x + y^2 + -z^2 - xz + yz + 2y + 2 + \frac{1}{5}xy = 0$

   (c)$3x^2 + 5y^2 + +9z^2 - xz - yz - 4x - 2y - 10 + xy = 0$

   (d)$4x^2 - 4y^2 - 2z^2 - xz - 2yz + 16x + 8y + 8 - xy = 0$

   (e)$5x^2 - 10x + 5y^2 + 4z^2 - xz - 7yz - 20y + 25 + xy = 0$

   (f)$10x^2 - 10y^2 - 70x - 50y - 80 + xy - 11z^2 - 3xz - 2yz = 0$

   (g)$3y^2 - 3x^2 - 6x - 27y - 24 + xy - 7z^2 - xz - 30xz = 0$

   (h)$x^2 + 2xy + y^2 + 5x + y - 9 = 0$

   (i)$3y^2 + 6xy + 3y^2 + 3x - 3y + 48 + 7z^2 - 3xz - 8yz = 0$

Referências Bibliográficas

[1] G. A. Golub e C. F. Van Loan. *Matrix Computations*, 3rd. Edition. The Johns Hopkins University Press, London, 1996.

[2] E. L. Lima. *Geometria Analítica e Álgebra Linear*, Coleção Matemática Universitária, IMPA, Rio de Janeiro, 2001.

[3] R. J. Santos. *Geometria Analítica e Álgebra Linear*, Imprensa Universitária, UFMG, Belo Horizonte, 2000.