

## CDI-II

### Variedades. Extremos Condicionados

Em termos simples, uma variedade será um conjunto definido localmente por um sistema de equações tais que o teorema da função implícita seja aplicável. Veremos que será possível descrever tais conjuntos de três formas geométricas diferentes.

#### 1 Variedades. Parametrizações

Seja  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  uma função de classe  $C^1$  e consideremos o respectivo conjunto de nível zero, ou seja, o conjunto

$$M = \{(x, y) \in \mathbb{R}^2 : F(x, y) = 0\}.$$

Seja  $(a, b) \in M$  tal que  $\frac{\partial F}{\partial y}(a, b) \neq 0$ .

Pelo Teorema da Função Implícita, localmente em torno do ponto  $(a, b)$  temos

$$F(x, y) = 0 \Leftrightarrow y = f(x),$$

em que  $f : ]a - \epsilon, a + \epsilon[ \rightarrow \mathbb{R}$ , com  $\epsilon > 0$ , é uma função de classe  $C^1$ .

Seja  $g : ]a - \epsilon, a + \epsilon[ \rightarrow \mathbb{R}^2$  a função definida do seguinte modo

$$g(x) = (x, f(x)).$$

É claro que  $g$  é de classe  $C^1$ . Note-se que  $g(a) = (a, f(a)) = (a, b)$  e  $g'(a) = (1, f'(a))$ . Note-se que a função  $g$  é injectiva. De facto, se  $x_1 \neq x_2$  então  $g(x_1) \neq g(x_2)$ .

Note-se também que temos

$$\nabla F(a, b) \neq (0, 0) ; \quad g'(a) \neq (0, 0).$$

Suponhamos que, localmente em torno do ponto  $(a, b)$ , um conjunto  $M \subset \mathbb{R}^2$  pode ser descrito por uma função injectiva  $g : ]t_0 - \epsilon, t_0 + \epsilon[ \rightarrow \mathbb{R}^2$ , de classe  $C^1$ , tal que

$$g(t_0) = (a, b) ; \quad g'(t_0) \neq (0, 0).$$

Dado que  $g'(t) = (x'(t), y'(t))$ , sem perda de generalidade, suponhamos que  $x'(t_0) \neq 0$ . Pelo Teorema da Função Inversa em  $\mathbb{R}$ , a função  $x = x(t)$  será localmente invertível em torno de  $t = t_0$ , ou seja,  $t = h(x)$  para alguma função de classe  $C^1$  designada por  $h$ .

Portanto, teremos

$$y = y(t) = y(h(x)) = f(x).$$

Fazendo  $F(x, y) = y - f(x)$ , concluímos que, localmente em torno do ponto  $(a, b)$ , o conjunto  $M$  será definido pela equação  $F(x, y) = 0$ .

Assim, temos três formas equivalentes de descrever localmente o mesmo conjunto (c.f. [2]).

- i) Como conjunto de nível zero de uma função  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ , de classe  $C^1$  e tal que  $\nabla F(x, y) \neq (0, 0)$ .
- ii) Como gráfico de uma função  $f$  de classe  $C^1$ , ou seja,  $y = f(x)$ .
- iii) Como a imagem de uma função injetiva  $g$ , de classe  $C^1$ , tal que  $(x, y) = g(t)$  com  $t \in \mathbb{R}$  e  $g'(t) \neq (0, 0)$ .

Um conjunto descrito desta forma designa-se por variedade de dimensão um e dizemos que a função  $g$  é uma **parametrização** desse conjunto.

Normalmente chamamos **variedade-1** a esse conjunto.

Localmente, em torno do ponto  $(a, b)$ , teremos

$$F(x, y) = 0 \Leftrightarrow y = f(x) \Leftrightarrow (x, y) = g(t),$$

e, portanto

$$F(g(t)) = 0$$

e pelo Teorema da Função Composta, obtemos

$$\nabla F(g(t_0)) \cdot g'(t_0) = 0,$$

ou seja,

$$\nabla F(a, b) \cdot g'(t_0) = 0,$$

Geometricamente, o vector gradiente  $\nabla F(a, b) = \left( \frac{\partial F}{\partial x}(a, b), \frac{\partial F}{\partial y}(a, b) \right)$  é um vector **normal** ao conjunto  $M$  no ponto  $(a, b)$  e, portanto, o vector  $g'(t_0) = (x'(t_0), y'(t_0))$  é um vector **tangente** a  $M$  no mesmo ponto.

Portanto, as diferentes formas de descrever uma variedade fornecem informações geométricas distintas.

Ao espaço linear gerado pelo vector  $\mathbf{N} = \nabla F(a, b)$  chamamos **espaço normal** a  $M$  no ponto  $(a, b)$ .

Ao espaço linear gerado pelo vector  $\mathbf{T} = g'(t_0)$  chamamos **espaço tangente** a  $M$  no ponto  $(a, b)$ .

É claro que a **recta tangente** a  $M$  no ponto  $\mathbf{P} = (a, b)$  é dada pela equação paramétrica:

$$\mathbf{X} - \mathbf{P} = \lambda \mathbf{T}, \quad \lambda \in \mathbb{R},$$

em que  $\mathbf{X} = (x, y)$ . (Ver figura 1).

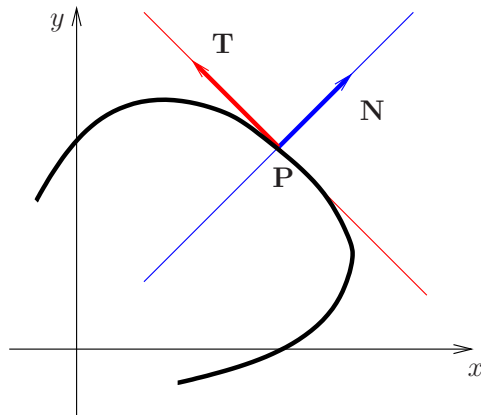


Figura 1: Recta tangente e recta normal em  $\mathbb{R}^2$

Do mesmo modo, a **recta normal** a  $M$  no ponto  $P = (a, b)$  é dada pela equação paramétrica:

$$\mathbf{X} - \mathbf{P} = \lambda \mathbf{N}, \lambda \in \mathbb{R}.$$

Note-se que os vectores  $\mathbf{T}$  e  $\mathbf{N}$  são ortogonais, ou seja,  $\mathbf{N} \cdot \mathbf{T} = 0$  e, portanto, a recta tangente no ponto  $P = (a, b)$  será dada pela equação cartesiana

$$(\mathbf{X} - \mathbf{P}) \cdot \mathbf{N} = 0$$

e a recta normal será dada pela equação cartesiana

$$(\mathbf{X} - \mathbf{P}) \cdot \mathbf{T} = 0.$$

Em  $\mathbb{R}^n$  com  $n \geq 2$ , estamos interessados em considerar conjuntos definidos por sistemas de  $m$  equações, ou seja, conjuntos  $M \subset \mathbb{R}^n$  da forma

$$M = \{x \in \mathbb{R}^n : F(x) = 0\}$$

em que  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , com  $m < n$ , é uma função de classe  $C^1$ .

Se o Teorema da Função Implícita for aplicável em  $M$  então dizemos que se trata de uma variedade. Quer isto dizer que, localmente em torno de cada um dos seus pontos,  $m$  variáveis serão expressas implicitamente como funções das restantes  $(n - m)$  variáveis, também designadas por **variáveis livres**. A tal conjunto chamaremos **variedade** de dimensão  $n - m$ .

Seja  $F(x) = (F_1(x), F_2(x), \dots, F_m(x))$ . Então o conjunto  $M$  será definido pelo sistema

$$\begin{cases} F_1(x_1, x_2, \dots, x_n) = 0 \\ F_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ F_m(x_1, x_2, \dots, x_n) = 0. \end{cases}$$

Note-se que o Teorema da Função Implícita é aplicável se as linhas da matriz que representa a derivada de  $F$ ,

$$DF(x) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(x) & \frac{\partial F_1}{\partial x_2}(x) & \cdots & \frac{\partial F_1}{\partial x_n}(x) \\ \frac{\partial F_2}{\partial x_1}(x) & \frac{\partial F_2}{\partial x_2}(x) & \cdots & \frac{\partial F_2}{\partial x_n}(x) \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F_m}{\partial x_1}(x) & \frac{\partial F_m}{\partial x_2}(x) & \cdots & \frac{\partial F_m}{\partial x_n}(x) \end{bmatrix}$$

forem linearmente independentes em cada um dos pontos de  $M$ .

Note-se também que as linhas da matriz  $DF(x)$  são os  $m$  vectores

$$\nabla F_1(x), \nabla F_2(x), \dots, \nabla F_m(x).$$

Sabendo que o gradiente de uma função escalar é perpendicular ao respectivo conjunto de nível no ponto considerado, as linhas da matriz  $DF(x)$  são vectores normais de  $M$ .

Ao espaço linear gerado por este conjunto de vectores chamamos **espaço normal** a  $M$  no ponto considerado.

Suponhamos que as  $(n - m)$  variáveis livres são  $(x_1, x_2, \dots, x_{n-m})$ .

Seja  $u = (x_1, x_2, \dots, x_{n-m})$  e  $v = (x_{n-m+1}, \dots, x_n)$ .

Então, localmente teremos

$$F(u, v) = 0 \Leftrightarrow v = f(u) = (f_1(u), f_2(u), \dots, f_m(u))$$

em que  $f = (f_1, f_2, \dots, f_m)$  é de classe  $C^1$ .

A função  $g(u) = (u, v) = (u, f(u))$  é de classe  $C^1$ , injectiva e a respectiva derivada

$$Dg(u) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ \frac{\partial f_1}{\partial u_1}(u) & \frac{\partial f_1}{\partial u_2}(u) & \cdots & \frac{\partial f_1}{\partial u_{n-m}}(u) \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_m}{\partial u_1}(u) & \frac{\partial f_m}{\partial u_2}(u) & \cdots & \frac{\partial f_m}{\partial u_{n-m}}(u) \end{bmatrix}$$

tem  $(n - m)$  colunas linearmente independentes.

À função  $g$  chamamos parametrização de  $M$ .

Note-se que, por definição de  $g$ , temos  $F(g(u)) = 0$  e, portanto,

$$DF(g(u))Dg(u) = 0$$

o que quer dizer que as colunas de  $Dg(u)$  são ortogonais às linhas de  $DF(g(u))$ .

Assim, o espaço gerado pelas colunas de  $Dg(u)$  é ortogonal ao espaço normal e será chamado **espaço tangente** a  $M$  no ponto considerado.

Assim, temos três formas equivalentes de descrever localmente o mesmo conjunto em torno de cada um dos seus pontos  $x \in M \subset \mathbb{R}^n$ .

- i) Como conjunto de nível zero de uma função  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , com  $m < n$ , de classe  $C^1$  e tal que as linhas da matriz  $DF(x)$  são linearmente independentes, ou seja, a matriz  $DF(x)$  tem característica  $m$ .
- ii) Como gráfico de uma função  $f$  de classe  $C^1$ , ou seja,  $v = f(u)$ .
- iii) Como a imagem de uma função injectiva  $g$ , de classe  $C^1$ , tal que  $x = g(t)$  com  $t \in \mathbb{R}^{n-m}$  e as colunas da matriz  $Dg(t)$  são linearmente independentes, ou seja, a matriz  $Dg(t)$  tem característica  $(n - m)$ .

Diz-se que  $M$  é uma variedade de dimensão  $(n - m)$  e usamos a notação variedade- $(n - m)$ .

**Exemplo 1.1** Consideremos a circunferência em  $C \subset \mathbb{R}^2$  dada pela equação

$$x^2 + y^2 = 1$$

e que se encontra representada na figura 2.

É claro que se trata do conjunto de nível zero da função  $F(x, y) = x^2 + y^2 - 1$ . Esta função é de classe  $C^1$  em  $\mathbb{R}^2$  e a respectiva derivada

$$\nabla F(x, y) = [2x \quad 2y]$$

é nula apenas na origem  $(x, y) = (0, 0)$ . No entanto, a origem não pertence à circunferência. Portanto, esta circunferência é uma variedade-1.

Consideremos o ponto  $P = (0, 1)$ . Dado que o vector  $N = \nabla F(0, 1) = (0, 2)$  é um vector normal em  $P$ , a recta normal à circunferência nesse ponto será dada na forma paramétrica por

$$(x, y) - (0, 1) = \lambda(0, 2) \Leftrightarrow \begin{cases} x = 0 \\ y - 1 = 2\lambda \end{cases}$$

e, portanto será dada pela equação  $x = 0$ . (Ver figura 2).

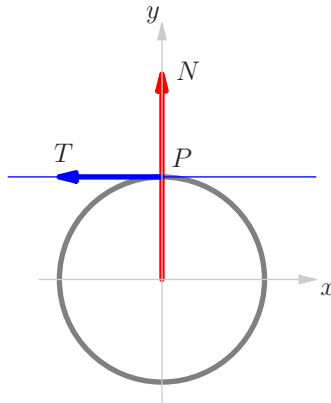


Figura 2: Circunferência  $C = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$

A recta tangente em  $P$  será dada por

$$(x, y - 1) \cdot (0, 2) = 0,$$

ou seja, pela equação  $y = 1$ . (Ver figura 2).

Note-se que para  $y > 0$  temos

$$x^2 + y^2 = 1 \Leftrightarrow y = \sqrt{1 - x^2}$$

e definindo  $g(x) = (x, \sqrt{1 - x^2})$  obtemos uma parametrização da parte da circunferência em que  $y > 0$ .

É claro que esta parametrização descreve apenas metade da circunferência.

Tendo em conta a simetria da circunferência podemos descrevê-la de outra forma. Note-se que os pontos de uma circunferência estão todos à mesma distância do centro. Se à distância ao centro associarmos o ângulo  $\theta$  tal como se ilustra na figura 3, obtemos novas coordenadas  $(r, \theta)$  que se relacionam com  $(x, y)$  da forma seguinte

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta. \end{cases}$$

em que  $r = \sqrt{x^2 + y^2}$ .

Nestas novas coordenadas, denominadas **coordenadas polares**, a circunferência dada por  $x^2 + y^2 = 1$  passa a ser descrita pela equação  $r = 1$  e, portanto podemos usar a variável  $\theta$  para descrever parametricamente a circunferência.

De facto, seja

$$g(\theta) = (\cos \theta, \sin \theta) \quad 0 < \theta < 2\pi.$$

Então, esta função é de classe  $C^1$ , injectiva e a respectiva derivada

$$g'(t) = \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}$$

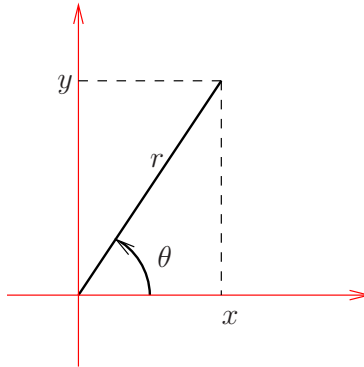


Figura 3: Coordenadas Polares  $(r, \theta)$

tem característica um. Para além disso a sua imagem é a circunferência sem o ponto  $(1, 0)$ , ou seja  $g(]0, 2\pi[) = C \setminus \{(1, 0)\}$ .

Note-se também que o vector  $g'(\frac{\pi}{2}) = (-1, 0)$  é o vector tangente  $T$  no ponto  $(0, 1)$  tal como se ilustra na figura 2.

Trata-se, portanto, de uma parametrização da circunferência. Note-se que esta parametrização descreve a circunferência excluindo um ponto apenas, ou seja, as coordenadas polares  $(r, \theta)$  são mais adequadas do que as coordenadas cartesianas  $(x, y)$ .

Para descrever completamente a circunferência deveremos ter outra parametrização que poderá ser dada pela função

$$h(\theta) = (\cos \theta, \sin \theta) \quad -\pi < \theta < \pi$$

que exclui apenas o ponto  $(-1, 0)$ .

Assim, as duas funções  $g$  e  $h$  descrevem completamente a circunferência  $C$ .

**Exemplo 1.2** Consideremos a esfera  $S \subset \mathbb{R}^3$  definida pela equação

$$x^2 + y^2 + z^2 = 1$$

que se encontra representada na figura 4.

Trata-se do conjunto de nível zero da função de classe  $C^1$  definida por

$$F(x, y, z) = x^2 + y^2 + z^2 - 1.$$

A derivada

$$\nabla F(x, y, z) = [2x \quad 2y \quad 2z]$$

tem característica um em todos os pontos de  $S$ , porque o caso contrário ocorre apenas na origem que não se encontra em  $S$ . Portanto,  $S$  é uma variedade-2.

O vector  $\nabla F(0, 1, 0) = (0, 2, 0)$  é normal a  $S$  no ponto  $(0, 1, 0)$ .

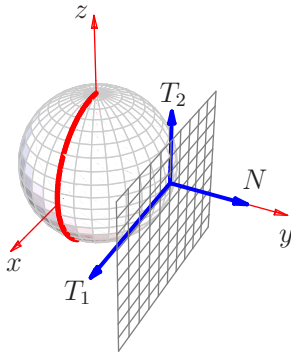


Figura 4: Esfera definida pela equação:  $x^2 + y^2 + z^2 = 1$

Os vectores tangentes a  $S$  no mesmo ponto resultam da resolução da equação

$$T \cdot N = 0.$$

Fazendo  $T = (\alpha, \beta, \gamma)$ , obtemos  $\beta = 0$  e, portanto,

$$T = (\alpha, 0, \gamma) = \alpha(1, 0, 0) + \gamma(0, 0, 1).$$

Assim, os vectores  $T_1 = (1, 0, 0)$  e  $T_2 = (0, 0, 1)$  geram o espaço tangente a  $S$  no ponto  $(0, 1, 0)$ .

Na figura 4 encontram-se representados os vectores  $N, T_1, T_2$ .

Fazendo  $X = (x, y, z)$  e  $P = (0, 1, 0)$ , o plano tangente a  $S$  no ponto  $(0, 1, 0)$ , será dado por

$$(X - P) \cdot N = 0,$$

ou seja,

$$(x, y - 1, z) \cdot (0, 2, 0) = 0 \Leftrightarrow y = 1,$$

e encontra-se representado na figura 4.

A recta normal a  $S$  no ponto  $P = (0, 1, 0)$ , será dada pelas equações

$$\begin{cases} (X - P) \cdot T_1 = 0 \\ (X - P) \cdot T_2 = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ z = 0, \end{cases}$$

ou seja, será o eixo  $Oy$ .

Note-se que para  $z > 0$  temos

$$x^2 + y^2 + z^2 = 1 \Leftrightarrow z = \sqrt{1 - x^2 - y^2}$$

e definindo  $g(x, y) = (x, y, \sqrt{1 - x^2 - y^2})$  obtemos uma parametrização da parte da esfera em que  $z > 0$ .

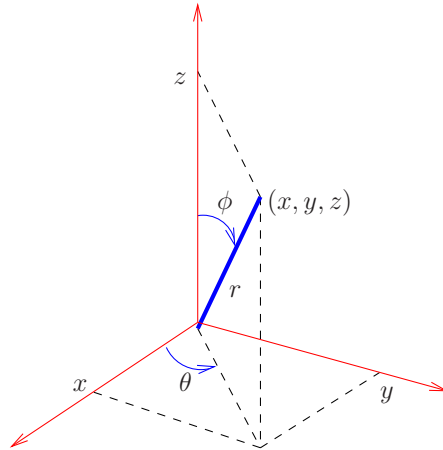


Figura 5: Coordenadas esféricas  $(r, \theta, \phi)$

É claro que esta parametrização descreve apenas metade da esfera.

Tendo em conta a simetria da esfera podemos descrevê-la de outra forma. Note-se que os pontos de uma esfera estão todos à mesma distância do centro. Se à distância ao centro associarmos os ângulos  $\theta$  e  $\phi$ , tal como se ilustra na figura 5, obtemos novas coordenadas  $(r, \theta, \phi)$  que se relacionam com  $(x, y, z)$  da forma seguinte

$$\begin{cases} x = r \operatorname{sen} \phi \cos \theta \\ y = r \operatorname{sen} \phi \operatorname{sen} \theta \\ z = r \cos \phi \end{cases}$$

em que  $r = \sqrt{x^2 + y^2 + z^2}$ .

Nestas novas coordenadas, denominadas **coordenadas esféricas**, a esfera dada por  $x^2 + y^2 + z^2 = 1$  passa a ser descrita pela equação  $r = 1$  e, portanto podemos usar as variáveis  $\theta, \phi$  para descrever parametricamente a esfera  $S$ .

De facto, seja

$$g(\theta, \phi) = (\operatorname{sen} \phi \cos \theta, \operatorname{sen} \phi \operatorname{sen} \theta, \cos \phi) \quad 0 < \theta < 2\pi; 0 < \phi < \pi$$

Então, esta função é de classe  $C^1$ , injectiva e a respectiva derivada

$$Dg(\theta, \phi) = \begin{bmatrix} -\operatorname{sen} \phi \operatorname{sen} \theta & \cos \phi \cos \theta \\ \operatorname{sen} \phi \cos \theta & \cos \phi \operatorname{sen} \theta \\ 0 & -\operatorname{sen} \phi \end{bmatrix}$$

tem característica dois. Para além disso a sua imagem é a esfera sem a linha em que  $x \geq 0, y = 0$ , ou seja

$$g([0, 2\pi[ \times ]0, \pi]) = S \setminus \{(x, y, z) : x \geq 0; y = 0\}.$$

Esta linha está representada a vermelho na figura 4.  
 Note-se também que as colunas da matriz

$$Dg\left(\frac{\pi}{2}, \frac{\pi}{2}\right) = \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ 0 & -1 \end{bmatrix}$$

são os vectores tangentes  $-T_1$  e  $-T_2$  no ponto  $(0, 1, 0)$ . (Ver figura 4).

Trata-se, portanto, de uma parametrização da esfera. Note-se que esta parametrização descreve a esfera excluindo uma linha apenas, ou seja, as coordenadas esféricas  $(r, \theta, \phi)$  são mais adequadas do que as coordenadas cartesianas  $(x, y, z)$ .

Para descrever completamente a esfera devemos considerar mais duas parametrizações. Consideremos o subconjunto de  $\mathbb{R}^2$  definido por

$$T = ]0, 2\pi[ \times ]0, \pi[$$

e as funções  $h, k : T \rightarrow \mathbb{R}^3$  definidas por

$$\begin{aligned} h(\theta, \phi) &= (\cos \phi, \sin \phi \cos \theta, \sin \phi \sin \theta) \\ k(\theta, \phi) &= (\sin \phi \sin \theta, \cos \phi, \sin \phi \cos \theta) \end{aligned}$$

Então, as funções  $g, h, k$  são de classe  $C^1$ , injectivas e se definirmos

$$\begin{aligned} G &= \{(x, y, z) : x \geq 0 ; y = 0\} \\ H &= \{(x, y, z) : y \geq 0 ; z = 0\} \\ K &= \{(x, y, z) : z \geq 0 ; x = 0\} \end{aligned}$$

cada uma das funções  $g, h, k$  estabelece uma bijecção entre o conjunto  $T \subset \mathbb{R}^2$  e as partes da esfera  $S \setminus G, S \setminus H, S \setminus K$ , respectivamente. As linhas  $G, H, K$  estão representadas na figura 6.

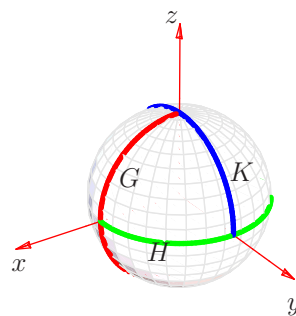


Figura 6: Parametrização da esfera

É fácil verificar que, tal como  $Dg(\theta, \phi)$ , as derivadas  $Dh(\theta, \phi)$  e de  $Dk(\theta, \phi)$  são matrizes com característica igual a dois.

Portanto, as funções  $g, h, k$  parametrizam a esfera  $S$ .

**Exemplo 1.3** Consideremos o cilindro  $C \subset \mathbb{R}^3$  definido por

$$C = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1; -1 < z < 1\}$$

que se encontra representado na figura 7.

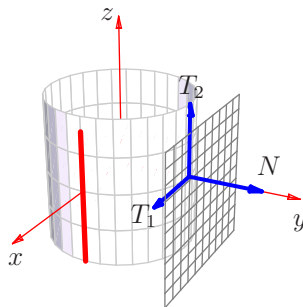


Figura 7: Cilindro definido por:  $x^2 + y^2 = 1; -1 < z < 1$

Trata-se do conjunto de nível zero da função de classe  $C^1$  definida por

$$F(x, y, z) = x^2 + y^2 - 1.$$

A derivada

$$\nabla F(x, y, z) = [2x \ 2y \ 0]$$

tem característica um em todos os pontos de  $S$ , porque o caso contrário ocorre apenas nos pontos da forma  $(0, 0, z)$  que não se encontram em  $C$ . Portanto,  $C$  é uma variedade-2.

O vector  $\nabla F(0, 1, 0) = (0, 2, 0)$  é normal a  $S$  no ponto  $(0, 1, 0)$ .

Os vectores tangentes a  $S$  no mesmo ponto resultam da resolução da equação

$$T \cdot N = 0.$$

Fazendo  $T = (\alpha, \beta, \gamma)$ , obtemos  $\beta = 0$  e, portanto,

$$T = (\alpha, 0, \gamma) = \alpha(1, 0, 0) + \gamma(0, 0, 1).$$

Assim, os vectores  $T_1 = (1, 0, 0)$  e  $T_2 = (0, 0, 1)$  geram o espaço tangente a  $S$  no ponto  $(0, 1, 0)$ .

Na figura 7 encontram-se representados os vectores  $N, T_1, T_2$ .

Fazendo  $X = (x, y, z)$  e  $P = (0, 1, 0)$ , o plano tangente a  $S$  no ponto  $(0, 1, 0)$ , será dado por

$$(X - P) \cdot N = 0,$$

ou seja,

$$(x, y - 1, z) \cdot (0, 2, 0) = 0 \Leftrightarrow y = 1,$$

e encontra-se representado na figura 7.

A recta normal a  $S$  no ponto  $P = (0, 1, 0)$ , será dada pelas equações

$$\begin{cases} (X - P) \cdot T_1 = 0 \\ (X - P) \cdot T_2 = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ z = 0, \end{cases}$$

ou seja, será o eixo  $Oy$ .

Note-se que para  $y > 0$  temos

$$x^2 + y^2 = 1 \Leftrightarrow y = \sqrt{1 - x^2}$$

e definindo  $g(x, z) = (x, \sqrt{1 - x^2}, z)$  obtemos uma parametrização da parte do cilindro em que  $y > 0$ .

É claro que esta parametrização descreve apenas metade do cilindro.

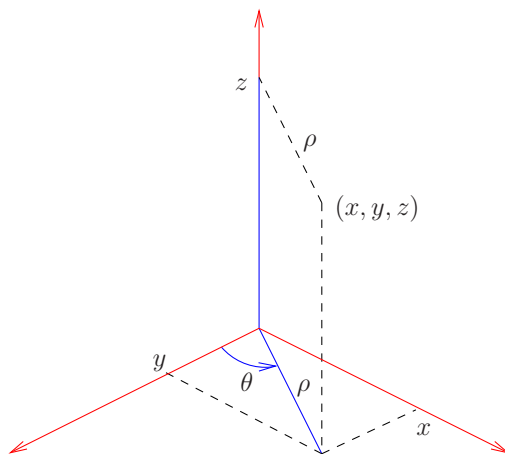


Figura 8: Coordenadas cilíndricas  $(\rho, \theta, z)$

Tendo em conta a simetria do cilindro podemos descrevê-lo de outra forma. Note-se que os pontos do cilindro  $C$  estão todos à mesma distância do eixo  $Oz$ . Se à distância ao eixo  $Oz$  associarmos o ângulo  $\theta$  e a variável  $z$ , tal como se ilustra na figura 8, obtemos novas coordenadas  $(\rho, \theta, z)$  que se relacionam com  $(x, y, z)$  da forma seguinte

$$\begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \\ z = z \end{cases}$$

em que  $\rho = \sqrt{x^2 + y^2}$ .

Nestas novas coordenadas, denominadas **coordenadas cilíndricas**, o cilindro dado por  $x^2 + y^2 = 1$  passa a ser descrito pela equação  $\rho = 1$  e, portanto podemos usar as variáveis  $\theta, z$  para o descrever parametricamente.

De facto, seja

$$g(\theta, z) = (\cos \theta, \sin \theta, z) \quad 0 < \theta < 2\pi; \quad -1 < z < 1$$

Então, esta função é de classe  $C^1$ , injectiva e a respectiva derivada

$$Dg(\theta, \phi) = \begin{bmatrix} -\sin \theta & 0 \\ \cos \theta & 0 \\ 0 & 1 \end{bmatrix}$$

tem característica dois. Para além disso a sua imagem é o cilindro sem a linha em que  $x = 1, y = 0$ , ou seja

$$g([0, 2\pi[ \times ]-1, 1[) = C \setminus \{(x, y, z) : x = 1; y = 0\}.$$

Esta linha está representada a vermelho na figura 7.

Note-se também que as colunas da matriz

$$Dg\left(\frac{\pi}{2}, 0\right) = \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

são os vectores tangentes  $-T_1$  e  $T_2$  no ponto  $(0, 1, 0)$ . (Ver figura 7).

Trata-se, portanto, de uma parametrização do cilindro. Note-se que esta parametrização descreve o cilindro excluindo uma linha apenas, ou seja, as coordenadas cilíndricas  $(\rho, \theta, z)$  são mais adequadas do que as coordenadas cartesianas  $(x, y, z)$ .

Para descrever completamente o cilindro devemos considerar mais uma parametrização.

Consideremos a função  $h : ]-\pi, \pi[ \times ]-1, 1[ \rightarrow \mathbb{R}^3$  definida por

$$h(\theta, z) = (\cos \theta, \sin \theta, z).$$

Então, a função  $h$  é de classe  $C^1$ , injectiva, a respectiva derivada é igual à derivada de  $g$  e, portanto, tem característica igual a dois.

Note-se que a imagem de  $h$  é o cilindro sem a linha vertical dada por  $x = -1; y = 0$ . Portanto, as funções  $g$  e  $h$  parametrizam o cilindro  $C$ .

## 2 Extremos Condicionados

Consideremos a função  $f(x, y) = x^2 + y^2$  e a elipse definida pela equação

$$x^2 + \frac{y^2}{4} = 1$$

e que se encontra representada na figura 9.

Dado que  $f(x, y)$  representa o quadrado da distância de um ponto  $(x, y)$  à origem, é claro que os pontos  $(0, 2)$  e  $(0, -2)$  são os máximos de  $f$  na elipse. Os pontos  $(1, 0)$  e

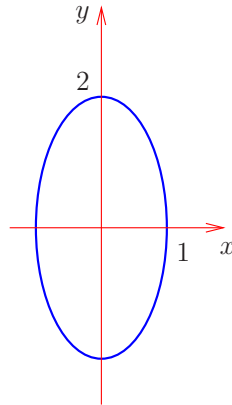


Figura 9: Elipse em  $\mathbb{R}^2$  dada por  $x^2 + \frac{y^2}{4} = 1$

$(-1, 0)$  são os mínimos de  $f$  sobre a elipse. Ou seja, se restringirmos a função  $f$  à elipse estes pontos são os respectivos extremos.

Note-se que a origem é o único ponto de estacionaridade da função  $f$  em  $\mathbb{R}^2$ . De facto, temos

$$\nabla f(x, y) = (2x, 2y) = (0, 0) \Leftrightarrow (x, y) = (0, 0).$$

Portanto os extremos de  $f$ , quando restringida à elipse, não se encontram no conjunto de pontos críticos de  $f$ . Assim, deveremos adoptar uma estratégia diferente para determinar os extremos de  $f$  sobre a elipse.

Seja  $\gamma(t) = (\cos t, 2 \sin t) = (x(t), y(t))$ , com  $-\frac{\pi}{6} < t < \frac{11\pi}{6}$ , uma parametrização da elipse.

A função composta  $f \circ \gamma$  é a restrição de  $f$  à elipse retirando o ponto  $(1, 0)$ . Trata-se de uma função real de variável real. De facto temos

$$\begin{array}{ccccc} \mathbb{R} & \xrightarrow{\gamma} & \mathbb{R}^2 & \xrightarrow{f} & \mathbb{R} \\ t & \mapsto & \gamma(t) & \mapsto & f(\gamma(t)). \end{array}$$

Um extremo da função composta  $f \circ \gamma$  é um zero da respectiva derivada,

$$\frac{d}{dt} f(\gamma(t)) = 0 \Leftrightarrow \nabla f(\gamma(t)) \cdot \gamma'(t) = 0,$$

ou seja,

$$(2 \cos t, 4 \sin t) \cdot (-\sin t, 2 \cos t) = 0 \Leftrightarrow 6 \sin t \cos t = 0 \Leftrightarrow \sin t = 0 \vee \cos t = 0$$

e, portanto, teremos

$$t = 0 \vee t = \frac{\pi}{2} \vee t = \pi \vee t = \frac{3\pi}{2}.$$

Assim, os pontos críticos de  $f$  restringida à elipse serão

$$\gamma(0) = (1, 0); \gamma\left(\frac{\pi}{2}\right) = (0, 2); \gamma(\pi) = (-1, 0); \gamma\left(\frac{3\pi}{2}\right) = (0, -2),$$

ou seja, exactamente os pontos determinados acima.

Note-se que  $\gamma'(t)$  é um vector tangente à elipse no ponto  $\gamma(t)$ . Dado que, num extremo, deveremos ter

$$\nabla f(\gamma(t)) \cdot \gamma'(t) = 0$$

concluimos que o vector  $\nabla f(\gamma(t))$  é ortogonal ao vector tangente  $\gamma'(t)$ .

Portanto, o vector  $\nabla f(x, y)$  pertence ao espaço normal à elipse no ponto  $(x, y)$ .

Consideremos a função

$$F(x, y) = x^2 + \frac{y^2}{4} - 1.$$

Então a elipse é o conjunto de nível zero de  $F$  e o vector  $\nabla F(x, y)$  gera o espaço normal à elipse no ponto  $(x, y)$ .

Assim, o vector  $\nabla f(x, y)$  é um múltiplo do vector  $\nabla F(x, y)$ , ou seja,

$$\nabla f(x, y) = \lambda \nabla F(x, y),$$

em que  $\lambda \in \mathbb{R}$ .

Deste modo, temos uma estratégia para determinar os extremos de  $f$  quando sujeitos à condição  $F = 0$ , que consiste em resolver o sistema

$$\begin{cases} \nabla f(x, y) = \lambda \nabla F(x, y) \\ F(x, y) = 0 \end{cases}$$

Este raciocínio pode ser aplicado à resolução de um problema mais geral que pode ser formulado do seguinte modo (c.f. [2, 3, 1]).

Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função de classe  $C^1$  e  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , com  $m < n$ , uma função também de classe  $C^1$ . Pretendemos determinar os extremos de  $f$  sujeitos ao sistema de equações (ou **condições**),  $F(x) = 0$ , ou seja,

$$\begin{cases} F_1(x_1, x_2, \dots, x_n) = 0 \\ F_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ F_m(x_1, x_2, \dots, x_n) = 0. \end{cases}$$

em que  $F_1, F_2, \dots, F_m$  são as componentes de  $F$ .

Dito de outro modo, trata-se de determinar os extremos de  $f$  restringida à variedade definida pelo sistema de equações  $F(x) = 0$ .

Este é o chamado problema dos **extremos condicionados**.

Tal como para a elipse, o vector  $\nabla f(x)$  deverá ser normal à variedade definida por  $M = \{x \in \mathbb{R}^n : F(x) = 0\}$ , ou seja, deverá ser uma combinação linear dos vectores que geram o espaço normal à variedade.

De facto, seja  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$  um caminho ou trajectória de classe  $C^1$  tal que

$$\gamma(0) = a ; \quad F(\gamma(t)) = 0 \quad \forall t \in \mathbb{R},$$

ou seja,  $\gamma$  define uma linha de pontos da variedade que passa no ponto  $a$ .

Então, a função composta  $f \circ \gamma : \mathbb{R} \rightarrow \mathbb{R}$  deverá apresentar um extremo em  $a$ , ou seja,

$$\frac{d}{dt}f(\gamma(t))|_{t=0} = 0 \Leftrightarrow \nabla f(a) \cdot \gamma'(0) = 0.$$

Dado que o vector  $\gamma'(0)$  é tangente a  $M$  no ponto  $a$ , concluímos que  $\nabla f(a)$  é um vector normal à variedade  $M$  nesse ponto e, portanto, será uma combinação linear dos vectores  $\nabla F_1(a), \nabla F_2(a), \dots, \nabla F_m(a)$ .

Assim, teremos

$$\begin{cases} \nabla f(x) = \lambda_1 \nabla F_1(x) + \lambda_2 \nabla F_2(x) + \dots + \lambda_m \nabla F_m(x) = 0 \\ F(x) = 0. \end{cases} \quad (1)$$

Note-se que este sistema apresenta  $(n + m)$  equações e  $(n + m)$  incógnitas e, em geral, não é linear.

Os escalares  $\lambda_1, \lambda_2, \dots, \lambda_m$  são os chamados **multiplicadores de Lagrange** e ao sistema 1 chamamos **método dos multiplicadores de Lagrange**.

**Exemplo 2.1** Para o caso considerado acima, temos

$$f(x, y) = x^2 + y^2 \quad ; \quad F(x, y) = x^2 + \frac{y^2}{4} - 1$$

e, portanto,

$$\begin{cases} \nabla f(x, y) = \lambda \nabla F(x, y) \\ F(x, y) = 0 \end{cases} \Leftrightarrow \begin{cases} 2x = 2\lambda x \\ 2y = \frac{\lambda y}{2} \\ x^2 + \frac{y^2}{4} = 1 \end{cases} \Leftrightarrow \begin{cases} x(1 - \lambda) = 0 \\ y(4 - \lambda) = 0 \\ x^2 + \frac{y^2}{4} = 1 \end{cases} \Leftrightarrow \begin{cases} x = 0 \vee \lambda = 1 \\ y = 0 \vee \lambda = 4 \\ x^2 + \frac{y^2}{4} = 1 \end{cases}$$

de onde obtemos os pontos  $(0, -2)$ ,  $(0, 2)$ ,  $(-1, 0)$ ,  $(1, 0)$ . Os dois primeiros são os mais afastados da origem e os outros dois são os mais próximos.

Note-se que o cálculo do escalar  $\lambda$  é irrelevante para o problema.

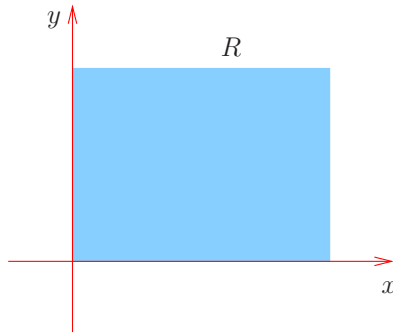


Figura 10: Rectângulo de comprimento  $x$  e largura  $y$

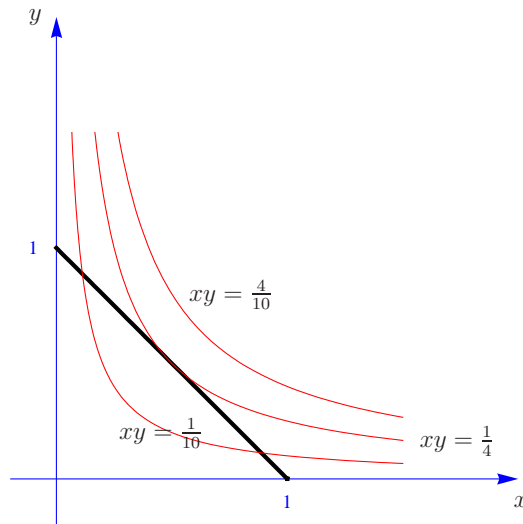


Figura 11: O rectângulo de perímetro 2 com área máxima é o quadrado de lado  $\frac{1}{2}$

**Exemplo 2.2** Consideremos o conjunto dos rectângulos em  $\mathbb{R}^2$  com perímetro igual a dois. Qual deles apresenta maior área?

Note-se que o perímetro fixo é uma condição ou restrição e pretendemos maximizar a área.

Podemos formular este problema, (ver figura 10), em termos do método dos multiplicadores de Lagrange fazendo  $f(x, y) = xy$  e  $F(x, y) = 2x + 2y - 2$ , ou seja, pretendemos determinar os extremos de  $f$  sujeitos à condição  $F(x, y) = 0 \Leftrightarrow x + y = 1$ .

Então teremos,

$$\begin{cases} \nabla f(x, y) = \lambda \nabla F(x, y) \\ F(x, y) = 0 \end{cases} \Leftrightarrow \begin{cases} y = \lambda \\ x = \lambda \\ x + y = 1 \end{cases} \Leftrightarrow \begin{cases} y = x \\ x = \lambda \\ 2x = 1 \end{cases}$$

e, portanto,  $y = x = \frac{1}{2}$ .

Trata-se de um quadrado de lado  $\frac{1}{2}$ , ou seja, um quadrado de área  $xy = \frac{1}{4}$ .

Na figura 11 estão representados o conjunto em que  $x + y = 1$ , ou seja, o conjunto dos retângulos de perímetro 2 e linhas em que  $xy = c$ ;  $c > 0$ , ou seja, área constante. Note-se que a área é máxima para  $c = \frac{1}{4}$ .

**Exemplo 2.3** Consideremos o conjunto  $L$  definido pelo sistema

$$\begin{cases} x^2 + y^2 + z^2 = 2 \\ y = x. \end{cases}$$

Quais os pontos de  $L$  mais próximos do ponto  $(0, 0, 1)$ ?

O conjunto  $L$  resulta da intersecção da esfera de raio  $\sqrt{2}$  e centro na origem com o plano vertical  $y = x$  e, portanto, é uma circunferência tal como se ilustra na figura 12.

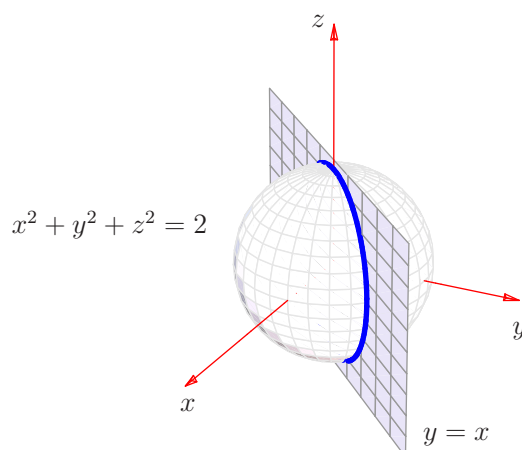


Figura 12: Circunferência em  $\mathbb{R}^3$  dada por  $x^2 + y^2 + z^2 = 2$ ;  $y = x$

Seja  $f(x, y, z) = x^2 + y^2 + (z - 1)^2$ . Esta é a função a minimizar em  $L$ . Note-se que  $L$  é um conjunto compacto em  $\mathbb{R}^3$  e, sendo  $f$  de classe  $C^1$ , terá mínimo nesse conjunto.

Note-se também que poderíamos usar a função  $\sqrt{x^2 + y^2 + (z - 1)^2}$  que é a distância de um ponto  $(x, y, z)$  ao ponto  $(0, 0, 1)$ . No entanto, no método dos multiplicadores de Lagrange temos de calcular as derivadas das funções envolvidas. É claro que essa tarefa é mais simples considerando o quadrado da distância em vez da distância propriamente dita. Note-se que a função distância, definida pela norma, não é diferenciável na origem porque se trata de uma raiz quadrada.

Assim, sejam  $F_1(x, y, z) = x^2 + y^2 + z^2 - 2$  e  $F_2(x, y, z) = y - x$ .

Portanto,

$$\begin{cases} \nabla f(x, y, z) = \lambda_1 \nabla F_1(x, y, z) + \lambda_2 \nabla F_2(x, y, z) \\ F_1(x, y, z) = 0 \\ F_2(x, y, z) = 0 \end{cases} \Leftrightarrow \begin{cases} 2x = 2\lambda_1 x - \lambda_2 \\ 2y = 2\lambda_1 y + \lambda_2 \\ 2(z - 1) = 2\lambda_1 z \\ x^2 + y^2 + z^2 = 2 \\ y = x \end{cases}$$

donde deduzimos

$$\begin{cases} 2x(1 - \lambda_1) = -\lambda_2 \\ 2y(1 - \lambda_1) = \lambda_2 \\ z(1 - \lambda_1) = 1 \\ x^2 + y^2 + z^2 = 2 \\ y = x. \end{cases}$$

Tendo em conta que  $y = x$ , da primeira e segunda equações concluimos que  $\lambda_2 = 0$ . Da primeira equação teremos

$$x(1 - \lambda_1) = 0 \Leftrightarrow x = 0 \vee \lambda_1 = 1.$$

Se  $\lambda_1 = 1$  então da terceira equação obtemos  $0 = 1$ . Assim,  $y = x = 0$  e da quarta equação teremos  $z = \sqrt{2}$  ou  $z = -\sqrt{2}$ .

Portanto, os pontos a considerar são  $(0, 0, -\sqrt{2})$  e  $(0, 0, \sqrt{2})$ . É claro que o mais próximo de  $(0, 0, 1)$  é o ponto  $(0, 0, \sqrt{2})$ .

**Exemplo 2.4** Consideremos a linha definida pelo sistema de equações

$$\begin{cases} z = x^2 + y^2 \\ x + y + z = 1. \end{cases}$$

e que se representa na figura 13. Trata-se da intersecção do plano definido por  $x + y + z = 1$  com o parabolóide dado por  $z = x^2 + y^2$ .

Pretendemos determinar o ponto desta linha que apresenta maior cota, ou seja, coordenada  $z$  mais elevada.

É fácil verificar que se trata de uma variedade de dimensão um, ou seja, uma linha em  $\mathbb{R}^3$ .

Pretendemos determinar os extremos da função  $f(x, y, z) = z$ , sujeitos à condição  $F(x, y, z) = (0, 0)$ .

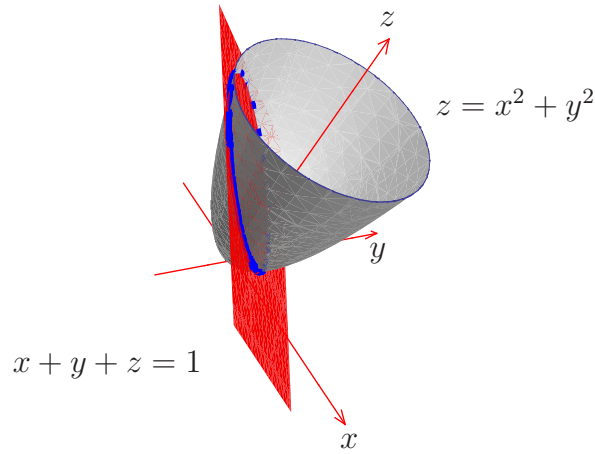


Figura 13: Linha em  $\mathbb{R}^3$  dada por  $z = x^2 + y^2$ ;  $x + y + z = 1$

Aplicando o método dos multiplicadores de Lagrange, obtemos

$$\begin{cases} 0 = -2\lambda_1 x + \lambda_2 \\ 0 = -2\lambda_1 y + \lambda_2 \\ 1 = \lambda_1 + \lambda_2 \\ z = x^2 + y^2 \\ x + y + z = 1 \end{cases} \Leftrightarrow \begin{cases} 2\lambda_1 x = \lambda_2 \\ 2\lambda_1 y = \lambda_2 \\ 1 = \lambda_1 + \lambda_2 \\ z = x^2 + y^2 \\ x + y + z = 1 \end{cases} \Leftrightarrow \begin{cases} 2\lambda_1(x - y) = 0 \\ 2\lambda_1 y = \lambda_2 \\ 1 = \lambda_1 + \lambda_2 \\ z = x^2 + y^2 \\ x + y + z = 1 \end{cases}$$

Da primeira equação teremos  $\lambda_1 = 0$  ou  $x = y$ .

No caso de  $\lambda_1 = 0$ , da segunda equação teremos  $\lambda_2 = 0$ . Substituindo estes valores na terceira equação, concluímos que este caso não pode ocorrer.

Para o caso em que  $x = y$ , da quarta e quinta equações, obtemos

$$2x^2 + 2x - 1 = 0$$

e, portanto,

$$x = \frac{-1 - \sqrt{3}}{2} \vee x = \frac{-1 + \sqrt{3}}{2}.$$

Dado que  $y = x$  e  $z = 1 - x - y$ , teremos os pontos

$$\left( \frac{-1 - \sqrt{3}}{2}, \frac{-1 - \sqrt{3}}{2}, 3 + \sqrt{3} \right); \left( \frac{-1 + \sqrt{3}}{2}, \frac{-1 + \sqrt{3}}{2}, 3 - \sqrt{3} \right).$$

Assim, o ponto de cota mais elevada é o primeiro destes dois. O outro será o de cota menos elevada.

**Exemplo 2.5** Quais os pontos da elipse definida pela equação  $x^2 + y^2 - xy = 3$  que se encontram mais afastados do eixo  $Ox$ ?

Facilmente se verifica que esta linha é uma variedade de dimensão um em  $\mathbb{R}^2$ .

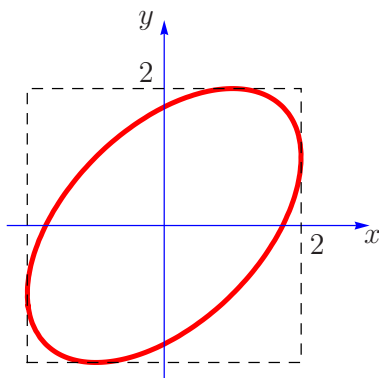


Figura 14: Linha em  $\mathbb{R}^2$  dada por  $x^2 + y^2 - xy = 3$

A distância de um ponto do plano de coordenadas  $(x, y)$  ao eixo  $Ox$  é dada por  $|y|$ . Consideremos então a função  $f(x, y) = y$ .

Aplicando o método dos multiplicadores de Lagrange, obtemos

$$\begin{cases} 0 = \lambda(2x - y) \\ 1 = \lambda(2y - x) \\ x^2 + y^2 - xy = 3. \end{cases}$$

Da primeira equação teremos  $\lambda = 0$  ou  $y = 2x$ . Fazendo  $\lambda = 0$  e substituindo na segunda equação teríamos  $1 = 0$ . Portanto, deveremos ter  $y = 2x$  e, da terceira equação obteremos  $x^2 = 1$ , ou seja, os pontos que resolvem o sistema são  $(-1, -2)$ ,  $(1, 2)$ .

Note-se que estes pontos estão ambos à distância dois do eixo  $Ox$ . Na figura 14 encontra-se representada esta elipse onde se pode constatar que os pontos mais afastados tanto do eixo  $Ox$  como do eixo  $Oy$  se encontram à distância dois.

\*\*\*

## Referências

- [1] Tom M. Apostol. *Calculus II*. Editorial Reverté, SA, 1977.
- [2] Luís T. Magalhães. *Integrais em Variedades e Aplicações*. Texto Editora, 1993.
- [3] J. E. Marsden and A. J. Tromba. *Vector Calculus*. W. H. Freeman and Company, 1998.