

Capítulo 12

Otimização com restrições

Um dos objetivos deste curso é o de tentar criar mecanismos que permitam resolver o problema de otimização

$$\begin{array}{l} \text{maximizar } f(\mathbf{x}) \\ \text{sujeito a } \mathbf{x} \in D, \end{array}$$

isto é, queremos encontrar (caso exista) um ponto \mathbf{p} no conjunto admissível D para o qual o valor da função-objetivo f seja o maior possível se comparado com o valor de f nos demais pontos de D .

No capítulo anterior vimos que se o problema de otimização possui uma solução \mathbf{p} no interior do conjunto admissível D então, pela regra de Fermat, \mathbf{p} deve ser um ponto crítico de f , isto é, $\nabla f(\mathbf{p}) = \mathbf{0}$.

Contudo, se uma solução \mathbf{p} do problema de otimização não está no interior do conjunto admissível D , então não é necessário que \mathbf{p} seja um ponto crítico de f . Considere, por exemplo, o seguinte problema de otimização:

$$\begin{array}{l} \text{maximizar } f(x_1, x_2) = x_1^2 + x_2^2 \\ \text{sujeito a } (x_1, x_2) \in D = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2/4 + x_2^2/9 = 1\}. \end{array}$$

Este problema de otimização possui duas soluções, a saber, os pontos

$$\mathbf{p} = (0, -3) \quad \text{e} \quad \mathbf{q} = (0, +3).$$

Por outro lado, \mathbf{p} e \mathbf{q} são pontos de fronteira de D e não são pontos críticos de f (veja a figura (12.1)). Se não podemos aplicar a regra de Fermat para estes pontos, como vamos caracterizá-los algebricamente? A resposta é dada pelo teorema dos multiplicadores de Lagrange ou, mais geralmente, pelo teorema de Karush-Kuhn-Tucker, que vamos estudar neste capítulo.

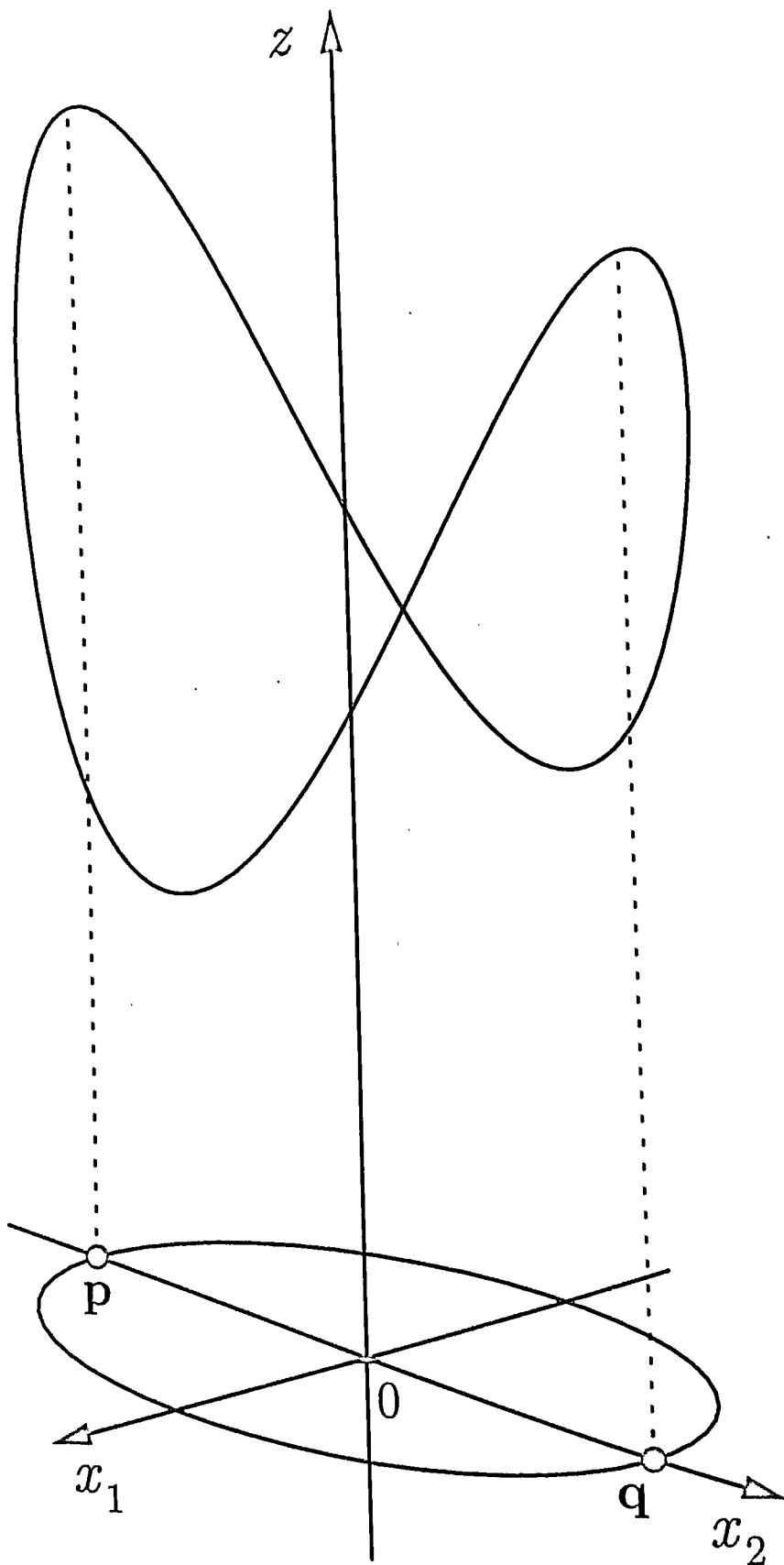


Figura 12.1: Os pontos $p = (0, -3)$ e $q = (0, +3)$ são extremos globais de f em $D = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2/4 + x_2^2/9 = 1\}$ mas p e q não são pontos críticos de f .

12.1

Vamos
duas va
de uma

Na figu
admiss
Vamos
sempre
na figu

Con
conjur
de ní
lado d
passa
arbitr
funçã
pode
curvas
e do l
que p
arbitr
funçã
pode
um p
nível
pode

Pe
é tan
curva
do la
passa
direi

12.1 Otimização com uma restrição em igualdade

Vamos começar com um exemplo bem simples: uma função-objetivo de duas variáveis e um conjunto admissível D definido como a curva de nível de uma função h , também de duas variáveis. Assim:

$$\begin{array}{l} \text{maximizar} \\ \text{sujeito a} \end{array} \quad f(x_1, x_2) \quad (x_1, x_2) \in D = \{(x_1, x_2) \in \mathbb{R}^2 \mid h(x_1, x_2) = c\}.$$

Na figura (12.2) estão desenhadas a curva de nível $h = c$ de h (o conjunto admissível D) e as curvas de nível $f = c_1$, $f = c_2$, $f = c_3$ e $f = c_4$ de f . Vamos supor que o valor da função f em uma de suas curvas de nível é sempre maior do que o valor nas curvas de nível abaixo desta. Sendo assim, na figura (12.2), temos $c_1 < c_2 < c_3 < c_4$.

Com este cenário, o ponto q não pode ser um extremo local de f no conjunto admissível D . Para ver isto, basta observar que existem curvas de nível f que interceptam D em pontos arbitrariamente próximos e do lado direito de q . Estas curvas de nível estão *acima* da curva de nível que passa pelo ponto q . Portanto, existem pontos do conjunto admissível D arbitrariamente próximos e do lado direito de q para os quais o valor da função f é *maior* do que o valor da função no ponto q . Conclusão: q não pode ser um ponto de máximo local de f em D . Analogamente, existem curvas de nível de f que interceptam D em pontos arbitrariamente próximos e do lado esquerdo de q . Estas curvas de nível estão *abaixo* da curva de nível que passa pelo ponto q . Portanto, existem pontos do conjunto admissível D arbitrariamente próximos e do lado esquerdo de q para os quais o valor da função f é *menor* do que o valor da função no ponto q . Conclusão: q não pode ser um ponto de mínimo local de f em D . Este argumento mostra que um ponto do conjunto admissível D onde a curva de nível de f e a curva de nível de h (que define o conjunto D) se interceptam *transversalmente* não pode ser um extremo local de f em D .

Por outro, na figura (12.2), a curva de nível de f que passa pelo ponto p é *tangente* à curva de nível de h (que define o conjunto admissível D). As curvas de nível que interceptam o conjunto D tanto do lado direito quanto do lado esquerdo do ponto p estão sempre *abaixo* da curva de nível de f que passa por p . Portanto, o valor de f em pontos do lado esquerdo e do lado direito de p é sempre *menor* do que o valor da função no ponto p . Conclusão:

o ponto p é um ponto de máximo local de f em D .

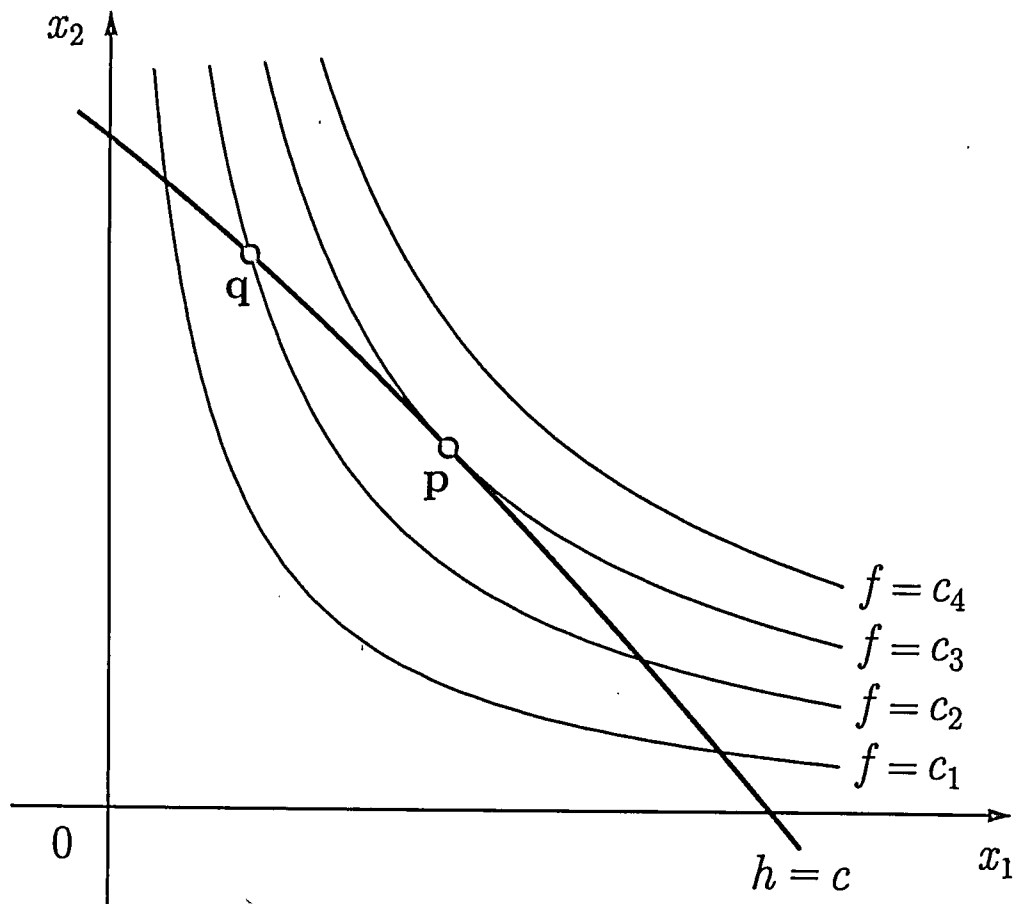


Figura 12.2: Se p é um ponto de máximo de f na curva de nível $h = c$, então a curva de nível de f que passa por p é tangente à curva de nível $h = c$.

Mas como caracterizar algebricamente um ponto p com esta propriedade? Se vetor gradiente $\nabla h(p)$ de h no ponto p é diferente de zero, pelo teorema (9.5) da página 333, sabemos que ele é perpendicular à curva de nível $h = c$ no ponto p e, analogamente, se o vetor gradiente $\nabla f(p)$ de f no ponto p é diferente de zero, então ele é perpendicular à curva de nível $f = c_3$ no ponto p . Como as curvas de nível $h = c$ e $f = c_3$ são *tangentes* em p , concluímos que os vetores $\nabla h(p)$ e $\nabla f(p)$ devem ser *paralelos* em p (veja a figura (12.3)). Desta maneira, existe um número real λ tal que

$$\nabla f(p) = \lambda \cdot \nabla h(p). \tag{12.1}$$

É a condição (12.1) que substitui a regra de Fermat como um "filtro" de primeira ordem para extremos de uma função f em um conjunto admissível D , construído como a curva de nível de uma função h .

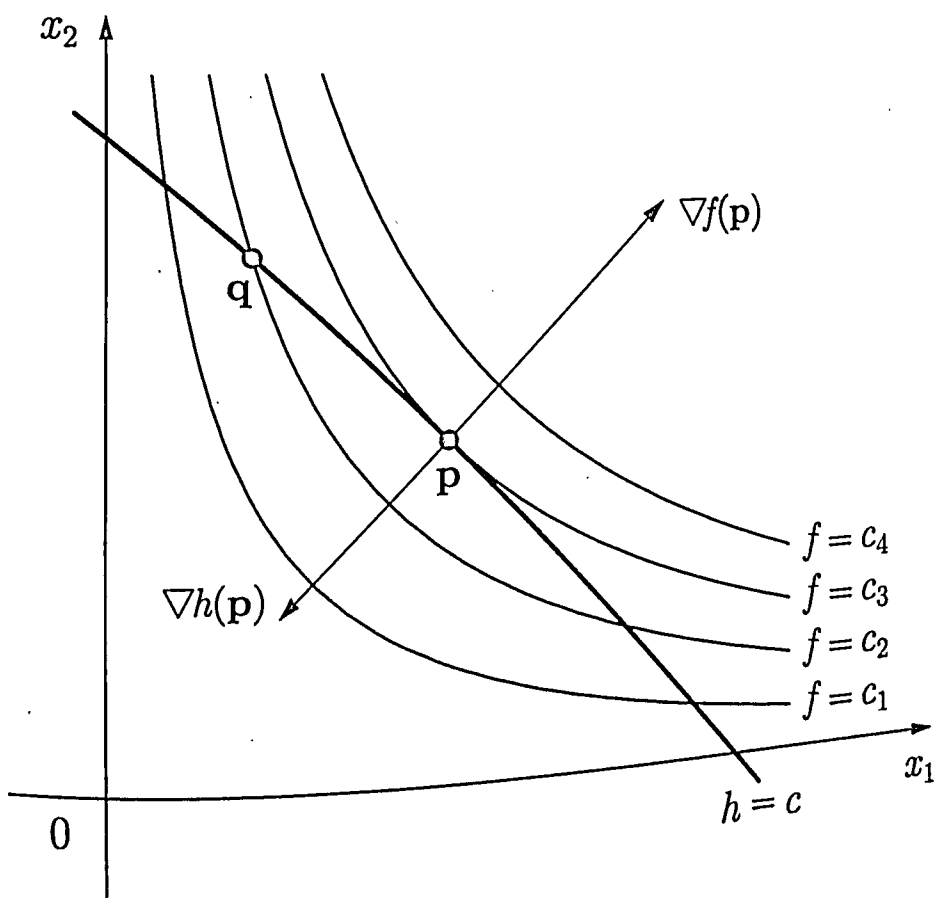
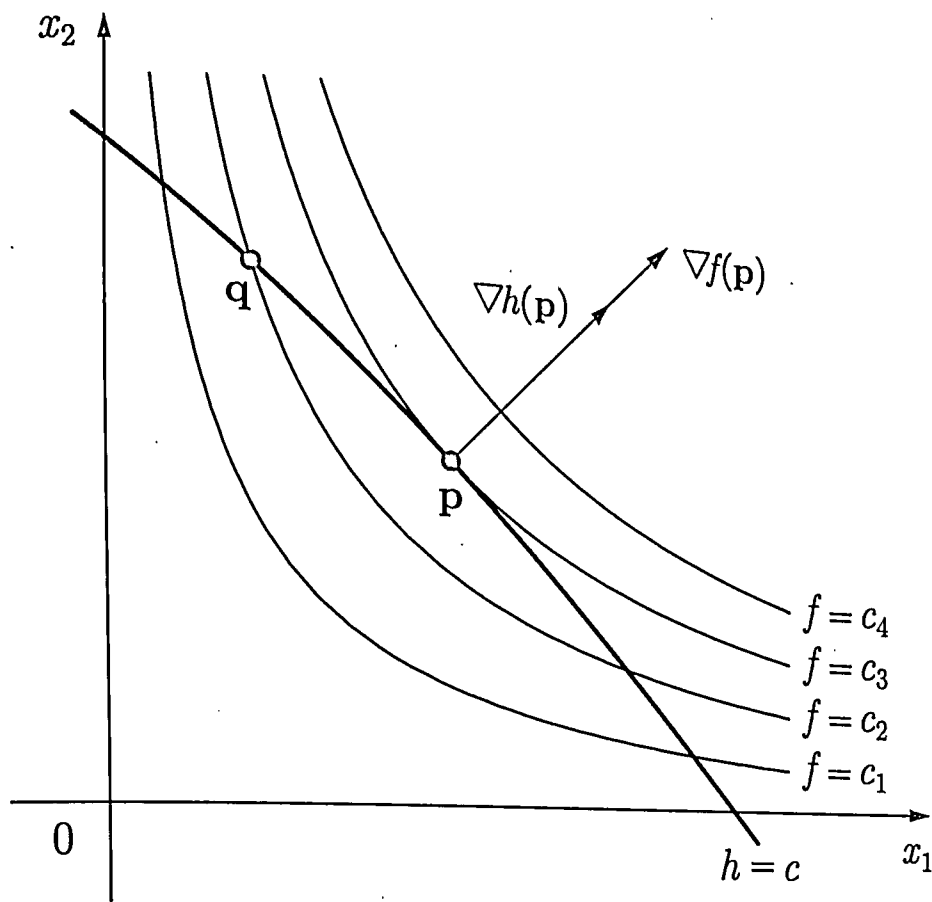


Figura 12.3: Os gradientes $\nabla f(p)$ e $\nabla h(p)$ são paralelos em um extremo local de f em D .

Teorema 12.1 (DOS MULTIPLICADORES DE LAGRANGE) Sejam f e h funções de classe C^1 de duas variáveis e seja $p = (p_1, p_2)$ uma solução (local) do problema de otimização

$$\begin{array}{l} \text{maximizar} \\ \text{sujeito a} \end{array} \quad \begin{array}{l} f(x_1, x_2) \\ (x_1, x_2) \in D = \{(x_1, x_2) \in \mathbb{R}^2 \mid h(x_1, x_2) = c\}. \end{array}$$

Suponha que $p = (p_1, p_2)$ satisfaça a seguinte *condição de regularidade*:

$$\nabla h(p) = \left(\frac{\partial h}{\partial x_1}(p_1, p_2), \frac{\partial h}{\partial x_2}(p_1, p_2) \right) \neq (0, 0).$$

Então existe um número real λ^* (o multiplicador de Lagrange) tal que

$$\begin{cases} \nabla f(p) = \lambda^* \cdot \nabla h(p), \\ h(p) = c, \end{cases} \quad (12.2)$$

isto é, o ponto (p_1, p_2, λ^*) é solução do sistema

$$\begin{cases} \frac{\partial f}{\partial x_1}(x_1, x_2) = \lambda \cdot \frac{\partial h}{\partial x_1}(x_1, x_2), \\ \frac{\partial f}{\partial x_2}(x_1, x_2) = \lambda \cdot \frac{\partial h}{\partial x_2}(x_1, x_2), \\ h(x_1, x_2) = c, \end{cases} \quad (12.3)$$

denominado de *condições de primeira ordem* para o ponto $p = (p_1, p_2)$.

Observação. A condição de regularidade $\nabla h(p) \neq 0$ exigida pelo teorema dos multiplicadores de Lagrange garante que podemos usar o teorema (9.5) para concluir que o vetor gradiente $\nabla h(p)$ é perpendicular à curva de nível $h = c$ no ponto p . Lembramos que este teorema, por sua vez, usa o teorema da função implícita para demonstrar este fato geométrico, daí a necessidade da condição de regularidade. Por outro lado, não é necessário que o vetor gradiente $\nabla f(p)$ de f no ponto p seja diferente de zero (apesar de termos usado este fato na justificativa geométrica que precede o teorema (12.1)). De fato, se p é um ponto de máximo de f no conjunto admissível D e $\nabla f(p) = 0$, basta considerarmos $\lambda = 0$ no sistema (12.2). Evidentemente, neste caso, as curvas de nível de f e h que passam por p não são necessariamente tangentes (veja o exercício resolvido (12.2) mais adiante). Para uma demonstração do

teorema

Existe
ordem (1

Agora, co

segue-se

samente s

muito int

problema

pelas solu

é o mesm

pelos can

três variá

Observa

no conjun

ao invés d

o teorema

caso, um

condição

ponto $(p_1,$ **Teore** h funçõ

(máxim

teorema dos multiplicadores de Lagrange, veja a referência [71].

Existe uma maneira conveniente de se escrever as condições de primeira ordem (12.3). Para isto, definimos o *lagrangeano*

$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda \cdot [h(x_1, x_2) - c].$$

Agora, como

$$\begin{cases} \frac{\partial L}{\partial x_1}(x_1, x_2, \lambda) = \frac{\partial f}{\partial x_1}(x_1, x_2) - \lambda \cdot \frac{\partial h}{\partial x_1}(x_1, x_2), \\ \frac{\partial L}{\partial x_2}(x_1, x_2, \lambda) = \frac{\partial f}{\partial x_2}(x_1, x_2) - \lambda \cdot \frac{\partial h}{\partial x_2}(x_1, x_2), \\ \frac{\partial L}{\partial \lambda}(x_1, x_2, \lambda) = -(h(x_1, x_2) - c), \end{cases} \quad (12.4)$$

segue-se que (x_1, x_2, λ) satisfaz as condições de primeira ordem (12.3) se, e somente se, (x_1, x_2, λ) é um *ponto crítico* do lagrangeano L . Este resultado é muito interessante! Ele diz que, procurar pelos candidatos à solução de um problema de otimização com uma restrição em duas variáveis, isto é, procurar pelas soluções do sistema formado pelas condições de primeira ordem (12.3), é o mesmo que procurar pelos pontos críticos do lagrangeano, isto é, procurar pelos candidatos à solução de um problema de otimização sem restrições em três variáveis.

Observação. Se estivéssemos interessados em minimizar a função-objetivo f no conjunto admissível

$$D = \{(x_1, x_2) \in \mathbb{R}^2 \mid h(x_1, x_2) = c\},$$

ao invés de maximizá-la, o mesmo argumento geométrico usado para motivar o teorema (12.1) poderia ser empregado para concluir que, também neste caso, um ponto $\mathbf{p} = (p_1, p_2)$ que é mínimo de f em D e que satisfaz a condição de regularidade $\nabla h(\mathbf{p}) \neq \mathbf{0}$ é tal que existe $\lambda \in \mathbb{R}$ para o qual o ponto (p_1, p_2, λ) satisfaz as condições de primeira ordem (12.3) ou ((12.4)).

Teorema 12.2 (DOS MULTIPLICADORES DE LAGRANGE) Sejam f e h funções de classe C^1 de duas variáveis e seja $\mathbf{p} = (p_1, p_2)$ um extremo (máximo ou mínimo) local de f no conjunto admissível

$$D = \{(x_1, x_2) \in \mathbb{R}^2 \mid h(x_1, x_2) = c\}.$$

Suponha que $\mathbf{p} = (p_1, p_2)$ satisfaça a seguinte condição de regularidade:

$$\nabla h(\mathbf{p}) = \left(\frac{\partial h}{\partial x_1}(p_1, p_2), \frac{\partial h}{\partial x_2}(p_1, p_2) \right) \neq (0, 0).$$

Então existe um número real λ^* (o multiplicador de Lagrange) tal que $(\mathbf{p}, \lambda) = (p_1, p_2, \lambda)$ é ponto crítico do lagrangeano

$$L(x_1, x_2, \lambda^*) = f(x_1, x_2) - \lambda^* \cdot [h(x_1, x_2) - c].$$

Em outras palavras:

$$\begin{cases} \frac{\partial L}{\partial x_1}(p_1, p_2, \lambda^*) = 0, \\ \frac{\partial L}{\partial x_2}(p_1, p_2, \lambda^*) = 0, \\ \frac{\partial L}{\partial \lambda}(p_1, p_2, \lambda^*) = 0. \end{cases}$$

Exercício resolvido 12.1 Resolva o seguinte problema de otimização

maximizar $f(x_1, x_2) = x_1 x_2$
 sujeito a $(x_1, x_2) \in D = \{(x_1, x_2) \in \mathbb{R}^2 \mid h(x_1, x_2) = x_1 + 4x_2 = 8\}$.

SOLUÇÃO: vamos subdividir a resolução deste problema em vários passos.

- PASSO 1. As funções f e h são de classe C^1 como soma e multiplicação de funções de classe C^1 .
- PASSO 2. Verificar a condição de regularidade.

A condição de regularidade exigida pelo teorema (12.2) diz que o gradiente de h deve ser diferente de zero em cada solução do problema de otimização. Mas como verificar a condição de regularidade se ainda não conhecemos a solução do problema de otimização? De fato, o que queremos é justamente encontrar a solução (ou soluções) do problema! A idéia é mostrar que *todos* os pontos do conjunto admissível (isto é, todos os pontos (x_1, x_2) em \mathbb{R}^2 tais que $h(x_1, x_2) = x_1 + 4x_2 = 8$) satisfazem a condição de regularidade! Com isto, caso exista alguma solução do problema de otimização, certamente ela irá satisfazer a condição

de re...
Com...

segue...
 $\nabla h(x$

• PASSO

$L(x_1,$

• PASSO

• PASSO

orden

De (1

concl

$x_2 =$

orden

• PASSO

Apes

suir u

$(x_1, x$

regra

orden

ponto

do sis

classi

no ca

para

de regularidade exigida pelo teorema dos multiplicadores de Lagrange.
Como

$$\nabla h(x_1, x_2) = \left(\frac{\partial h}{\partial x_1}(x_1, x_2), \frac{\partial h}{\partial x_2}(x_1, x_2) \right) = (1, 4)$$

segue-se que $\nabla h(x_1, x_2) \neq (0, 0)$ para todo (x_1, x_2) em \mathbb{R}^2 , em particular, $\nabla h(x_1, x_2) \neq (0, 0)$ para todo ponto (x_1, x_2) do conjunto admissível.

- PASSO 3. Escrever o lagrangeano.

$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda \cdot [h(x_1, x_2) - c] = x_1 x_2 - \lambda \cdot (x_1 + 4x_2 - 8).$$

- PASSO 4. Escrever as condições de primeira ordem.

$$\begin{cases} \frac{\partial L}{\partial x_1}(x_1, x_2, \lambda) = 0, \\ \frac{\partial L}{\partial x_2}(x_1, x_2, \lambda) = 0, \\ \frac{\partial L}{\partial \lambda}(x_1, x_2, \lambda) = 0, \end{cases} \Rightarrow \begin{cases} x_2 - \lambda = 0, & (1) \\ x_1 - 4\lambda = 0, & (2) \\ x_1 + 4x_2 = 8. & (3) \end{cases}$$

- PASSO 5. Resolver o sistema correspondente às condições de primeira ordem.

De (1) e (2) temos $x_1 = 4\lambda$ e $x_2 = \lambda$. Substituindo estes valores em (3) concluímos que $4\lambda + 4\lambda = 8$, isto é, $\lambda = 1$. Portanto, $x_1 = 4\lambda = 4$ e $x_2 = \lambda = 1$. Assim, o sistema correspondente às condições de primeira ordem possui uma única solução: $(x_1, x_2, \lambda) = (4, 1, 1)$.

- PASSO 6. Conclusão.

Apesar do sistema correspondente às condições de primeira ordem possuir uma única solução, não podemos afirmar de imediato que o ponto $(x_1, x_2) = (4, 1)$ é solução do problema de otimização! A exemplo da regra de Fermat, nem todo ponto que satisfaz as condições de primeira ordem é solução do problema de otimização! O que dizer então do ponto $(4, 1)$? Mais adiante construiremos um classificador das soluções do sistema correspondente às condições de primeira ordem (isto é, um classificador dos pontos críticos do lagrangeano) da mesma maneira que, no capítulo anterior, construímos um classificador de pontos críticos para o caso de problemas de otimização sem restrições. Por outro lado,

com o desenho das curvas de nível de f e do conjunto admissível (veja a figura (12.4)), não é difícil de se convencer que, de fato, $(4, 1)$ é solução do problema de otimização. ■

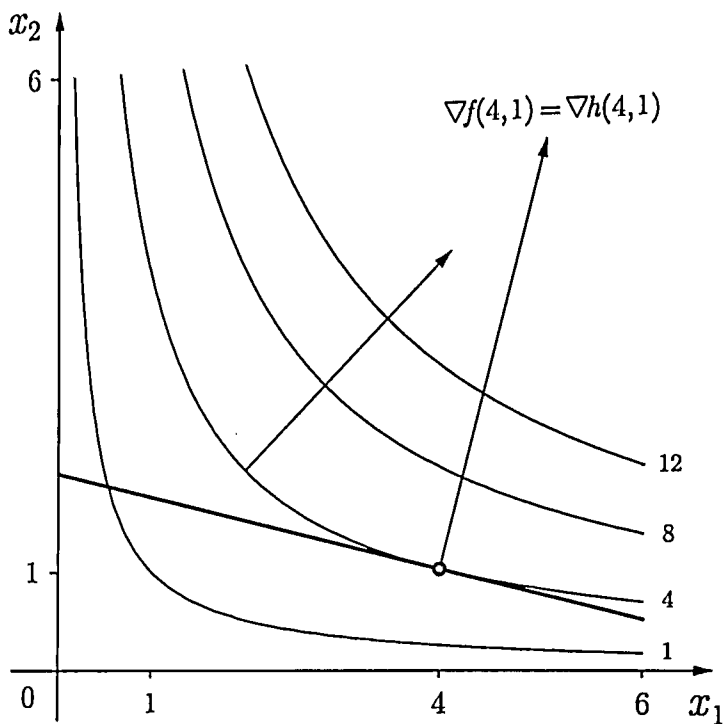


Figura 12.4: Desenho das curvas de nível de f (os números indicam o valor do nível) e do conjunto admissível (curva com traçado mais forte).

Exercício resolvido 12.2 Resolva o seguinte problema de otimização

maximizar $f(x_1, x_2) = x_1^2 x_2$ sujeito a $(x_1, x_2) \in D = \{(x_1, x_2) \in \mathbb{R}^2 \mid h(x_1, x_2) = 2x_1^2 + x_2^2 = 3\}$.

SOLUÇÃO: vamos subdividir a resolução deste problema em vários passos.

- PASSO 1. As funções f e h são de classe C^1 como soma e multiplicação de funções de classe C^1 .
- PASSO 2. Verificar a condição de regularidade.

A condição de regularidade exigida pelo teorema (12.2) diz que o gradiente de h deve ser diferente de zero em cada solução do problema de otimização. Mas como verificar a condição de regularidade se ainda não conhecemos a solução do problema de otimização? De fato, o que queremos é justamente encontrar a solução (ou soluções) do problema!

A idéia
 todos o
 zem a
 do pro
 de regu
 Como

segue-s
 Como o
 $h(0, 0)$
 (x_1, x_2)

• PASSO

$L(x_1, x_2)$

• PASSO

$$\left\{ \begin{array}{l} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \end{array} \right.$$

• PASSO
 ordem.

De (1) t

Se $x_1 =$

assim, d

são soluç

Se $x_2 =$

obtemos

A idéia é mostrar que *todos* os pontos do conjunto admissível (isto é, todos os pontos (x_1, x_2) em \mathbb{R}^2 tais que $h(x_1, x_2) = 2x_1^2 + x_2^2 = 3$) satisfazem a condição de regularidade! Com isto, caso exista alguma solução do problema de otimização, certamente ela irá satisfazer a condição de regularidade exigida pelo teorema dos multiplicadores de Lagrange. Como

$$\nabla h(x_1, x_2) = \left(\frac{\partial h}{\partial x_1}(x_1, x_2), \frac{\partial h}{\partial x_2}(x_1, x_2) \right) = (4x_1, 2x_2)$$

segue-se que $\nabla h(x_1, x_2) = (0, 0)$ se, e somente se, $(x_1, x_2) = (0, 0)$. Como o ponto $(0, 0)$ não pertence ao conjunto admissível (uma vez que $h(0, 0) = 0 < 3$) segue-se que $\nabla h(x_1, x_2) \neq (0, 0)$ para todo ponto (x_1, x_2) do conjunto admissível.

- PASSO 3. Escrever o lagrangeano.

$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda \cdot [h(x_1, x_2) - c] = x_1^2 x_2 - \lambda \cdot (2x_1^2 + x_2^2 - 3).$$

- PASSO 4. Escrever as condições de primeira ordem.

$$\begin{cases} \frac{\partial L}{\partial x_1}(x_1, x_2, \lambda) = 0, \\ \frac{\partial L}{\partial x_2}(x_1, x_2, \lambda) = 0, \\ \frac{\partial L}{\partial \lambda}(x_1, x_2, \lambda) = 0, \end{cases} \Rightarrow \begin{cases} 2x_1 x_2 - 4\lambda x_1 = 0, & (1) \\ x_1^2 - 2\lambda x_2 = 0, & (2) \\ 2x_1^2 + x_2^2 = 3. & (3) \end{cases}$$

- PASSO 5. Resolver o sistema correspondente às condições de primeira ordem.

De (1) temos que $2x_1(x_2 - 2\lambda) = 0$. Logo $x_1 = 0$ ou $x_2 = 2\lambda$.

Se $x_1 = 0$ então, de (3), $x_2^2 = 3$, isto é, $x_2 = -\sqrt{3}$ ou $x_2 = +\sqrt{3}$. Sendo assim, de (2), temos $\lambda = 0$. Concluimos então que

$$(0, -\sqrt{3}, 0) \quad \text{e} \quad (0, +\sqrt{3}, 0)$$

são soluções do sistema. Se $x_2 = 2\lambda$, então $x_2^2 = 4\lambda^2$ e, de (2), $x_1^2 = 4\lambda^2$. Substituindo em (3) obtemos $8\lambda^2 + 4\lambda^2 - 3 = 0$, isto é, $\lambda^2 = 1/4$. Assim, $\lambda = -1/2$ ou

$\lambda = +1/2$. Se $\lambda = +1/2$, então $x_2 = +1$ e $x_1^2 = 1$, isto é, $x_1 = \pm 1$. Se $\lambda = -1/2$, então $x_2 = -1$ e $x_1^2 = 1$, isto é, $x_1 = \pm 1$. Desta maneira,

$$(-1, +1, +1/2), \quad (+1, +1, +1/2), \quad (-1, -1, -1/2) \quad \text{e} \\ (+1, -1, -1/2)$$

também são soluções do sistema associado às condições de primeira ordem.

• PASSO 6. Conclusão.

A função objetivo f é contínua (pois é de classe C^1) e o conjunto admissível é compacto (pois é uma elipse, veja a figura (12.5)). Pelo teorema de Weierstrass, f possui máximos e mínimos globais no conjunto admissível. Pelo teorema dos multiplicadores de Lagrange, qualquer solução do problema de otimização que estamos estudando deve ser solução do sistema associado às condições de primeira ordem (a condição de regularidade é fundamental aqui). Desta maneira, os pontos de máximo de $f(x_1, x_2) = x_1^2 x_2$ sujeitos à restrição $h(x_1, x_2) = 2x_1^2 + x_2^2 = 3$ estão entre

$$(0, -\sqrt{3}), \quad (0, +\sqrt{3}), \quad (-1, +1), \quad (+1, +1), \quad (-1, -1), \quad \text{e} \\ (+1, -1).$$

Para encontrá-los, basta calcular a função-objetivo nestes pontos e selecionar os de maior valor. Como

$$f(0, -\sqrt{3}) = f(0, +\sqrt{3}) = 0, \\ f(-1, -1) = f(+1, -1) = -1, \\ f(+1, +1) = f(-1, +1) = -1,$$

segue-se que os pontos $(+1, +1)$ e $(-1, +1)$ são soluções do problema de otimização.

Observe que, na figura (12.5), os vetores gradientes de f e h são paralelos nos pontos $(-1, -1)$, $(-1, +1)$, $(+1, -1)$ e $(+1, +1)$. As curvas de nível de f e h que passam por estes pontos são, de fato, tangentes. Por outro lado, o vetor gradiente de f se anula nos pontos $(0, -\sqrt{3})$ e $(0, +\sqrt{3})$, enquanto que o vetor gradiente de h não. As curvas de nível de f (o eixo y) e h (a elipse $2x_1^2 + x_2^2 = 3$) não são tangentes nestes pontos! Mesmo assim, os pontos $(x_1, x_2, \lambda) = (0, -\sqrt{3}, 0)$ e $(x_1, x_2, \lambda) = (0, +\sqrt{3}, 0)$ satisfazem as condições

de prime
é um pon
em D .

Figur

A co

Para voc
que cons

de primeira ordem e, pela figura (12.5), não é difícil de ver que $(0, -\sqrt{3})$ é um ponto de máximo local e $(0, +\sqrt{3})$ é um ponto de mínimo local de f em D . □

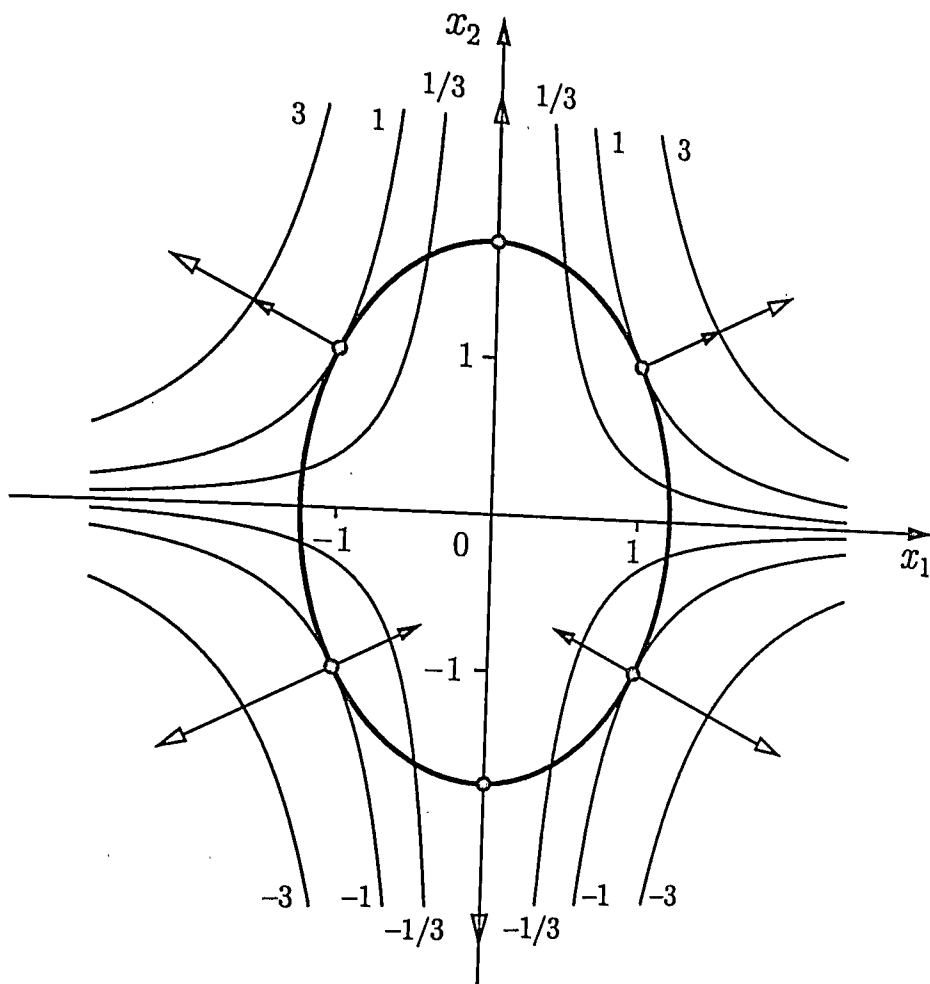


Figura 12.5: Desenho das curvas de nível de f (os números indicam o valor do nível) e do conjunto admissível (curva com traçado mais forte). Os vetores com seta branca indicam o gradiente de h enquanto que os vetores com seta preta indicam o gradiente de f .

CUIDADO!

CUIDADO!

CUIDADO!

A condição de regularidade não pode ser omitida no teorema dos multiplicadores de Lagrange!

Para você se convencer deste fato, considere o problema de otimização que consiste em

$$\begin{array}{l} \text{maximizar} \quad f(x_1, x_2) = x_2 \\ \text{sujeito a} \quad h(x_1, x_2) = x_1^2 + x_2^3 = 0. \end{array}$$

As funções f e h são de classe C^1 e o lagrangeano para o problema de otimização é dado por

$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda \cdot [h(x_1, x_2) - c] = x_2 - \lambda \cdot (x_1^2 + x_2^3),$$

de modo que as condições de primeira ordem para o problema de otimização são dadas por

$$\begin{cases} \frac{\partial L}{\partial x_1}(x_1, x_2, \lambda) = 0, \\ \frac{\partial L}{\partial x_2}(x_1, x_2, \lambda) = 0, \\ \frac{\partial L}{\partial \lambda}(x_1, x_2, \lambda) = 0, \end{cases} \Rightarrow \begin{cases} -2\lambda x_1 = 0, & (1) \\ 1 - 3\lambda x_2^2 = 0, & (2) \\ x_1^2 + x_2^3 = 0. & (3) \end{cases}$$

Não existe ponto (x_1, x_2, λ) em \mathbb{R}^3 que satisfaça as condições de primeira ordem! De fato: de (1) temos que $\lambda = 0$ ou $x_1 = 0$. Se $\lambda = 0$, de (2) temos que $1 - 2 \cdot 0 \cdot x_1 = 0$, isto é, $1 = 0$, um absurdo. Por outro lado, se $x_1 = 0$, de (3) temos que $x_2^3 = 0$, isto é, $x_2 = 0$. Portanto, de (2), segue-se que $1 - 2 \cdot \lambda \cdot 0 = 0$, isto é, $1 = 0$, também um absurdo.

Como o sistema associado às condições de primeira ordem não possui solução, podemos então concluir que o problema de otimização também não possui solução? A resposta a esta pergunta depende da *condição de regularidade*! Neste exemplo, o problema de otimização possui solução! Para encontrá-la, basta observar que

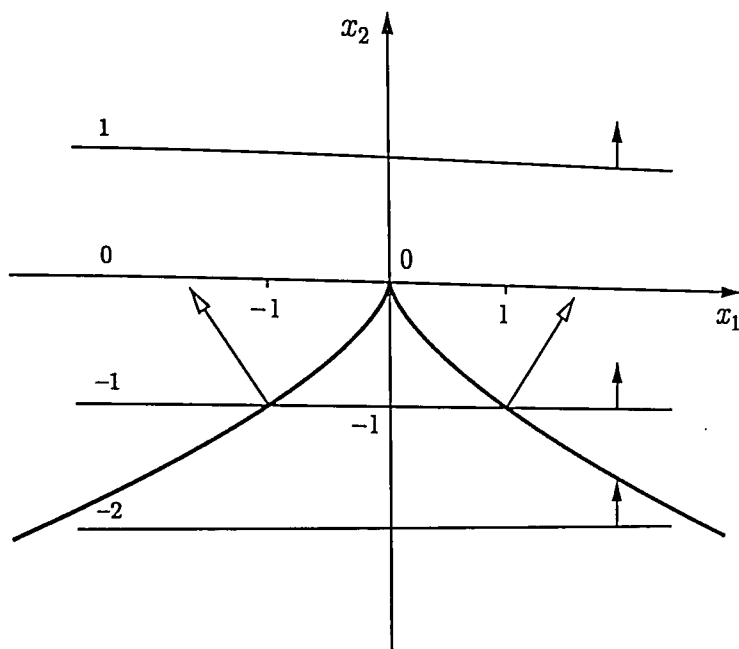
$$h(x_1, x_2) = x_1^2 + x_2^3 = 0 \Leftrightarrow x_2 = -\sqrt[3]{x_1^2},$$

isto é, um ponto (x_1, x_2) satisfaz a restrição $x_1^2 + x_2^3 = 0$ se, e somente se, $x_2 = -\sqrt[3]{x_1^2}$. Uma vez que $x_1^2 \geq 0$ para todo $x_1 \in \mathbb{R}$, segue-se que todos os pontos (x_1, x_2) do conjunto admissível

$$D = \{(x_1, x_2) \in \mathbb{R}^2 \mid h(x_1, x_2) = x_1^2 + x_2^3 = 0\}$$

são tais que $x_2 \leq 0$. Como nossa função-objetivo é $f(x_1, x_2) = x_2$, resolver o problema de otimização consiste em procurar o ponto do

conjunto admissível D com a maior ordenada. Desta maneira, não é difícil ver que $(0, 0)$ é a solução do problema de otimização. Na figura abaixo temos um desenho com as curvas de nível de f (os números indicam o valor do nível) e do conjunto admissível (curva com o traçado mais forte). Os vetores com seta branca indicam o gradiente de h enquanto que os vetores de seta preta indicam o gradiente de f .



Com esta figura, é fácil visualizar porque o ponto $(0, 0)$ é máximo global de $f(x_1, x_2) = x_2$ no conjunto admissível D formado pelos pontos que satisfazem a restrição $h(x_1, x_2) = x_1^2 + x_2^3 = 0$. Também é fácil visualizar porque não existem pontos que satisfazem as condições de primeira ordem: não existem pontos (x_1, x_2) na curva $x_1^2 + x_2^3 = 0$ para os quais $\nabla f(x_1, x_2) = \lambda \cdot \nabla h(x_1, x_2)$. Observe que, na origem $(0, 0)$, $\nabla f(0, 0) = (0, 1)$ enquanto que $\nabla h(0, 0) = (0, 0)$. Desta maneira, $\nabla f(0, 0) = (0, 1) \neq (0, 0) = \lambda \cdot (0, 0) = \lambda \cdot \nabla h(0, 0)$, para todo $\lambda \in \mathbb{R}$.

Este exemplo compromete a veracidade do teorema dos multiplicadores de Lagrange? A resposta é não! O teorema dos multiplicadores de Lagrange garante que qualquer extremo p de uma função f de classe C^1 em um conjunto admissível formado por uma curva de nível de uma função h também de classe C^1 irá satisfazer as condições de primeira ordem para algum $\lambda \in \mathbb{R}$ desde que p satisfaça a condição de regularidade, isto é, desde que $\nabla h(p) \neq \mathbf{0}$. A solução do problema de otimização acima não satisfaz a condição de regularidade pois $\nabla h(0, 0) = (0, 0)$. Moral da história: a condição de regularidade não pode ser omitida no teorema (12.2).

Vimos que não existe $\lambda \in \mathbb{R}$ tal que $\nabla f(0,0) = \lambda \cdot \nabla h(0,0)$, uma vez que $\nabla f(0,0) = (0,1)$ e $\nabla h(0,0) = (0,0)$. Mas se usássemos dois multiplicadores, um para a função-objetivo e outro para a função-restrição, isto é, se trocássemos a condição

$$\nabla f(\mathbf{p}) = \lambda^* \cdot \nabla h(\mathbf{p})$$

por

$$\lambda_0^* \cdot \nabla f(\mathbf{p}) = \lambda_1^* \cdot \nabla h(\mathbf{p}),$$

é fácil ver que a escolha $\lambda_0^* = 0$ faz com que $\lambda_0^* \cdot \nabla f(0,0) = \lambda_1^* \cdot \nabla h(0,0)$ para as funções f e h de nosso exemplo, independentemente da escolha de λ_1^* . Mais formalmente, se f e h são funções de classe C^1 e \mathbf{p} é um extremo de f sujeito a $h(\mathbf{x}) = c$, então existem multiplicadores λ_0^* e λ_1^* tais que

$$\begin{cases} \lambda_0^* \cdot \nabla f(\mathbf{p}) = \lambda_1^* \cdot \nabla h(\mathbf{p}), \\ h(\mathbf{p}) = c, \\ (\lambda_0^*, \lambda_1^*) \neq (0,0), \\ \lambda_0^* \in \{0,1\}. \end{cases}$$

Esta teorema é devido a Fritz John. A condição $(\lambda_0^*, \lambda_1^*) \neq (0,0)$ diz que os multiplicadores λ_0^* e λ_1^* não são simultaneamente nulos. Evidentemente, a situação desejável é que $\lambda_0^* = 1$, pois caso contrário, a função-objetivo desapareceria completamente das condições de primeira ordem. A hipótese de regularidade, $\nabla h(\mathbf{p}) \neq \mathbf{0}$, garante que podemos tomar λ_0^* igual a 1 no teorema de Fritz John.

O teorema (12.2) generaliza-se facilmente para funções que dependam de mais do que duas variáveis.

Teorema 12.3 (DOS MULTIPLICADORES DE LAGRANGE) Sejam f e h funções de classe C^1 de n variáveis e seja \mathbf{p} um extremo (máximo ou mínimo) local de f no conjunto admissível

$$D = \{\mathbf{x} \in \mathbb{R}^n \mid h(\mathbf{x}) = c\}.$$

Sup

Ent
(p,

Em

Observ

falsa! N

tal que

exempl

extremo

ponto (

ao prob

 $h(x, y, z$

12.2

Vamo

com o u

interess

conjunto

formado

restriçõe

é do que

de \mathbb{R}^n pa $D =$

Suponha que \mathbf{p} satisfaça a seguinte condição de regularidade:

$$\nabla h(\mathbf{p}) \neq \mathbf{0}.$$

Então existe um número real λ^* (o multiplicador de Lagrange) tal que $(\mathbf{p}, \lambda^*) \in \mathbb{R}^{n+1}$ é ponto crítico do lagrangeano

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda \cdot [h(\mathbf{x}) - c].$$

Em outras palavras:

$$\frac{\partial L}{\partial x_1}(\mathbf{p}, \lambda^*) = 0, \quad \dots, \quad \frac{\partial L}{\partial x_n}(\mathbf{p}, \lambda^*) = 0, \quad \frac{\partial L}{\partial \lambda}(\mathbf{p}, \lambda^*) = 0.$$

Observação. A recíproca do teorema dos multiplicadores de Lagrange é falsa! Nem todo ponto (\mathbf{p}, λ^*) que satisfaz as condições de primeira ordem é tal que \mathbf{p} é um extremo de f sujeito à restrição $h(\mathbf{x}) = c$. Como um contra-exemplo, considere o problema de otimização que consiste em encontrar os extremos de $f(x, y, z) = x^2 - y^2 + z$ sujeito à restrição $h(x, y, z) = z = 0$. O ponto $(\mathbf{p}, \lambda^*) = (0, 0, 0)$ satisfaz as condições de primeira ordem associadas ao problema (exercício) mas $\mathbf{p} = (0, 0)$, não é um extremo de f sujeito a $h(x, y, z) = z = 0$ (exercício).

12.2 Otimização com várias restrições em igualdade

Vamos estudar agora o caso em que o conjunto admissível é construído com o uso de várias restrições em igualdade. Mais precisamente, estamos interessados em encontrar os extremos de uma função f de n variáveis no conjunto admissível

$$D = \{\mathbf{x} \in \mathbb{R}^n \mid h_1(\mathbf{x}) = c_1, \dots, h_m(\mathbf{x}) = c_m\}$$

formado por todos os pontos \mathbf{x} de \mathbb{R}^n que satisfazem simultaneamente as m restrições $h_1(\mathbf{x}) = c_1, \dots, h_m(\mathbf{x}) = c_m$. Note que o conjunto D nada mais é do que o conjunto de nível da função vetorial $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \dots, h_m(\mathbf{x}))$ de \mathbb{R}^n para \mathbb{R}^m associado ao nível $\mathbf{c} = (c_1, \dots, c_m)$:

$$D = \mathcal{F}_{\mathbf{c}} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \dots, h_m(\mathbf{x})) = (c_1, \dots, c_m) = \mathbf{c}\}.$$

No caso de uma única restrição em igualdade ($m = 1$), vimos que a hipótese $\nabla h(\mathbf{p}) \neq \mathbf{0}$ é fundamental, no sentido que ela nos garante que a curva de nível $h(\mathbf{x}) = c$ pode ser escrita como o gráfico de uma função em uma vizinhança de ponto \mathbf{p} . No caso de várias restrições em igualdade, para podermos escrever o conjunto de nível $D = \mathcal{F}_c$ como o gráfico de uma função (vetorial), segundo o teorema (9.6) da página 343, é suficiente que a matriz jacobiana

$$\begin{bmatrix} \boxed{\nabla h_1(\mathbf{p})} \\ \vdots \\ \boxed{\nabla h_m(\mathbf{p})} \end{bmatrix}_{m \times n} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1}(\mathbf{p}) & \cdots & \frac{\partial h_1}{\partial x_n}(\mathbf{p}) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1}(\mathbf{p}) & \cdots & \frac{\partial h_m}{\partial x_n}(\mathbf{p}) \end{bmatrix}_{m \times n}$$

de h no ponto \mathbf{p} possua uma submatriz $m \times m$ inversível, isto é, existe uma escolha de variáveis $x_{i_1}, x_{i_2}, \dots, x_{i_m}$ (as variáveis independentes no teorema da função implícita) para as quais a matriz

$$\begin{bmatrix} \frac{\partial h_1}{\partial x_{i_1}}(\mathbf{p}) & \cdots & \frac{\partial h_1}{\partial x_{i_m}}(\mathbf{p}) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_{i_1}}(\mathbf{p}) & \cdots & \frac{\partial h_m}{\partial x_{i_m}}(\mathbf{p}) \end{bmatrix}_{m \times m}$$

é inversível. Uma maneira prática de fazê-lo, ao invés de tentar todas as escolhas possíveis, é a seguinte: basta verificar que o número de linhas não-nulas da matriz escalonada equivalente à matriz jacobiana de h no ponto \mathbf{p} é igual a m (o número de restrições). Dada uma matriz A , o número de linhas não-nulas da matriz escalonada equivalente a A é denominado *posto* de A . Pode-se mostrar que este número independe da maneira como A foi escalonada. Lembramos que uma matriz está escalonada se ela satisfaz as seguintes propriedades:

- (1) Se uma linha da matriz é nula, então todas as linhas subsequentes de A são nulas.
- (2) Se uma linha da matriz é não-nula, então o número de zeros no início desta linha é sempre *maior* do que o número de zeros no início da linha precedente (caso exista).

Com estas observações, o teorema dos multiplicadores de Lagrange generaliza-se facilmente para o caso de várias restrições em igualdade.

Teorema 12.4 (DOS MULTIPLICADORES DE LAGRANGE) Sejam f, h_1, \dots, h_m funções de classe C^1 de n variáveis e seja \mathbf{p} um extremo (máximo ou mínimo) local de f no conjunto admissível

$$D = \{\mathbf{x} \in \mathbb{R}^n \mid h_1(\mathbf{x}) = c_1, \dots, h_m(\mathbf{x}) = c_m\}.$$

Suponha que \mathbf{p} satisfaça a seguinte condição de regularidade: o posto da matriz jacobiana

$$\begin{bmatrix} \nabla h_1(\mathbf{p}) \\ \vdots \\ \nabla h_m(\mathbf{p}) \end{bmatrix}_{m \times n} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1}(\mathbf{p}) & \dots & \frac{\partial h_1}{\partial x_n}(\mathbf{p}) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1}(\mathbf{p}) & \dots & \frac{\partial h_m}{\partial x_n}(\mathbf{p}) \end{bmatrix}_{m \times n}$$

é igual a m (o número de restrições). Então existem números reais $\lambda_1^*, \dots, \lambda_m^*$ (os multiplicadores de Lagrange) tais que

$$\begin{cases} \nabla f(\mathbf{p}) = \lambda_1^* \cdot \nabla h_1(\mathbf{p}) + \dots + \lambda_m^* \cdot \nabla h_m(\mathbf{p}), \\ h_1(\mathbf{p}) = c_1, \\ \vdots \\ h_m(\mathbf{p}) = c_m. \end{cases}$$

isto é, o vetor gradiente $\nabla f(\mathbf{p})$ é uma combinação linear dos vetores gradiente $\nabla h_1(\mathbf{p}), \dots, \nabla h_m(\mathbf{p})$. O sistema acima é denominado de condições de primeira ordem para o problema de otimização. Equivalentemente, o ponto $(\mathbf{p}, \boldsymbol{\lambda}^*) = (p_1, \dots, p_n, \lambda_1^*, \dots, \lambda_m^*)$ é ponto crítico do lagrangeano

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \lambda_1 \cdot [h_1(\mathbf{x}) - c_1] - \dots - \lambda_m \cdot [h_m(\mathbf{x}) - c_m],$$

isto é,

$$\begin{aligned} \frac{\partial L}{\partial x_1}(\mathbf{p}, \boldsymbol{\lambda}^*) = 0, \quad \dots, \quad \frac{\partial L}{\partial x_n}(\mathbf{p}, \boldsymbol{\lambda}^*) = 0, \\ \frac{\partial L}{\partial \lambda_1}(\mathbf{p}, \boldsymbol{\lambda}^*) = 0, \quad \dots, \quad \frac{\partial L}{\partial \lambda_m}(\mathbf{p}, \boldsymbol{\lambda}^*) = 0. \end{aligned}$$

Exercício resolvido 12.3 Resolva o seguinte problema de otimização

$$\begin{array}{l} \text{maximizar} \quad xyz \\ \text{sujeito a} \quad x^2 + y^2 = 1, \\ \quad \quad \quad x + z = 1. \end{array}$$

SOLUÇÃO: vamos subdividir a resolução deste problema em vários passos.

- PASSO 1. As funções definidas por $f(x, y, z) = xyz$, $h_1(x, y, z) = x^2 + y^2$ e $h_2(x, y, z) = x + z$ são de classe C^1 como soma e multiplicação de funções de classe C^1 .
- PASSO 2. Verificar a condição de regularidade.

A condição de regularidade exigida pelo teorema (12.4) diz que o posto de $Dh(\mathbf{p})$, com $\mathbf{h}(\mathbf{p}) = (h_1(\mathbf{p}), h_2(\mathbf{p}))$, deve ser igual a 2 (o número de restrições) em cada solução do problema de otimização. Mas como verificar a condição de regularidade se ainda não conhecemos a solução do problema de otimização? De fato, o que queremos é justamente encontrar a solução (ou soluções) do problema! A idéia é mostrar que *todos* os pontos do conjunto admissível (isto é, todos os pontos (x, y, z) em \mathbb{R}^3 tais que $h_1(x, y, z) = 1$ e $h_2(x, y, z) = 1$) satisfazem a condição de regularidade! Com isto, caso exista alguma solução do problema de otimização, certamente ela irá satisfazer a condição de regularidade exigida pelo teorema dos multiplicadores de Lagrange. A matriz jacobiana $Dh(x, y, z)$ é dada por

$$\begin{bmatrix} \frac{\partial h_1}{\partial x}(x, y, z) & \frac{\partial h_1}{\partial y}(x, y, z) & \frac{\partial h_1}{\partial z}(x, y, z) \\ \frac{\partial h_2}{\partial x}(x, y, z) & \frac{\partial h_2}{\partial y}(x, y, z) & \frac{\partial h_2}{\partial z}(x, y, z) \end{bmatrix} = \begin{bmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Se $x = 0$, então $y^2 = 1$, isto é, $y = \pm 1$. Neste caso a matriz jacobiana

$$\begin{bmatrix} 0 & \pm 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & \pm 2 & 0 \end{bmatrix}$$

tem posto 2. Se $x \neq 0$, subtraindo-se $2x$ vezes a linha 2 da linha 1 da matriz jacobiana e substituindo o resultado na linha 2, concluímos que

$$\begin{bmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 2x & 2y & 0 \\ 0 & 2y & -2x \end{bmatrix}$$

também tem posto 2. Desta maneira, todos os pontos do conjunto admissível satisfazem a condição de regularidade.

o PASSO 3. Escrever o lagrangeano.

$$\begin{aligned} L(x, y, z, \lambda_1, \lambda_2) &= f(x, y, z) - \lambda_1 [h_1(x, y, z) - c_1] - \lambda_2 [h_2(x, y, z) - c_2] \\ &= xyz - \lambda_1 (x^2 + y^2 - 1) - \lambda_2 (x + z - 1). \end{aligned}$$

o PASSO 4. Escrever as condições de primeira ordem.

$$\begin{cases} \frac{\partial L}{\partial x}(x, y, z, \lambda_1, \lambda_2) = 0, \\ \frac{\partial L}{\partial y}(x, y, z, \lambda_1, \lambda_2) = 0, \\ \frac{\partial L}{\partial z}(x, y, z, \lambda_1, \lambda_2) = 0, \\ \frac{\partial L}{\partial \lambda_1}(x, y, z, \lambda_1, \lambda_2) = 0, \\ \frac{\partial L}{\partial \lambda_2}(x, y, z, \lambda_1, \lambda_2) = 0, \end{cases} \Rightarrow \begin{cases} yz - 2\lambda_1 x - \lambda_2 = 0, & (1) \\ xz - 2\lambda_1 y = 0, & (2) \\ xy - \lambda_2 = 0, & (3) \\ x^2 + y^2 = 1, & (4) \\ x + z = 1. & (5) \end{cases}$$

o PASSO 5. Resolver o sistema correspondente às condições de primeira ordem.

Se $y = 0$, de (2) temos $xz = 0$ e de (3) temos $\lambda_2 = 0$. De (4) temos $x^2 = 1$, isto é, $x = -1$ ou $x = +1$. Se $x = -1$, de (5), temos $z = 1 - x = 2$, mas isto contradiz o fato de que $xz = 0$. Se $x = +1$, de (5), temos $z = 1 - x = 0$. Como $\lambda_2 = 0$ e $x = +1$, de (1) temos $\lambda_1 = 0$. Desta maneira, a única solução do sistema para o caso $y = 0$ é o ponto

$$(x, y, z, \lambda_1, \lambda_2) = (1, 0, 0, 0, 0).$$

Se $y \neq 0$, de (2) e (3) temos que

$$\lambda_1 = \frac{xz}{2y} \quad \text{e} \quad \lambda_2 = xy.$$

Substituindo estes valores na equação (1), segue-se que

$$yz - \frac{xz}{y}x - xy = 0 \Rightarrow y^2z - x^2z - xy^2 = 0.$$

De (4) temos $y^2 = 1 - x^2$ e, de (5), $z = 1 - x$. Substituindo estes valores na expressão acima, concluímos que

$$(1 - x)^2(1 - x) - x^2(1 - x) - x(1 - x^2) = 0.$$

Uma vez que $1 - x^2 = (1 - x)(1 + x)$, podemos colocar $(1 - x)$ em evidência na expressão acima, de forma que

$$(1 - x)(1 - x - 3x^2) = 0.$$

Sendo assim, $1 - x = 0$ ou $1 - x - 3x^2 = 0$, isto é,

$$x = 1 \quad \text{ou} \quad x = \frac{-1 - \sqrt{13}}{6} \quad \text{ou} \quad x = \frac{-1 + \sqrt{13}}{6}.$$

Com os valores possíveis de x , podemos facilmente calcular os valores de y , z , λ_1 e λ_2 com as fórmulas

$$y^2 = 1 - x^2, \quad z = 1 - x, \quad \lambda_1 = \frac{xz}{2y} \quad \text{e} \quad \lambda_2 = xy.$$

Portanto, as soluções do sistema associado às condições de primeira ordem para o caso $y \neq 0$ são

$$\begin{aligned} & \left(\frac{-1 - \sqrt{13}}{6}, -\frac{\sqrt{22 - 2\sqrt{13}}}{6}, \frac{7 + \sqrt{13}}{6}, +\frac{\sqrt{82 + 22\sqrt{13}}}{12}, +\frac{\sqrt{16 + \sqrt{13}}}{9} \right), \\ & \left(\frac{-1 - \sqrt{13}}{6}, +\frac{\sqrt{22 - 2\sqrt{13}}}{6}, \frac{7 + \sqrt{13}}{6}, -\frac{\sqrt{82 + 22\sqrt{13}}}{12}, -\frac{\sqrt{16 + \sqrt{13}}}{9} \right), \\ & \left(\frac{-1 + \sqrt{13}}{6}, -\frac{\sqrt{22 + 2\sqrt{13}}}{6}, \frac{7 - \sqrt{13}}{6}, -\frac{\sqrt{82 - 22\sqrt{13}}}{12}, -\frac{\sqrt{16 - \sqrt{13}}}{9} \right), \\ & \left(\frac{-1 + \sqrt{13}}{6}, +\frac{\sqrt{22 + 2\sqrt{13}}}{6}, \frac{7 - \sqrt{13}}{6}, +\frac{\sqrt{82 - 22\sqrt{13}}}{12}, +\frac{\sqrt{16 - \sqrt{13}}}{9} \right). \end{aligned}$$

• PASSO 6. Conclusão.

A função-objetivo f é contínua (pois é de classe C^1) e o conjunto ad-

missível é compacto (pois é uma elipse, resultante da interseção do cilindro $x^2 + y^2 = 1$ com o plano $x + z = 1$, veja a figura (12.6)). Pelo teorema de Weierstrass, f possui máximos e mínimos globais no conjunto admissível.

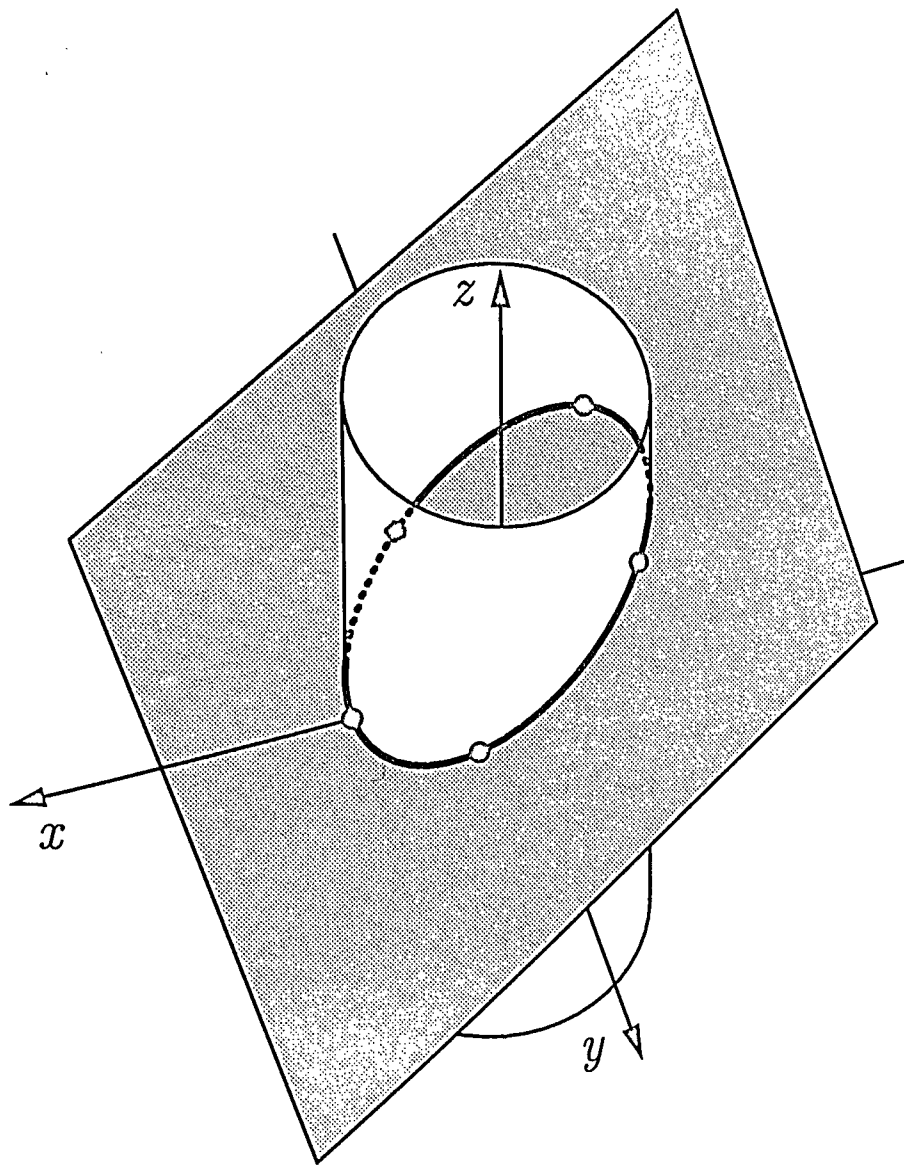


Figura 12.6: O conjunto admissível é a elipse resultante da interseção do cilindro $x^2 + y^2 = 1$ com o plano $x + z = 1$.

Agora, pelo teorema dos multiplicadores de Lagrange, qualquer solução do problema de otimização que estamos estudando deve ser solução do sistema associado às condições de primeira ordem (a condição de regularidade é fundamental aqui pois, com ela, temos a garantia de que qualquer solução do problema de otimização também será solução do sistema associado às condições de primeira ordem, para valores adequados dos multiplicadores de Lagrange).

le
10
la
10
25
le
15
15
15
e

3
13
13
13

ad-

Desta maneira, os pontos de máximo de $f(x, y, z) = xyz$ sujeitos às restrições $h_1(x, y, z) = x^2 + y^2 = 1$ e $h_2(x, y, z) = x + z = 1$ estão entre

$$\begin{aligned} & \left(\frac{-1 - \sqrt{13}}{6}, -\frac{\sqrt{22 - 2\sqrt{13}}}{6}, \frac{7 + \sqrt{13}}{6} \right), \\ & \left(\frac{-1 - \sqrt{13}}{6}, +\frac{\sqrt{22 - 2\sqrt{13}}}{6}, \frac{7 + \sqrt{13}}{6} \right), \\ & \left(\frac{-1 + \sqrt{13}}{6}, -\frac{\sqrt{22 + 2\sqrt{13}}}{6}, \frac{7 - \sqrt{13}}{6} \right), \\ & \left(\frac{-1 + \sqrt{13}}{6}, +\frac{\sqrt{22 + 2\sqrt{13}}}{6}, \frac{7 - \sqrt{13}}{6} \right), \\ & (1, 0, 0). \end{aligned}$$

Para encontrá-los, basta calcular a função-objetivo nestes pontos e selecionar os de maior valor. Temos

$$\begin{aligned} f\left(\frac{-1 - \sqrt{13}}{6}, -\frac{\sqrt{22 - 2\sqrt{13}}}{6}, \frac{7 + \sqrt{13}}{6}\right) &= +0.869\dots, \\ f\left(\frac{-1 - \sqrt{13}}{6}, +\frac{\sqrt{22 - 2\sqrt{13}}}{6}, \frac{7 + \sqrt{13}}{6}\right) &= -0.869\dots, \\ f\left(\frac{-1 + \sqrt{13}}{6}, -\frac{\sqrt{22 + 2\sqrt{13}}}{6}, \frac{7 - \sqrt{13}}{6}\right) &= -0.221\dots, \\ f\left(\frac{-1 + \sqrt{13}}{6}, +\frac{\sqrt{22 + 2\sqrt{13}}}{6}, \frac{7 - \sqrt{13}}{6}\right) &= +0.221\dots, \\ f(1, 0, 0) &= 0. \end{aligned}$$

Conseqüentemente, o ponto

$$\left(\frac{-1 - \sqrt{13}}{6}, -\frac{\sqrt{22 - 2\sqrt{13}}}{6}, \frac{7 + \sqrt{13}}{6} \right)$$

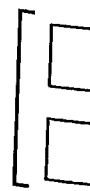
é a solução do problema de otimização, isto é, ele é ponto de *máximo global* de f no conjunto admissível

$$D = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = 1 \text{ e } x + z = 1\}.$$

No

é p

e

são,
Pod
no cVamo
ordem no
conjuntoSeja p
as restriçõ
dores λ_1 e
Se p satistem posto
 $h_1 = c_1$ e $h_2 = c_2$
geométrica

Note que, por outro lado,

$$\left(\frac{-1 - \sqrt{13}}{6}, + \frac{\sqrt{22 - 2\sqrt{13}}}{6}, \frac{7 + \sqrt{13}}{6} \right)$$

é ponto de mínimo global de f em D . Pode-se mostrar (exercício) que

$$\left(\frac{-1 + \sqrt{13}}{6}, + \frac{\sqrt{22 + 2\sqrt{13}}}{6}, \frac{7 - \sqrt{13}}{6} \right)$$

e

$$\left(\frac{-1 + \sqrt{13}}{6}, - \frac{\sqrt{22 + 2\sqrt{13}}}{6}, \frac{7 - \sqrt{13}}{6} \right)$$

são, respectivamente, pontos de mínimo *local* e máximo local de f em D . Pode-se mostrar também que o ponto $(1, 0, 0)$ não é extremo local de f no conjunto admissível D . □

Vamos agora tentar interpretar geometricamente as condições de primeira ordem no caso particular de uma função-objetivo de três variáveis e com um conjunto admissível construído através de duas funções-restrições h_1 e h_2 :

$$\begin{cases} \nabla f(x, y, z) = \lambda_1 \cdot \nabla h_1(x, y, z) + \lambda_2 \cdot \nabla h_2(x, y, z), \\ h_1(x, y, z) = c_1, \\ h_2(x, y, z) = c_2. \end{cases}$$

Seja $p = (p_1, p_2, p_3)$ um ponto do conjunto admissível (isto é, p satisfaz as restrições $h_1(p) = c_1$ e $h_2(p) = c_2$) e suponha que *não existam* multiplicadores λ_1 e λ_2 para os quais $(p_1, p_2, p_3, \lambda_1, \lambda_2)$ seja solução do sistema acima. Se p satisfaz a condição de regularidade, isto é, se a matriz jacobiana

$$\begin{bmatrix} \nabla h_1(p) \\ \nabla h_2(p) \end{bmatrix}_{2 \times 3} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1}(p) & \frac{\partial h_1}{\partial x_2}(p) & \frac{\partial h_1}{\partial x_3}(p) \\ \frac{\partial h_2}{\partial x_1}(p) & \frac{\partial h_2}{\partial x_2}(p) & \frac{\partial h_2}{\partial x_3}(p) \end{bmatrix}_{2 \times 3}$$

tem posto 2, então p não pode ser um extremo local de f sujeito às restrições $h_1 = c_1$ e $h_2 = c_2$. O fato de p ser um ponto regular tem várias implicações geométricas.

1. Se p satisfaz a condição de regularidade então, em particular, $\nabla h_1(p) \neq 0$. Assim, pelo teorema da função implícita, sabemos que o conjunto de nível $h_1 = c_1$ é gráfico de uma função de duas variáveis em uma vizinhança do ponto p . Geometricamente, isto significa que o conjunto de nível $h_1 = c_1$ é, em uma vizinhança do ponto p , uma superfície em \mathbb{R}^3 (como um "lençol" deformado) e não um objeto geométrico degenerado (como um ponto, uma curva ou todo o \mathbb{R}^3).
2. Analogamente, se p satisfaz a condição de regularidade então, em particular, $\nabla h_2(p) \neq 0$ e, portanto, geometricamente, o conjunto de nível $h_2 = c_2$, em uma vizinhança do ponto p , também é uma superfície em \mathbb{R}^3 "legítima" (outro "lençol" deformado).
3. O conjunto admissível é formado pelos pontos em \mathbb{R}^3 que satisfazem as restrições $h_1 = c_1$ e $h_2 = c_2$ *simultaneamente*. Geometricamente, isto significa que o conjunto admissível é a interseção das duas superfícies (a interseção dos dois "lençóis" deformados) $h_1 = c_1$ e $h_2 = c_2$. Observe que a interseção de duas superfícies pode ser uma curva, um ponto ou mesmo uma superfície ou o conjunto vazio (pense em alguns exemplos). Contudo, novamente, a condição de regularidade impõe características geométricas: *o conjunto admissível deve ser uma curva em uma vizinhança do ponto p* . Para ver isto, você pode aplicar a versão geral do teorema da função implícita (teorema (9.6)) ou seguir o seguinte argumento geométrico informal: como vimos, a condição de regularidade nos garante que $\nabla h_1(p) \neq 0$ e $\nabla h_2(p) \neq 0$. Mais ainda, ela garante que o posto da matriz jacobiana

$$\begin{bmatrix} \nabla h_1(p) \\ \nabla h_2(p) \end{bmatrix}_{2 \times 3}$$

é igual a 2. Isto significa que as duas linhas desta matriz não são uma múltipla da outra. Geometricamente, isto significa que os vetores gradientes $\nabla h_1(p)$ e $\nabla h_2(p)$ (as linhas desta matriz) *não são paralelos*. Assim, as duas superfícies de nível *não* podem se tocar (uma situação que seria favorável a uma interseção formada por um ponto ou uma superfície). Assim, as duas superfícies de nível devem se interceptar "transversalmente", gerando uma curva.

No exercício resolvido (12.3), as restrições são o cilindro $h_1(x, y, z) = x^2 + y^2 = 1$ e o plano $h_2(x, y, z) = x + z = 1$. Estas duas superfícies se

12.2 Ot
inter
4. Ago
isto s
veto
dient
risco
 $\nabla f(p)$
das f
conju
aume
Um p
gerac
então
à cur
deriv
tange
opost
direci
casos
Por e
(0, -1
 ∇f
segue-
 $\nabla h_2(p)$
ponto
octant
no qu
 $p =$
restric
o vetor
isto é,

interceptam “transversalmente”, gerando uma elipse.

4. Agora, se não existem multiplicadores λ_1 e λ_2 para os quais

$$\nabla f(\mathbf{p}) = \lambda_1 \cdot \nabla h_1(\mathbf{p}) + \lambda_2 \cdot \nabla h_2(\mathbf{p}),$$

isto significa que o vetor gradiente $\nabla f(\mathbf{p})$ não está no plano gerado pelos vetores $\nabla h_1(\mathbf{p})$ e $\nabla h_2(\mathbf{p})$. Se “caminharmos” na direção do vetor gradiente $\nabla f(\mathbf{p})$, certamente o valor da função aumentará, mas corremos o risco de sairmos do conjunto admissível. Contudo, como o vetor gradiente $\nabla f(\mathbf{p})$ não está no plano gerado pelos vetores gradientes $\nabla h_1(\mathbf{p})$ e $\nabla h_2(\mathbf{p})$ das funções restrições, se “caminharmos” ao longo da curva formada pelo conjunto admissível, a partir do ponto \mathbf{p} , ainda assim o valor da função aumentará ou diminuirá conforme seguimos numa ou noutra direção.

Um pouco mais formalmente, se o vetor gradiente $\nabla f(\mathbf{p})$ não está no plano gerado pelos vetores gradientes $\nabla h_1(\mathbf{p})$ e $\nabla h_2(\mathbf{p})$ das funções restrições, então a derivada direcional de f no ponto \mathbf{p} na direção do vetor tangente à curva formada pelo conjunto admissível é diferente de zero. Se esta derivada direcional for positiva, se “caminharmos” na direção do vetor tangente, o valor da função aumentará e, se “caminharmos” na direção oposta do vetor tangente, o valor da função diminuirá. Caso a derivada direcional seja negativa, ocorre o contrário. Em qualquer um dos dois casos, \mathbf{p} não é um extremo local de f com as restrições $h_1 = c_1$ e $h_2 = c_2$.

Por exemplo, no exercício resolvido (12.3), é fácil ver que o ponto $\mathbf{p} = (0, -1, 1)$ é admissível. Contudo, como

$$\nabla f(\mathbf{p}) = (-1, 0, 0), \quad \nabla h_1(\mathbf{p}) = (0, -2, 0) \quad \text{e} \quad \nabla h_2(\mathbf{p}) = (1, 0, 1),$$

segue-se que $\nabla f(\mathbf{p})$ não está no plano gerado pelos vetores $\nabla h_1(\mathbf{p})$ e $\nabla h_2(\mathbf{p})$. Com isto, se “caminharmos” ao longo da elipse, a partir do ponto $\mathbf{p} = (0, -1, 1)$, o valor da função aumentará se entrarmos no terceiro octante (região do \mathbb{R}^3 onde $x \leq 0$, $y \leq 0$ e $z \geq 0$) e diminuirá se entrarmos no quarto octante (região do \mathbb{R}^3 onde $x \geq 0$, $y \leq 0$ e $z \geq 0$). Assim, $\mathbf{p} = (0, -1, 1)$ não é um extremo local de $f(x, y, z) = xyz$ sujeito às restrições $h_1(x, y, z) = x^2 + y^2 = 1$ e $h_2(x, y, z) = x + z = 1$. Observe que o vetor $\nabla f(\mathbf{p})$ não está no plano gerado pelos vetores $\nabla h_1(\mathbf{p})$ e $\nabla h_2(\mathbf{p})$, isto é, não existem multiplicadores λ_1 e λ_2 para os quais

$$\nabla f(\mathbf{p}) = \lambda_1 \cdot \nabla h_1(\mathbf{p}) + \lambda_2 \cdot \nabla h_2(\mathbf{p}).$$

Desta maneira, $\mathbf{p} = (0, -1, 1)$ não satisfaz as condições de primeira ordem. Observe também que \mathbf{p} satisfaz a condição de regularidade pois a matriz jacobiana

$$\begin{bmatrix} \nabla h_1(0, -1, 1) \\ \nabla h_2(0, -1, 1) \end{bmatrix}_{2 \times 3} = \begin{bmatrix} 0 & -2 & 0 \\ 1 & 0 & 1 \end{bmatrix}_{2 \times 3}$$

tem posto 2.

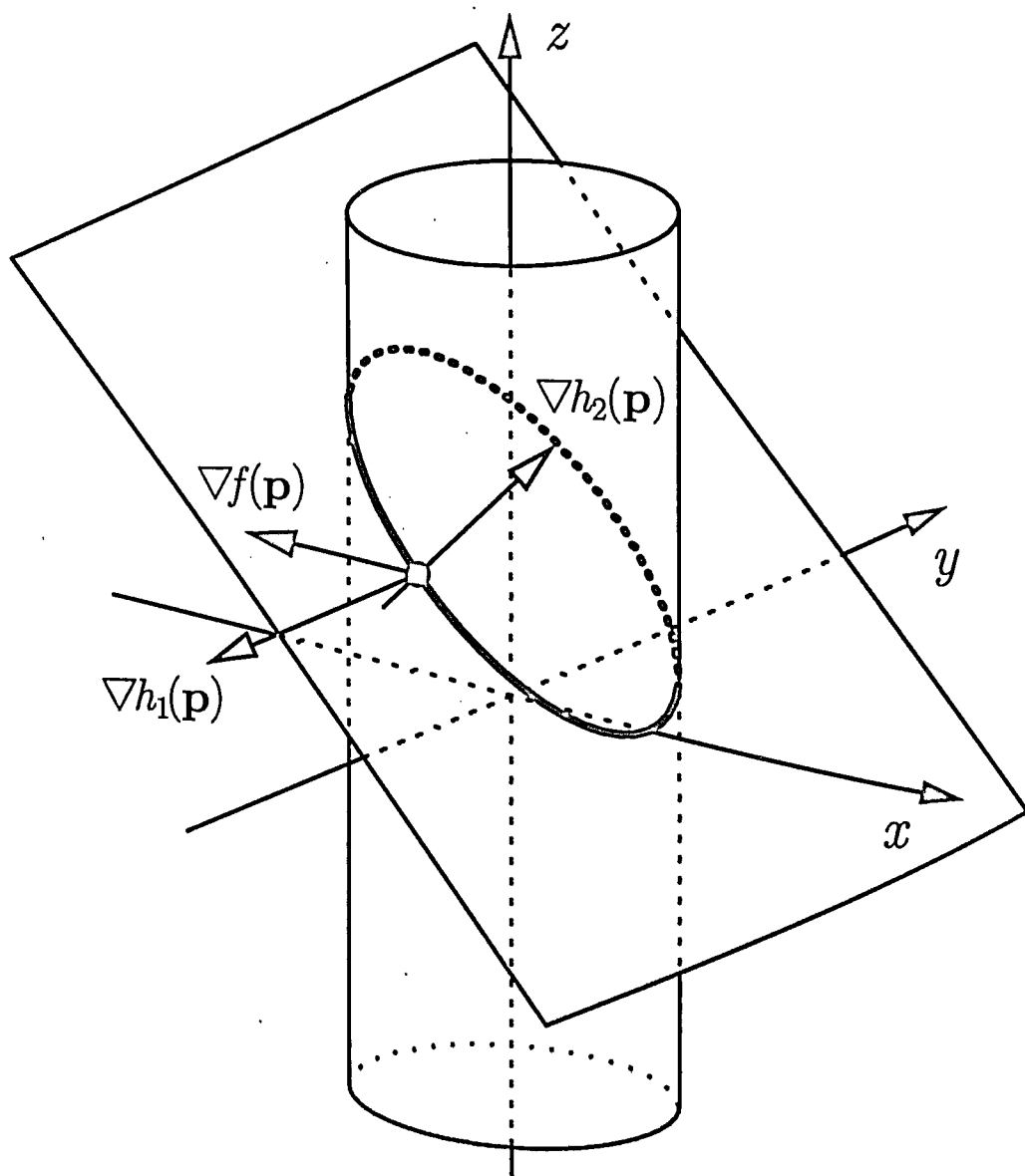


Figura 12.7: O ponto $\mathbf{p} = (0, -1, 1)$ não é um extremo local da função $f(x, y, z) = xyz$ sujeito às restrições $h_1(x, y, z) = x^2 + y^2 = 1$ e $h_2(x, y, z) = x + z = 1$. Observe que o vetor $\nabla f(\mathbf{p})$ não está no plano gerado pelos vetores $\nabla h_1(\mathbf{p})$ e $\nabla h_2(\mathbf{p})$.

Observ
no teore
necessár

Teor
class
local

Entã

A con
 λ_0^*, λ_1^*

Evidente
função-ob
dem. A h
no teorem

12.3 C

Nas du
o conjunt
de nível, is

Para encon
uma dada

Observação. Novamente, a condição de regularidade *não* pode ser omitida no teorema dos multiplicadores de Lagrange. Como antes, para omití-la, é necessário acrescentar um multiplicador para a função-objetivo.

Teorema 12.5 (FRITZ JOHN) Sejam f, h_1, \dots, h_m funções de classe C^1 de n variáveis e seja \mathbf{p} um extremo (máximo ou mínimo) local de f no conjunto admissível

$$D = \{x \in \mathbb{R}^n \mid h_1(x) = c_1, \dots, h_m(x) = c_m\}.$$

Então existem multiplicadores $\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*$ tais que

$$\begin{cases} \lambda_0^* \cdot \nabla f(\mathbf{p}) = \lambda_1^* \cdot \nabla h_1(\mathbf{p}) + \dots + \lambda_m^* \cdot \nabla h_m(\mathbf{p}), \\ h_1(\mathbf{p}) = c_1, \\ \vdots \\ h_m(\mathbf{p}) = c_m, \\ (\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*) \neq (0, 0, \dots, 0), \\ \lambda_0^* \in \{0, 1\}. \end{cases}$$

A condição $(\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*) \neq (0, 0, \dots, 0)$ diz que os multiplicadores $\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*$ não são simultaneamente nulos.

Evidentemente, a situação desejável é que $\lambda_0^* = 1$, pois caso contrário, a função-objetivo desapareceria completamente das condições de primeira ordem. A hipótese de regularidade garante que podemos tomar λ_0^* igual a 1 no teorema de Fritz John.

12.3 Otimização com uma restrição em desigualdade

Nas duas seções anteriores estudamos problemas de otimização nos quais o conjunto admissível é construído como a interseção de várias "hiperfícies" de nível, isto é, com o uso de *igualdades*:

$$D = \{x \in \mathbb{R}^n \mid h_1(x) = c_1, \dots, h_m(x) = c_m\}.$$

Para encontrar os candidatos a pontos de máximo ou pontos de mínimo de uma dada função-objetivo, em um tal conjunto admissível, vimos que basta

calcular os pontos críticos do lagrangeano correspondente, não esquecendo de verificar a condição de regularidade. Contudo, a grande maioria dos problemas de otimização em engenharia e economia são modelados de tal forma que o conjunto admissível é construído com o uso de *desigualdades*:

$$D = \{x \in \mathbb{R}^n \mid g_1(x) \leq b_1, \dots, g_k(x) \leq b_k\}.$$

A partir de agora, construiremos a teoria necessária para caracterizar algebricamente os candidatos a extremo de uma função em um conjunto admissível deste tipo. Vamos começar com um caso bem simples:

$$\begin{array}{l} \text{maximizar} \quad f(x_1, x_2) \\ \text{sujeito a} \quad (x_1, x_2) \in D = \{(x_1, x_2) \in \mathbb{R}^2 \mid g(x_1, x_2) \leq b\}. \end{array}$$

Suponha que p seja solução do problema de otimização, isto é, suponha que $p \in D$ seja um ponto de máximo de f no conjunto admissível D . Temos duas possibilidades: $g(p) = b$ ou $g(p) < b$.

1. Caso $g(p) = b$.

Nesta situação, dizemos que a restrição g está *ativa* no ponto p . Geometricamente, isto significa dizer que o ponto p está na *fronteira* do conjunto admissível (veja a figura (12.8)). Observe que, aqui, estamos em uma situação semelhante ao caso de uma única restrição em igualdade. Desta maneira, se p é ponto de máximo local de f em D , então vale a mesma conclusão do teorema dos multiplicadores de Lagrange, isto é, os gradientes de f e g no ponto p são paralelos:

$$\nabla f(p) = \mu^* \cdot \nabla g(p).$$

Mas, agora, vale uma condição extra:

$$\mu^* \geq 0,$$

isto é, o multiplicador μ^* deve ser maior ou igual a zero. Vamos justificar este fato com argumentos geométricos. Considere a figura (12.8). Se p é máximo local de f em D , então as curvas de nível de f em uma vizinhança de p devem ser tais que

$$c_1 < c_2 < c_3 < c_4.$$

Fi

Dest
do cor
não-n
outro
da reg
maior
a mes
com μ
no cas
condiç

2. Caso g

Nesta
Geome
conjun
em um
curáva
conjun
de f en

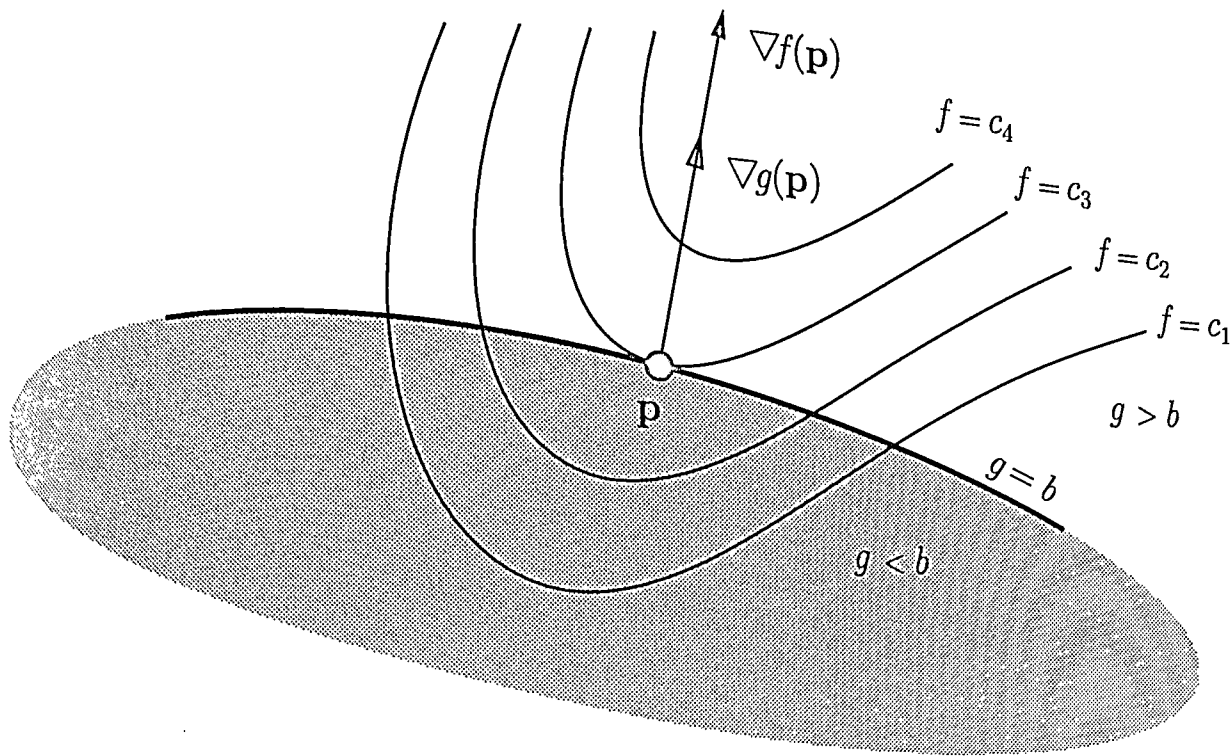


Figura 12.8: Caso em que a restrição g está ativa na solução p do problema de otimização.

Desta maneira, o vetor gradiente de f em p deve apontar para “fora” do conjunto admissível (lembre-se que o gradiente de uma função, quando não-nulo, fornece a direção de maior crescimento da função no ponto). Por outro lado, o vetor gradiente de g em p também deve apontar para “fora” da região, uma vez que $g(p) = b$ e, “fora” da região, o valor da função é maior do que b . Moral da história: os vetores $\nabla f(p)$ e $\nabla g(p)$ possuem a mesma direção e o mesmo sentido, sendo assim, $\nabla f(p) = \mu^* \cdot \nabla g(p)$, com $\mu^* \geq 0$. Evidentemente, devemos ter o mesmo cuidado que tivemos no caso de uma única restrição em igualdade: o ponto p deve satisfazer a condição de regularidade, isto é, o ponto p deve ser tal que $\nabla g(p) \neq 0$.

2. Caso $g(p) < b$.

Nesta situação, dizemos que a restrição g não está ativa no ponto p . Geometricamente, isto significa dizer que o ponto p está no interior do conjunto admissível (veja a figura (12.9)). Observe que, aqui, estamos em uma situação semelhante à que estudamos no capítulo 11, onde procurávamos por extremos sem restrições, isto é, extremos no interior do conjunto admissível. Desta maneira, se p é um ponto de máximo local de f em D , então p deve satisfazer a regra de Fermat:

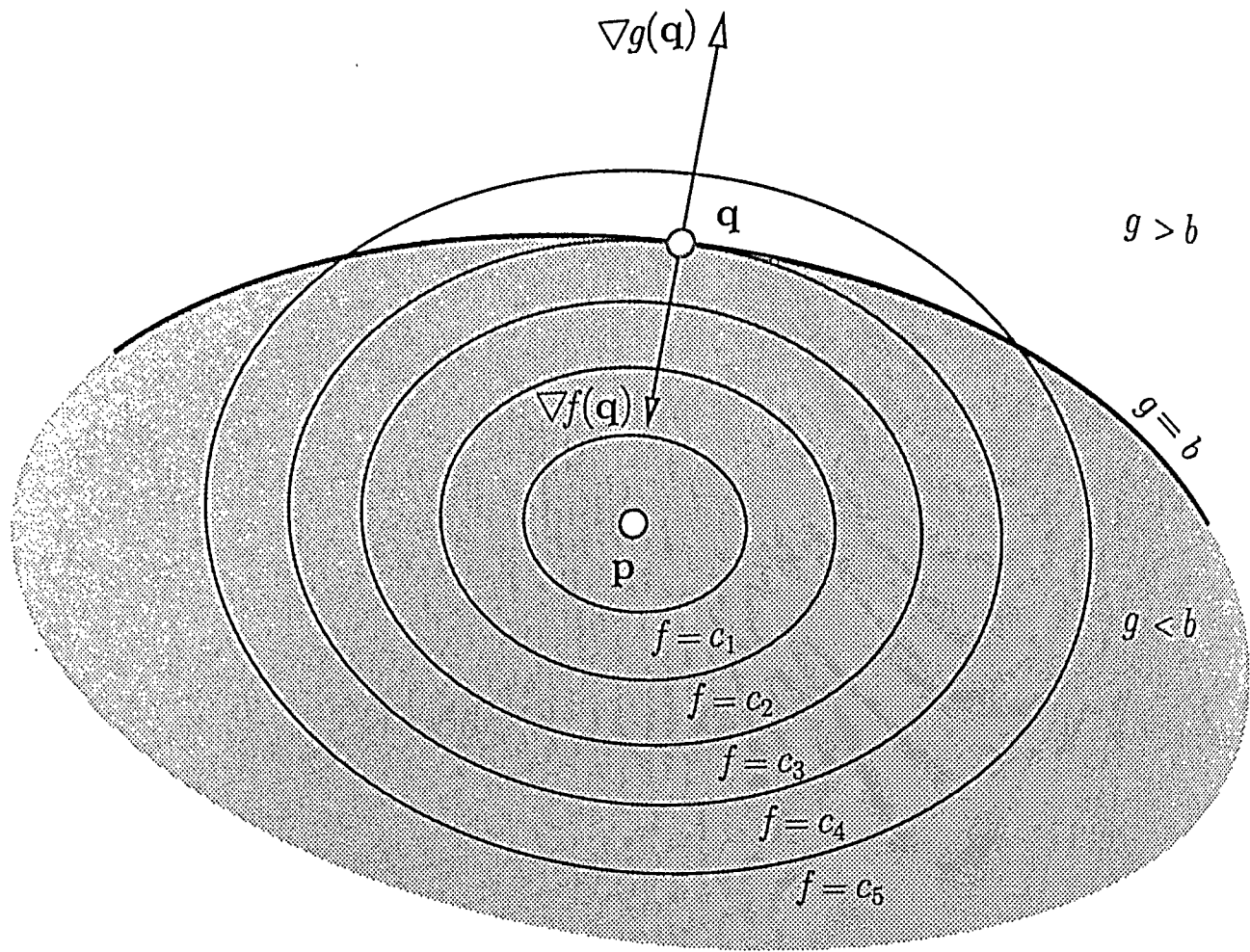


Figura 12.9: Caso em que a restrição g não está ativa na solução p do problema de otimização.

$$\nabla f(p) = 0,$$

isto é, p deve ser um ponto crítico de f . Observe que, na figura (12.9), os vetores gradientes $\nabla f(q)$ e $\nabla g(q)$ são paralelos mas possuem *sentidos opostos*, isto é, $\nabla f(q) = \mu^* \cdot \nabla g(q)$ mas com $\mu^* \leq 0$. O ponto q não é máximo local de f no conjunto admissível D pois se, a partir do ponto q , “caminharmos” na direção do vetor gradiente $\nabla f(q)$, continuaremos dentro do conjunto admissível e o valor da função f aumentará.

Resumindo: se p é um extremo local de f em D , então ou $g(p) = b$ ou $g(p) < b$, isto é, a restrição g ou está ou não está ativa no ponto p . Se g está ativa no ponto p , então vale que

$$\begin{cases} \nabla f(p) = \mu^* \cdot \nabla g(p), \\ \mu^* \geq 0, \\ g(p) = b, \end{cases}$$

lembr
dade,
está a

Ex
bem s

A “es

De fa
 $\nabla f(p)$
(segur
ativa e
caso).
caso p
admiss
regula
teorema

Te
seja

Sup
se a

lembrando que, neste caso, supõe-se que \mathbf{p} satisfaz a condição de regularidade, isto é, supõe-se que $\nabla g(\mathbf{p}) \neq \mathbf{0}$. Por outro lado, se a restrição g não está ativa no ponto \mathbf{p} , então vale que

$$\begin{cases} \nabla f(\mathbf{p}) = \mathbf{0}, \\ g(\mathbf{p}) < b. \end{cases}$$

Existe uma maneira de se unificar estes dois casos, seguindo uma idéia bem simples, devida a Karush-Kuhn-Tucker: basta considerar o sistema

$$\begin{cases} \nabla f(\mathbf{p}) = \mu^* \cdot \nabla g(\mathbf{p}), \\ \mu^* \cdot [g(\mathbf{p}) - b] = 0, \\ \mu^* \geq 0, \\ g(\mathbf{p}) \leq b. \end{cases}$$

A "esperteza" aqui está na segunda linha do sistema:

$$\mu^* \cdot [g(\mathbf{p}) - b] = 0.$$

De fato, se $\mu^* \cdot [g(\mathbf{p}) - b] = 0$, então $\mu^* = 0$ ou $g(\mathbf{p}) = b$. Se $\mu^* = 0$, então $\nabla f(\mathbf{p}) = \mu^* \cdot \nabla g(\mathbf{p}) = 0 \cdot \nabla g(\mathbf{p}) = \mathbf{0}$, isto é, \mathbf{p} é um ponto crítico de f (segundo caso). Por outro lado, se $\mu^* > 0$, então $g(\mathbf{p}) = b$ (a restrição g está ativa em \mathbf{p}) e, portanto, vale que $\nabla f(\mathbf{p}) = \mu^* \cdot \nabla g(\mathbf{p})$, com $\mu^* > 0$ (primeiro caso). Observe que se $\mu^* = 0$ e $g(\mathbf{p}) = b$, então estamos considerando o caso particular em que f possui um ponto crítico na fronteira do conjunto admissível. Novamente, no caso $g(\mathbf{p}) = b$, devemos impor a condição de regularidade: $\nabla g(\mathbf{p}) \neq \mathbf{0}$. Vamos resumir todas estas idéias no próximo teorema.

Teorema 12.6 Sejam f e g funções de classe C^1 de duas variáveis e seja $\mathbf{p} = (p_1, p_2)$ uma solução (local) do problema de otimização

$$\begin{array}{l} \text{maximizar} \quad f(x_1, x_2) \\ \text{sujeito a} \quad (x_1, x_2) \in D = \{(x_1, x_2) \in \mathbb{R}^2 \mid g(x_1, x_2) \leq b\}. \end{array}$$

Suponha que $\mathbf{p} = (p_1, p_2)$ satisfaça a seguinte condição de regularidade: se a restrição g está ativa em \mathbf{p} , isto é, se $g(\mathbf{p}) = b$, então

$$\nabla g(\mathbf{p}) = \left(\frac{\partial g}{\partial x_1}(p_1, p_2), \frac{\partial g}{\partial x_2}(p_1, p_2) \right) \neq (0, 0).$$

Então existe um número real μ^* tal que $(\mathbf{p}, \mu^*) = (p_1, p_2, \mu^*)$ satisfaz as condições de primeira ordem

$$\begin{cases} \nabla f(\mathbf{p}) = \mu^* \cdot \nabla g(\mathbf{p}), \\ \mu^* \cdot [g(\mathbf{p}) - b] = 0, \\ \mu^* \geq 0, \\ g(\mathbf{p}) \leq b. \end{cases}$$

Este sistema também pode ser escrito, de maneira equivalente, em termos do *lagrangeano*

$$L(x_1, x_2, \mu) = f(x_1, x_2) - \mu \cdot [g(x_1, x_2) - b],$$

da seguinte maneira:

$$\begin{cases} \frac{\partial L}{\partial x_1}(p_1, p_2, \mu^*) = 0, \\ \frac{\partial L}{\partial x_2}(p_1, p_2, \mu^*) = 0, \\ \mu^* \cdot [g(p_1, p_2) - b] = 0, \\ \mu^* \geq 0, \\ g(p_1, p_2) \leq b. \end{cases}$$

Observe que a condição $g(p_1, p_2) \leq b$ é equivalente à condição $(\partial L / \partial \mu)(p_1, p_2, \mu^*) \geq 0$.

Observe as semelhanças e diferenças entre o teorema (12.2), que trata do problema de otimização com uma restrição em *igualdade*, e o teorema (12.6), que trata do problema de otimização com uma restrição em *desigualdade*:

1. Os dois teoremas usam o mesmo lagrangeano L e os dois teoremas pedem que as derivadas parciais de L com relação a x_1 e x_2 sejam iguais a zero.
2. A condição

$$\frac{\partial L}{\partial \lambda}(x_1, x_2, \lambda) = -[h(x_1, x_2) - c] = 0,$$

para a caso com uma restrição em igualdade, não vale mais quando consideramos a situação com uma restrição em desigualdade pois, neste caso, a restrição não precisa estar ativa no ponto de máximo. No lugar, usamos as duas condições

se
e

3. O
de
or

4. N
tr
de
o

5. Pa

bl

fu

de

di

fig

em

fun

a r

é r

na

ma

Exer

SOLU

• P

de

• P

A

di

$$\mu \cdot [g(x_1, x_2) - b] = 0 \quad \text{e} \quad \frac{\partial L}{\partial \mu}(x_1, x_2, \mu) = -[g(x_1, x_2) - b] \geq 0,$$

sendo que esta segunda condição nada mais é do que a própria restrição em desigualdade.

3. Os dois teoremas exigem uma condição de regularidade. Contudo, no caso de uma restrição em desigualdade, só precisamos verificá-la para pontos onde a restrição está ativa.
4. Não existem restrições para o sinal do multiplicador no caso de uma restrição em igualdade. Por outro lado, para o problema de *maximização* de uma função f sujeito a uma restrição em desigualdade do tipo $g \leq b$, o multiplicador deve ser não-negativo.
5. Para problemas de otimização com restrições em igualdade (e mesmo problemas de otimização sem restrições), as condições de primeira ordem funcionam tanto para problemas de maximização quanto para problemas de minimização. Por outro lado, o argumento de que os vetores gradientes $\nabla f(\mathbf{p})$ e $\nabla g(\mathbf{p})$ possuem a mesma direção e sentido, dado na figura (12.8), só é válido para problemas de *maximização* com restrições em desigualdade do tipo $g \leq b$. Se, por exemplo, queremos minimizar uma função f com uma restrição do tipo $g \leq b$, então $\nabla f(\mathbf{p})$ e $\nabla g(\mathbf{p})$ possuem a mesma direção mas com *sentidos opostos* em um ponto de mínimo \mathbf{p} que é regular e está na fronteira do conjunto admissível (o caso do ponto \mathbf{q} na figura (12.9)). Falaremos mais sobre as diferenças entre problemas de maximização e de minimização em uma seção subsequente.

Exercício resolvido 12.4 Resolva o seguinte problema de otimização

$$\begin{array}{l} \text{maximizar} \quad f(x_1, x_2) = x_1 x_2 \\ \text{sujeito a} \quad g(x_1, x_2) = x_1^2 + x_2^2 \leq 1. \end{array}$$

SOLUÇÃO: vamos subdividir a resolução deste problema em vários passos.

- PASSO 1. As funções f e g são de classe C^1 como soma e multiplicação de funções de classe C^1 .
- PASSO 2. Verificar a condição de regularidade.
A condição de regularidade exigida pelo teorema (12.6) diz que o gradiente de g deve ser diferente de zero em cada solução do problema de

otimização para a qual a restrição g está ativa. Mas como verificar a condição de regularidade se ainda não conhecemos a solução do problema de otimização? De fato, o que queremos é justamente encontrar a solução (ou soluções) do problema! A idéia é mostrar que *todos* os pontos do conjunto admissível para os quais a restrição g está ativa (isto é, todos os pontos (x_1, x_2) em \mathbb{R}^2 tais que $g(x_1, x_2) = x_1^2 + x_2^2 = 1$) satisfazem a condição de regularidade! Com isto, caso exista alguma solução do problema de otimização, certamente ela irá satisfazer a condição de regularidade exigida pelo teorema (12.6). Se $g(x_1, x_2) = x_1^2 + x_2^2 = 1$, então x_1 e x_2 não podem se anular *simultaneamente*. Portanto, como

$$\nabla g(x_1, x_2) = \left(\frac{\partial g}{\partial x_1}(x_1, x_2), \frac{\partial g}{\partial x_2}(x_1, x_2) \right) = (2x_1, 2x_2)$$

segue-se que $\nabla g(x_1, x_2) \neq (0, 0)$ em todo (x_1, x_2) para o qual a restrição g está ativa.

- PASSO 3. Escrever o lagrangeano.

$$L(x_1, x_2, \mu) = f(x_1, x_2) - \mu \cdot [g(x_1, x_2) - b] = x_1 x_2 - \mu \cdot (x_1^2 + x_2^2 - 1).$$

- PASSO 4. Escrever as condições de primeira ordem.

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial x_1}(x_1, x_2, \mu) = 0, \\ \frac{\partial L}{\partial x_2}(x_1, x_2, \mu) = 0, \\ \mu \cdot [g(x_1, x_2) - b] = 0, \\ \mu \geq 0, \\ g(x_1, x_2) \leq b, \end{array} \right. \Rightarrow \left\{ \begin{array}{l} x_2 - 2\mu x_1 = 0, \quad (1) \\ x_1 - 2\mu x_2 = 0, \quad (2) \\ \mu \cdot [x_1^2 + x_2^2 - 1] = 0, \quad (3) \\ \mu \geq 0, \quad (4) \\ x_1^2 + x_2^2 \leq 1. \quad (5) \end{array} \right.$$

- PASSO 5. Resolver o sistema correspondente às condições de primeira ordem.

De (4), temos dois casos: $\mu = 0$ ou $\mu > 0$. Se $\mu = 0$, então de (1) e (2), concluímos que $x = 0$ e $y = 0$. Logo, o ponto

$$(x_1, x_2, \mu) = (0, 0, 0)$$

satisfaz as condições de primeira ordem. Por outro lado, se $\mu > 0$ então, de (3), segue-se que $x_1^2 + x_2^2 = 1$. Conseqüentemente, devemos ter $x_1 \neq 0$.

De fato: Se $x_1 = 0$ então, de (1), teríamos $x_2 = 0$ e, portanto, $x_1^2 + x_2^2$ seria igual a 0 e não igual a 1. Analogamente, podemos mostrar que $x_2 \neq 0$. Agora, uma vez que $x_1 \neq 0$ e $x_2 \neq 0$, então de (1) e (2), temos

$$\mu = \frac{x_2}{2x_1} = \frac{x_1}{2x_2},$$

de modo que $x_1^2 = x_2^2$. Logo, $x_1^2 + x_2^2 = 2x_2^2 = 1$, isto é,

$$x_2 = -\frac{\sqrt{2}}{2} \quad \text{ou} \quad x_2 = +\frac{\sqrt{2}}{2}.$$

Dado que $x_1^2 = x_2^2$, segue-se que

$$x_1 = -\frac{\sqrt{2}}{2} \quad \text{ou} \quad x_1 = +\frac{\sqrt{2}}{2}.$$

Finalmente, como $\mu = x_2/x_1 = x_1/x_2$ e o multiplicador μ não pode ser negativo (pela condição (4)), os pontos que satisfazem as condições de primeira ordem, no caso $\mu > 0$, são

$$\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, +\frac{1}{2}\right) \quad \text{e} \quad \left(+\frac{\sqrt{2}}{2}, +\frac{\sqrt{2}}{2}, +\frac{1}{2}\right).$$

Sendo assim, apenas três pontos satisfazem as condições de primeira ordem: $(0, 0, 0)$, $(-\sqrt{2}/2, -\sqrt{2}/2, +1/2)$ e $(+\sqrt{2}/2, +\sqrt{2}/2, +1/2)$.

• PASSO 6. Conclusão.

A função-objetivo f é contínua (pois é de classe C^1) e o conjunto admissível é compacto (pois é um disco de centro na origem e raio 1, veja a figura (12.10)). Pelo teorema de Weierstrass, f possui máximos e mínimos globais no conjunto admissível. Agora, pelo teorema (12.6), qualquer solução do problema de otimização que estamos estudando deve ser solução do sistema associado às condições de primeira ordem (a condição de regularidade é fundamental aqui pois, com ela, temos a garantia que qualquer solução do problema de otimização será também solução do sistema associado às condições de primeira ordem para valores adequados dos multiplicadores). Desta maneira, os pontos de máximo de $f(x_1, x_2) = x_1x_2$ sujeito à restrição $g(x_1, x_2) = x_1^2 + x_2^2 \leq 1$ estão entre $(0, 0)$, $(-\sqrt{2}/2, -\sqrt{2}/2)$ e $(+\sqrt{2}/2, +\sqrt{2}/2)$. Para encontrá-los, basta calcular a função-objetivo nestes pontos e selecionar os de maior valor. Temos

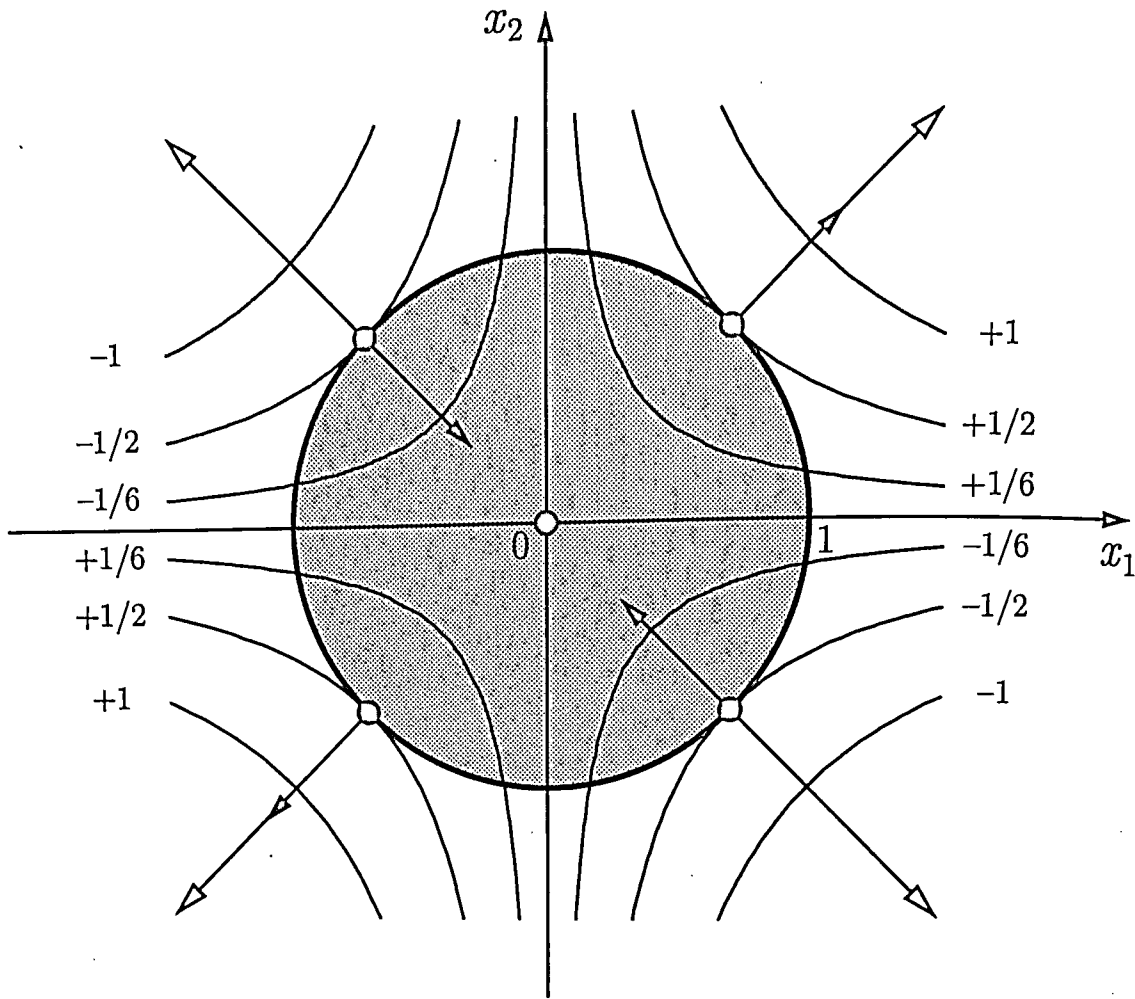


Figura 12.10: Desenho das curvas de nível de f (os números indicam o valor do nível) e do conjunto admissível (o disco de centro na origem e raio 1). Com uma escala menor, os vetores com seta branca indicam o gradiente de g enquanto que os vetores com seta preta indicam o gradiente de f .

$$f(0, 0) = 0, \quad f\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = \frac{1}{2} \quad \text{e} \quad f\left(+\frac{\sqrt{2}}{2}, +\frac{\sqrt{2}}{2}\right) = \frac{1}{2}$$

Conseqüentemente, os pontos

$$\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) \quad \text{e} \quad \left(+\frac{\sqrt{2}}{2}, +\frac{\sqrt{2}}{2}\right)$$

são soluções do problema de otimização, isto é, eles são pontos de *máximo global* de f no conjunto admissível

$$D = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}.$$

Note que, na figura (12.10), os pontos $(-\sqrt{2}/2, -\sqrt{2}/2)$ e $(+\sqrt{2}/2, +\sqrt{2}/2)$ estão na fronteira do conjunto admissível, de modo que a restrição g está

ativa n
mizaçã
com o
sentido
de mod
é máxi
 $(-\sqrt{2}/2)$
as cond
pontos.
em D .

Observ
pelo acr
condiçõe
em uma

12.4

O teo
e várias
ponto p
ativa em

Teore
defini
conjun

onde
mos re
Supon
da ma

ativa nestes pontos. Uma vez que eles são soluções do problema de otimização e satisfazem a condição de regularidade, segue-se que, de acordo com o teorema (12.6), os gradientes de f e g possuem a mesma direção e sentido nestes pontos. O ponto $(0, 0)$ está no interior do conjunto admissível, de modo que a restrição g não está ativa neste ponto. Note que $(0, 0)$ não é máximo e nem mínimo local de f em D (por que?). Agora, os pontos $(-\sqrt{2}/2, +\sqrt{2}/2)$ e $(+\sqrt{2}/2, -\sqrt{2}/2)$, apesar de admissíveis, não satisfazem as condições de primeira ordem, pois o multiplicador é negativo para estes pontos. De fato, pode-se mostrar que eles são pontos de *mínimo* global de f em D .

Observação. Como antes, a condição de regularidade pode ser substituída pelo acréscimo de um multiplicador para a função-objetivo, resultando nas condições de Fritz John. Trataremos mais sobre condições de regularidade em uma seção subsequente.

12.4 Otimização com várias restrições em desigualdade

O teorema (12.6) generaliza-se facilmente para o caso de várias variáveis e várias restrições. Lembre-se que uma restrição $g(x) \leq b$ está *ativa* em um ponto p se $g(p) = b$. Se $g(p) < b$, então dizemos que a restrição *não está ativa* em p .

Teorema 12.7 Sejam f, g_1, \dots, g_k funções de classe C^1 de n variáveis definidas em um aberto de \mathbb{R}^n e seja $p \in \mathbb{R}^n$ um máximo local de f no conjunto admissível

$$D = \{x \mid g_1(x) \leq b_1, \dots, g_k(x) \leq b_k\},$$

onde $x = (x_1, \dots, x_n)$. Caso alguma restrição esteja ativa em p , vamos renomeá-las de forma que elas sejam as l primeiras: g_1, \dots, g_l . Suponha que p satisfaça a seguinte condição de regularidade: o posto da matriz jacobiana

$$\begin{bmatrix} \nabla g_1(\mathbf{p}) \\ \vdots \\ \nabla g_l(\mathbf{p}) \end{bmatrix}_{l \times n} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{p}) & \cdots & \frac{\partial g_1}{\partial x_n}(\mathbf{p}) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_l}{\partial x_1}(\mathbf{p}) & \cdots & \frac{\partial g_l}{\partial x_n}(\mathbf{p}) \end{bmatrix}_{l \times n}$$

é igual a l (o número de restrições ativas em \mathbf{p}). Então existem multiplicadores μ_1^*, \dots, μ_k^* tais que

$$\left\{ \begin{array}{l} \nabla f(\mathbf{p}) = \mu_1^* \cdot \nabla g_1(\mathbf{p}) + \cdots + \mu_k^* \cdot \nabla g_k(\mathbf{p}), \\ \mu_1^* \cdot [g_1(\mathbf{p}) - b_1] = 0, \\ \vdots \\ \mu_k^* \cdot [g_k(\mathbf{p}) - b_k] = 0, \\ \mu_1^* \geq 0, \\ \vdots \\ \mu_k^* \geq 0, \\ g_1(\mathbf{p}) \leq b_1, \\ \vdots \\ g_k(\mathbf{p}) \leq b_k, \end{array} \right.$$

ou, equivalentemente,

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial x_1}(\mathbf{p}, \boldsymbol{\mu}^*) = 0, \quad \dots, \quad \frac{\partial L}{\partial x_n}(\mathbf{p}, \boldsymbol{\mu}^*) = 0, \\ \mu_1^* \cdot [g_1(\mathbf{p}) - b_1] = 0, \\ \vdots \\ \mu_k^* \cdot [g_k(\mathbf{p}) - b_k] = 0, \\ \mu_1^* \geq 0, \\ \vdots \\ \mu_k^* \geq 0, \\ g_1(\mathbf{p}) \leq b_1, \\ \vdots \\ g_k(\mathbf{p}) \leq b_k, \end{array} \right.$$

onde $\boldsymbol{\mu}^* = (\mu_1^*, \dots, \mu_k^*)$ e L é o lagrangeano:

Estes s
o pont

A condiçã
condições
apenas as
ativas não
Sendo ass
ativas do
da matriz
máximo.

Exercício
em

SOLUÇÃO

- PASS
Para
mizaç
guma
basta
missi
Assim

$$L(\mathbf{x}, \boldsymbol{\mu}) = f(\mathbf{x}) - \mu_1 \cdot [g_1(\mathbf{x}) - b_1] - \dots - \mu_k \cdot [g_k(\mathbf{x}) - b_k].$$

Estes sistemas são denominados de condições de primeira ordem para o ponto de máximo (local) p .

A condição de regularidade do teorema (12.7) é a generalização natural das condições de regularidade dos teoremas (12.4) e (12.6). Esta condição envolve apenas as restrições que estão ativas, uma vez que restrições que não estão ativas não desempenham nenhum papel nas condições de primeira ordem. Sendo assim, com relação à condição de regularidade, tratamos as restrições ativas do mesmo modo que tratamos as restrições em igualdade: o posto da matriz jacobiana formada pelos gradientes das restrições ativas deve ser máximo.

Exercício resolvido 12.5 Resolva o problema de otimização que consiste em

maximizar	xyz
sujeito a	$x + y + z \leq 1,$
	$x \geq 0,$
	$y \geq 0,$
	$z \geq 0.$

SOLUÇÃO: vamos subdividir a resolução deste problema em vários passos.

• PASSO 0. Escrever o problema na forma-padrão.

Para aplicarmos o teorema (12.7), devemos ter um problema de maximização com todas as restrições em desigualdade na forma \leq . Caso alguma restrição seja da forma \geq , é muito fácil convertê-la para a forma \leq : basta multiplicá-la por -1 . Fazendo isto, não alteramos o conjunto admissível e podemos usar as condições de primeira ordem do teorema. Assim, na forma-padrão, temos

maximizar	$f(x, y, z)$	=	xyz
sujeito a	$g_1(x, y, z)$	=	$x + y + z \leq 0,$
	$g_2(x, y, z)$	=	$-x \leq 1,$
	$g_3(x, y, z)$	=	$-y \leq 0,$
	$g_4(x, y, z)$	=	$-z \leq 0.$

de no da no es de os os es e

• PASSO 1. As funções f e g são de classe C^1 como soma e multiplicação de funções de classe C^1 .

• PASSO 2. Verificar a condição de regularidade.

A condição de regularidade exigida pelo teorema (12.7) diz que o posto da matriz formada pelos vetores gradientes das restrições que estão *ativas* na solução do problema deve ser máximo. Mas como verificar a condição de regularidade se ainda não conhecemos a solução do problema de otimização? De fato, o que queremos é justamente encontrar a solução (ou soluções) do problema! A idéia é mostrar que *todos* os pontos do conjunto admissível satisfazem a condição de regularidade! Evidentemente, dependendo do ponto, podemos ter apenas uma restrição ativa, apenas duas restrições ativas e assim por diante. Em cada caso, devemos mostrar que o posto da matriz é igual ao número de restrições ativas. Com isto, caso exista alguma solução do problema de otimização, certamente ela irá satisfazer a condição de regularidade exigida pelo teorema (12.7).

Primeiro, vamos representar geometricamente o conjunto admissível. As restrições $x \geq 0$, $y \geq 0$ e $z \geq 0$ dizem que estamos no primeiro octante enquanto que a restrição $x + y + z \leq 1$ diz que, ao mesmo tempo, devemos estar abaixo do plano que passa pelos pontos $(1, 0, 0)$, $(0, 1, 0)$ e $(0, 0, 1)$. Desta maneira, é fácil ver que o conjunto admissível é formado pelo tetraedro sólido cujos vértices são $\mathbf{v}_1 = (0, 0, 0)$, $\mathbf{v}_2 = (1, 0, 0)$, $\mathbf{v}_3 = (0, 1, 0)$ e $\mathbf{v}_4 = (0, 0, 1)$.

Uma vez que temos o desenho do conjunto admissível, é fácil identificar geometricamente os pontos com apenas uma restrição ativa, apenas duas restrições ativas e assim por diante. Por exemplo, dizer que apenas a restrição g_1 está ativa em um ponto (x, y, z) significa dizer que o ponto está no interior da face triangular com vértices \mathbf{v}_2 , \mathbf{v}_3 e \mathbf{v}_4 . Por outro lado, os pontos no interior do segmento que une os vértices \mathbf{v}_2 e \mathbf{v}_3 possuem apenas duas restrições ativas: g_1 e g_4 . O vértice \mathbf{v}_4 é o único ponto onde apenas as restrições g_1 , g_2 e g_3 estão ativas. Moral da história: para o tetraedro, nos vértices temos apenas três restrições ativas, nos pontos das arestas (que não são vértices) temos apenas duas restrições ativas e nos pontos das faces (que não são vértices e não estão nas arestas) temos apenas uma única restrição ativa. Pontos do interior do tetraedro correspondem a pontos onde *nenhuma* restrição está ativa.

Van
faze

1.

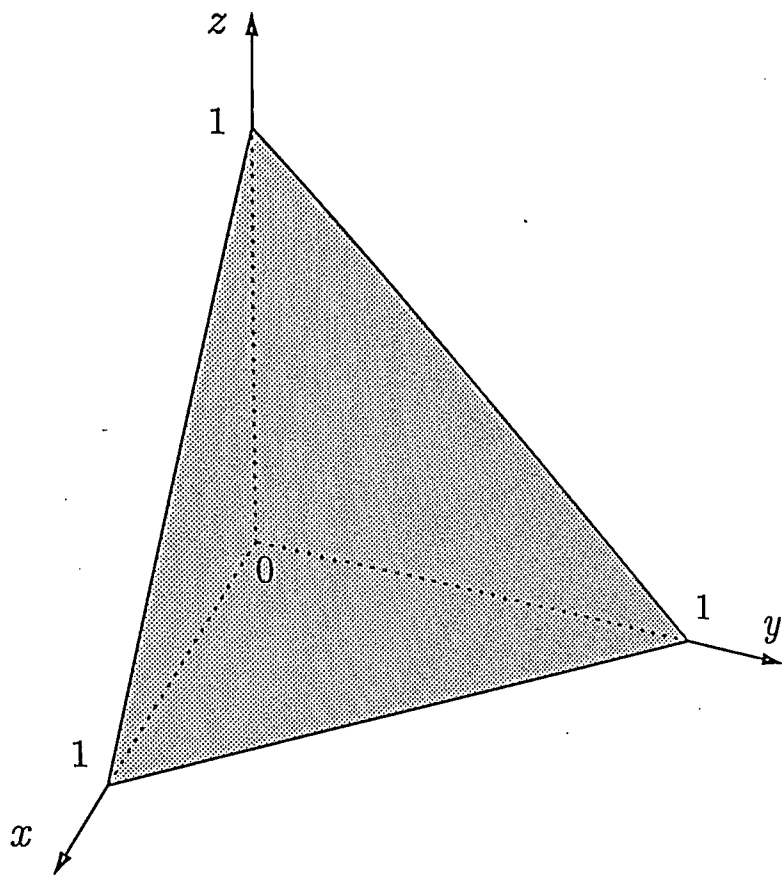


Figura 12.11: Conjunto admissível para o exercício resolvido (12.4).

Vamos agora mostrar que todos os pontos do conjunto admissível satisfazem a condição de regularidade, de acordo com os vários casos.

1. Uma restrição ativa apenas: temos quatro casos.

(a) Se $g_1(x, y, z) = 1$ então

$$\left[\nabla g_1(x, y, z) \right] = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

tem posto 1.

(b) Se $g_2(x, y, z) = 0$ então

$$\left[\nabla g_2(x, y, z) \right] = \begin{bmatrix} -1 & 0 & 0 \end{bmatrix}$$

tem posto 1.

(c) Se $g_3(x, y, z) = 0$ então

$$\left[\nabla g_3(x, y, z) \right] = \begin{bmatrix} 0 & -1 & 0 \end{bmatrix}$$

tem posto 1.

(d) Se $g_4(x, y, z) = 0$ então

$$\left[\nabla g_4(x, y, z) \right] = \begin{bmatrix} 0 & 0 & -1 \end{bmatrix}$$

tem posto 1.

2. Duas restrições ativas apenas: temos seis casos.

(a) Se $g_1(x, y, z) = 1$ e $g_2(x, y, z) = 0$, então

$$\begin{bmatrix} \nabla g_1(x, y, z) \\ \nabla g_2(x, y, z) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

tem posto 2.

(b) Se $g_1(x, y, z) = 1$ e $g_3(x, y, z) = 0$, então

$$\begin{bmatrix} \nabla g_1(x, y, z) \\ \nabla g_3(x, y, z) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

tem posto 2.

(c) Se $g_1(x, y, z) = 1$ e $g_4(x, y, z) = 0$, então

$$\begin{bmatrix} \nabla g_1(x, y, z) \\ \nabla g_4(x, y, z) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

tem posto 2.

(d) Se $g_2(x, y, z) = 0$ e $g_3(x, y, z) = 0$, então

$$\begin{bmatrix} \nabla g_2(x, y, z) \\ \nabla g_3(x, y, z) \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

tem posto 2.

(e) Se $g_2(x, y, z) = 0$ e $g_4(x, y, z) = 0$, então

$$\begin{bmatrix} \nabla g_2(x, y, z) \\ \nabla g_4(x, y, z) \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

tem posto 2.

(f) Se $g_3(x, y, z) = 1$ e $g_4(x, y, z) = 0$, então

$$\begin{bmatrix} \nabla g_3(x, y, z) \\ \nabla g_4(x, y, z) \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

tem posto 2.

3. Três restrições ativas apenas: temos quatro casos.

(a) Se $g_1(x, y, z) = 1$, $g_2(x, y, z) = 0$ e $g_3(x, y, z) = 0$, então

$$\begin{bmatrix} \nabla g_1(x, y, z) \\ \nabla g_2(x, y, z) \\ \nabla g_3(x, y, z) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

tem posto 3.

(b) Se $g_1(x, y, z) = 1$, $g_2(x, y, z) = 0$ e $g_4(x, y, z) = 0$, então

$$\begin{bmatrix} \nabla g_1(x, y, z) \\ \nabla g_2(x, y, z) \\ \nabla g_4(x, y, z) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

tem posto 3.

(c) Se $g_1(x, y, z) = 1$, $g_3(x, y, z) = 0$ e $g_4(x, y, z) = 0$, então

$$\begin{bmatrix} \nabla g_1(x, y, z) \\ \nabla g_3(x, y, z) \\ \nabla g_4(x, y, z) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

tem posto 3.

(d) Se $g_2(x, y, z) = 1$, $g_3(x, y, z) = 0$ e $g_4(x, y, z) = 0$, então

$$\begin{bmatrix} \nabla g_2(x, y, z) \\ \nabla g_3(x, y, z) \\ \nabla g_4(x, y, z) \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

tem posto 3.

de
no
da
no
ies
de
os
es
e

4. Quatro restrições ativas: não temos caso algum pois não existem pontos com quatro restrições ativas. Para vê-lo, basta observar que se $g_1(x, y, z) = 1$, $g_2(x, y, z) = 0$, $g_3(x, y, z) = 0$ e $g_4(x, y, z) = 0$, para algum ponto (x, y, z) , então teríamos $x = 0$, $y = 0$ e $z = 0$, de modo que $x + y + z = 0 = 1$, um absurdo.

• PASSO 3. Escrever o lagrangeano.

$$\begin{aligned} L(x, y, z, \mu_1, \mu_2, \mu_3, \mu_4) &= f(x, y, z) - \mu_1 \cdot [g_1(x, y, z) - b_1] \\ &\quad - \mu_2 \cdot [g_2(x, y, z) - b_2] \\ &\quad - \mu_3 \cdot [g_3(x, y, z) - b_3] \\ &\quad - \mu_4 \cdot [g_4(x, y, z) - b_4] \\ &= xyz - \mu_1 (x + y + z - 1) \\ &\quad + \mu_2 x + \mu_3 y + \mu_4 z. \end{aligned}$$

• PASSO 4. Escrever as condições de primeira ordem.

Temos

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial x}(x, y, z, \mu_1, \mu_2, \mu_3, \mu_4) = 0, \\ \frac{\partial L}{\partial y}(x, y, z, \mu_1, \mu_2, \mu_3, \mu_4) = 0, \\ \frac{\partial L}{\partial z}(x, y, z, \mu_1, \mu_2, \mu_3, \mu_4) = 0, \\ \mu_1 \cdot [g_1(x, y, z) - b_1] = 0, \\ \mu_2 \cdot [g_2(x, y, z) - b_2] = 0, \\ \mu_3 \cdot [g_3(x, y, z) - b_3] = 0, \\ \mu_4 \cdot [g_4(x, y, z) - b_4] = 0, \\ \mu_1 \geq 0, \\ \mu_2 \geq 0, \\ \mu_3 \geq 0, \\ \mu_4 \geq 0, \\ g_1(x, y, z) \leq b_1, \\ g_2(x, y, z) \leq b_2, \\ g_3(x, y, z) \leq b_3, \\ g_4(x, y, z) \leq b_4. \end{array} \right.$$

Calculando as derivadas parciais do lagrangeano com relação a x, y e z , obtemos o seguinte sistema

• P
or
Da

De
Se
(9)

Por
igu
que
que
que
vért

Se μ
do c
 $\mu_1 =$
 $y =$

$$\left\{ \begin{array}{ll} yz - \mu_1 + \mu_2 = 0, & (1) \\ xz - \mu_1 + \mu_3 = 0, & (2) \\ xy - \mu_1 + \mu_4 = 0, & (3) \\ \mu_1 [x + y + z - 1] = 0, & (4) \\ \mu_2 x = 0, & (5) \\ \mu_3 y = 0, & (6) \\ \mu_4 z = 0, & (7) \\ \mu_1 \geq 0, & (8) \\ \mu_2 \geq 0, & (9) \\ \mu_3 \geq 0, & (10) \\ \mu_4 \geq 0, & (11) \\ x + y + z \leq 1, & (12) \\ x \geq 0, & (13) \\ y \geq 0, & (14) \\ z \geq 0. & (15) \end{array} \right.$$

nde
no
ada
no
ses
de
vos
ios
res
Se

- PASSO 5. Resolver o sistema correspondente às condições de primeira ordem.

Das condições (1), (2) e (3) temos

$$\mu_1 = yz + \mu_2 = xz + \mu_3 = xy + \mu_4.$$

De (8), temos dois casos: $\mu_1 = 0$ ou $\mu_1 > 0$.

Se $\mu_1 = 0$, como todas as variáveis são maiores ou iguais a zero, por (8), (9), (10), (11), (13), (14) e (15), segue-se que

$$yz = xz = xy = 0 \quad \text{e} \quad \mu_2 = \mu_3 = \mu_4 = 0.$$

Portanto, pelo menos duas entre as três variáveis x , y e z devem ser iguais a zero. Por outro lado, como $x + y + z \leq 1$, por (12), concluímos que a variável restante deve estar entre 0 e 1. Sendo assim, os pontos que satisfazem as condições de primeira ordem, neste caso, são aqueles que estão em cada uma das três arestas que ligam $(0, 0, 0)$ aos demais vértices do tetraedro.

Se $\mu_1 > 0$, por (4), devemos ter $x + y + z = 1$. Nesta situação, x é maior do que zero pois, se $x = 0$, como $\mu_1 = xz + \mu_3 = xy + \mu_4$, teríamos $\mu_1 = \mu_3 = \mu_4 > 0$ e, conseqüentemente, por (6) e (7), teríamos também $y = 0$ e $z = 0$. Mas então $x + y + z$ seria igual a 0 e não igual a 1,

uma contradição. Da mesma maneira podemos demonstrar que $y > 0$ e $z > 0$. Logo, por (5), (6) e (7), segue-se que

$$\mu_2 = \mu_3 = \mu_4 = 0 \quad \text{e} \quad \mu_1 = yz = xz = xy.$$

Dividindo-se por x , y e z (isto é possível pois $x > 0$, $y > 0$ e $z > 0$), concluímos que

$$x = y = z$$

Como $x + y + z = 1$, temos $x = y = z = 1/3$ e $\mu_1 = 1/9$. Logo,

$$(x, y, z, \mu_1, \mu_2, \mu_3, \mu_4) = (1/3, 1/3, 1/3, 1/9, 0, 0, 0)$$

é o único ponto que satisfaz as condições de primeira ordem, neste segundo caso.

• PASSO 6. Conclusão.

A função-objetivo f é contínua (pois é de classe C^1) e o conjunto admissível é compacto (veja a figura (12.11)). Pelo teorema de Weierstrass, f possui máximos e mínimos globais no conjunto admissível. Agora, pelo teorema (12.7), qualquer solução do problema de otimização que estamos estudando deve ser solução do sistema associado às condições de primeira ordem (a condição de regularidade é fundamental aqui pois, com ela, temos a garantia que qualquer solução do problema de otimização será também solução do sistema associado às condições de primeira ordem para valores adequados dos multiplicadores). Para encontrá-los, basta calcular a função-objetivo nestes pontos e selecionar os de maior valor. No passo 5 vimos que os únicos pontos do conjunto admissível que satisfazem as condições de primeira ordem, para valores adequados dos multiplicadores, são os pontos sobre as três arestas que ligam $(0, 0, 0)$ aos demais vértices do tetraedro, para os quais o valor da função f é 0, e o ponto

$$(1/3, 1/3, 1/3),$$

para o qual o valor de f é $1/27$. Sendo assim, dentre todos os valores de x , y e z que satisfazem as restrições, $x + y + z \leq 1$, $x \geq 0$, $y \geq 0$ e $z \geq 0$, a escolha $x = 1/3$, $y = 1/3$ e $z = 1/3$ é que torna o valor da expressão xyz o maior possível. ■

Como vo
trar os p
temos qu
resolvido
Neste pr
trando to
valores d
de uma c
subcasos.
os subcas

Mas c
ordem em
o sistema
no proble
em design
com 48 c

1. Mais d
blema
na tent
Este ti
e Mate

2. As con
implem
usados

Um erro
de prim
nada ex
não. Po
primeiro
equação
pressão
consider
a expres

Como você pode observar no passo 5 deste exercício resolvido, para encontrar os pontos que satisfazem as condições de primeira ordem, quase sempre temos que fazê-lo considerando-se vários casos. Por exemplo, no exercício resolvido (12.5), a condição $\mu_1 \geq 0$ gerou dois casos: $\mu_1 = 0$ ou $\mu_1 > 0$. Neste processo, você deve fazer um estudo completo de cada caso, encontrando todos os pontos que satisfazem as relações estabelecidas (*incluindo os valores dos multiplicadores*) ou mostrando que não existem soluções através de uma contradição. Frequentemente, casos precisam ser subdivididos em subcasos. Por exemplo, no exercício resolvido (12.5), tivemos que considerar os subcasos $x = 0$ e $x > 0$ dentro do caso $\mu_1 > 0$.

Mas como encontrar os pontos que satisfazem as condições de primeira ordem em uma situação com várias variáveis e várias restrições? Certamente o sistema resultante terá muitas condições e muitas variáveis. Por exemplo, no problema da dieta (página 26), com funções de 9 variáveis e 13 restrições em desigualdade, as condições de primeira ordem resultam em um sistema com 48 condições e 22 variáveis. Neste sentido, temos duas observações.

1. Mais do que uma possível ferramenta de cálculo das soluções de um problema de otimização, as condições de primeira ordem são ferramentas úteis na tentativa de se obter propriedades importantes do problema em estudo. Este tipo de análise teórica é muito freqüente em Economia, Engenharia e Matemática.
2. As condições de primeira ordem podem sugerir algoritmos que possam ser implementados em um computador. Os *métodos de ponto interior*, muito usados e discutidos atualmente, são um exemplo disto.

CUIDADO! CUIDADO! CUIDADO!

Um erro muito freqüente na resolução do sistema associado às condições de primeira ordem é o de se "dividir" uma equação por uma determinada expressão sem se preocupar com o fato desta expressão ser zero ou não. Por exemplo, no exercício resolvido (12.5), tivemos que garantir primeiro que x é maior do que zero para concluir que $y = z$, a partir da equação $xy = xz$. Se você vai "dividir" uma equação por uma certa expressão e você não sabe se ela é igual a zero ou não, então *você terá que considerar dois casos*: quando a expressão é diferente de zero e quando a expressão é igual a zero.

ide no ida no ses de los los res e

12.5 Otimização com restrições mistas

É fácil combinar os resultados dos teoremas (12.4) e (12.7) para obter um teorema geral, com restrições em igualdade e em desigualdade.

Teorema 12.8 (DE KARUSH-KUHN-TUCKER) Sejam $f, h_1, \dots, h_m, g_1, \dots, g_k$ funções de classe C^1 de n variáveis definidas em um aberto de \mathbb{R}^n e seja $\mathbf{p} \in \mathbb{R}^n$ um máximo local de f no conjunto admissível

$$D = \{\mathbf{x} \mid h_1(\mathbf{x}) = c_1, \dots, h_m(\mathbf{x}) = c_m, g_1(\mathbf{x}) \leq b_1, \dots, g_k(\mathbf{x}) \leq b_k\},$$

formado com m restrições em igualdade e k restrições em desigualdade, onde $\mathbf{x} = (x_1, \dots, x_n)$. Caso alguma restrição em desigualdade esteja ativa em \mathbf{p} , vamos renomeá-las de forma que elas sejam as l primeiras: g_1, \dots, g_l . Suponha que \mathbf{p} satisfaça a seguinte condição de regularidade: o posto da matriz jacobiana

$$\begin{bmatrix} \nabla h_1(\mathbf{p}) \\ \vdots \\ \nabla h_m(\mathbf{p}) \\ \nabla g_1(\mathbf{p}) \\ \vdots \\ \nabla g_l(\mathbf{p}) \end{bmatrix}_{(m+l) \times n} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1}(\mathbf{p}) & \cdots & \frac{\partial h_1}{\partial x_n}(\mathbf{p}) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1}(\mathbf{p}) & \cdots & \frac{\partial h_m}{\partial x_n}(\mathbf{p}) \\ \frac{\partial g_1}{\partial x_1}(\mathbf{p}) & \cdots & \frac{\partial g_1}{\partial x_n}(\mathbf{p}) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_l}{\partial x_1}(\mathbf{p}) & \cdots & \frac{\partial g_l}{\partial x_n}(\mathbf{p}) \end{bmatrix}_{(m+l) \times n}$$

é igual a $m + l$ (o número de restrições ativas em \mathbf{p}). Considere o *lagrangeano* definido por

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) - \lambda_1 \cdot [h_1(\mathbf{x}) - c_1] - \cdots - \lambda_m \cdot [h_m(\mathbf{x}) - c_m] - \mu_1 \cdot [g_1(\mathbf{x}) - b_1] - \cdots - \mu_k \cdot [g_k(\mathbf{x}) - b_k],$$

onde

$$\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m) \quad \text{e} \quad \boldsymbol{\mu} = (\mu_1, \dots, \mu_k).$$

Então existem multiplicadores $\lambda_1^*, \dots, \lambda_m^*, \mu_1^*, \dots, \mu_k^*$ tais que

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial x_1}(\mathbf{p}, \boldsymbol{\mu}^*) = 0, \quad \dots, \quad \frac{\partial L}{\partial x_n}(\mathbf{p}, \boldsymbol{\mu}^*) = 0, \\ h_1(\mathbf{p}) = c_1, \\ \vdots \\ h_m(\mathbf{p}) = c_m, \\ \mu_1^* \cdot [g_1(\mathbf{p}) - b_1] = 0, \\ \vdots \\ \mu_k^* \cdot [g_k(\mathbf{p}) - b_k] = 0, \\ \mu_1^* \geq 0, \\ \vdots \\ \mu_k^* \geq 0, \\ g_1(\mathbf{p}) \leq b_1, \\ \vdots \\ g_k(\mathbf{p}) \leq b_k, \end{array} \right.$$

ou, equivalentemente,

$$\left\{ \begin{array}{l} \nabla f(\mathbf{p}) = \sum_{i=1}^m \lambda_i^* \cdot \nabla h_i(\mathbf{p}) + \sum_{j=1}^k \mu_j^* \cdot \nabla g_j(\mathbf{p}), \\ h_1(\mathbf{p}) = c_1, \\ \vdots \\ h_m(\mathbf{p}) = c_m, \\ \mu_1^* \cdot [g_1(\mathbf{p}) - b_1] = 0, \\ \vdots \\ \mu_k^* \cdot [g_k(\mathbf{p}) - b_k] = 0, \\ \mu_1^* \geq 0, \\ \vdots \\ \mu_k^* \geq 0, \\ g_1(\mathbf{p}) \leq b_1, \\ \vdots \\ g_k(\mathbf{p}) \leq b_k. \end{array} \right.$$

Estes sistemas são denominados de *condições de primeira ordem* para o ponto de máximo (local) \mathbf{p} .

ide
no
da
no
ses
de
los
ios
res
5 e

Exercício resolvido 12.6 Resolva o problema de otimização que consiste em

$$\begin{array}{ll} \text{maximizar} & x - y^2 \\ \text{sujeito a} & x^2 + y^2 = 4, \\ & x \geq 0, \\ & y \geq 0. \end{array}$$

SOLUÇÃO: vamos subdividir a resolução deste problema em vários passos.

- PASSO 0. Escrever o problema na forma-padrão.

Para aplicarmos o teorema (12.8), devemos ter um problema de *maximização* com *todas* as restrições em desigualdade na forma \leq . Caso alguma restrição seja da forma \geq , é muito fácil convertê-la para a forma \leq : basta multiplicá-la por -1 . Fazendo isto, *não* alteramos o conjunto admissível e podemos usar as condições de primeira ordem do teorema. Assim, na forma-padrão, temos

$$\begin{array}{llll} \text{maximizar} & f(x, y) & = & x - y^2 \\ \text{sujeito a} & h_1(x, y) & = & x^2 + y^2 = 4, \\ & g_1(x, y) & = & -x \leq 0, \\ & g_2(x, y) & = & -y \leq 0. \end{array}$$

- PASSO 1. As funções f e g são de classe C^1 como soma e multiplicação de funções de classe C^1 .
- PASSO 2. Verificar a condição de regularidade.

A condição de regularidade exigida pelo teorema (12.8) (teorema de Karush-Kuhn-Tucker) diz que o posto da matriz formada pelos vetores gradientes das restrições que estão *ativas* na solução do problema deve ser máximo. *Lembre-se que as restrições em igualdade estão sempre ativas em qualquer ponto do conjunto admissível.* Mas como verificar a condição de regularidade se ainda não conhecemos a solução do problema de otimização? De fato, o que queremos é justamente encontrar a solução (ou soluções) do problema! A idéia é mostrar que *todos* os pontos do conjunto admissível satisfazem a condição de regularidade! Evidentemente, dependendo do ponto, podemos ter apenas uma restrição ativa ou apenas duas restrições ativas. Em cada caso, devemos mostrar

que o posto da matriz é sempre igual ao número de restrições ativas. Com isto, caso exista alguma solução do problema de otimização, certamente ela irá satisfazer a condição de regularidade exigida pelo teorema de Karush-Kuhn-Tucker. Vamos agora mostrar que todos os pontos do conjunto admissível satisfazem a condição de regularidade, de acordo com os vários casos.

1. Uma restrição ativa apenas: temos o único caso.

(a) Se $h_1(x, y) = 4$, então

$$\begin{bmatrix} \nabla h_1(x, y) \end{bmatrix} = \begin{bmatrix} 2x & 2y \end{bmatrix}$$

tem posto 1 pois se $h_1(x, y) = x^2 + y^2 = 4$, então x e y não podem anular-se simultaneamente.

Lembre-se que a restrição em igualdade $h_1(x, y) = 4$ está sempre ativa. Desta maneira, não existem pontos do conjunto admissível onde a restrição $g_1 \leq 0$ ou a restrição $g_2 \leq 0$ seja a única restrição ativa. Por este motivo, não se deve considerar estes casos no estudo da condição de regularidade.

2. Duas restrições ativas apenas: temos dois casos.

(a) Se $h_1(x, y) = 4$ e $g_1(x, y) = 0$, então $x^2 + y^2 = 4$ e $x = 0$. Como $y \geq 0$, segue-se que o único ponto do conjunto admissível com estas duas restrições ativas é o ponto $(0, 2)$. Sendo assim, a matriz

$$\begin{bmatrix} \nabla h_1(0, 2) \\ \nabla g_1(0, 2) \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ -1 & 0 \end{bmatrix} \sim \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$$

tem posto 2.

(b) Se $h_1(x, y) = 4$ e $g_2(x, y) = 0$, então $x^2 + y^2 = 4$ e $y = 0$. Como $x \geq 0$, segue-se que o único ponto do conjunto admissível com estas duas restrições ativas é o ponto $(2, 0)$. Sendo assim, a matriz

$$\begin{bmatrix} \nabla h_1(2, 0) \\ \nabla g_2(2, 0) \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$$

tem posto 2.

ade
no
da
no
ses
de
los
ios
res
se

3. Três restrições ativas: não temos caso algum pois não existem pontos com três restrições ativas. Para vê-lo, basta observar que se $h_1(x, y) = 4$, $g_1(x, y) = 0$ e $g_2(x, y) = 0$, para algum ponto (x, y) , então teríamos $x = 0$ e $y = 0$, de modo que $x^2 + y^2 = 0 = 4$, um absurdo.

• PASSO 3. Escrever o lagrangeano.

$$\begin{aligned} L(x, y, \lambda, \mu_1, \mu_2) &= f(x, y) - \lambda_1 \cdot [h_1(x, y) - c_1] \\ &\quad - \mu_1 \cdot [g_1(x, y) - b_1] \\ &\quad - \mu_2 \cdot [g_2(x, y) - b_2] \\ &= x - y^2 - \lambda_1 (x^2 + y^2 - 4) + \mu_1 x + \mu_2 y. \end{aligned}$$

• PASSO 4. Escrever as condições de primeira ordem.

Temos

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial x}(x, y, \lambda_1, \mu_1, \mu_2) = 0, \\ \frac{\partial L}{\partial y}(x, y, \lambda_1, \mu_1, \mu_2) = 0, \\ h_1(x, y) = c_1 \\ \mu_1 \cdot [g_1(x, y) - b_1] = 0, \\ \mu_2 \cdot [g_2(x, y) - b_2] = 0, \\ \mu_1 \geq 0, \\ \mu_2 \geq 0, \\ g_1(x, y) \leq b_1, \\ g_2(x, y) \leq b_2. \end{array} \right.$$

Calculando as derivadas parciais do lagrangeano com relação a x e y , obtemos o seguinte sistema

$$\left\{ \begin{array}{ll} 1 - 2\lambda_1 x + \mu_1 = 0, & (1) \\ -2y - 2\lambda_1 y + \mu_2 = 0, & (2) \\ x^2 + y^2 = 4, & (3) \\ \mu_1 x = 0, & (4) \\ \mu_2 y = 0, & (5) \\ \mu_1 \geq 0, \mu_2 \geq 0, & (6) \\ x \geq 0, y \geq 0. & (7) \end{array} \right.$$

• PASSO 5. Resolver o sistema correspondente às condições de primeira ordem.

De (1) temos $1 + \mu_1 = 2\lambda_1 x$. Como, por (6), $\mu_1 \geq 0$, segue-se que $2\lambda_1 x > 0$. Dado que, por (7), $x \geq 0$, concluímos que $\lambda_1 > 0$ e $x > 0$. Sendo assim, por (4), temos

$$\mu_1 = 0.$$

De (2) temos $2y(1 + \lambda_1) = \mu_2$. Logo, se $\mu_2 > 0$, então $y > 0$ e, portanto, $\mu_2 y > 0$, o que contradiz (5). Desta maneira,

$$\mu_2 = 0 \quad \text{e} \quad y = \mu_2 / (2(1 + \lambda_1)) = 0.$$

Por (3) e (7) temos

$$x = 2 \quad \text{e} \quad \lambda_1 = (1 + \mu_1) / (2x) = 1/4.$$

Desta maneira, o único ponto que satisfaz as condições de primeira ordem é

$$(x, y, \lambda_1, \mu_1, \mu_2) = (2, 0, 1/4, 0, 0).$$

• PASSO 6. Conclusão.

A função-objetivo f é contínua (pois é de classe C^1) e o conjunto admissível é compacto (veja a figura (12.12)).

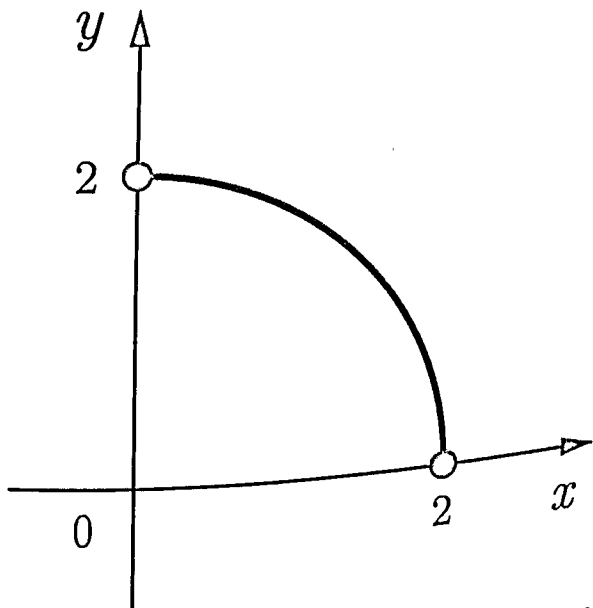


Figura 12.12: Conjunto admissível para o exercício resolvido (12.6).

Pelo teorema de Weierstrass, f possui máximos e mínimos globais no conjunto admissível. Agora, pelo teorema de Karush-Kuhn-Tucker, qualquer solução do problema de otimização que estamos estudando deve ser solução do sistema associado às condições de primeira ordem (a condição de regularidade é fundamental aqui pois, com ela, temos a garantia que qualquer solução do problema de otimização será também solução do sistema associado às condições de primeira ordem para valores adequados dos multiplicadores). Para encontrá-los, basta calcular a função-objetivo nestes pontos e selecionar os de maior valor. Portanto, como existe um único ponto que satisfaz as condições de primeira, ele é a solução do problema de otimização, isto é, dentre todos os valores de x e y que satisfazem as restrições, $x^2 + y^2 = 4$, $x \geq 0$ e $y \geq 0$, a escolha $x = 2$ e $y = 0$ é a que torna o valor da expressão $x - y^2$ o maior possível. \square

Observação. Os teoremas (12.1), (12.2), (12.3), (12.4), (12.6) e (12.7) são casos particulares do teorema de Karush-Kuhn-Tucker.

12.6 Alternativas para a condição de regularidade

A condição de regularidade do teorema de Karush-Kuhn-Tucker nos fornece a garantia de que *qualquer* solução do problema de otimização será recuperada quando resolvemos o sistema associado às condições de primeira ordem. Evidentemente, caso algum ponto não satisfaça a condição de regularidade, ele deve ser incluído como um candidato em potencial para a solução do problema de otimização. Dito de outra forma, se no passo 2 você descobriu que algum ponto admissível não satisfaz a condição de regularidade, você deve incluí-lo no passo 6, junto com os pontos obtidos no passo 5.

Como vimos, para omitirmos a condição de regularidade, basta incluir um multiplicador para a função-objetivo. Esta idéia, devida a Fritz John, é dada no próximo teorema em sua forma mais geral.

Teorema 12.9 (DE FRITZ JOHN) Sejam $f, h_1, \dots, h_m, g_1, \dots, g_k$ funções de classe C^1 de n variáveis definidas em um aberto de \mathbb{R}^n e seja $\mathbf{p} \in \mathbb{R}^n$ um máximo local de f no conjunto admissível

$$D = \{\mathbf{x} \mid h_1(\mathbf{x}) = c_1, \dots, h_m(\mathbf{x}) = c_m, g_1(\mathbf{x}) \leq b_1, \dots, g_k(\mathbf{x}) \leq b_k\},$$

form
onde
 μ_1^*

Estes
Fritz

Observe q
as mesma
Claramen
função-ob
Uma man
é o de exig
que estejan
outra alter

Teorema
funções

formado com m restrições em igualdade e k restrições em desigualdade, onde $\mathbf{x} = (x_1, \dots, x_n)$. Então existem multiplicadores $\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*, \mu_1^*, \dots, \mu_k^*$ tais que

$$\left\{ \begin{array}{l} \lambda_0^* \cdot \nabla f(\mathbf{p}) = \sum_{i=1}^m \lambda_i^* \cdot \nabla h_i(\mathbf{p}) + \sum_{j=1}^k \mu_j^* \cdot \nabla g_j(\mathbf{p}), \\ h_1(\mathbf{p}) = c_1, \\ \vdots \\ h_m(\mathbf{p}) = c_m, \\ \mu_1^* \cdot [g_1(\mathbf{p}) - b_1] = 0, \\ \vdots \\ \mu_k^* \cdot [g_k(\mathbf{p}) - b_k] = 0, \\ \mu_1^* \geq 0, \\ \vdots \\ \mu_k^* \geq 0, \\ g_1(\mathbf{p}) \leq b_1, \\ \vdots \\ g_k(\mathbf{p}) \leq b_k, \\ (\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*, \mu_1^*, \dots, \mu_k^*) \neq (0, 0, \dots, 0, 0, \dots, 0), \\ \lambda_0^* \in \{0, 1\}. \end{array} \right.$$

Estes sistemas são denominados de *condições de primeira ordem de Fritz John* para o ponto de máximo (local) \mathbf{p} .

Observe que quando $\lambda_0^* = 1$, as condições de primeira ordem de Fritz John são as mesmas condições de primeira ordem do teorema de Karush-Kuhn-Tucker. Claramente, é preferível uma situação onde $\lambda_0^* = 1$ pois, caso contrário, a função-objetivo desaparece completamente das condições de primeira ordem. Uma maneira de garantir que λ_0^* seja igual a 1 no teorema de Fritz John, é o de exigir que o posto da matriz formada pelos gradientes das restrições que estejam ativas na solução do problema de otimização seja máximo. Uma outra alternativa é dada no próximo teorema.

Teorema 12.10 Se h_1, \dots, h_m são funções lineares e g_1, \dots, g_k são funções côncavas, então $\lambda_0^* = 1$ no teorema de Fritz John.

ade
no
ada
no
ises
de
Vos
ios
res
5 e

A demonstração deste teorema é uma aplicação do lema de Farkas, que não faremos aqui. Para ver uma aplicação deste teorema, considere o exercício resolvido (12.5). Todas as funções-restrições deste exercício são lineares. O teorema (12.10) então nos permite omitir o passo 2 na resolução do exercício e, mesmo assim, podemos concluir que a solução do problema de otimização está entre os pontos que resolvem as condições de primeira ordem, obtidos no passo 5.

12.7 Problemas de minimização

O teorema (12.8) (de Karush-Kuhn-Tucker) só pode ser aplicado em problemas de *maximização* onde as restrições em desigualdade devem estar na forma \leq . Vamos ver que, de fato, basta considerar este caso, pois qualquer problema de otimização pode ser convertido para esta forma-padrão. Por exemplo, já vimos que, se alguma restrição em desigualdade é do tipo \geq , basta multiplicá-la por -1 e teremos uma restrição da forma \leq , uma modificação que *não altera* o conjunto admissível.

Mas como converter problemas de minimização em problemas de maximização? A resposta a esta pergunta foi dada no exercício [12] da página 361. Lá, vimos que p é um ponto de mínimo de uma função f em um conjunto admissível D se, e somente se, p é um ponto de máximo da função $-f$ no mesmo conjunto admissível. Assim, para resolver um problema de minimização, basta resolver um problema de maximização trocando-se a função-objetivo f original por $-f$. Por exemplo, para se resolver o problema de minimização

$$\begin{array}{ll} \text{minimizar} & 2y - x^2 \\ \text{sujeito a} & x^2 + y^2 \leq 1, \\ & x \geq 0, \\ & y \geq 0, \end{array}$$

basta resolver o problema de maximização

$$\begin{array}{ll} \text{maximizar} & -2y + x^2 \\ \text{sujeito a} & x^2 + y^2 \leq 1, \\ & -x \leq 0, \\ & -y \leq 0, \end{array}$$

com o teorema de Karush-Kuhn-Tucker. Outra alternativa para se minimizar uma função, com restrições em desigualdade da forma \leq , é a de se manter a mesma função-objetivo mas, na construção do lagrangeano, usar o sinal de + no lugar do sinal de - para os termos associados às desigualdades.

12.8 Condições de segunda ordem*

Da mesma maneira que, nos problemas de otimização sem restrições que estudamos no capítulo 11, desenvolvemos ferramentas para classificar pontos críticos, vamos agora estabelecer teoremas que permitam classificar as soluções das condições de primeira ordem como máximo local, mínimo local ou nenhum dos dois.

Condições de segunda ordem para otimização com restrições em igualdade

Sejam f, h_1, \dots, h_m funções de classe C^2 definidas em \mathbb{R}^n e p um ponto para o qual existem números reais $\lambda_1^*, \dots, \lambda_m^*$ que satisfazem as condições de primeira ordem

$$\begin{cases} \nabla f(p) = \lambda_1^* \cdot \nabla h_1(p) + \dots + \lambda_m^* \cdot \nabla h_m(p), \\ h_1(p) = c_1, \\ \vdots \\ h_m(p) = c_m. \end{cases}$$

Defina $h(x) = (h_1(x), \dots, h_m(x))$, o lagrangeano

$$L(x, \lambda) = f(x) - \lambda_1 \cdot [h_1(x) - c_1] - \dots - \lambda_m \cdot [h_m(x) - c_m].$$

seja a matriz hessiana orlada

$$D_x^2 L(p, \lambda^*) = \begin{bmatrix} \frac{\partial^2 L}{\partial x_1^2}(p, \lambda^*) & \dots & \frac{\partial^2 L}{\partial x_1 \partial x_n}(p, \lambda^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial x_n \partial x_1}(p, \lambda^*) & \dots & \frac{\partial^2 L}{\partial x_n^2}(p, \lambda^*) \end{bmatrix}_{n \times n}$$

ade
no
ada
no
ses
de
vos
ios
res
Se

Considere também

$$H(p, \lambda) = \left[\begin{array}{c|c} \mathbf{0}_{m \times m} & Dh(p) \\ \hline Dh(p)^T & D_x^2 L(p, \lambda^*) \end{array} \right]_{(n+m) \times (n+m)},$$

onde $\mathbf{0}_{m \times m}$ é a matriz nula $m \times m$ e $Dh(p)$ é a matriz jacobiana de h em p de tamanho $m \times m$. Se

$$D = \{x \in \mathbb{R}^n \mid h_1(x) = c_1, \dots, h_m(x) = c_m\},$$

temos os casos abaixo.

(a) Se para todo $v \neq 0$ com $Dh(p) \cdot v = 0$ tem-se que

$$Q(v) = v^T \cdot D_x^2 L(p, \lambda^*) \cdot v < 0,$$

então p é um ponto de máximo local de f no conjunto admissível D . A fim de que isto ocorra, é suficiente que $\det(H(p, \lambda))$ tenha o mesmo sinal de $(-1)^n$ e os últimos $n - m$ menores principais líderes de $H(p, \lambda)$ alternem de sinal.

(b) Se para todo $v \neq 0$ com $Dh(p) \cdot v = 0$ tem-se que

$$Q(v) = v^T \cdot D_x^2 L(p, \lambda^*) \cdot v > 0,$$

então p é um ponto de mínimo local de f no conjunto admissível D . A fim de que isto ocorra, é suficiente que $\det(H(p, \lambda))$ e os últimos $n - m$ menores principais líderes de $H(p, \lambda)$ tenham o sinal de $(-1)^m$.

(c) Se existem vetores $v_1 \neq 0$ e $v_2 \neq 0$ tais que $Dh(v_1) = 0, Dh(v_2) = 0,$

$$Q(v_1) = v_1^T \cdot D_x^2 L(p, \lambda^*) \cdot v_1 < 0 \quad \text{e} \quad Q(v_2) = v_2^T \cdot D_x^2 L(p, \lambda^*) \cdot v_2 > 0,$$

então p não é máximo local e nem mínimo local de f no conjunto admissível D . A fim de que isto ocorra, é suficiente que as condições (a) e (b) acima sejam violadas por alguns menores principais líderes não-nulos.

Exemplo 12.1 No exercício resolvido (12.2), estudamos o problema de otimização que consiste em

maxi
suj

Vimos q
regulari

$$L(x_1,$$

e que os
ordem s

$$(-1, +$$

Usando
possuía
tremo, d

são pont

são pont
(0, $-\sqrt{3}$)
ções de p

Para c
de segund

$$H(x_1, x$$

isto é,

$$\begin{array}{l} \text{maximizar} \\ \text{sujeito a } (x_1, x_2) \in D = \{(x_1, x_2) \in \mathbb{R}^2 \mid h(x_1, x_2) = 2x_1^2 + x_2^2 = 3\}. \end{array}$$

Vimos que todos os pontos do conjunto admissível satisfazem a condição de regularidade, que o lagrangeano para o problema é

$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda \cdot [h(x_1, x_2) - c] = x_1^2 x_2 - \lambda \cdot (2x_1^2 + x_2^2 - 3),$$

e que os únicos pontos (x_1, x_2, λ) que satisfazem as condições de primeira ordem são

$$\begin{aligned} &(-1, +1, +1/2), \quad (+1, +1, +1/2), \quad (-1, -1, -1/2), \quad (+1, -1, -1/2), \\ &\quad (0, -\sqrt{3}, 0), \quad \text{e} \quad (0, +\sqrt{3}, 0). \end{aligned}$$

Usando o teorema de Weierstrass, sabíamos que o problema de otimização possuía solução. Avaliando a função-objetivo nos únicos candidatos a extremo, descobrimos que

$$(-1, -1) \quad \text{e} \quad (+1, -1),$$

são pontos de mínimo global e que

$$(-1, +1) \quad \text{e} \quad (+1, +1)$$

são pontos de máximo global de f em D . Mas o que dizer dos pontos $(0, -\sqrt{3})$ e $(0, +\sqrt{3})$? Lembre-se que nem todo ponto que satisfaz as condições de primeira ordem é um extremo local de f em D .

Para classificar os pontos $(0, -\sqrt{3})$ e $(0, +\sqrt{3})$ vamos usar as condições de segunda ordem. A matriz hessiana orlada do problema é dada por

$$H(x_1, x_2, \lambda) = \begin{bmatrix} 0 & \frac{\partial h}{\partial x_1}(x_1, x_2, \lambda) & \frac{\partial h}{\partial x_2}(x_1, x_2, \lambda) \\ \frac{\partial h}{\partial x_1}(x_1, x_2, \lambda) & \frac{\partial L}{\partial x_1 \partial x_1}(x_1, x_2, \lambda) & \frac{\partial L}{\partial x_1 \partial x_2}(x_1, x_2, \lambda) \\ \frac{\partial h}{\partial x_2}(x_1, x_2, \lambda) & \frac{\partial L}{\partial x_2 \partial x_1}(x_1, x_2, \lambda) & \frac{\partial L}{\partial x_2 \partial x_2}(x_1, x_2, \lambda) \end{bmatrix},$$

isto é,

ide
no
ada
no
ses
de
Vos
ios
res
Se

$$H(x_1, x_2, \lambda) = \begin{bmatrix} 0 & 4x_1 & 2x_2 \\ 4x_1 & 2x_2 - 4\lambda & 2x_1 \\ 2x_2 & 2x_1 & -2\lambda \end{bmatrix},$$

Neste problema de otimização temos $n = 2$ variáveis e $m = 1$ restrição, de modo que $n - m = 1$. Pelas condições de segunda ordem, precisamos verificar o sinal de apenas 1 determinante, a saber, o determinante de $H(x_1, x_2, \lambda)$. Se $\det(H(x_1, x_2, \lambda))$ possui o mesmo sinal de $(-1)^n = +1$, isto é, se $\det(H(x_1, x_2, \lambda)) > 0$ em um candidato (x_1, x_2, λ) , então (x_1, x_2, λ) é um ponto de máximo local de f em D . Se $\det(H(x_1, x_2, \lambda))$ possui o mesmo sinal de $(-1)^m = -1$, isto é, se $\det(H(x_1, x_2, \lambda)) < 0$ em um candidato (x_1, x_2, λ) , então (x_1, x_2, λ) é um ponto de mínimo local de f em D .

No ponto $(x_1, x_2, \lambda) = (0, -\sqrt{3}, 0)$ temos

$$H(0, -\sqrt{3}, 0) = \begin{bmatrix} 0 & 0 & -2\sqrt{3} \\ 0 & -2\sqrt{3} & 0 \\ -2\sqrt{3} & 0 & 0 \end{bmatrix},$$

cujo determinante é $+24\sqrt{3} > 0$, de modo que $(0, -\sqrt{3})$ é um ponto de máximo local de f em D .

No ponto $(x_1, x_2, \lambda) = (0, +\sqrt{3}, 0)$ temos

$$H(0, +\sqrt{3}, 0) = \begin{bmatrix} 0 & 0 & +2\sqrt{3} \\ 0 & +2\sqrt{3} & 0 \\ +2\sqrt{3} & 0 & 0 \end{bmatrix},$$

cujo determinante é $-24\sqrt{3} < 0$, de modo que $(0, +\sqrt{3})$ é um ponto de mínimo local de f em D .

Estes resultados coincidem com aqueles obtidos no exercício resolvido (12.2), através da análise geométrica das curvas de nível de f e h . ▀

Condições de segunda ordem para otimização com restrições mistas

Sejam $f, h_1, \dots, h_m, g_1, \dots, g_k$ funções de classe C^2 definidas em \mathbb{R}^n e p um ponto para o qual existem números reais $\lambda_1^*, \dots, \lambda_m^*, \mu_1^*, \dots, \mu_k^*$ que satisfazem o sistema

Defina $h(x)$ $L(x, \lambda)$

a matriz

 $D_x^2 L(p)$

e

 $D = \{x \in$ Caso algum
las de form
 $g_A(x) = (g_1,$

$$\left\{ \begin{array}{l} \nabla f(\mathbf{p}) = \mu_1^* \cdot \nabla g_1(\mathbf{p}) + \dots + \mu_k^* \cdot \nabla g_k(\mathbf{p}), \\ h_1(\mathbf{p}) = c_1, \\ \vdots \\ h_m(\mathbf{p}) = c_m, \\ \mu_1^* \cdot [g_1(\mathbf{p}) - b_1] = 0, \\ \vdots \\ \mu_k^* \cdot [g_k(\mathbf{p}) - b_k] = 0, \\ \mu_1^* \geq 0, \\ \vdots \\ \mu_k^* \geq 0, \\ g_1(\mathbf{p}) \leq b_1, \\ \vdots \\ g_k(\mathbf{p}) \leq b_k. \end{array} \right.$$

Defina $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \dots, h_m(\mathbf{x}))$, $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_k(\mathbf{x}))$, o lagrangeano

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) - \lambda_1 \cdot [h_1(\mathbf{x}) - c_1] - \dots - \lambda_m \cdot [h_m(\mathbf{x}) - c_m] \\ - \mu_1 \cdot [g_1(\mathbf{x}) - b_1] - \dots - \mu_k \cdot [g_k(\mathbf{x}) - b_k],$$

a matriz

$$D_{\mathbf{x}}^2 L(\mathbf{p}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \begin{bmatrix} \frac{\partial^2 L}{\partial x_1^2}(\mathbf{p}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) & \dots & \frac{\partial^2 L}{\partial x_1 \partial x_n}(\mathbf{p}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial x_n \partial x_1}(\mathbf{p}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) & \dots & \frac{\partial^2 L}{\partial x_n^2}(\mathbf{p}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \end{bmatrix}_{n \times n}$$

e

$$D = \{\mathbf{x} \in \mathbb{R}^n \mid h_1(\mathbf{x}) = c_1, \dots, h_m(\mathbf{x}) = c_m, g_1(\mathbf{x}) \leq b_1, \dots, g_k(\mathbf{x}) \leq b_k\}.$$

Caso alguma restrição em desigualdade esteja ativa em \mathbf{p} , vamos renomeá-las de forma que elas sejam as l primeiras: g_1, \dots, g_l . Finalmente, defina $\mathcal{S}_A(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_l(\mathbf{x}))$ e

ade
no
ada
no
ses
de
Vos
ios
res
S e

$$H(\mathbf{p}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \begin{bmatrix} \mathbf{0}_{m \times m} & \mathbf{0}_{m \times l} & Dh(\mathbf{p}) \\ \mathbf{0}_{l \times m} & \mathbf{0}_{l \times l} & Dg_A(\mathbf{p}) \\ Dh(\mathbf{p})^T & Dg_A(\mathbf{p})^T & D_x^2 L(\mathbf{p}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \end{bmatrix}_{(n+m+l) \times (n+m+l)}$$

Se para todo $\mathbf{v} \neq \mathbf{0}$ com $Dh(\mathbf{p}) \cdot \mathbf{v} = \mathbf{0}$ e $Dg_A(\mathbf{p}) \cdot \mathbf{v} = \mathbf{0}$ tem-se

$$Q(\mathbf{v}) = \mathbf{v}^T \cdot D_x^2 L(\mathbf{p}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \cdot \mathbf{v} < 0,$$

então \mathbf{p} é um ponto de máximo local de f no conjunto admissível D . A fim de que isto ocorra, é suficiente que $\det(H(\mathbf{p}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*))$ tenha o mesmo sinal de $(-1)^n$ e os últimos $n - (m + l)$ menores principais líderes de $H(\mathbf{p}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ alternem de sinal.

12.9 Questões de globalidade

As condições de segunda ordem estabelecidas na seção anterior estabelecem um critério para decidir se um ponto que satisfaz as condições de primeira ordem é um extremo *local*. Elas nada afirmam sobre a *globalidade* do ponto. A exemplo de problemas de otimização sem restrições (capítulo 11), é preciso impor propriedades adicionais nas funções envolvidas, a fim de se obter globalidade.

Mais uma vez, a noção de convexidade desempenhará um papel importante. De fato, vale o seguinte teorema de "suficiência":

Teorema 12.11 Sejam $f, h_1, \dots, h_m, g_1, \dots, g_k$ funções de classe C^1 de n variáveis definidas em um aberto *convexo* U de \mathbb{R}^n e seja $\mathbf{p} \in \mathbb{R}^n$ um ponto no conjunto admissível

$$D = \{\mathbf{x} \mid h_1(\mathbf{x}) = c_1, \dots, h_m(\mathbf{x}) = c_m, g_1(\mathbf{x}) \leq b_1, \dots, g_k(\mathbf{x}) \leq b_k\},$$

formado com m restrições em igualdade e k restrições em desigualdade, onde $\mathbf{x} = (x_1, \dots, x_n)$. Se

- (1) a função f é uma *côncava* em U ,
- (2) as funções h_1, \dots, h_m são *lineares*,

(3)

(4)

então

12.10

[01] Na

esca

de c

do v

(a)

(b)

- (3) as funções g_1, \dots, g_k são convexas em U e
- (4) p satisfaz as condições de primeira ordem,

$$\left\{ \begin{array}{l} \nabla f(\mathbf{p}) = \sum_{i=1}^m \lambda_i^* \cdot \nabla h_i(\mathbf{p}) + \sum_{j=1}^k \mu_j^* \cdot \nabla g_j(\mathbf{p}), \\ h_1(\mathbf{p}) = c_1, \\ \vdots \\ h_m(\mathbf{p}) = c_m, \\ \mu_1^* \cdot [g_1(\mathbf{p}) - b_1] = 0, \\ \vdots \\ \mu_k^* \cdot [g_k(\mathbf{p}) - b_k] = 0, \\ \mu_1^* \geq 0, \\ \vdots \\ \mu_k^* \geq 0, \\ g_1(\mathbf{p}) \leq b_1, \\ \vdots \\ g_k(\mathbf{p}) \leq b_k, \end{array} \right.$$

para valores adequados dos multiplicadores λ_i^* e μ_j^* ,
então p é um ponto de *mínimo global* de f em D

ade
no
ada
no
ses
de
Vos
ios
res
Se

12.10 Exercícios

[01] Na figura (12.13) encontram-se algumas curvas de nível de uma função escalar f e uma curva de nível D de uma função escalar h , com f e h de classe C^∞ definidas em \mathbb{R}^2 . Os vetores indicam a direção e sentido do vetor gradiente.

- (a) O ponto p pode ser extremo local de f no conjunto admissível D ? Em caso afirmativo, o ponto p seria máximo ou mínimo local? É possível garantir a globalidade?
- (b) O ponto q pode ser extremo local de f no conjunto admissível D ? Em caso afirmativo, o ponto q seria máximo ou mínimo local? É possível garantir a globalidade?

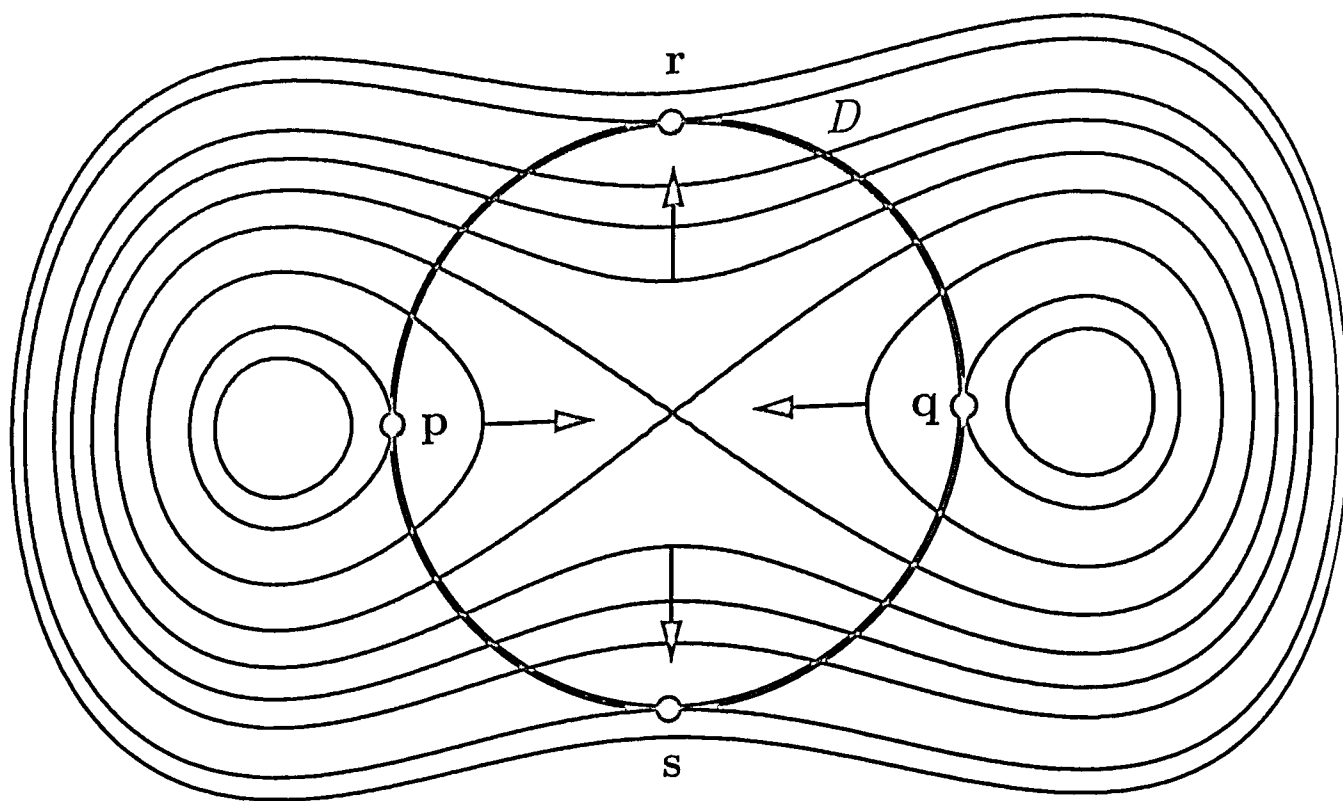


Figura 12.13: Desenho das curvas de nível de f e do conjunto admissível D (curva com traçado mais forte).

- (c) O ponto r pode ser extremo local de f no conjunto admissível D ? Em caso afirmativo, o ponto r seria máximo ou mínimo local? É possível garantir a globalidade?
- (d) O ponto s pode ser extremo local de f no conjunto admissível D ? Em caso afirmativo, o ponto s seria máximo ou mínimo local? É possível garantir a globalidade?
- [02] Use o teorema dos multiplicadores de Lagrange para encontrar os extremos globais de $f(x, y) = x + y$ no conjunto admissível

$$D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.$$

- [03] Sejam $f(x, y) = x^2 + y$ e

$$D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.$$

- (a) Use o teorema dos multiplicadores de Lagrange para encontrar os extremos globais de f no conjunto admissível D .
- (b) Mostre que o ponto $(\sqrt{2}/2, \sqrt{2}/2)$ não é um extremo global de f em D . Justifique sua resposta cuidadosamente.

- (c) Mostre que o ponto $(\sqrt{2}/2, \sqrt{2}/2)$ não é um extremo local de f em D . Justifique sua resposta cuidadosamente.
- (d) Faça um esboço de algumas curvas de nível de f e do conjunto admissível D (em um mesmo sistema cartesiano) para estabelecer uma interpretação geométrica das soluções obtidas no item (a). Use esta figura para determinar se o ponto $(0, 1)$ é máximo local, mínimo local ou nenhuma destas coisas para o problema de otimização.

- [04] Use o teorema dos multiplicadores de Lagrange para encontrar os extremos globais de $f(x, y) = x + y^2$ no conjunto admissível

$$D = \{(x, y) \in \mathbb{R}^2 \mid (x - 2)^2 + y^2 = 1\}.$$

- [05] Considere a função-objetivo $f(x_1, x_2, x_3) = x_1x_2 + x_2x_3 + x_1x_3$ e o conjunto admissível

$$D = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 3\}.$$

- (a) Mostre que $(1, 2, 0)$ não é extremo global de f em D . Justifique sua resposta cuidadosamente.
- (b) Mostre que $(1, 2, 0)$ não é extremo local de f em D . Justifique sua resposta cuidadosamente.
- (c) Mostre que todos os pontos do conjunto admissível satisfazem a condição de regularidade exigida pelo teorema dos multiplicadores de Lagrange.
- (d) Mostre que $(1, 1, 1)$ é o único ponto admissível que satisfaz as condições de primeira ordem para um valor apropriado do multiplicador de Lagrange.
- (e) Mostre que o conjunto admissível não é compacto. Justifique sua resposta cuidadosamente.
- (f) Mostre que f não possui mínimos globais em D . Justifique sua resposta cuidadosamente.

Como o conjunto admissível D não é compacto, não sabemos, *a priori*, se f possui máximos globais em D . Se este for o caso, pelos itens (c) e (d), o ponto de máximo global deve ser igual a $(1, 1, 1)$. Para demonstrar que, de fato, $(1, 1, 1)$ é ponto de máximo global de f em D , considere

ade
no
ada
no
ses
de
Vos
ios
res
S e

$$g(x_1, x_2) = f(x_1, x_2, 3 - x_1 - x_2) = x_1 x_2 + (x_1 + x_2)(3 - x_1 - x_2)$$

e os itens a seguir.

(g) Mostre que (x^*, y^*, z^*) , com $z^* = 3 - x^* - y^*$, é ponto de máximo global de f em D se, e somente se, (x^*, y^*) é ponto de máximo global de g em \mathbb{R}^2 .

(h) Use as técnicas do capítulo 11 para demonstrar que $(x^*, y^*) = (1, 1)$ é ponto de máximo global de g em \mathbb{R}^2 .

(i) Use os itens (g) e (h) para concluir que $(1, 1, 1)$ é ponto de máximo global de f em D .

[06] Encontre a distância máxima e mínima da origem $(0, 0)$ até a elipse $x^2 + xy + y^2 = 3$. Sugestão: use $x^2 + y^2$ no lugar de $\sqrt{x^2 + y^2}$ como função-objetivo. Veja o exercício [11] da página 361 para saber porque podemos fazer isto.

[07] Encontre o ponto (x^*, y^*) da parábola $y = x^2$ que está mais próximo do ponto $(2, 1)$. Você pode assumir que tal ponto existe.

[08] Calcule o posto das matrizes abaixo.

$$(a) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad (b) \begin{bmatrix} 2 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad (c) \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad (d) \begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & 1 & 5 & 1 \\ 0 & 2 & 3 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

[09] Considere a matriz

$$M = \begin{bmatrix} a & 0 & 0 \\ 1 & b & 0 \end{bmatrix}.$$

Para que valores de a e b a matriz M tem posto 2, posto 1 e posto 0?

[10] Verdadeira ou falsa? Se A é uma matriz $m \times n$, com $m > n$, então o posto de A é menor do que m . Apresente uma demonstração caso a sentença seja verdadeira ou um contra-exemplo caso ela seja falsa.

[11] Verdadeira ou falsa? Seja

$$A = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}_{2 \times 3}$$

uma matriz tal que $a \neq 0$. Se o posto de A é 1, então existe uma constante k tal que

$$(d, e, f) = k \cdot (a, b, c).$$

Apresente uma demonstração caso a sentença seja verdadeira ou um contra-exemplo caso ela seja falsa.

[12] Encontre o ponto (x^*, y^*, z^*) mais próximo da origem $(0, 0, 0)$ que está simultaneamente nos planos $3x - y + z = 5$ e $x + y + z = 1$. Você pode assumir que tal ponto existe.

[13] Encontre os extremos de $f(x, y, z) = x + y + z^2$ sujeitos às restrições $x^2 + y^2 + z^2 = 1$ e $y = 0$.

[14] Considere o conjunto restrição D definido por

$$\begin{cases} x_1 & \geq 0, \\ x_2 & \geq 0, \\ x_2 - (x_1 - 1)^2 & \leq 0. \end{cases}$$

Resolva as questões abaixo.

- (a) O ponto $(0, 5)$ é admissível? E o ponto $(1, 0)$?
- (b) Quais são as restrições que estão ativas no ponto $(7, 7)$? E no ponto $(5, 16)$? E no ponto $(7, 0)$? E no ponto $(1, 0)$?
- (c) Verifique que o ponto $(1, 0)$ não satisfaz a condição de regularidade exigida pelo teorema de Karush-Kuhn-Tucker.
- (d) Faça um esboço do conjunto admissível D .

[15] Considere o conjunto admissível D definido por

$$\begin{cases} x_1 & \geq 0, \\ x_1 & \leq 2, \\ x_2 & \geq 0, \\ x_2 & \leq 2, \\ (x_1 - 1)^2 + (y_1 - 1)^2 & \geq 1. \end{cases}$$

Resolva as questões abaixo.

- (a) O ponto $(1, 1)$ é admissível? E o ponto $(1, 0)$? E o ponto $(0, 0)$? E o ponto $(1 - \sqrt{2}/2, 1 - \sqrt{2}/2)$?

tade
 no
 cada
 no
 sses
 de
 Nos
 rios
 pres
 ES e

- (b) Quais são as restrições que estão ativas no ponto $(2, 2)$? E no ponto $(1, 2)$? E no ponto $(1, 0)$? E no ponto $(1/4, 1/4)$?
- (c) O ponto $(1, 0)$ satisfaz a condição de regularidade exigida pelo teorema de Karush-Kuhn-Tucker? E o ponto $(2, 0)$?
- (d) As restrições $x_1 \geq 0$, $x_1 \leq 2$, $x_2 \geq 0$, $x_2 \leq 2$ e $(x_1 - 1)^2 + (x_2 - 1)^2 \geq 1$ satisfazem a hipótese do teorema (12.10)?
- (e) Faça um esboço do conjunto admissível D . Ele é um conjunto compacto?

[16] Apresente uma prova geométrica para o fato de que o gradiente de f e o gradiente de g possuem o mesmo sentido em um ponto que minimiza $f(x, y)$ sujeito a $g(x, y) \geq b$.

[17] Associe as quatro sentenças a seguir

- (a) (x^*, y^*) é um mínimo de $f(x, y)$ sujeito à restrição $h(x, y) \leq c$ que está ativa neste ponto, que satisfaz a condição de regularidade $\nabla h(x^*, y^*) \neq (0, 0)$.
- (b) (x^*, y^*) é um mínimo de $f(x, y)$ sujeito à restrição $h(x, y) \geq c$ que está ativa neste ponto, que satisfaz a condição de regularidade $\nabla h(x^*, y^*) \neq (0, 0)$.
- (c) (x^*, y^*) é um máximo de $f(x, y)$ sujeito à restrição $h(x, y) \leq c$ que está ativa neste ponto, que satisfaz a condição de regularidade $\nabla h(x^*, y^*) \neq (0, 0)$.
- (d) (x^*, y^*) é um máximo de $f(x, y)$ sujeito à restrição $h(x, y) \geq c$ que está ativa neste ponto, que satisfaz a condição de regularidade $\nabla h(x^*, y^*) \neq (0, 0)$.

com as outras duas abaixo

(1) Existe $\lambda^* \geq 0$ tal que $\nabla f(x^*, y^*) = \lambda^* \cdot \nabla h(x^*, y^*)$.

(2) Existe $\lambda^* \leq 0$ tal que $\nabla f(x^*, y^*) = \lambda^* \cdot \nabla h(x^*, y^*)$.

Aqui f e h são funções de classe C^1 . Justifique sua resposta.

[18] Considere o problema de otimização que consiste em

maximizar	$x + 2y$
sujeito a	$x^2 + y^2 \leq 1,$
	$x \geq 0,$
	$y \geq 0.$

- (a) Verifique que todos os pontos do conjunto admissível do problema de otimização satisfazem a condição de regularidade exigida pelo teorema de Karush-Kuhn-Tucker.
- (b) Escreva o lagrangeano e as condições de primeira ordem (sistema lagrangeano) para o problema de otimização.
- (c) Encontre todos os pontos que satisfazem as condições de primeira ordem do problema de otimização.
- (d) O problema de otimização possui solução? Em caso afirmativo, encontre-a! Justifique cuidadosamente sua resposta.

[19] Considere o problema de otimização que consiste em

$$\begin{array}{ll} \text{maximizar} & x + y + 50z \\ \text{sujeito a} & z - x^2 - y^2 + 4 \geq 0, \\ & z + x^2 + y^2 - 4 \leq 0, \\ & x \geq 0, \\ & y \geq 0. \end{array}$$

- (a) Escreva as condições de primeira ordem (sistema lagrangeano) para o problema de otimização.
- (b) Quais restrições estão ativas no ponto $p = (0, 0, 4)$?
- (c) Mostre que $p = (0, 0, 4)$ não é solução do problema de otimização. Justifique cuidadosamente sua resposta.

[20] Considere o problema de otimização que consiste em

$$\begin{array}{ll} \text{maximizar} & x + y + z \\ \text{sujeito a} & z - x^2 - y^2 + 8 \geq 0, \\ & z + x^2 + y^2 \leq 0. \end{array}$$

- (a) Faça um esboço do conjunto admissível do problema de otimização.
- (b) Mostre que todos os pontos do conjunto admissível do problema de otimização satisfazem a condição de regularidade exigida pelo teorema de Karush-Kuhn-Tucker.
- (c) Escreva o lagrangeano e as condições de primeira ordem para o problema de otimização.
- (d) O problema de otimização possui solução? Em caso afirmativo, encontre-a! Justifique cuidadosamente sua resposta.

lade
no
ada
no
ses
de
Nos
rios
res
5 e

[21] Considere o problema de otimização que consiste em

$$\begin{array}{ll} \text{maximizar} & x + 2y + 4z \\ \text{sujeito a} & x^2 + y^2 + z^2 \leq 64, \\ & x^2 + y^2 + z^2 \geq 21, \\ & x \geq 0, \\ & y \geq 0, \\ & z \geq 0. \end{array}$$

- (a) Mostre que o problema de otimização não possui soluções no interior do conjunto admissível. Justifique sua resposta.
- (b) Mostre que o problema de otimização não possui soluções onde apenas a restrição $x^2 + y^2 + z^2 \geq 21$ está ativa.

[22] Considere o problema de otimização que consiste em

$$\begin{array}{ll} \text{maximizar} & 2x + y + 2z \\ \text{sujeito a} & x + 2y + z = 1, \\ & x^2 + y^2 \leq 1. \end{array}$$

- (a) Escreva o lagrangeano e as condições de primeira ordem para o problema de otimização.
- (b) Encontre todos os pontos que satisfazem as condições de primeira ordem do problema de otimização.
- (c) Mostre que todos os pontos do conjunto admissível do problema de otimização satisfazem a condição de regularidade exigida pelo teorema de Karush-Kuhn-Tucker.
- (d) O problema de otimização possui solução? Em caso afirmativo, encontre-a! Justifique cuidadosamente sua resposta.

[23] Considere o problema de otimização que consiste em

$$\begin{array}{ll} \text{maximizar} & x + 250y + 200z \\ \text{sujeito a} & x^2 + y^2 + z^2 \leq 3, \\ & x^2 + y^2 - z^2 \leq 1, \\ & x \geq 0, \\ & y \geq 0, \\ & z \geq 0. \end{array}$$

[24] C

(a)

(b)

(c)

(d)

(e)

(f)

[25] Verif
funç

sujei

Justi

[26] Maxi
 $y \geq 0$ [27] Maxim
 $y \geq 0$ [28] Maxim
 $x \geq 0$

Mostre que $\mathbf{p} = (0, \sqrt{2}, 1)$ não é um máximo local do problema de otimização. Justifique cuidadosamente sua resposta.

[24] Considere o problema de otimização que consiste em

maximizar	$x + y + z$
sujeito a	$x^2 + y^2 + z^2 = 3,$
	$x \geq 0,$
	$y \geq 0,$
	$z \geq 0.$

- (a) Faça um esboço do desenho do conjunto admissível e diga se ele é compacto ou não.
- (b) O problema de otimização admite solução? Justifique sua resposta!
- (c) Verifique a condição de regularidade exigida pelo teorema de Karush-Kuhn-Tucker.
- (d) Escreva as condições de primeira ordem deste problema de otimização.
- (e) Resolva o sistema obtido a partir das condições de primeira ordem. Especifique cuidadosamente os seus cálculos!
- (f) Verdadeira ou falsa? O ponto $(x, y, z) = (1, 1, 1)$ é a solução do problema de otimização. Justifique cuidadosamente sua resposta!

[25] Verdadeira ou falsa? O ponto $(x, y) = (1, 1)$ não é mínimo global da função

$$f(x, y) = x^2 - 2\sqrt{3}xy - y^2$$

sujeito à restrição

$$h(x, y) = x^2 + y^2 = 2.$$

Justifique cuidadosamente sua resposta.

[26] Maximize $f(x, y) = x^2 + y^2$ sujeito às restrições $2x + y \leq 2$, $x \geq 0$ e $y \geq 0$.

[27] Maximize $f(x, y) = 2y^2 - x$ sujeito às restrições $x^2 + y^2 \leq 1$, $x \geq 0$ e $y \geq 0$.

[28] Maximize $3xy - x^3$ sujeito às restrições $2x - y = -5$, $5x + 2y \geq 27$, $x \geq 0$ e $y \geq 0$.

dade
o no
cada
r no
sses
s de
Nos
ários
ores
ES e

[29] Minimize $x^2 - 2y$ sujeito às restrições $x^2 + y^2 \leq 1$, $x \geq 0$ e $y \geq 0$.

[30] Minimize $2x^2 + 2y^2 - 2xy - 9y$ sujeito a $4x + 3y \leq 10$, $y - 4x^2 \geq -2$, $x \geq 0$ e $y \geq 0$.

[31] Considere o problema de maximizar $f(x, y, z) = xyz + z$ sujeito a $x^2 + y^2 + z \leq 6$, $x \geq 0$, $y \geq 0$ e $z \geq 0$.

- (a) Escreva as condições de primeira ordem associadas a este problema.
- (b) Determine se a restrição $x^2 + y^2 + z \leq 6$ está ativa ou não na solução (ou soluções) do problema.
- (c) Encontre a solução das condições de primeira ordem que inclui o caso $x = 0$.
- (d) Encontre três equações nas variáveis x , y e z que devem ser satisfeitas se $x \neq 0$ na solução.
- (e) Mostre que $x = 1$, $y = 1$ e $z = 4$ satisfazem estas equações.

[32] Considere o problema de otimização que consiste em

$$\begin{array}{ll} \text{maximizar} & x^2 + y^2 + z^2 \\ \text{sujeito a} & x \geq 0, \\ & x \leq 1, \\ & y \geq 0, \\ & y \leq 1. \\ & z \geq 0, \\ & z \leq 1. \end{array}$$

- (a) Faça um esboço do desenho do conjunto admissível.
- (b) Verifique a condição de regularidade exigida pelo teorema de Karush-Kuhn-Tucker.
- (c) Escreva as condições de primeira ordem deste problema de otimização.
- (d) O problema de otimização possui solução? Em caso afirmativo, encontre-a! Justifique cuidadosamente sua resposta.

[33] Uma caixa de madeira sem tampa deve ser construída de forma a conter $\alpha \text{ cm}^3$, onde α é um número positivo dado. Ignorando-se a espessura da madeira, como a caixa deve ser construída a fim de se utilizar a menor quantidade de madeira (medida pela soma das áreas dos lados da caixa)?

[34] En
(x_1
 $p_1 x$

[35] Uti
do
c c

[36] Ma
pos
det

Est
siti
núm
núm
func

com

[37] Dete

no c
satis

onde

[34] Encontre a expressão geral (em termos de todos os parâmetros) do ponto (x_1, x_2) que maximiza a Cobb-Douglas $U(x_1, x_2) = kx_1^\alpha x_2^{1-\alpha}$ sujeito a $p_1x_1 + p_2x_2 = I$.

[35] Utilize multiplicadores de Lagrange para encontrar o ponto (x^*, y^*, z^*) do plano $ax + by + cz = d$ mais próximo da origem $(0, 0, 0)$, com a, b e c constantes não simultaneamente nulas.

[36] Maximize $x^2y^2z^2$ sujeito a $x^2 + y^2 + z^2 = c$, onde c é uma constante real positiva fixa. Qual é o valor máximo da função-objetivo no conjunto determinado pela restrição? Mostre que para todo $x, y, z \in \mathbb{R}$ ocorre

$$\sqrt[3]{x^2 \cdot y^2 \cdot z^2} \leq \frac{x^2 + y^2 + z^2}{3}.$$

Esta desigualdade afirma que a *média geométrica* de três números positivos é sempre menor ou igual do que a *média aritmética* destes três números. Mais ainda, estas médias são iguais se, e somente se, os três números $(x^2, y^2$ e $z^2)$ são iguais. Naturalmente, a mesma demonstração funciona para um conjunto de n números positivos:

$$\sqrt[n]{x_1^2 \cdot x_2^2 \cdot \dots \cdot x_n^2} \leq \frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}$$

com igualdade se, e somente se, $x_1^2 = x_2^2 = \dots = x_n^2$.

[37] Determine o valor máximo da função

$$z = f(x_1, x_2, \dots, x_n) = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$$

no conjunto admissível formado pelos pontos (x_1, x_2, \dots, x_n) em \mathbb{R}^n que satisfazem as restrições

$$\begin{cases} x_1 + x_2 + \dots + x_n = c, \\ x_1 \geq 0, \\ x_2 \geq 0, \\ \vdots \\ x_n \geq 0, \end{cases}$$

onde c é uma constante real. Deduza então a desigualdade

$$\sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} \leq \frac{x_1 + x_2 + \dots + x_n}{n},$$

idade
no
cada
no
sses
de
Nos
rios
ores
ES e

entre as médias geométrica e aritmética dos números x_1, x_2, \dots, x_n . Compare com o exercício anterior.

- [38] Verdadeira ou falsa? Sejam $f: \mathbb{R}^n \rightarrow \mathbb{R}$ e $h: \mathbb{R}^n \rightarrow \mathbb{R}$ funções de classe C^∞ . Suponha que o sistema associado às condições de primeira ordem

$$\begin{cases} \nabla f(\mathbf{x}) = \lambda \cdot \nabla h(\mathbf{x}), \\ h(\mathbf{x}) = 0, \end{cases}$$

possua uma única solução $(\mathbf{x}^*, \lambda^*)$. Então \mathbf{x}^* é um extremo global de f no conjunto admissível

$$D = \{\mathbf{x} \in \mathbb{R}^n \mid h(\mathbf{x}) = 0\}.$$

Apresente uma demonstração caso a sentença seja verdadeira ou um contra-exemplo caso ela seja falsa.

- [39] Diga se cada uma das sentenças abaixo é verdadeira ou falsa, apresentando uma demonstração caso ela seja verdadeira ou um contra-exemplo caso ela seja falsa.

- (a) Sejam $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ e $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ funções de classe C^∞ . Suponha que o sistema

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = \lambda \cdot \frac{\partial h}{\partial x}(x, y), \\ \frac{\partial f}{\partial y}(x, y) = \lambda \cdot \frac{\partial h}{\partial y}(x, y), \\ h(x, y) = 0, \end{cases}$$

não possua soluções. Então f não possui extremos globais no conjunto

$$C = \{(x, y) \in \mathbb{R}^2 \mid h(x, y) = 0\}.$$

- (b) Sejam $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ e $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ funções de classe C^∞ . Suponha que o sistema

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = \lambda \cdot \frac{\partial h}{\partial x}(x, y), \\ \frac{\partial f}{\partial y}(x, y) = \lambda \cdot \frac{\partial h}{\partial y}(x, y), \\ h(x, y) = 0, \end{cases}$$

(c)

[40] Seja
p =Most
remaEm c
pode
(x*, y*)

[41] Cons

não possua soluções e que

$$\nabla h(x, y) \neq (0, 0) \text{ para todo } (x, y) \in \mathbb{R}^2.$$

Então f não possui extremos globais no conjunto

$$C = \{(x, y) \in \mathbb{R}^2 \mid h(x, y) = 0\}.$$

(c) Sejam $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ e $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ funções de classe C^∞ . Suponha que o sistema

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = \lambda \cdot \frac{\partial h}{\partial x}(x, y), \\ \frac{\partial f}{\partial y}(x, y) = \lambda \cdot \frac{\partial h}{\partial y}(x, y), \\ h(x, y) = 0, \end{cases}$$

possua uma única solução (x^*, y^*, λ^*) com

$$\nabla h(x^*, y^*) \neq (0, 0).$$

Então (x^*, y^*) é um extremo local de f no conjunto

$$C = \{(x, y) \in \mathbb{R}^2 \mid h(x, y) = 0\}.$$

[40] Sejam $f: \mathbb{R}^n \rightarrow \mathbb{R}$ e $h: \mathbb{R}^n \rightarrow \mathbb{R}$ funções de classe C^∞ . Suponha que $\mathbf{p} = (\mathbf{x}^*, \lambda_1^*)$ e $\mathbf{q} = (\mathbf{x}^*, \lambda_2^*)$ satisfaçam as condições de primeira ordem

$$\begin{cases} \nabla f(\mathbf{x}) = \lambda \cdot \nabla h(\mathbf{x}), \\ h(\mathbf{x}) = 0. \end{cases}$$

Mostre que se \mathbf{x}^* satisfaz a condição de regularidade exigida pelo teorema dos multiplicadores de Lagrange, então $\lambda_1^* = \lambda_2^*$.

Em outras palavras, se \mathbf{x}^* satisfaz a condição de regularidade, então não podem existir dois multiplicadores λ_1^* e λ_2^* diferentes tais que $(\mathbf{x}^*, \lambda_1^*)$ e $(\mathbf{x}^*, \lambda_2^*)$ satisfaçam as condições de primeira ordem.

[41] Considere

problema de otimização 1:

$$\begin{array}{l} \text{Maximizar } f(x, y) \\ \text{sujeito a } h_1(x, y) = 0. \end{array}$$

problema de otimização 2:

$$\begin{array}{l} \text{Maximizar } f(x, y) \\ \text{sujeito a } h_2(x, y) = 0. \end{array}$$

tade
o no
cada
r no
sses
i de
Nos
rios
ores
ES e

onde $f(x, y) = x + y$, $h_1(x, y) = x^2 + y^2 - 1$ e $h_2(x, y) = (x^2 + y^2 - 1)^2$.

- (a) Mostre que todos os pontos do conjunto admissível do problema de otimização 1 satisfazem a condição de regularidade exigida pelo teorema dos multiplicadores de Lagrange.
- (b) Mostre que todos os pontos do conjunto admissível do problema de otimização 2 *não* satisfazem a condição de regularidade exigida pelo teorema dos multiplicadores de Lagrange.
- (c) Verdadeira ou falsa? Um ponto $p \in \mathbb{R}^2$ é solução do problema de otimização 1 se, e somente se, $p \in \mathbb{R}^2$ é solução do problema de otimização 2. Justifique cuidadosamente sua resposta.
- (d) Escreva as condições de primeira ordem (sistema lagrangeano) para o problema de otimização 2 e mostre que o sistema resultante não possui soluções.
- (e) Encontre todos os pontos que resolvem o problema de otimização 2. Justifique cuidadosamente sua resposta.

[42] Verdadeira ou falsa? Sejam $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ e $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ funções de classe C^∞ . Considere o conjunto

$$D = \{(x, y) \in \mathbb{R}^2 \mid h(x, y) = 0\}.$$

Se todo ponto $(x, y) \in D$ satisfaz $\nabla h(x, y) = (0, 0)$, isto é, se todo ponto de D não satisfaz a condição de regularidade, então f *não* possui extremos globais em D . Apresente uma justificativa caso a sentença seja verdadeira ou um contra-exemplo caso ela seja falsa.

[43] Verdadeira ou falsa? Se $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ e $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ são funções de classe C^∞ e o conjunto

$$C = \{(x, y) \in \mathbb{R}^2 \mid h(x, y) = 0\}$$

não é compacto, então f não possui extremos globais em C . Justifique cuidadosamente sua resposta, apresentando uma demonstração caso a sentença seja verdadeira ou um contra-exemplo caso ela seja falsa.

[44] Diga se cada uma das sentenças abaixo é verdadeira ou falsa, apresentando uma demonstração caso ela seja verdadeira ou um contra-exemplo caso ela seja falsa.

- (a) Se o conjunto de pontos (x, y) que satisfazem a desigualdade

$$g(x, y) \leq 1$$

é um conjunto compacto, então o conjunto de pontos que satisfazem a desigualdade

$$-g(x, y) \geq -1$$

não é compacto. Aqui g é uma função de classe C^∞ definida em \mathbb{R}^2 .

- (b) Sejam $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ e $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ funções de classe C^∞ . Suponha que o sistema

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

não possua soluções. Então f não possui extremos globais no conjunto $C = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 0\}$.

- (c) Sejam $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ e $g: \mathbb{R}^3 \rightarrow \mathbb{R}$ funções de classe C^∞ . Suponha que o sistema

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

possua uma única solução $(x, y, z, \lambda) = (a^*, b^*, c^*, \lambda^*)$ para a qual

$$\nabla g(a^*, b^*, c^*) \neq (0, 0, 0).$$

Então $(x, y, z) = (a^*, b^*, c^*)$ é um extremo global de $w = f(x, y, z)$ sujeito à restrição $g(x, y, z) = 0$.

[45] Resolva o problema de otimização, que consiste em

$$\begin{aligned} \text{maximizar } f(x_1, \dots, x_n, y_1, \dots, y_n) &= \sum_{i=1}^n x_i \cdot y_i \\ \text{sujeito a } h_1(x_1, \dots, x_n, y_1, \dots, y_n) &= \sum_{i=1}^n x_i^2 = 1, \\ h_2(x_1, \dots, x_n, y_1, \dots, y_n) &= \sum_{i=1}^n y_i^2 = 1. \end{aligned}$$

A partir das soluções obtidas, deduza a desigualdade de Cauchy-Schwarz

$$\sum_{i=1}^n a_i \cdot b_i \leq \sqrt{\sum_{i=1}^n a_i^2} \cdot \sqrt{\sum_{i=1}^n b_i^2},$$

tade
no
cada
no
sses
i de
Nos
rios
ores
ES e

escrevendo $x_i = a_i / \sum_{i=1}^n a_i^2$ e $y_i = b_i / \sum_{i=1}^n b_i^2$. Compare com as provas sugeridas no exercício [46] da página 78 e no exercício [25] da página 158.

[46] Se a, b e c são constantes positivas, ache os pontos (x, y, z) do elipsóide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

que estão mais perto da origem $0 = (0, 0, 0)$.

[47] (O significado do multiplicador de Lagrange) Considere o problema de otimização que consiste em

Maximizar $f(x, y) = x + y$
 sujeito a $h(x, y) = x^2 + y^2 = a$.

- (a) Resolva o problema de otimização para o caso $a^* = 1$. Mostre que, neste caso, o valor máximo de f é $z^* = \sqrt{2}$ e que o multiplicador de Lagrange associado ao ponto de máximo é $\lambda^* = \sqrt{2}/2$.
- (b) Resolva o problema de otimização para o caso $a > 0$.
- (c) Note que, para cada $a > 0$, você encontrará um ponto de máximo que depende de a e, conseqüentemente, um valor máximo de f no conjunto admissível que também depende de a . Mostre que

$$z = g(a) = \sqrt{2} \sqrt{a}$$

é a função que estabelece a dependência do valor máximo de f no conjunto admissível em função do parâmetro a .

(d) Mostre que

$$\frac{dg}{da}(a^*) = \lambda^*$$

O que os cálculos acima mostram é a evidência de um resultado que vale em geral.

Teorema 12.12 Sejam f e h funções de classe C^∞ . Para um dado valor do parâmetro a , seja $(x^*(a), y^*(a))$ a solução do problema de otimização

Maximizar $f(x, y)$
 sujeito a $h(x, y) = a$,

e seja λ^* o multiplicador de Lagrange correspondente. Suponha ainda que $x^* = x^*(a)$, $y^* = y^*(a)$ e $\lambda^* = \lambda^*(a)$ sejam funções de classe C^1 tais que $(x^*(a), y^*(a))$ satisfaça a condição de regularidade para cada valor de a . Então

$$\lambda^*(a) = \frac{d}{da} f(x^*(a), y^*(a)).$$

Desta maneira, o multiplicador mede a *sensitividade* do valor ótimo da função objetivo com relação a variações de um parâmetro que determina o conjunto admissível. Sendo assim, ele fornece uma medida natural do valor de recursos escassos em problemas econômicos de maximização. Evidentemente, o teorema acima pode ser generalizado para o caso de várias restrições dependendo de várias variáveis. Veja a referência [71]. Economistas chamam $\lambda^*(a)$ de *preço sombra* do parâmetro a .

idade
o não
cada
r no
esses
s de
Nos
ários
ores
ES e