Instituto Superior Técnico
Departamento de Matemática
Prof. Gabriel Pires

Extremos Condicionados. Multiplicadores de Lagrange.

1 Extremos Condicionados

Considere-se o problema de determinar os pontos da elipse, definida pela equação

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

e representada na Figura 1, que estão mais próximos (ou os que estão mais afastados) da origem.

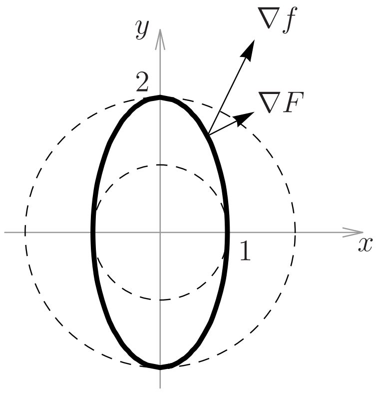

Figura 1: Extremos de $f(x, y) = x^2 + y^2$ sobre a elipse $x^2 + \frac{y^2}{4} = 1$

A distância de um ponto $(x, y) \in \mathbb{R}^2$ à origem é a norma $||(x, y)|| = \sqrt{x^2 + y^2}$. Dado que a norma não é diferenciável na origem é preferível usar o quadrado da distância, ou seja, a função $f : \mathbb{R}^2 \to \mathbb{R}$, de classe C^1, definida por $f(x, y) = x^2 + y^2$.

Note-se que os pontos de mínimo ou de máximo da norma são os mesmos que os do respetivo quadrado.

Assim, o problema proposto é a determinação dos pontos de extremo da função f que se encontram na elipse.

Da Figura 1, é claro que $(0, 2)$ e $(0, -2)$ são os pontos de máximo de f na elipse. Do mesmo modo, $(1, 0)$ e $(-1, 0)$ são os pontos de mínimo de f sobre a elipse. Ou seja, se a função f for restringida à elipse, estes são os pontos onde se encontram os extremos.

No entanto, a origem é o único ponto de estacionaridade da função f em $\mathbb{R}^2$. De facto,

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

Os pontos críticos de f não pertencem à elipse.

Assim, é necessário seguir uma estratégia diferente para determinar os extremos de f sobre a elipse. A função f deverá ser cuidadosamente restringida à elipse.

Note-se que, resolvendo a equação da elipse para $y > 0$, se tem $y = g(x) = 2\sqrt{1 - x^2}$, ou seja, para $y > 0$, a elipse é o gráfico da função $g :]-1, 1[\to \mathbb{R}$ que é de classe C^1.

O gráfico da função g é o conjunto

$$\{(x, g(x)) \in \mathbb{R}^2 : -1 < x < 1\}.$$

Portanto a **restrição da função** f à elipse, para $y > 0$, é dada por $f(x, g(x))$.

A elipse é o conjunto de nível zero da função $F : \mathbb{R}^2 \to \mathbb{R}$, definida por

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

ou seja, é o conjunto

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

No caso geral, pode não ser possível resolver a equação $F(x, y) = 0$ em ordem a y ou em ordem a x. Ver-se-á de seguida que a existência da função g pode ser garantida pelo teorema da função implícita.

Note-se que, para qualquer $(x, y) \in E$, se tem

$$DF(x, y) = \begin{bmatrix} 2x & \frac{y}{2} \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \end{bmatrix}.$$

De facto, o contrário ocorre apenas no ponto $(0, 0)$ que não pertence à elipse.

Assim, pelo **teorema da função implícita**, localmente em torno de qualquer ponto da elipse, uma das variáveis x ou y é função, de classe C^1, da outra.

Seja $(a, b) \in E$ um ponto de extremo da função f, ou seja, um ponto de extremo de f sujeito à condição $F(x, y) = 0$.

Supondo que $\dfrac{\partial F}{\partial y}(a, b) \neq 0$ então, em alguma vizinhança do ponto $(a, b) \in E$, fica garantido que a variável y é função, de classe C^1, da variável x, ou seja, $y = g(x)$, e tem-se

$$F(x, g(x)) = 0.$$

Sendo $b = g(a)$ é claro que a função definida por $h(x) = f(x, g(x))$ tem um extremo no ponto $x = a$.

É muito importante observar que a função h é a restrição de f à elipse E, numa vizinhança do ponto (a, b). De facto, $h(x)$ é o valor da função f no ponto $(x, g(x))$ que pertence à elipse.

Usando a derivada da função composta, obtém-se

$$0 = h'(a) = \frac{\partial f}{\partial x}(a, b) + \frac{\partial f}{\partial y}(a, b) g'(a), \quad \frac{\partial F}{\partial x}(a, b) + \frac{\partial F}{\partial y}(a, b) g'(a) = 0$$

e, portanto,

$$\frac{\partial f}{\partial x}(a, b) = \frac{\partial f}{\partial y}(a, b) \left(\frac{\partial F}{\partial y}(a, b) \right)^{-1} \frac{\partial F}{\partial x}(a, b).$$

Fazendo $\lambda = \dfrac{\partial f}{\partial y}(a,b)\left(\dfrac{\partial F}{\partial y}(a,b)\right)^{-1}$, conclui-se que

$$\frac{\partial f}{\partial x}(a,b) = \lambda \frac{\partial F}{\partial x}(a,b).$$

Por outro lado é claro que

$$\frac{\partial f}{\partial y}(a,b) = \frac{\partial f}{\partial y}(a,b)\left(\frac{\partial F}{\partial y}(a,b)\right)^{-1}\frac{\partial F}{\partial y}(a,b) = \lambda \frac{\partial F}{\partial y}(a,b).$$

Assim, o ponto (a,b) deverá satisfazer o sistema de equações seguinte

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

Portanto, os extremos de f pertencentes ao conjunto em que $F = 0$ são aqueles onde os vetores ∇f e ∇F são colineares.

Resolvendo o sistema para o caso da elipse tem-se

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

cuja solução é o conjunto de 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 tal como ficou claro da observação da Figura 1.

***

Este procedimento é generalizável como se verá de seguida.

Sejam $f : \mathbb{R}^n \to \mathbb{R}$ e $F : \mathbb{R}^n \to \mathbb{R}^m$, com $m < n$, funções de classe C^1.

O problema dos extremos condicionados consiste na determinação dos extremos da função escalar f restringida ao conjunto $M = \{x \in \mathbb{R}^n : F(x) = 0\}$.

Dito de outro modo, pretende-se determinar os extremos de f sujeitos ao sistema de equações (ou **condições**):

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

em que $F_1, F_2, \ldots, F_m$ são as componentes da função F.

Seja então $a \in M$ um ponto em que a função f tem um extremo.

Suponha-se que uma das submatrizes quadradas $m \times m$ da matriz $DF(a)$ tem determinante não nulo, ou, equivalentemente, $DF(a)$ tem característica máxima, ou ainda, as linhas de $DF(a)$ são linearmente independentes.

Sem perda de generalidade pode ser assumido que o determinante da submatriz quadrada, definida pelas últimas m colunas de $DF(a)$, é não nulo.

Assim, é possível renomear as variáveis de tal forma que se tem $\mathbb{R}^n = \mathbb{R}^{n-m} \times \mathbb{R}^m$ e

$$M = \{(x,y) \in R^{n-m} \times \mathbb{R}^m : F(x,y) = 0\}.$$

Seja então $(a,b) \in M$ o ponto de extremo da função f.

Sendo $\det D_y F(a,b) \neq 0$ e usando o **teorema da função implícita** conclui-se que existe uma função $g : U \to \mathbb{R}^m$, de classe C^1, sendo $U \subset \mathbb{R}^{n-m}$ um aberto que contém o ponto a, e existe um aberto em $\mathbb{R}^n$, que contém o ponto (a,b), em que se verifica a equivalência

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

ou seja, a equação $F(x,y) = 0$ define implicitamente y como função de x, sendo $b = g(a)$.

Assim sendo, é claro que o ponto a é um ponto de extremo da função $h : U \to \mathbb{R}$, de classe C^1, definida por

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

A função h é de facto a restrição de f ao conjunto M, numa vizinhança do ponto (a,b), porque $h(x)$ é o valor de f no ponto $(x, g(x))$ que pertence ao conjunto M.

Usando o teorema da derivada da função composta, tem-se

$$0 = \nabla h(a) = D_x f(a,b) + D_y f(a,b) Dg(a). \tag{1}$$

Por outro lado, sendo $F(x, g(x)) = 0$, tem-se

$$D_x F(a,b) + D_y F(a,b) Dg(a) = 0,$$

ou seja,

$$Dg(a) = -\left(D_y F(a,b)\right)^{-1} D_x F(a,b).$$

Substituindo em (1) obtém-se

$$D_x f(a,b) = D_y f(a,b) \left(D_y F(a,b)\right)^{-1} D_x F(a,b). \tag{2}$$

Note-se que $D_y f(a,b) \left(D_y F(a,b)\right)^{-1}$ é uma matriz linha com m colunas.

Seja $\Lambda = \begin{bmatrix} \lambda_1 \, \lambda_2 \, \cdots \lambda_m \end{bmatrix}$ essa matriz.

Note-se também que

$$D_y f(a,b) = D_y f(a,b) \left(D_y F(a,b)\right)^{-1} D_y F(a,b) = \Lambda D_y F(a,b). \tag{3}$$

Finalmente, de (2) e (3) obtém-se

$$\nabla f(a,b) = \Lambda DF(a,b).$$

Voltando à notação inicial, se $a \in M$ for um ponto de extremo de f existem escalares $\lambda_1,\ \lambda_2,\ \ldots \lambda_m$, os chamados **multiplicadores de Lagrange**, tais que

$$\begin{cases} \nabla f(a) = \lambda_1 \nabla F_1(a) + \lambda_2 \nabla F_2(a) + \ldots + \lambda_m \nabla F_m(a) \\ \\ F(a) = 0 \end{cases} \tag{4}$$

Este sistema apresenta $(n+m)$ equações e $(n+m)$ incógnitas e, em geral, não é linear.

Os extremos da função f no conjunto $M = \{x \in \mathbb{R}^n : F(x) = 0\}$ encontram-se no conjunto de soluções do sistema (4.)

À resolução do problema de extremos condicionados usando o sistema de equações (4) chama-se **método dos multiplicadores de Lagrange**.

***

Nota 1 *Seja $c \in \mathbb{R}$ o valor extremo da função f no conjunto definido por $F = 0$. Os pontos em que isso acontece serão dados pelo sistema*

$$\begin{cases} f(x) = c \\ F(x) = 0, \end{cases} \tag{5}$$

ou seja, estarão na intersecção do conjunto de nível c da função f com o conjunto de nível zero de F.

Recorde-se que o vetor $\nabla f(x)$ é normal ao conjunto de nível dado por $f(x) = c$. Do mesmo modo, os vetores $\nabla F_1(x),\ \nabla f_2(x), \ldots, \nabla F_m(x)$ são normais ao conjunto descrito por $F(x) = 0$.

Assim, pode concluir-se que num ponto de extremo condicionado de f, os conjuntos de nível descritos por $f(x) = c$ e por $F(x) = 0$ são tangentes nesse ponto.

Portanto, geometricamente a determinação de um extremo condicionado consiste em encontrar o conjunto de nível de f que é tangente ao conjunto de nível zero de F.

No caso do exemplo da elipse, os conjuntos de nível da função $f(x,y) = x^2 + y^2$ são circunferências centradas na origem.

Na Figura 1 estão representadas as circunferências de raios 1 e 2, respetivamente, e a elipse. Note-se que estas circunferências intersetam a elipse nos pontos de extremo. Por outro lado, os vetores ∇f e ∇F são colineares nesses pontos, ou seja, nesses pontos as circunferências são tangentes à elipse.

***

2 Exemplos

Exemplo 2.1 No conjunto dos retângulos com perímetro igual a 2, qual deles apresenta maior área?

Neste caso, o perímetro fixo é uma condição ou restrição, e a área é a grandeza a maximizar.

Este problema pode ser formulado (ver Figura 2) em termos do método dos multiplicadores de Lagrange, fazendo $f(x,y) = xy$ e $F(x,y) = 2x + 2y - 2$, ou seja, pretende-se determinar os extremos de f sujeitos à condição $F(x,y) = 0 \Leftrightarrow x + y = 1$.

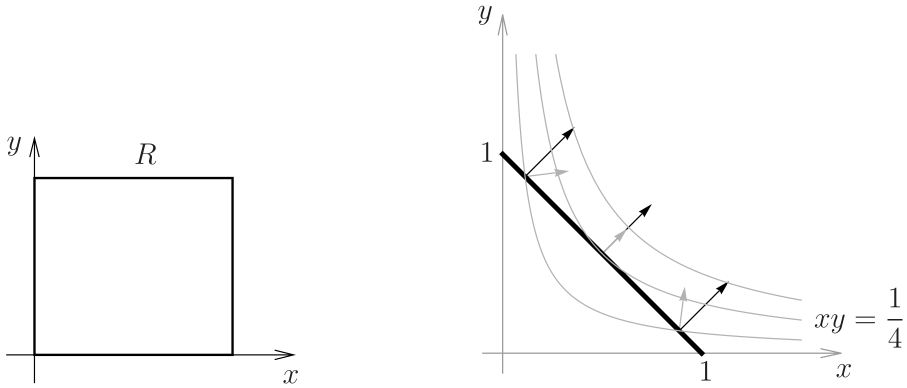

Figura 2: Retângulo de perímetro fixo com área máxima

Então:

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

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

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

Na Figura 2 está ilustrado o método dos multiplicadores de Lagrange. De facto, nela estão representados os conjuntos de nível da função f, ou seja, $xy = c\ ;\ c > 0$ (área), o conjunto de nível zero de F, ou seja, $x + y = 1$ (perímetro), e os vetores ∇f e ∇F nos pontos de intersecção.

O máximo de $f(x,y) = xy$ é atingido quando o respetivo gradiente $\nabla f(x,y) = (y,x)$ é múltiplo (colinear) do gradiente $\nabla F(x,y)$, e isso acontece no ponto em que a linha $xy = \frac{1}{4}$ é tangente à reta $x + y = 1$. Nos pontos de intersecção das linhas $xy = c \neq \dfrac{1}{4}$ com a linha $x+y = 1$, os vetores $\nabla f(x,y)$ e $\nabla F(x,y)$ não são colineares, ou seja, nesses pontos tais linhas não são tangentes.

Exemplo 2.2 Seja L o conjunto definido pelo sistema de equações

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

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 3.

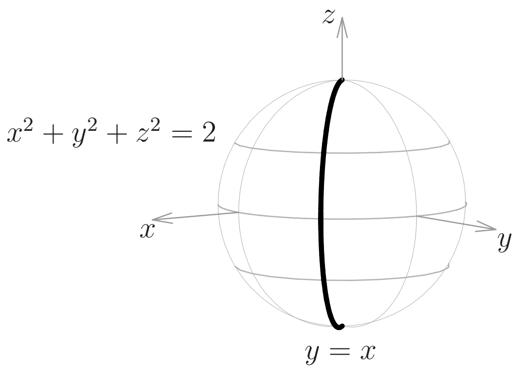

Figura 3: Circunferência em $\mathbb{R}^3$ dada por $x^2+y^2+z^2=2\ ;\ y=x$

Seja $f(x,y,z)=x^2+y^2+(z-1)^2$. Esta é a função a minimizar em L. É claro que L é um conjunto compacto em $\mathbb{R}^3$ e, sendo f contínua, tem mínimo nesse conjunto.

A função $\sqrt{x^2+y^2+(z-1)^2}$ é a distância de um ponto (x,y,z) ao ponto $(0,0,1)$. No entanto, esta função não é diferenciável no ponto $(0,0,1)$ e, dado que no método dos multiplicadores de Lagrange as funções envolvidas devem ser de classe C^1, não é elegível para a solução do problema.

O quadrado da distância definido pela função $f(x,y,z)=x^2+y^2+(z-1)^2$, de classe C^1, resolve o mesmo problema.

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 se deduz

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

Tendo em conta que $y = x$, da primeira e segunda equações deduz-se que $\lambda_2 = 0$. Da primeira equação obtém-se

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

Se $\lambda_1 = 1$, então da terceira equação obtém-se $0 = 1$. Assim, $y = x = 0$, e da quarta equação deduz-se $z = \sqrt{2}$ 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 2.3 Na Figura 4 encontra-se representada a linha definida pelo sistema de equações

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

Esta linha resulta da intersecção do plano definido por $x + y + z = 1$ com o parabolóide dado por $z = x^2 + y^2$.

Pretende-se determinar o ponto desta linha que apresenta maior cota, ou seja, coordenada z mais elevada.

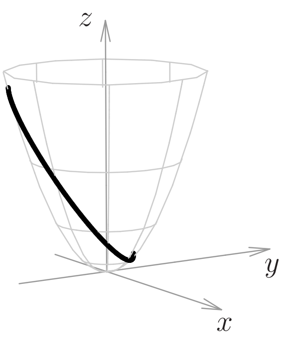

Figura 4: Linha em $\mathbb{R}^3$ dada por $z = x^2 + y^2$; $x + y + z = 1$

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

Aplicando o método dos multiplicadores de Lagrange, obtém-se

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

Da primeira equação é claro que $\lambda_1 = 0$ ou $x = y$.

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

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

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

e, portanto,

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

Sendo $y = x$ e $z = 1 - x - y$, os pontos que resolvem o sistema são:

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

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

Exemplo 2.4 Quais os pontos da elipse definida pela equação $x^2 + y^2 - xy = 3$ que se encontram mais afastados do eixo Ox?

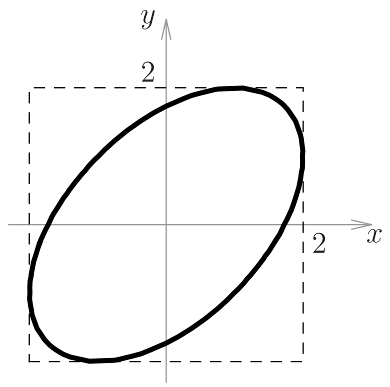

Figura 5: Linha em $\mathbb{R}^2$ dada por $x^2 + y^2 - xy = 3$

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

Aplicando o método dos multiplicadores de Lagrange, obtém-se o sistema

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

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

Note-se que estes pontos estão ambos à distância 2 do eixo Ox. Na Figura 5 encontra-se representada esta elipse onde se pode verificar que os pontos mais afastados tanto do eixo Ox como do eixo Oy se encontram à distância 2.

Note-se, também, que nesses pontos as linhas retas $y = 2$ e $y = -2$ são tangentes à elipse.

Do mesmo modo, os pontos da elipse mais afastados do eixo Oy podem ser determinados usando a função $f(x, y) = x$.

Exemplo 2.5 Seja $f(x, y) = x + y$ e $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 2\}$.

É claro que a função f é contínua, e D é um subconjunto compacto de $\mathbb{R}^2$. Pelo teorema de Weierstrass, a função f tem máximo e mínimo em D.

Os extremos de f em D ou estão no interior e são pontos críticos ou estão na fronteira. Sendo a fronteira dada pela equação $x^2 + y^2 = 2$, o método dos multiplicadores de Lagrange permite determinar os extremos que aí se encontram.

Sendo $\nabla f(x, y) = (1, 1)$, a função f não tem pontos críticos e, portanto, os extremos estão na fronteira, ou seja, na circunferência definida pela equação $x^2 + y^2 = 2$. Os pontos correspondentes são as soluções do sistema

$$\begin{cases} 1 = 2\lambda x \\ 1 = 2\lambda y \\ x^2 + y^2 = 2, \end{cases}$$

de onde se obtém $2\lambda(x - y) = 0$. O caso em que $\lambda = 0$ não ocorre, porque daí viria $1 = 0$. Portanto, $y = x$ e, da terceira equação, deduz-se que $x^2 = 1 \Leftrightarrow x = -1 \ \vee \ x = 1$.

Sendo $f(-1, -1) = -2$ e $f(1, 1) = 2$, conclui-se que $(-1, -1)$ é o ponto de mínimo e $(1, 1)$ é o ponto de máximo.

3 Exercícios

1. Determine os extremos da função $f(x, y) = 2x(y - 1)$ no círculo de raio igual a 1 e centro na origem.

2. Determine os extremos da função $f(x, y) = x^3 + y$ no conjunto limitado pela elipse $x^2 + \dfrac{y^2}{2} = 1$.

3. Determine os extremos da função $f(x,y,z) = (x+z)(y+z)$ na esfera $x^2+y^2+z^2=1$.

4. Calcule a distância do ponto $(0,1,2)$ ao conjunto dado pela equação $z = x^2+y^2$.

5. Mostre que no conjunto definido pelas equações $x^2+y^2+z^2=1\ ;\ z=1+y$ existe um ponto com maior coordenada x. Determine esse ponto.

6. Determine os pontos da linha definida pelas equações $z = x^2-y^2\ ;\ x^2+y^2=1$ em que a reta tangente é horizontal.

7. Determine os extremos da função $f(x,y) = xy$ no conjunto

$$T = \{(x,y)\in\mathbb{R}^2 : x+y=1\ ;\ x\geq 0\ ;\ y\geq 0\}$$

e mostre que $\sqrt{xy} \leq \dfrac{x+y}{2}$.