Capítulo 4

Continuidade, noções de topologia e o teorema de Weierstrass

No capítulo anterior estudamos o conceito de função de várias variáveis, que é o objeto matemático por excelência para se modelar problemas. Vimos também as diversas maneiras de representá-lo geometricamente: seja através de seu gráfico ou seja através de suas hiperfícies de nível. Com estes recursos geométricos vamos estudar uma primeira propriedade interessante que funções podem ter: continuidade.

4.1 Por que funções contínuas são importantes?

O conceito de continuidade não deve ser novo para você pois você já deve tê-lo estudado em Cálculo I. De qualquer modo, vamos fazer uma revisão do assunto.

Sejam $f\colon D \subset \mathbb{R} \to \mathbb{R}$ uma função de uma variável e $p \in D$ um ponto do domínio D de f. A função f associa p no domínio a $f(p)$ na imagem. Considere agora $x \in D$ uma aproximação de p. O que podemos dizer a respeito de $f(x)$ (*o valor de* f *na aproximação*) e $f(p)$ (*o valor de* f *no ponto*)? Para uma função contínua é possível "amarrar" estes quatro valores: se f é *contínua em* p, então $f(x)$ pode ser considerado uma aproximação de $f(p)$, tão boa quanto se queira, desde que x esteja suficientemente próximo de p. Em outras palavras, para garantir que $f(x)$ esteja suficientemente próximo de $f(p)$ (na imagem) basta tomar x suficientemente próximo de p (no domínio).

Como formalizar matematicamente este conceito? Através de seqüências numéricas: construir aproximações de p é construir seqüências x_n que convergem para p e, para funções contínuas em p, espera-se que $f(x_n)$ convirja para $f(p)$ (veja figura (4.1)). Mais formalmente:

Definição 4.1 (Continuidade) Seja $f: D \subset \mathbb{R} \to \mathbb{R}$. Dizemos que f é *contínua* em um ponto p de seu domínio D se *para toda* seqüência

$$x_n \to p$$

no domínio, tem-se

$$f(x_n) \to f(p)$$

na imagem. Se f não é contínua em $p \in D$, então dizemos que f é *descontínua* em p. Se f é contínua em *todos* os pontos de seu domínio, então dizemos simplesmente que f é contínua.

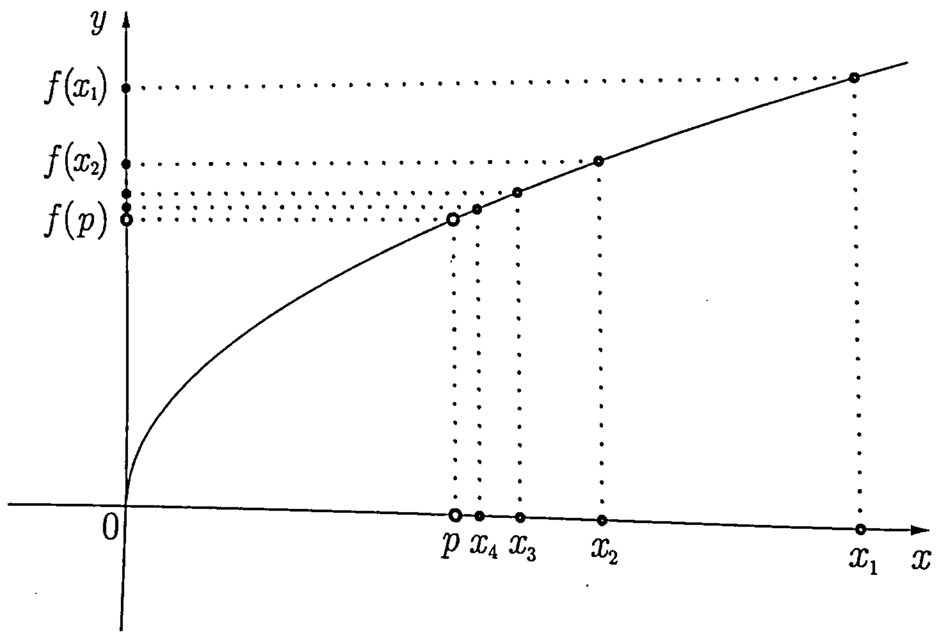

Figura 4.1: Ilustração do conceito de continuidade em 1 variável através do gráfico de f.

Mas por que funções contínuas, isto é, funções que satisfazem a definição acima, são importantes? Vamos citar duas propriedades de interesse.

Propriedade 1:

O valor de uma função contínua em um ponto p é "estável" por pequenas variações em p. Para exemplificar, considere a seguinte situação: você está modelando matematicamente um determinado fenômeno (digamos, o quanto uma ponte suspensa é resistente ao vento) e é necessário calcular a função f resultante em pontos da forma $2 \cdot k \cdot \pi$, com $k = 0, 1, \ldots, 99$. Se você está utilizando um computador para fazer os cálculos, certamente você não conseguirá calcular a função exatamente nestes pontos pois, sendo o computador uma máquina com precisão finita, ele não conseguirá representar o número real π que possui infinitas casas decimais não-periódicas. O melhor que se consegue fazer é calcular a função f em aproximações destes pontos, digamos, em $6.28 \cdot k$ ou $6.2831 \cdot k$ (isto sem considerar os erros de arredondamento efetuados durante os cálculos).

Você passaria por uma ponte suspensa que foi projetada utilizando aproximações e não os valores originais? Se a função que descreve o fenômeno é contínua, então você pode ficar mais tranqüilo pois você sabe que, para pontos suficientemente próximos de $2 \cdot k \cdot \pi$, o valor da função nestes pontos estará tão próximo do valor correto $f(2 \cdot k \cdot \pi)$ quanto se queira. Basta então escolher uma tolerância e prosseguir com os cálculos.

Propriedade 2:

Nem todo problema de otimização possui uma solução, isto é, nem toda função definida em um conjunto possui um máximo global ou um mínimo global neste conjunto. Uma importante pergunta é: como saber, *a priori*, se um problema de otimização possui ou não uma solução? Um teorema que resolve esta questão envolve o conceito de continuidade.

Teorema 4.1 (WEIERSTRASS) Toda função $f: [a, b] \subset \mathbb{R} \to \mathbb{R}$ *contínua* definida no *intervalo fechado e limitado* $[a, b]$ possui pelo menos um máximo global e pelo menos um mínimo global em $[a, b]$.

Observe que o teorema exige duas coisas: (1) a continuidade da função f e, (2), que ela esteja definida em um intervalo fechado e limitado da forma $[a, b]$. Se uma ou as duas destas condições não se verificarem, nada se pode garantir. Faça algumas figuras e tente pensar em exemplos e contra-exemplos!

4.2 Continuidade em várias variáveis

A idéia é a mesma! Vamos começar com duas variáveis: considere uma função $f: D \subset \mathbb{R}^2 \to \mathbb{R}$ e (a, b) um ponto do domínio D de f. A função associa (a,b) no domínio a $f(a,b)$ na imagem. Considere agora (x,y) uma aproximação de (a,b), isto é, x é uma aproximação de a e y é uma aproximação de b. O que podemos dizer a respeito de $f(x,y)$ (*o valor de f calculado na aproximação*) e $f(a,b)$ (*o valor de f no ponto*)? Como antes, para uma função contínua em (a,b), vale que $f(x,y)$ pode ser considerado uma aproximação de $f(a,b)$, tão boa quanto se queira, desde que (x,y) esteja suficientemente próximo de (a,b), isto é, desde que x esteja suficientemente próximo de a e y esteja suficientemente próximo de b (veja a figura (4.2)). A definição formal, como antes, faz o uso de seqüências.

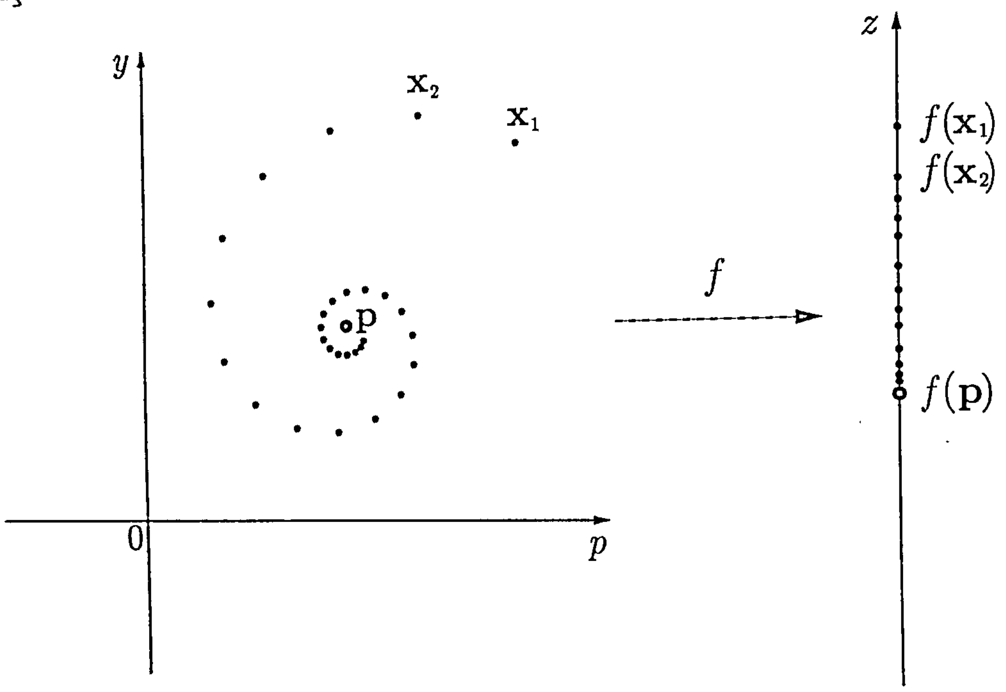

Figura 4.2: Ilustração do conceito de continuidade em 2 variáveis através do domínio/contradomínio de f.

Definição 4.2 (Continuidade) Seja $f: D \subset \mathbb{R}^2 \to \mathbb{R}$. Dizemos que f é *contínua* em um ponto $\mathrm{p} = (a,b)$ de seu domínio D se *para toda* seqüência

$$\mathrm{x}_n = (x_n, y_n) \to (a,b) = \mathrm{p}$$

no domínio, isto é, se *para todo* par de seqüências numéricas

$$x_n \to a \quad \text{e} \quad y_n \to b,$$

com $(x_n, y_n) \in D$, tem-se

$$f(x_n, y_n) \to f(a, b)$$

na imagem. Se f não é contínua em $\mathbf{p} \in D$, então dizemos que f é *descontínua* em $\mathbf{p}$. Se f é contínua em *todos* os pontos de seu domínio, então dizemos simplesmente que f é contínua. Note que $z_n = f(x_n, y_n)$ é uma seqüência de números reais!

Talvez a melhor maneira de entender como uma função contínua se parece é visualizar o gráfico de uma função que *não é contínua.*

Exemplo 4.1 A função

$$z = f(x, y) = \begin{cases} 2 & \text{se } y > 0, \\ 1 & \text{se } y \leq 0, \end{cases}$$

não é contínua em $\mathbf{p} = (0, 0)$ pois a seqüência

$$\mathbf{x}_n = (x_n, y_n) = \left(0, \frac{1}{n}\right)$$

converge para $\mathbf{p} = (0, 0)$ mas

$$f(\mathbf{x}_n) = f(x_n, y_n) = f\left(0, \frac{1}{n}\right) = 2$$

converge para 2, que é diferente de $1 = f(\mathbf{p})$ (veja a figura (4.3)). Em verdade, f é descontínua em todos os pontos da forma $(a, 0)$ (o eixo x) e contínua nos demais pontos. □

CUIDADO! CUIDADO! CUIDADO!

Para mostrar que uma função é contínua em um ponto (a, b), é preciso provar que para *todas* as seqüências (x_n, y_n) que convergem para (a, b), tem-se $f(x_n, y_n)$ convergente para $f(a, b)$. Estabelecer a convergência para uma ou duas seqüências particulares não é suficiente para demonstrar a continuidade da função no ponto. No exemplo anterior, a seqüência $(x_n, y_n) = (1/n, 0)$ converge para $(0, 0)$ e $f(x_n, y_n) = f(1/n, 0) = 1$ converge para $1 = f(0, 0)$ mas, mesmo assim, f é descontínua em $(0, 0)$.

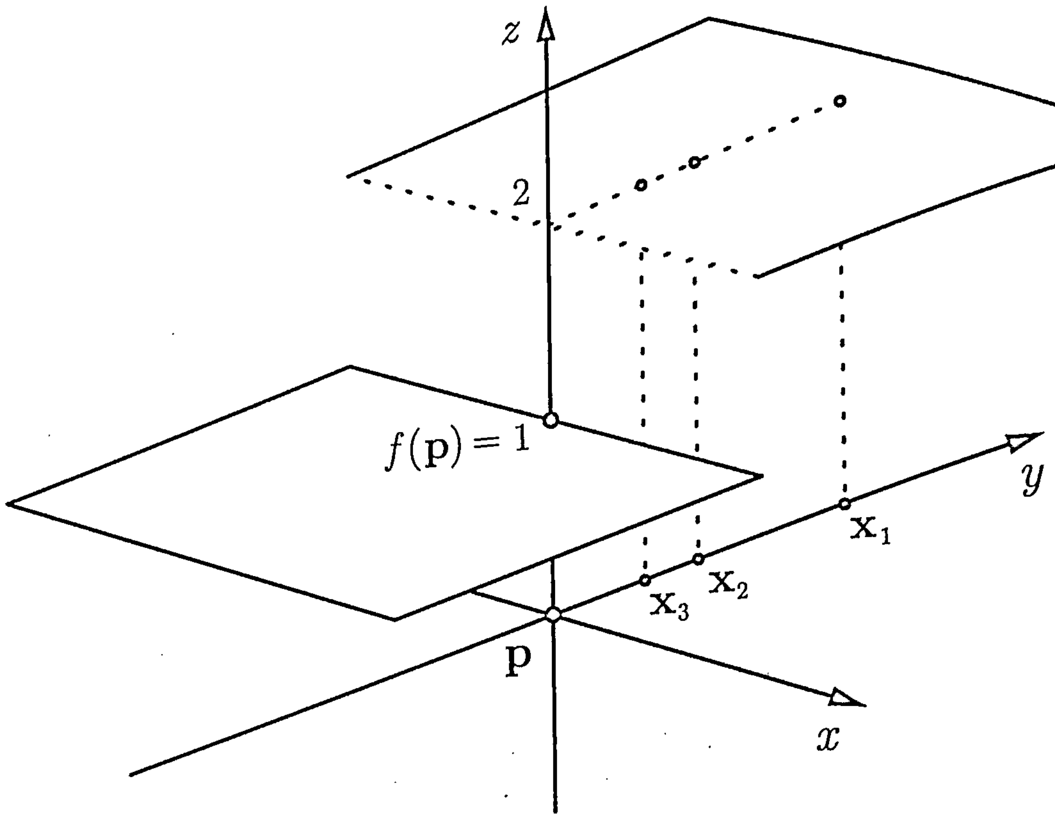

Figura 4.3: O gráfico de uma função descontínua.

Exercício resolvido 4.1 A função

$$f(x,y)=\begin{cases}\dfrac{x\cdot y}{x^2+y^2} & \text{se } (x,y)\neq(0,0),\\ 1/2 & \text{se } (x,y)=(0,0),\end{cases}$$

é contínua em $(0,0)$?

Solução: Inicialmente, observe que

$$(x_n,y_n)=\left(\frac{1}{n},\frac{1}{n}\right)\to(0,0)$$

e

$$f(x_n,y_n)=\frac{x_n\cdot y_n}{x_n^2+y_n^2}=\frac{\frac{1}{n}\cdot\frac{1}{n}}{\left(\frac{1}{n}\right)^2+\left(\frac{1}{n}\right)^2}=\frac{\frac{1}{n^2}}{\frac{2}{n^2}}=\frac{1}{2}\to\frac{1}{2}=f(0,0).$$

Contudo, isto não garante que f seja contínua em $(0,0)$ pois pode existir uma outra seqüência (x_n,y_n) que converge para $(0,0)$ mas com $f(x_n,y_n)$ não convergindo para $1/2=f(0,0)$. De fato, tal seqüência existe pois

$$(x_n,y_n)=\left(\frac{1}{n},\frac{2}{n}\right)\to(0,0)$$

e

$$f(x_n, y_n) = \frac{x_n \cdot y_n}{x_n^2 + y_n^2} = \frac{\frac{1}{n} \cdot \frac{2}{n}}{\left(\frac{1}{n}\right)^2 + \left(\frac{2}{n}\right)^2} = \frac{\frac{2}{n^2}}{\frac{5}{n^2}} = \frac{2}{5} \to \frac{2}{5} \neq \frac{1}{2} = f(0,0).$$

Sendo assim, f *não é* contínua em $(0,0)$. Observe que os pontos da seqüência $(1/n, 1/n)$ (que converge para $(0,0)$) estão sobre a reta $y = x$ e que os pontos da seqüência $(1/n, 2/n)$ (que também converge para $(0,0)$) estão sobre a reta $y = 2x$. □

Vamos ver agora exemplos de como provar que funções são contínuas.

Exercício resolvido 4.2 Mostre que a função $z = f(x,y) = x$ é contínua no ponto $(2,1)$.

SOLUÇÃO: Dada uma seqüência qualquer (x_n, y_n) que converge para $(2,1)$, devemos mostrar que $f(x_n, y_n)$ converge para $f(2,1) = 2$. Como $(x_n, y_n) \to (2,1)$, sabemos que

$$x_n \to 2$$

e $y_n \to 1$. Então,

$$f(x_n, y_n) = x_n \to 2 = f(2,1).$$ □

Exercício resolvido 4.3 Mostre que a função $z = f(x,y) = x$ é contínua.

SOLUÇÃO: Devemos mostrar que f é contínua em todos os pontos (a,b) de $\mathbb{R}^2$. Seja então (x_n, y_n) uma seqüência qualquer que converge para (a,b). Devemos mostrar que $f(x_n, y_n)$ converge para $f(a,b)$. Como $(x_n, y_n) \to (a,b)$ segue-se que

$$x_n \to a$$

e $y_n \to b$. Portanto,

$$f(x_n, y_n) = x_n \to a = f(a,b).$$ □

A definição de continuidade se estende naturalmente para funções que dependem de três ou mais variáveis:

Definição 4.3 (Continuidade) Seja $f\colon D \subset \mathbb{R}^m \to \mathbb{R}$. Dizemos que f é contínua em um ponto $\mathrm{p} = (p_1, p_2, \ldots, p_m)$ de seu domínio D se *para toda* seqüência

$$\mathrm{x}_n = (x_{1,n}, x_{2,n}, \ldots, x_{m,n}) \to (p_1, p_2, \ldots, p_m) = \mathrm{p}$$

no domínio, isto é, *para toda* coleção de m seqüências numéricas

$$x_{1,n} \to p_1, \quad x_{2,n} \to p_2, \quad \ldots, \quad x_{m,n} \to p_m,$$

com $(x_{1,n}, x_{2,n}, \ldots, x_{m,n}) \in D$, tem-se

$$f(\mathrm{x}_n) = f(x_{1,n}, x_{2,n}, \ldots, x_{m,n}) \to f(p_1, p_2, \ldots, p_m) = f(\mathrm{p})$$

na imagem. Se f não é contínua em $\mathrm{p} \in D$, então dizemos que f é *descontínua* em p. Se f é contínua em *todos* os pontos de seu domínio, então dizemos simplesmente que f é contínua. Finalmente, observe que $z_n = f(x_{1,n}, x_{2,n}, \ldots, x_{m,n})$ é uma seqüência de números reais!

O teorema seguinte estabelece uma maneira fácil de se criar e de se identificar funções contínuas a partir de outras funções contínuas.

Teorema 4.2 Sejam $f\colon D_f \subset \mathbb{R}^m \to \mathbb{R}$ e $g\colon D_g \subset \mathbb{R}^m \to \mathbb{R}$ funções contínuas em $\mathrm{p} \in D_f \cap D_g$. Então

$$f + g, \quad f - g \quad \text{e} \quad f \cdot g$$

são contínuas em p. Se $g(\mathrm{p}) \neq 0$, então

$$f/g$$

é contínua em p. Mais ainda se $h\colon D_h \subset \mathbb{R} \to \mathbb{R}$ é contínua em $f(\mathrm{p}) \in D_h$, então

$$h \circ f$$

é contínua em p.

Demonstração: A demonstração deste teorema é uma aplicação direta de um teorema análogo para seqüências que você já estudou em Cálculo I. Para referência, vamos provar aqui que a soma de duas funções contínuas também é uma função contínua. Os demais casos ($f - g$, $f \cdot g$, f/g e $h \circ f$) são demonstrados analogamente e vamos deixá-los como exercício.

Seja x_n uma seqüência convergindo para $\mathrm{p} \in D_f \cap D_g$. Então, pela continuidade de f e g, temos

$$f(\mathrm{x}_n) \to f(\mathrm{p}) \qquad \text{e} \qquad g(\mathrm{x}_n) \to g(\mathrm{p}).$$

Em Cálculo I você aprendeu que a soma de seqüências numéricas convergentes converge para a soma de seus respectivos limites e, como $f(\mathrm{x}_n)$ e $g(\mathrm{x}_n)$ são seqüências numéricas (isto é, seqüências em $\mathbb{R}$) que convergem respectivamente para $f(\mathrm{p})$ e $g(\mathrm{q})$, segue-se que

$$(f+g)(\mathrm{x}_n) = f(\mathrm{x}_n) + g(\mathrm{x}_n) \to f(\mathrm{p}) + g(\mathrm{p}) = (f+g)(\mathrm{p}).$$

Isto mostra que $f + g$ é contínua em p. □

Com este teorema é possível justificar, por exemplo, a continuidade da função

$$f(x, y) = \operatorname{sen}\left(x^2 + y^2\right).$$

De fato: x e y são funções contínuas (veja o exercício resolvido (4.3) e o exercício [01]). Logo, x^2 e y^2 são funções contínuas como produto de funções contínuas. Conseqüentemente, $x^2 + y^2$ é contínua como soma de funções contínuas. Finalmente, como a função seno é contínua, segue-se que $f(x, y) = \operatorname{sen}\left(x^2 + y^2\right)$ é contínua como composição de funções contínuas.

4.3 O teorema de Weierstrass no caso de n variáveis

Como vimos na seção (4.1), funções *contínuas* de uma variável definidas em um *intervalo fechado e limitado* $[a, b]$ possuem pelo menos um máximo global e pelo menos um mínimo global em $[a, b]$. Uma pergunta natural é se existe um resultado parecido para funções de duas ou mais variáveis. A resposta é sim! É de se esperar que a continuidade da função ainda seja exigida mas o que faria o papel de um intervalo fechado e limitado $[a, b]$ no caso de n variáveis?

Para responder a esta pergunta, vamos estudar várias situações e tentar intuir as propriedades que um conjunto admissível D deveria possuir a fim de que uma função contínua definida em D possua pelo menos um máximo global e pelo menos um mínimo global em D.

Considere

$$\begin{array}{rrcl} f: & D \subset \mathbb{R}^2 & \to & \mathbb{R} \\ & (x,y) & \mapsto & z = f(x,y) = x^2 + y^2 \end{array},$$

onde D pode ser um dos seis conjuntos abaixo.

(1) $D = \mathbb{R}^2$,

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

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

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

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

(6) $D = \{(x,y) \in \mathbb{R}^2 \mid 1 < x^2 + y^2 < 4\}$.

Pergunta: em quais dos seis casos a função-objetivo f possui pelo menos um máximo global e pelo menos um mínimo global no conjunto admissível D?

Observe que estamos usando uma mesma expressão algébrica para a função-objetivo f e mudando apenas o conjunto admissível D.

O desenho dos vários conjuntos admissíveis e respectivos gráficos de $z = f(x,y) = x^2 + y^2$ encontram-se nas figuras (4.4), (4.5), (4.6), (4.7), (4.8), (4.9) e (4.10). Note que as figuras (4.5) e (4.9) representam apenas uma parte do gráfico de f.

No caso (1), f possui um único ponto de mínimo global, $(0,0)$, mas f não possui pontos de máximo global em $D = \mathbb{R}^2$. De fato, a partir dos cálculos feitos no exemplo (3.1) da página 81, é fácil de ver que as curvas de nível (não-vazias) de f são circunferências de centro na origem. Se nos aproximamos de $(0,0)$, então o valor da função diminui e quando nos afastamos de $(0,0)$, o valor da função aumenta (nos pontos da circunferência de centro em $(0,0)$ e raio r, o valor da função é r^2). Como o ponto mais próximo de $(0,0)$ é o próprio ponto $(0,0)$, segue-se que $(0,0)$ é ponto de mínimo global de f em $\mathbb{R}^2$. Por outro lado, como podemos obter pontos em $D = \mathbb{R}^2$ *arbitrariamente distantes* de $(0,0)$, isto é, pontos em $D = \mathbb{R}^2$ para os quais o valor de f é tão grande quanto se queira, concluímos que f não possui pontos de máximo global em $D = \mathbb{R}^2$.

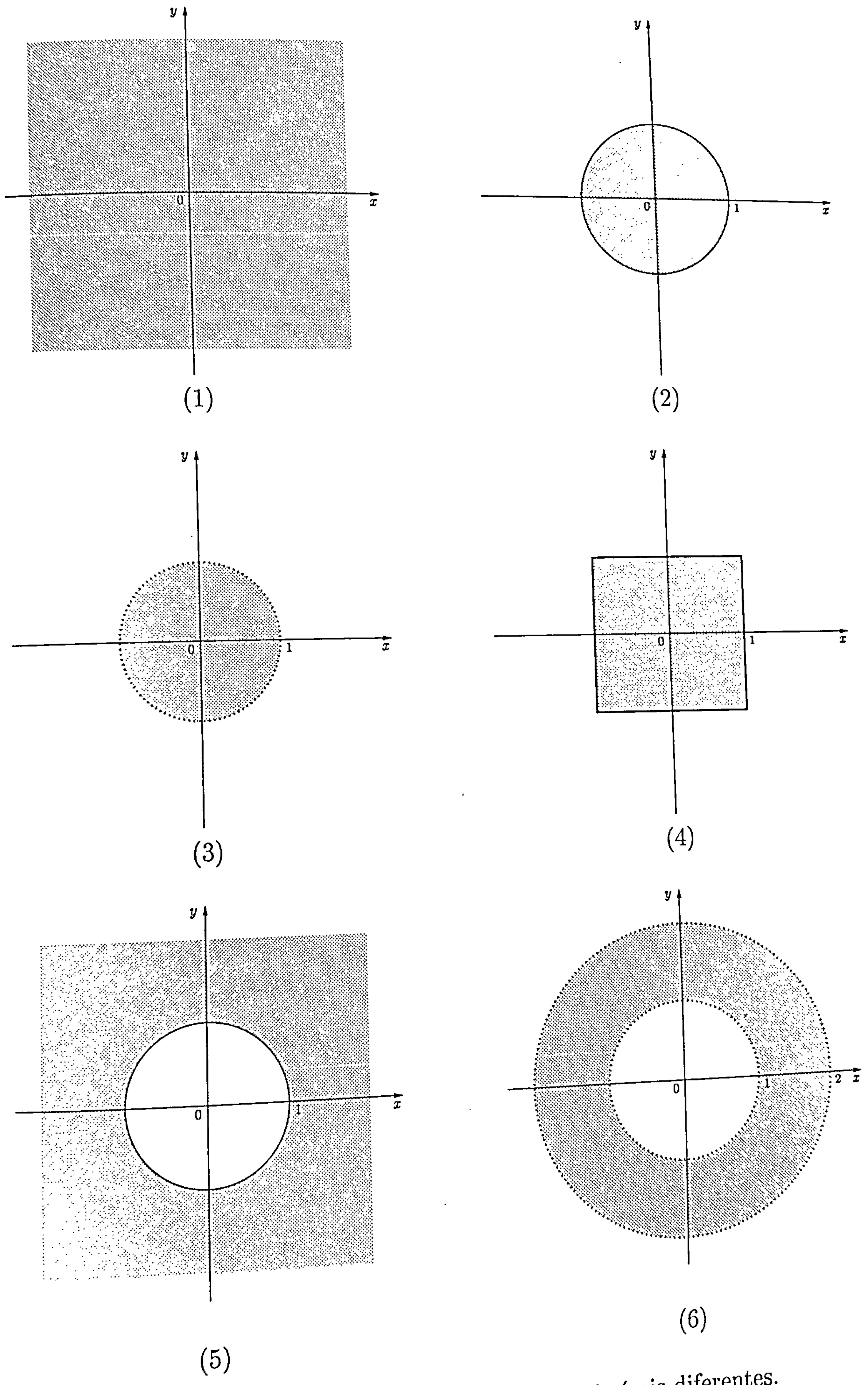

Figura 4.4: Desenho de seis conjuntos admissíveis diferentes.

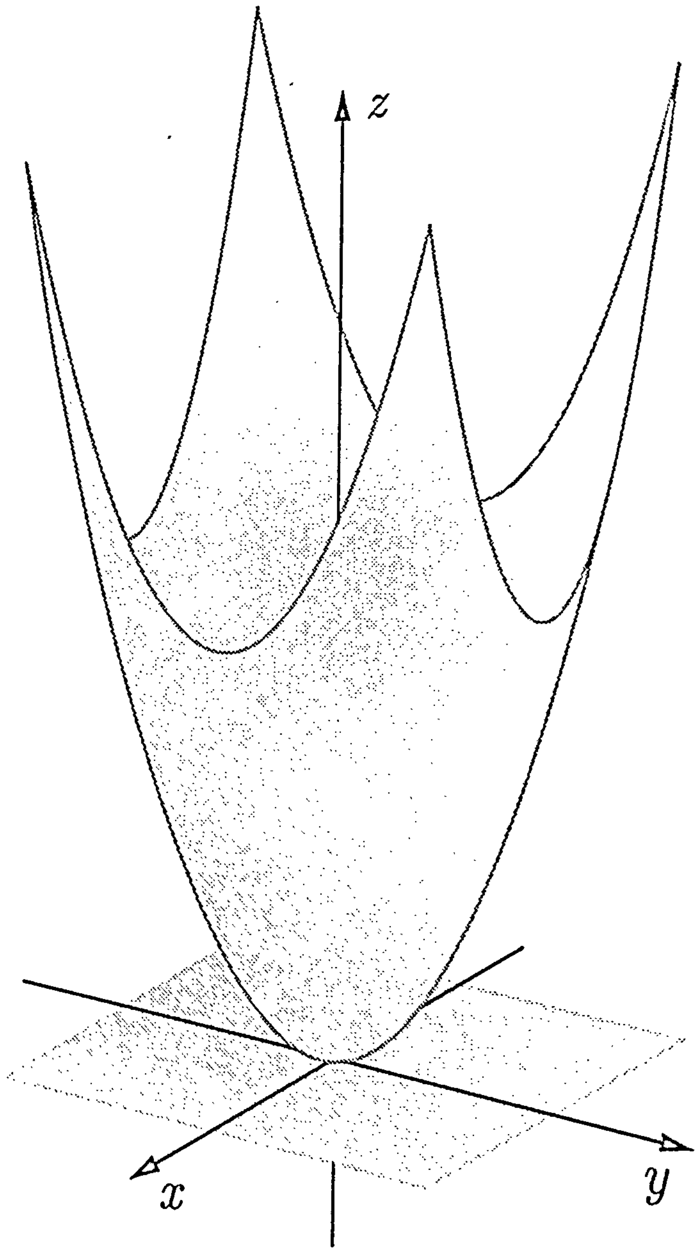

Figura 4.5: O gráfico de f definida em $D = \mathbb{R}^2$.

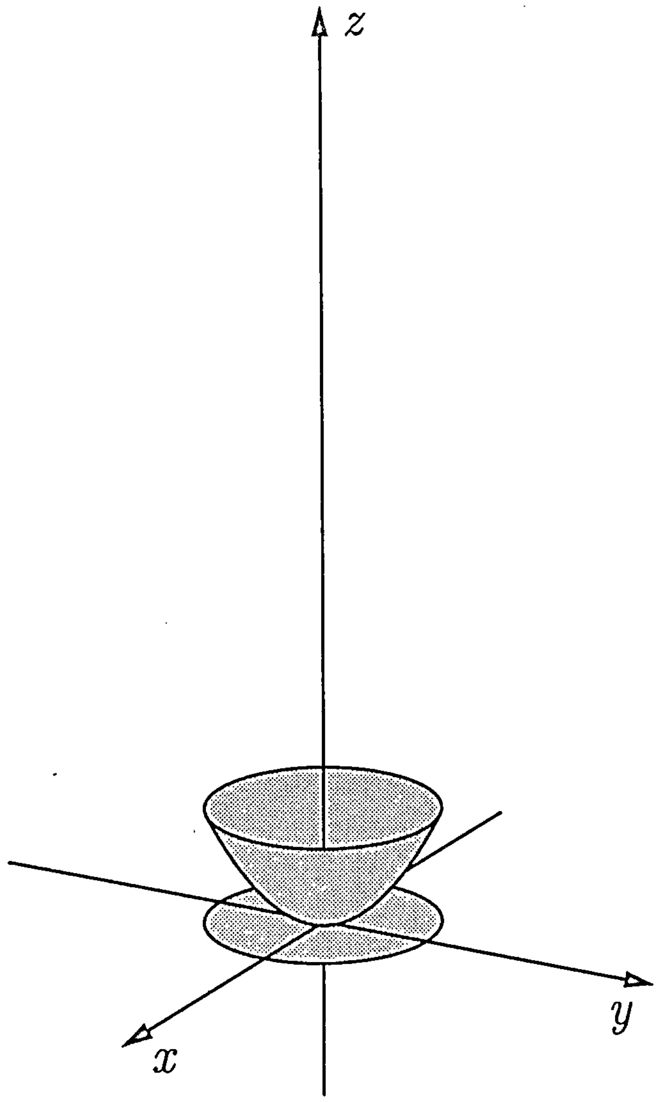

Figura 4.6: O gráfico de f definida em $D = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$.

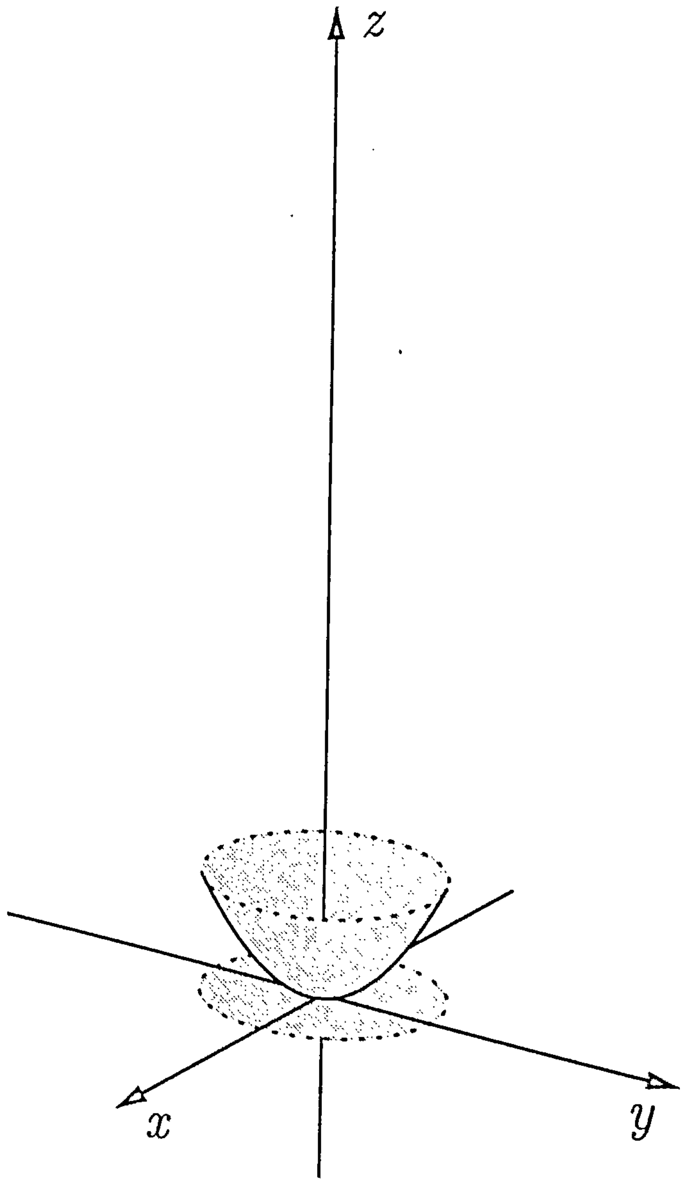

Figura 4.7: O gráfico de f definida em $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$.

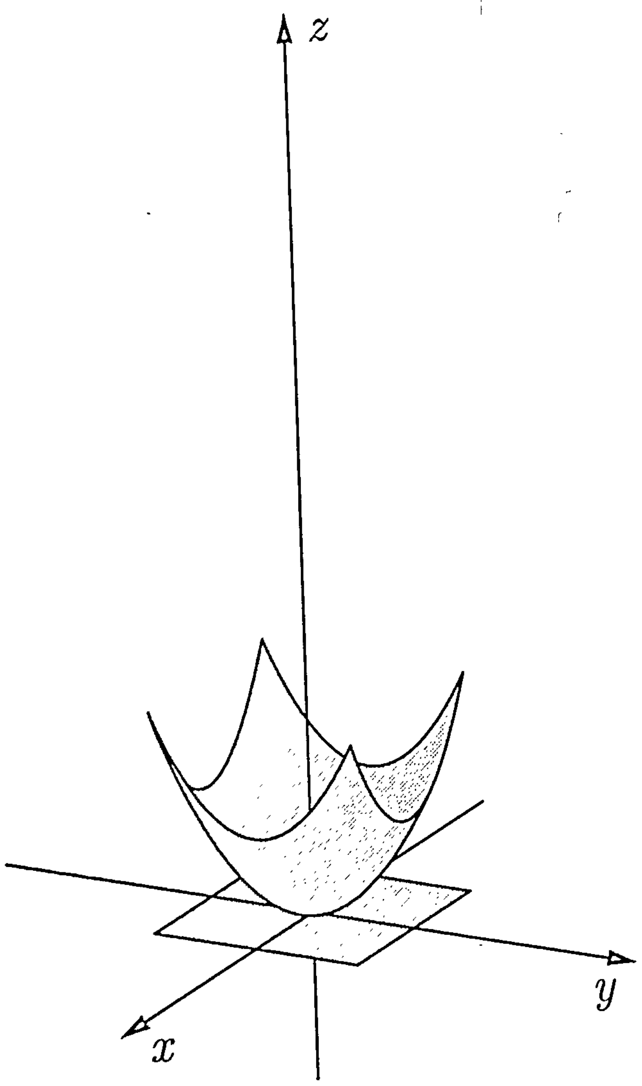

Figura 4.8: O gráfico de f definida em $D = \{(x, y) \in \mathbb{R}^2 \mid -1 \leq x \leq +1 \text{ e } -1 \leq y \leq +1\}$.

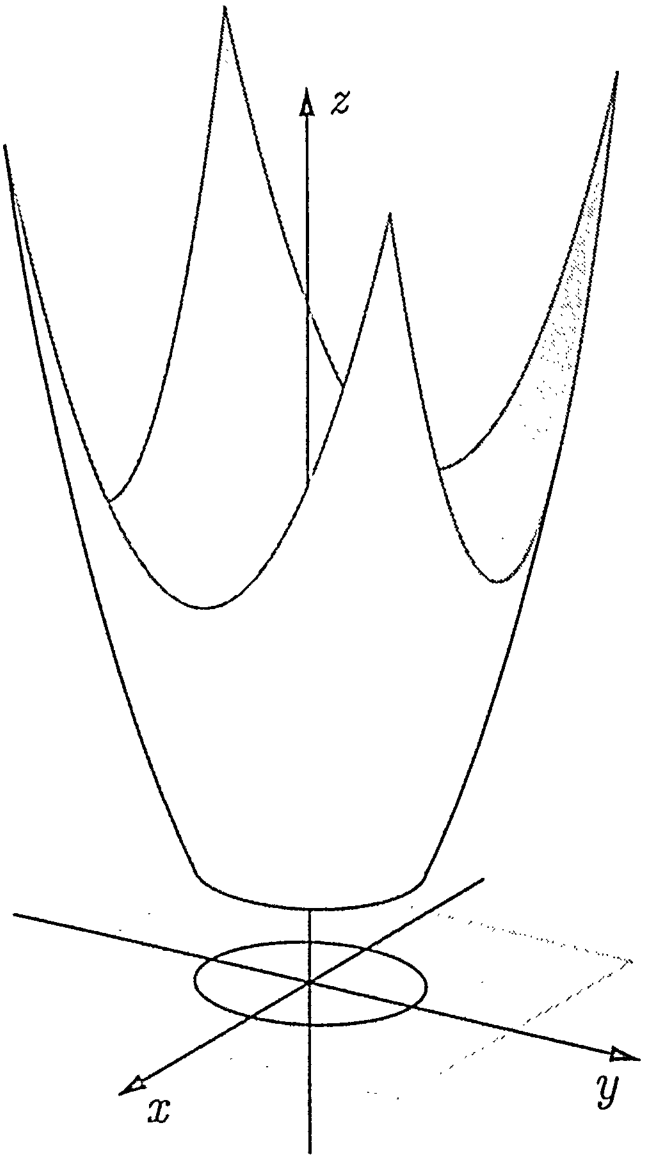

Figura 4.9: O gráfico de f definida em $D = \{(x, y) \in \mathbb{R}^2 \mid 1 \leq x^2 + y^2\}$.

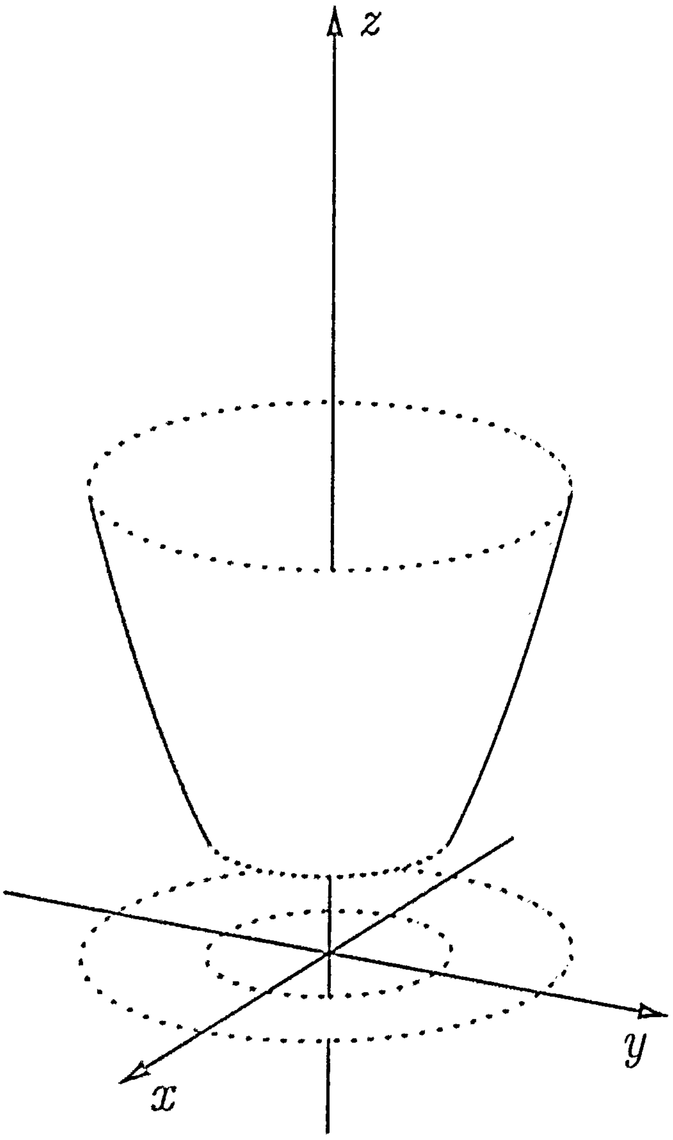

Figura 4.10: O gráfico de f definida em $D = \{(x, y) \in \mathbb{R}^2 \mid 1 < x^2 + y^2 < 4\}$.

No caso (2), $(0,0)$ é ainda o único ponto de mínimo global de f em $D = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \leq 1\}$. Como *não podemos tomar pontos arbitrariamente distantes* de $(0,0)$, segue-se que f não assume valores arbitrariamente grandes. De fato, os pontos de D mais distantes de $(0,0)$ são os pontos da circunferência de centro em $(0,0)$ e raio 1 sendo, portanto, os pontos de máximo global de f em $D = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \leq 1\}$.

No caso (3), $(0,0)$ é ainda o único ponto de mínimo global de f em $D = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 < 1\}$. Como no caso (2), não podemos tomar pontos arbitrariamente distantes de $(0,0)$ em D e, portanto, f não pode assumir valores arbitrariamente grandes em D. Mas, agora, não existe um ponto em D que seja considerado o ponto "mais distante" de $(0,0)$. Dado qualquer ponto em D, é sempre possível conseguir um outro ponto em D ainda mais distante de $(0,0)$. Quanto mais nos aproximamos da "fronteira" $\{(x,y) \in \mathbb{R}^2 \mid x^2+y^1 = 1\}$, mais o valor da função f aumenta, aproximando-se cada vez mais do valor 1. Mas não existem pontos de D que realizem este valor. Sendo assim, f não possui pontos de máximo global no conjunto $D = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 < 1\}$.

No caso (4), $(0,0)$ é o único ponto de mínimo global de f no quadrado $D = \{(x,y) \in \mathbb{R}^2 \mid -1 \leq x \leq +1 \text{ e } -1 \leq y \leq +1\}$. Por outro lado, os pontos $(-1,-1)$, $(-1,+1)$, $(+1,-1)$ e $(+1,+1)$ são os pontos de máximo global de f em D. *Justifique!*

No caso (5), os pontos da circunferência de centro em $(0,0)$ e raio 1 são os pontos de mínimo global de f em $D = \{(x,y) \in \mathbb{R}^2 \mid 1 \leq x^2+y^2\}$. Não existem pontos de máximo global de f em D. *Justifique!*

No caso (6), f não possui pontos de mínimo global e nem pontos de máximo global em $D = \{(x,y) \in \mathbb{R}^2 \mid 1 < x^2+y^2 < 4\}$. *Justifique!*

A fim de garantirmos a existência de máximos e mínimos globais, além da continuidade da função-objetivo, estes exemplos sugerem que o conjunto admissível D também deve ter certas propriedades especiais:

- D deve ser *limitado*, isto é, D não pode possuir pontos arbitrariamente distantes da origem (veja os casos (1) e (5)) e
- D deve conter seus pontos de *"fronteira"* (veja os casos (3) e (6)).

Surpreendentemente, estas duas propriedades (junto com a continuidade da função-objetivo) são *suficientes* para se estabelecer a existência de máximos e mínimos globais.

Para se formalizar estas propriedades (e estabelecer outras de interesse), vamos precisar de alguns conceitos topológicos: distância euclidiana, bola aberta, bola fechada, ponto interior, conjunto aberto, ponto de fronteira, conjunto fechado, conjunto limitado e conjunto compacto.

Distância euclidiana

Definição 4.4 (DISTÂNCIA EUCLIDIANA EM $\mathbb{R}^n$) No plano $\mathbb{R}^2$, a *distância euclidiana* entre dois pontos (a, b) e (x, y) pode ser calculada pelo teorema de Pitágoras que fornece o tamanho do segmento de reta que une os dois pontos:

$$d((x,y),(a,b)) = \sqrt{(x-a)^2 + (y-b)^2}.$$

No espaço euclidiano tridimensional $\mathbb{R}^3$, ainda com o auxílio do famoso teorema, podemos encontrar a *distância euclidiana* entre dois pontos (a, b, c) e (x, y, z) calculando-se o tamanho do segmento de reta que une os dois pontos:

$$d((x,y,z),(a,b,c)) = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}.$$

Mais geralmente, no espaço euclidiano $\mathbb{R}^n$, a *distância euclidiana* entre dois pontos $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ e $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ é calculada por:

$$d(\mathbf{x},\mathbf{p}) = \sqrt{(x_1-p_1)^2 + \cdots + (x_n-p_n)^2} = \sqrt{\sum_{i=1}^{n}(x_i-p_i)^2}.$$

Vamos ver alguns exemplos.

(a) A distância entre os pontos $(1, 2)$ e $(3, 4)$ é $\sqrt{(3-1)^2 + (4-2)^2} = \sqrt{8}$.

(b) A distância entre os vértices $(0, 0)$ e $(1, 1)$ do quadrado $Q_2 = [0, 1] \times [0, 1]$ é $\sqrt{(1-0))^2 + (1-0))^2} = \sqrt{2}$.

(c) A distância entre os vértices $(0, 0, \ldots, 0)$ e $(1, 1, \ldots, 1)$ do cubo n-dimensional $Q_n = [0, 1] \times [0, 1] \times \cdots \times [0, 1]$ é $\sqrt{n}$.

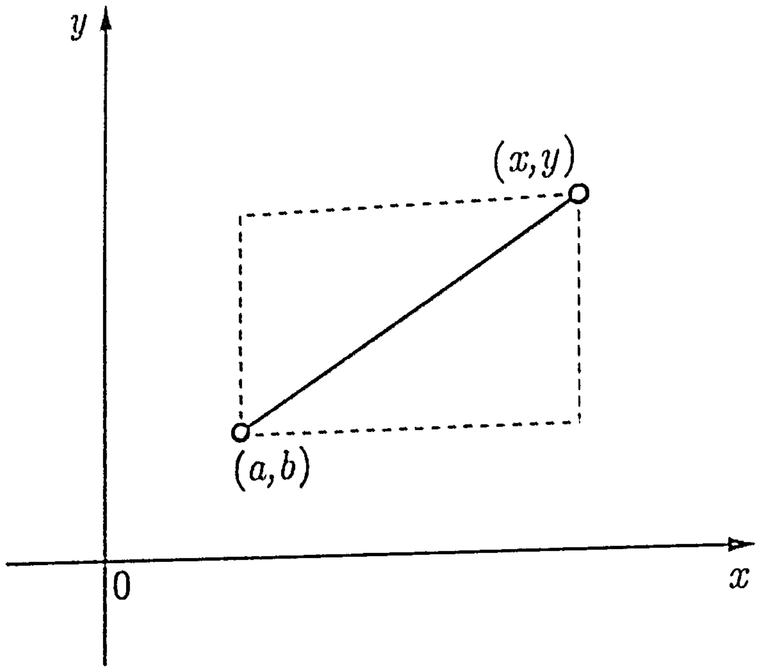

Figura 4.11: Distância euclidiana entre dois pontos de $\mathbb{R}^2$.

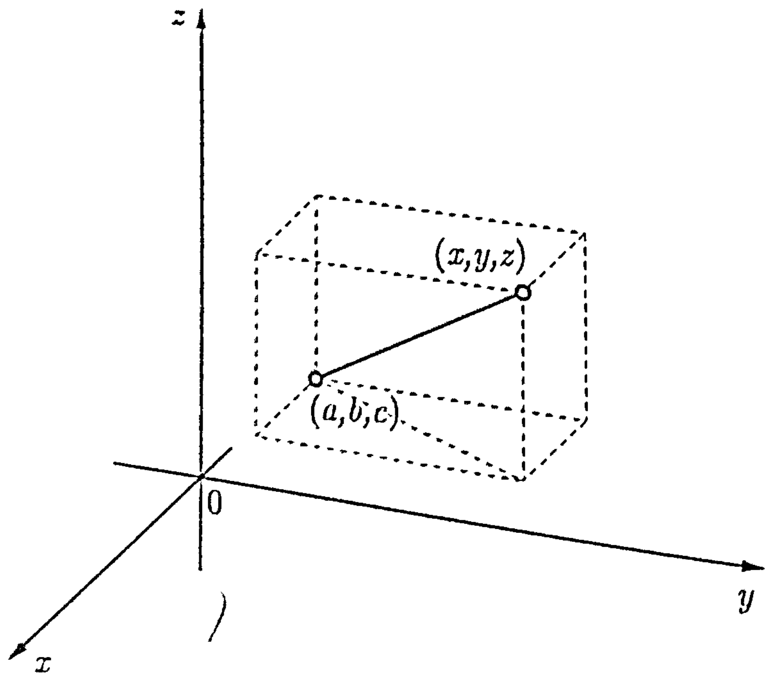

Figura 4.12: Distância euclidiana entre dois pontos de $\mathbb{R}^3$.

O próximo teorema estabelece algumas propriedades da distância euclidiana.

Teorema 4.3 (Propriedades da função distância) Considere p, q e r pontos em $\mathbb{R}^n$. Então

(1) $d(\mathrm{p},\mathrm{q}) \geq 0$ e $d(\mathrm{p},\mathrm{p}) = 0$,

(2) $d(\mathrm{p},\mathrm{q}) = d(\mathrm{q},\mathrm{p})$ e

(3) (Desigualdade triangular) $d(\mathrm{p},\mathrm{q}) \leq d(\mathrm{p},\mathrm{r}) + d(\mathrm{r},\mathrm{q})$,

onde $d(\mathrm{p},\mathrm{q})$ é a distância euclidiana entre os pontos p e q.

A primeira propriedade diz que a distância entre dois pontos é sempre ≥ 0 e que a distância de um ponto p até ele mesmo é zero, a segunda propriedade afirma que a distância de p até q é igual à distância de q até p e, finalmente, a terceira propriedade (a desigualdade triangular) afirma que dado um triângulo de vértices p, q e r, a medida de um dos seus lados ($d(\mathrm{p},\mathrm{q})$) é sempre menor ou igual que à soma das medidas dos outros dois lados ($d(\mathrm{p},\mathrm{r}) + d(\mathrm{r}),\mathrm{q})$).

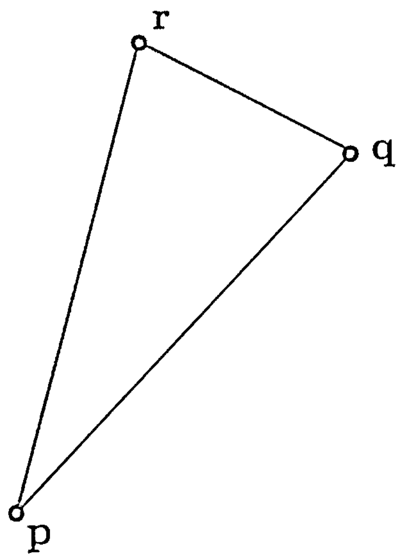

Figura 4.13: A desigualdade triangular: $d(\mathrm{p},\mathrm{q}) \leq d(\mathrm{p},\mathrm{r}) + d(\mathrm{r},\mathrm{q})$.

As propriedades (1) e (2) são fáceis de se demonstrar. No exercício [25] indicamos como obter a desigualdade triangular (propriedade (3)) a partir da desigualdade de Cauchy-Schwarz.

Bola aberta

Agora que sabemos como medir distâncias em $\mathbb{R}^n$, uma pergunta muito natural que se põe é: dado um ponto p em $\mathbb{R}^n$, quais são todos os pontos x em $\mathbb{R}^n$ que estão a uma distância menor do que um número r (fixo) com relação a p? No caso do plano $(n = 2)$, temos

$$d(\mathrm{p}, \mathrm{x}) < r \Leftrightarrow d((p_1, p_2), (x_1, x_2)) < r \Leftrightarrow \sqrt{(x_1 - p_1)^2 + (x_2 - p_2)^2} < r \Leftrightarrow (x_1 - p_1)^2 + (x_2 - p_2)^2 < r^2,$$

isto é, um ponto $\mathrm{x} = (x_1, x_2)$ está a uma distância menor do que r com relação a $\mathrm{p} = (p_1, p_2)$ se, e somente se, x está no disco sem "casca"

$$\left\{(x_1, x_2) \in \mathbb{R}^2 \mid (x_1 - p_1)^2 + (x_2 - p_2)^2 < r^2\right\},$$

de centro em p e raio r. Não é difícil de ver que no caso tridimensional $(n = 3)$, um ponto $\mathrm{x} = (x_1, x_2, x_3)$ está a uma distância menor do que r com relação a $\mathrm{p} = (p_1, p_2, p_3)$ se, e somente se, x está na bola aberta

$$\left\{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid (x_1 - p_1)^2 + (x_2 - p_2)^2 + (x_3 - p_3)^2 < r^2\right\},$$

de centro em p e raio r. Isto motiva a definição de uma bola aberta n-dimensional:

Definição 4.5 (Bola aberta) A *bola aberta* de centro em $\mathrm{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n$ de raio $r > 0$ é o conjunto

$$B_r(\mathrm{p}) = \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid (x_1 - p_1)^2 + \cdots + (x_n - p_n)^2 < r^2\right\}$$

formado por todos os pontos de $\mathbb{R}^n$, cuja distância até p é menor do que r.

Bola fechada

Analogamente, podemos nos perguntar quais são todos os pontos x em $\mathbb{R}^n$ que estão a uma distância menor ou igual a um número r (fixo) com relação a um ponto $\mathrm{p} \in \mathbb{R}^n$ dado. Não é difícil ver que a resposta é uma bola aberta com sua "casca": a bola fechada.

Definição 4.6 (Bola fechada) A *bola fechada* de centro em $\mathrm{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n$ de raio $r \geq 0$ é o conjunto

$$\overline{B_r(\mathrm{p})} = \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid (x_1 - p_1)^2 + \cdots + (x_n - p_n)^2 \leq r^2\right\}$$

formado por todos os pontos de $\mathbb{R}^n$, cuja distância até p é menor ou igual a r.

Conjunto limitado

Com o conceito de bola fechada, podemos definir o que é um conjunto limitado:

Definição 4.7 (Conjunto limitado) Um subconjunto D de $\mathbb{R}^n$ é dito ser *limitado* se existe uma bola fechada $\overline{B_r(\mathbf{0})}$ de centro em $\mathbf{0} = (0, \ldots, 0)$ e raio r tal que $D \subset \overline{B_r(\mathbf{0})}$.

Em outras palavras, um subconjunto D de $\mathbb{R}^n$ é limitado se é possível colocá-lo dentro de alguma bola fechada de centro na origem $\mathbf{0} = (0, 0, \ldots, 0)$.

Considere os exemplos a seguir.

(a) Os conjuntos dos casos (2), (3), (4) e (6) da página 130 são limitados. Os demais conjuntos não são limitados.

(b) O quadrado $Q_2 = [0, 1] \times [0, 1]$ é um subconjunto limitado de $\mathbb{R}^2$ pois ele está contido na bola fechada $\overline{B_{\sqrt{2}}(0, 0)}$ de centro em $(0, 0)$ e raio $\sqrt{2}$. Mais geralmente, o cubo n-dimensional $Q_n = [0, 1] \times [0, 1] \times \cdots [0, 1]$ é um subconjunto limitado de $\mathbb{R}^n$ pois ele está contido na bola fechada $\overline{B_{\sqrt{n}}(0, 0, \ldots, 0)}$ de centro em $(0, 0, \ldots, 0)$ e raio $\sqrt{n}$.

(c) O semiplano superior $\left\{(x, y) \in \mathbb{R}^2 \mid y \geq 0\right\}$ não é um subconjunto limitado de $\mathbb{R}^2$ pois não é possível colocá-lo em nenhuma bola fechada.

Ponto de fronteira

Vamos agora tornar mais formal a noção de "fronteira", um conceito que guiou nossa intuição na discussão sobre a existência de extremos globais de uma função contínua. Intuitivamente falando, um ponto $\mathbf{p}$ está na fronteira de um conjunto D se é possível aproximá-lo por pontos que estão em D e por pontos que não estão em D. Mais formalmente,

Definição 4.8 (Fronteira de um conjunto) Um ponto $\mathbf{p}$ é *ponto de fronteira* de um conjunto D se *toda* bola aberta com centro $\mathbf{p}$ contém pontos que estão em D e pontos que não estão em D. A *fronteira* de D é o conjunto formado por todos os pontos de fronteira de D.

Por exemplo, qualquer ponto da circunferência

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

é um ponto da fronteira da bola aberta $B_1(0,0)$. Observe que um ponto da fronteira de um conjunto não precisa estar no conjunto! Com relação aos conjuntos definidos na página 130, temos o que se segue.

(1) $D = \mathbb{R}^2$ não possui pontos de fronteira.

(2) Os pontos de fronteira de $D = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ são os pontos do círculo $\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$.

(3) Os pontos de fronteira de $D = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ são os pontos do círculo $\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$.

(4) Os pontos de fronteira de

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

são os pontos do conjunto

$$(\{+1\} \times [-1,+1]) \cup (\{-1\} \times [-1,+1]) \cup$$
$$([-1,+1] \times \{-1\}) \cup ([-1,+1] \times \{+1\}),$$

isto é, os pontos dos quatro lados de D.

(5) Os pontos de fronteira de $D = \{(x,y) \in \mathbb{R}^2 \mid 1 \leq x^2 + y^2\}$ são os pontos do círculo $\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$.

(6) Os pontos de fronteira de $D = \{(x,y) \in \mathbb{R}^2 \mid 1 < x^2 + y^2 < 4\}$ são os pontos do círculo $\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ e do círculo $\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 4\}$.

Conjunto fechado

O conceito que vamos definir agora — o de conjunto fechado — possui um papel fundamental na geometria e análise dos espaços euclidianos.

Definição 4.9 (Conjunto fechado) Um subconjunto D de $\mathbb{R}^n$ é dito ser *fechado* em $\mathbb{R}^n$ se *todos* os pontos de fronteira de D pertencem a D.

Considere os exemplos a seguir.

(a) Os conjuntos dos casos (1), (2), (4) e (5) da página 130 são fechados. Os demais conjuntos não são fechados.

(b) A bola fechada $\overline{B_r(\mathrm{p})}$ é um subconjunto fechado em $\mathbb{R}^n$.

(c) O quadrado $Q_2 = [0,1] \times [0,1]$ é um subconjunto fechado em $\mathbb{R}^2$. Mais geralmente, o cubo n-dimensional $Q_n = [0,1] \times [0,1] \times \cdots [0,1]$ é um subconjunto fechado em $\mathbb{R}^n$.

(d) O semiplano superior $\{(x,y) \in \mathbb{R}^2 \mid y \geq 0\}$ é um subconjunto fechado em $\mathbb{R}^2$.

(e) A bola aberta $B_1(0,0) = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ (note a desigualdade estrita) de centro em $(0,0)$ e raio 1 não é um subconjunto fechado em $\mathbb{R}^2$, pois os pontos de fronteira de $B_1(0,0)$ (isto é, os pontos (x,y) em $\mathbb{R}^2$ tais que $x^2 + y^2 = 1$) não pertencem a $B_1(0,0)$.

(f) A bola "furada" $X = \{(x,y) \in \mathbb{R}^2 \mid 0 < x^2 + y^2 \leq 1\}$ de centro em $(0,0)$ e raio 1 não é um subconjunto fechado em $\mathbb{R}^2$, pois $(0,0)$ é um ponto de fronteira de X que *não pertence* a X. Observe, contudo, que existem pontos de fronteira de X que pertencem a X.

A próxima proposição estabelece uma outra caracterização de conjuntos fechados: se p é um ponto qualquer e se existem pontos em um conjunto fechado D arbitrariamente próximos de p, então p também deve estar em D.

Proposição 4.1 Um subconjunto D de $\mathbb{R}^n$ é *fechado* em $\mathbb{R}^n$ se, e somente se, *toda* seqüência *convergente* de pontos x_n *em* D tem seu limite também em D, isto é, se $\mathrm{x}_n \in D$ é uma seqüência tal que $\mathrm{x}_n \to \mathrm{p}$, então $\mathrm{p} \in D$.

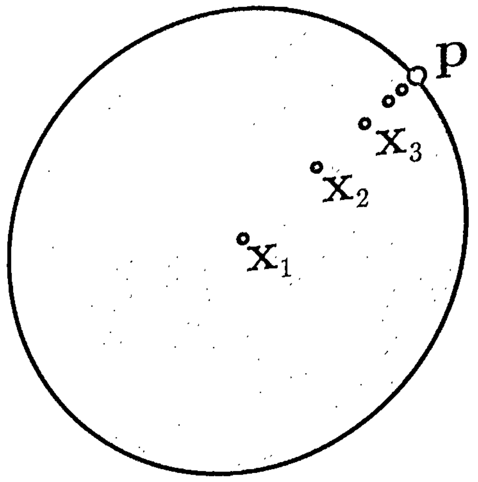

Figura 4.14: D é fechado, $x_n \in D$ e $x_n \to p$, logo, $p \in D$.

Existe uma maneira muito simples de se identificar conjuntos fechados:

Proposição 4.2 Se $h\colon \mathbb{R}^n \to \mathbb{R}$ é *contínua*, então o conjunto

$$\{x \in \mathbb{R}^n \mid h(x) \leq c\},$$

é fechado em $\mathbb{R}^n$, onde c é uma constante real.

A demonstração segue-se facilmente da proposição (4.1) e será deixada como exercício (veja o exercício [11] no final do capítulo).

Conjunto compacto

Podemos definir agora o objeto em $\mathbb{R}^n$ "equivalente" a um intervalo fechado e limitado $[a, b]$ na reta $\mathbb{R}$:

Definição 4.10 (Conjunto compacto) Um subconjunto K de $\mathbb{R}^n$ é dito ser *compacto* se ele é *limitado* e *fechado* em $\mathbb{R}^n$.

Assim, a bola fechada $\overline{B_1(0,0)}$ e o quadrado Q_2 são exemplos de conjuntos compactos. A bola aberta $B_1(0,0)$ não é compacta porque ela não é fechada (apesar de ser limitada). O semiplano superior $\{(x, y) \in \mathbb{R}^2 \mid y \geq 0\}$ não é compacto porque ele não é limitado (apesar de ser fechado).

Os conjuntos dos casos (2) e (4) da página 130 são compactos. Os demais conjuntos não são compactos.

O teorema de Weierstrass

Com tudo isto, podemos finalmente enunciar o teorema de Weierstrass no caso de funções que dependam de n variáveis:

Teorema 4.4 (WEIERSTRASS) Toda função $f \colon K \subset \mathbb{R}^n \to \mathbb{R}$ *contínua* definida no *conjunto compacto* K (não-vazio) possui pelo menos um máximo global e pelo menos um mínimo global em K.

Como vimos, os conjuntos dos casos (2) e (4) da página 130 são compactos e uma vez que $z = f(x, y) = x^2 + y^2$ é uma função contínua, podemos usar o teorema de Weierstrass para garantir a existência de máximos e mínimos globais de f nestes conjuntos.

CUIDADO! CUIDADO! CUIDADO!

O teorema de Weierstrass diz que uma função *contínua* f definida em um conjunto *compacto* K *não-vazio* possui pelo menos um máximo global e pelo menos um mínimo global em K. E se f não for contínua ou K não for compacto? Nesta situação, o teorema de Weierstrass não se aplica: f pode possuir ou não extremos globais em K. Por exemplo, $D = \mathbb{R}^2$ não é um conjunto compacto, mas $z = g(x, y) = \cos(xy)$ possui máximos e mínimos globais em D. Por outro lado, $z = h(x, y) = x + y$ não possui máximo e nem mínimo global em D. Cada caso é um caso. Você precisa usar outras ferramentas para decidir se f possui ou não extremos globais.

O teorema (4.2) fornece uma maneira simples de se identificar funções contínuas. Mas como identificar conjuntos compactos? Freqüentemente, em problemas de otimização, o conjunto admissível é modelado da seguinte forma:

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

onde $h_1, \ldots, h_m$ são funções de $\mathbb{R}^n$ para $\mathbb{R}$. Agora, se as funções $h_1, \ldots, h_m$ são *contínuas*, então pela proposição (4.2), os conjuntos

$$D_1 = \{\mathrm{x} \in \mathbb{R}^n \mid h_1(\mathrm{x}) \leq c_1\}, \quad \ldots, \quad D_m = \{\mathrm{x} \in \mathbb{R}^n \mid h_m(\mathrm{x}) \leq c_m\}$$

são fechados. Como a interseção de um número finito de conjuntos fechados é ainda um conjunto fechado (veja o exercício [15]), segue-se que

$$D = D_1 \cap \cdots \cap D_m$$

também é fechado. Infelizmente, ainda não se conhece um critério simples que permita dizer se um conjunto da forma

$$\{\mathrm{x} \in \mathbb{R}^n \mid h(\mathrm{x}) \leq c\}$$

é limitado ou não. Veja, entretanto, o exercício [12].

Ponto interior

Intuitivamente, um ponto $\mathbf{p}$ é interior a um subconjunto D de $\mathbb{R}^n$ se $\mathbf{p}$ pertence a D e qualquer "pertubação" suficientemente pequena de $\mathbf{p}$ ainda está em D.

Definição 4.11 Um ponto $\mathbf{p}$ é *interior* a um subconjunto D de $\mathbb{R}^n$ se existe uma bola aberta de centro $\mathbf{p}$ contida em D. O conjunto de todos os pontos interiores de um conjunto D é denominado *conjunto interior* de D.

Por exemplo, a origem $(0, 0)$ é um ponto interior da bola fechada $\overline{B_1(0, 0)}$ de centro na origem $(0, 0)$ e raio 1, pois podemos conseguir facilmente uma outra bola aberta de centro em $(0, 0)$ contida em $\overline{B_1(0, 0)}$. Por outro lado, o ponto $(\sqrt{2}/2, \sqrt{2}/2)$ não é um ponto interior de $\overline{B_1(0, 0)}$, pois sempre existirá um ponto que não pertence a $\overline{B_1(0, 0)}$ em *qualquer* bola aberta de centro em $(\sqrt{2}/2, \sqrt{2}/2)$ (veja a figura (4.15)).

Não é difícil de se convencer que o conjunto interior da bola fechada unitária $\overline{B_1(0, 0)}$ é a bola aberta unitária $B_1(0, 0)$. Alertamos ainda que o interior de um conjunto pode ser o conjunto vazio. Isto acontece, por exemplo, para uma reta em $\mathbb{R}^2$ ou um plano em $\mathbb{R}^3$.

Conjunto aberto

Um conceito fortemente relacionado com a definição de ponto interior é o conceito de conjunto aberto.

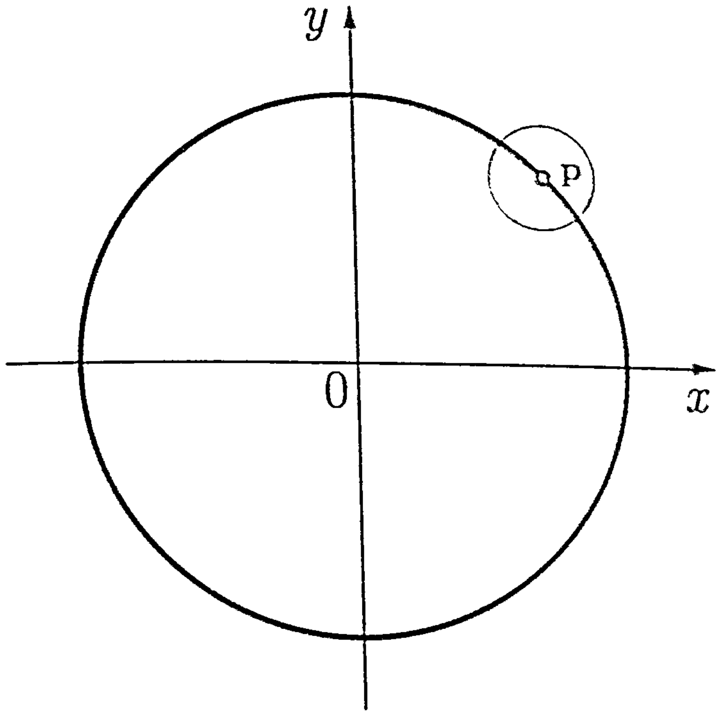

Figura 4.15: O ponto $\mathrm{p} = (\sqrt{2}/2, \sqrt{2}/2)$ não é um ponto interior da bola fechada $\overline{B_1(0,0)}$.

Definição 4.12 (Conjunto aberto) Dizemos que $D \subset \mathbb{R}^n$ é *aberto* se todos os seus pontos são interiores a D, isto é, se para todo ponto $\mathrm{p} \in D$ existe uma bola aberta $B_r(\mathrm{p})$ de centro em p e raio $r > 0$ contida em D.

A palavra "aberto" tem a conotação de "sem fronteira": a partir de qualquer ponto podemos sempre nos mover a uma pequena distância em *qualquer* direção e ainda ficar dentro do conjunto. A definição de um conjunto aberto torna esta idéia precisa: cada elemento de um conjunto aberto possui uma bola com centro neste elemento contida no conjunto.

Exemplo 4.2 A bola aberta $B_1(0)$ de centro em $0 = (0,0)$ e raio 1 é um conjunto aberto. De fato: seja $\mathrm{p} = (x, y)$ um ponto de $B_1(0)$. Devemos mostrar que existe uma bola aberta $B_r(\mathrm{p})$ de centro em p contida em $B_1(0)$. Seja

$$r = 1 - d(\mathrm{p}, 0).$$

(veja a figura (4.16)). Para mostrar que $B_r(\mathrm{p})$ está contida em $B_1(0)$, considere q um ponto em $B_r(\mathrm{p})$. Pela desigualdade triangular (veja o teorema (4.3)),

$$d(\mathrm{q}, 0) \leq d(\mathrm{q}, \mathrm{p}) + d(\mathrm{p}, 0) = d(\mathrm{p}, \mathrm{q}) + d(\mathrm{p}, 0).$$

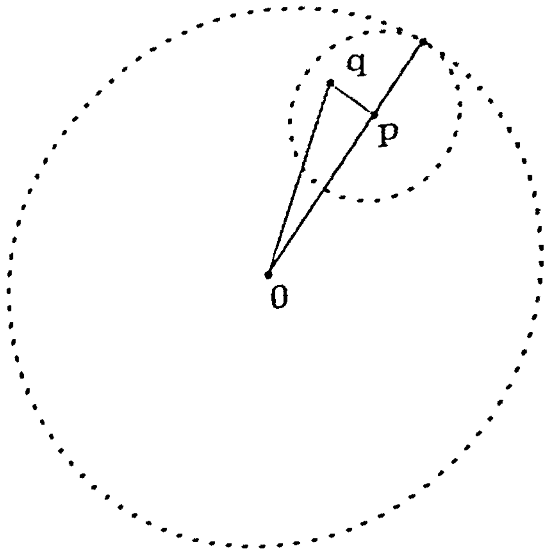

Figura 4.16: A bola $B_{1-d(\mathrm{p},0)}(\mathrm{p})$ está contida na bola $B_1(0)$

Como $d(\mathrm{p},\mathrm{q}) < 1 - d(\mathrm{p},0)$ vem que

$$d(\mathrm{q},0) < 1 - d(\mathrm{p},0) + d(\mathrm{p},0) = 1$$ ◻

A definição de um conjunto aberto também implica que estes conjuntos devem ser "cheios", pois um conjunto aberto em $\mathbb{R}^n$ contém uma bola aberta n-dimensional com centro em cada um de seus pontos. Conseqüentemente, uma reta em $\mathbb{R}^2$ não é um conjunto aberto (veja a figura (4.17)). Analogamente, retas e planos em $\mathbb{R}^3$ não são conjuntos abertos.

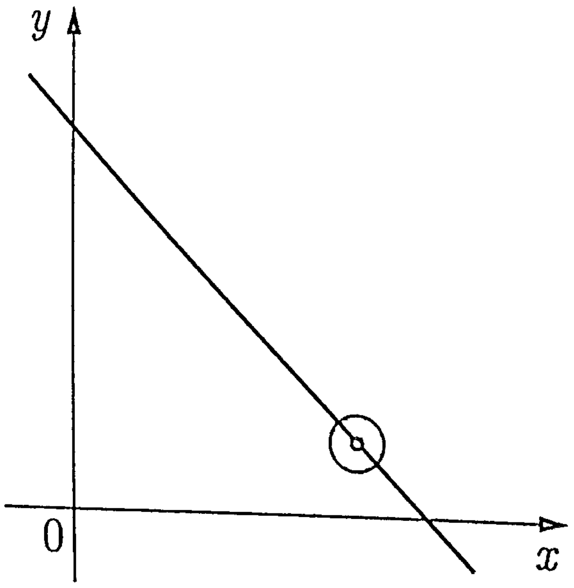

Figura 4.17: Toda bola aberta com centro em um ponto da reta contém pontos fora da reta.

Existe uma proposição análoga à proposição (4.2) para conjuntos abertos:

Proposição 4.3 Se $h \colon \mathbb{R}^n \to \mathbb{R}$ é *contínua*, então o conjunto

$$\{\mathrm{x} \in \mathbb{R}^n \mid h(\mathrm{x}) < c\},$$

é aberto em $\mathbb{R}^n$, onde c é uma constante real.

4.4 Exercícios

[01] Dê três exemplos de funções contínuas e três exemplos de funções descontínuas definidas em $\mathbb{R}^2$.

[02] No exercício resolvido (4.3) mostramos que $f(x, y) = x$ é uma função contínua. Mostre agora que $g(x, y) = y$ e $h(x, y) = k =$ constante são funções contínuas.

[03] Complete a demonstração do teorema (4.2) da página 128.

[04] Mostre que toda função polinomial

$$p(x, y) = \sum_{i=0}^{k} a_i \cdot x^{k-i} \cdot y^i = \\ a_0 \cdot x^k + a_1 \cdot x^{k-1} \cdot y + \cdots a_{k-1} \cdot x \cdot y^{k-1} + a_k \cdot y^k$$

em duas variáveis x e y é contínua em $\mathbb{R}^2$.

[05] Mostre que a função

$$f(x, y) = \begin{cases} \dfrac{x^2 \cdot y}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0), \\ 0 & \text{se } (x, y) = (0, 0). \end{cases}$$

é contínua em $(0, 0)$.

[06] Enuncie outros teoremas importantes envolvendo continuidade que você aprendeu no curso de Cálculo I.

*[07] Seja $f \colon D \subset \mathbb{R}^n \to \mathbb{R}$ uma função contínua em $\mathrm{p} \in D$. Mostre que se $f(\mathrm{p}) > 0$, então existe uma bola aberta $B_r(\mathrm{p})$ de centro em p e raio $r > 0$ tal que $f(\mathrm{x}) > 0$ para todo $\mathrm{x} \in B_r(\mathrm{p}) \cap D$.

[08] Resolva as questões abaixo.

(a) Dê três exemplos de conjuntos limitados e três exemplos de conjuntos não-limitados.

(b) Dê três exemplos de conjuntos abertos e três exemplos de conjuntos que não são abertos.

(c) Dê três exemplos de conjuntos fechados e três exemplos de conjuntos que não são fechados.

[09] Represente geometricamente cada um dos subconjuntos de $\mathbb{R}^2$ abaixo e diga se ele é um conjunto fechado ou não, aberto ou não, limitado ou não, compacto ou não. Determine também o interior e a fronteira de cada um destes subconjuntos.

(a) $\{(x,y) \in \mathbb{R}^2 \mid -1 < x < +1 \text{ e } y = 0\}$.

(b) $\{(x,y) \in \mathbb{R}^2 \mid x \text{ e } y \text{ são números inteiros}\}$.

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

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

(e) $\{(x,y) \in \mathbb{R}^2 \mid x = 0 \text{ ou } y = 0\}$.

(f) $\{(x,y) \in \mathbb{R}^2 \mid x = 0 \text{ e } y \geq 0\}$.

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

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

(i) $\{(x,y) \in \mathbb{R}^2 \mid 1 \leq x \leq 2\}$.

(j) $\{(x,y) \in \mathbb{R}^2 \mid x = 0 \text{ ou } y = 0 \text{ mas } x^2 + y^2 \neq 0\}$.

[10] O conjunto $\{(x,y,z) \in \mathbb{R}^3 \mid x \geq 0, y \geq 0, z \geq 0 \text{ e } x + y + z \leq 1\}$ é compacto? Justifique sua resposta. *Dica: represente geometricamente este conjunto em* $\mathbb{R}^3$.

[11] Mostre que se $h\colon \mathbb{R}^n \to \mathbb{R}$ é uma função contínua então os conjuntos $\{\mathrm{x} \in \mathbb{R}^n \mid h(\mathrm{x}) = c\}$ — o nosso já conhecido conjunto de nível da função h — e $\{\mathrm{x} \in \mathbb{R}^n \mid h(\mathrm{x}) \leq c\}$ são conjuntos fechados para qualquer escolha do número c.

[12] Mostre que o conjunto

$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1 \leq x_1 \leq b_1, \ldots, a_m \leq x_m \leq b_m\}$$

é limitado, onde $a_1, \ldots, a_m, b_1, \ldots, b_m$ são constantes reais.

[13] Mostre que o teorema de Weierstrass não é mais válido se removermos a hipótese de continuidade ou de compacidade do domínio da função f exibindo:

(a) Uma função descontínua definida em um subconjunto K compacto em $\mathbb{R}^2$ que não possui nem máximo global e nem mínimo global em K.

(b) Uma função contínua definida em um subconjunto F fechado mas não-limitado (e, portanto, não-compacto) em $\mathbb{R}^2$ que não possui nem máximo global e nem mínimo global em F.

(c) Uma função contínua definida em um subconjunto L limitado mas não-fechado (e, portanto, não-compacto) em $\mathbb{R}^2$ que não possui nem máximo global e nem mínimo global em L.

[14] Use o teorema de Weierstrass para dizer se é possível garantir, *a priori*, se cada um dos problemas de otimização abaixo possui ou não uma solução.

(a) Maximizar $f(x, y) = x + y$
sujeito às restrições: $x \geq 0, y \geq 0$ e $x^2 + y^2 \leq 1$.

(b) Maximizar $f(x, y) = x^2 - y^2$
sujeito às restrições: $x \geq 0$ e $y \geq 0$.

(c) Minimizar $f(x, y) = x \cdot y$
sujeito às restrições: $x \geq 0, y \geq 0$ e $x + y = 1$.

(d) Maximizar $f(x, y) = \operatorname{sen}(x^2 + y^2)$
sujeito à restrição: $x^2 + y^2 < 1$.

(e) Maximizar $f(x, y) = \operatorname{sen}(x^2 + y^2)$
sujeito à restrição: $x^2 + y^2 \leq 1$.

(f) Maximizar $f(x, y) = \operatorname{sen}(x^2 + y^2)$
sujeito à restrição: $x^2 + y^2 = 1$.

(g) Maximizar $f(x, y, z) = x + 2y + 3z$
sujeito às restrições: $x \geq 0, y \geq 0, z \geq 0$ e $x + y + z \leq 1$.

(h) Maximizar $f(x, y, z) = x + 2y + 3z$
sujeito às restrições: $x \geq 0, y \geq 0$ e $z \leq x^2 + y^2$.

(i) Maximizar $f(x, y, z) = x + 2y + 3z$
sujeito às restrições: $x \geq 0, y \geq 0, z \geq 0, x \leq 1, y \leq 1$ e $z \leq 1$.

*[15] Diga se cada uma das sentenças abaixo é verdadeira ou falsa, justificando sua resposta.

(a) Qualquer conjunto *finito* é um conjunto limitado.

(b) A interseção de dois conjuntos fechados é um conjunto fechado.

(c) A interseção de um número *infinito* de conjuntos fechados é um conjunto fechado.

(d) A união de dois conjuntos fechados é um conjunto fechado.

(e) A união de um número *finito* de conjuntos fechados é um conjunto fechado.

(f) A união de um número *infinito* de conjuntos fechados é um conjunto fechado.

(g) Qualquer conjunto *finito* é um conjunto fechado.

(h) A união de dois conjuntos abertos é um conjunto aberto.

(i) A união de um número *finito* de conjuntos abertos é um conjunto aberto.

(j) A união de um número *infinito* de conjuntos abertos é um conjunto aberto.

(k) A interseção de dois conjuntos abertos é um conjunto aberto.

(l) A interseção de um número *finito* de conjuntos abertos é um conjunto aberto.

(m) A interseção de um número *infinito* de conjuntos abertos é um conjunto aberto.

(n) Um subconjunto fechado de um conjunto compacto é um conjunto compacto.

(o) A união de um número *finito* de conjuntos compactos é um conjunto compacto.

(p) A união de um número *infinito* de conjuntos compactos é um conjunto compacto.

(q) A interseção de um número *finito* de conjuntos compactos é um conjunto compacto.

(r) A interseção de um número *infinito* de conjuntos compactos é um conjunto compacto.

[16] Dê um exemplo de um conjunto que é aberto e fechado e dê um exemplo de um conjunto que não nem aberto e nem fechado.

*[17] Seja $\mathbb{Q} \subset \mathbb{R}$ o subconjunto dos números racionais. $\mathbb{Q}$ é aberto? $\mathbb{Q}$ é fechado? $\mathbb{Q}$ é limitado? $\mathbb{Q}$ é compacto? Qual é o interior de $\mathbb{Q}$? Qual é a fronteira de $\mathbb{Q}$? Responda a estas mesmas perguntas para o subconjunto dos números irracionais.

*[18] Seja $f\colon \mathbb{R}^n \to \mathbb{R}$ uma função contínua em $\mathbf{p} = (p_1, p_2, \ldots, p_n)$. Considere a função $g\colon \mathbb{R} \to \mathbb{R}$ definida por $g(t) = f(t, p_2, \ldots, p_n)$. Mostre que g é contínua em p_1. Este resultado implica que se f é contínua, então sua *restrição* em qualquer reta paralela a um eixo coordenado é também contínua. Contudo, a recíproca deste resultado é falsa. Considere a função

$$z = f(x,y) = \begin{cases} \dfrac{x \cdot y^2}{x^2 + y^4} & \text{se } (x,y) \neq (0,0), \\ 0 & \text{se } (x,y) = (0,0). \end{cases}$$

(a) Mostre que $f_1(t) = f(t,p)$ e $f_2(t) = f(p,t)$ são funções contínuas em t para cada valor de p fixo.

(b) Mostre que f, por sua vez, não é contínua em $(0,0)$.

(Limite de funções) Intuitivamente falando, dizemos que uma função $f\colon D_f \subset \mathbb{R}^n \to \mathbb{R}$ possui limite $L \in \mathbb{R}$ em $\mathbf{p}$ (não necessariamente no domínio D_f de f) se $f(\mathbf{x})$ pode ser considerado uma aproximação de L tão boa quanto se queira desde que $\mathbf{x}$ esteja suficientemente próximo de $\mathbf{p}$. Mais formalmente:

Definição 4.13 Dizemos que $L \in \mathbb{R}$ é o limite de $f\colon D_f \subset \mathbb{R}^n \to \mathbb{R}$ em $\mathbf{p} \in \mathbb{R}^n$ se para toda seqüência $\mathbf{x}_n \to \mathbf{p}$, com $\mathbf{x}_n \in D_f$ e $\mathbf{x}_n \neq \mathbf{p}$, tem-se $f(\mathbf{x}_n) \to L$. Neste caso, escrevemos

$$\lim_{\mathbf{x} \to \mathbf{p}} f(\mathbf{x}) = L.$$

Estamos considerando os pontos $\mathbf{p}$ para os quais existe pelo menos uma seqüência $\mathbf{x}_n \in D_f$ que converge para $\mathbf{p}$ (nesta situação, dizemos que $\mathbf{p}$ é um *ponto de acumulação* de D_f).

[19] Calcule os limites abaixo, caso eles existam. Justifique sua resposta.

(a) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2y+y^3}{x^2+y^2}$.

(b) $\displaystyle\lim_{(x,y)\to(1,1)} \frac{x+y}{(x-1)^2+1}$.

(c) $\displaystyle\lim_{(x,y)\to(2,3)} \frac{x^2y+y^3+x^2+y^2}{\sqrt{x^2+y^2}}$.

(d) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{4\,x^2+3\,y^2+x^3y^3}{x^2+y^2+x^4y^4}$.

(e) $\displaystyle\lim_{(x,y)\to(1,1)} (e^x\cos(\pi y))$.

(f) $\displaystyle\lim_{(x,y)\to(0,1)} (e^{xy}\cos(\pi xy))$.

(g) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{1}{1+\ln(1+1/(x^2+y^2))}$.

(Seqüências em $\mathbb{R}^m$) Como vimos, uma seqüência em $\mathbb{R}^m$ é uma aplicação $\mathbf{x}\colon \mathbb{N}\to\mathbb{R}^m$ que a cada $n\in\mathbb{N}$ associa uma n-upla $\mathbf{x}_n=(x_{1,n},x_{2,n},\ldots,x_{m,n})$ em $\mathbb{R}^m$. Observe que $\mathbf{x}_n$ nada mais é do que uma coleção de N seqüências numéricas:

$$\begin{aligned} x_1\colon \mathbb{N} &\to \mathbb{R} \\ n &\mapsto x_{1,n} \end{aligned}$$

$$\begin{aligned} x_2\colon \mathbb{N} &\to \mathbb{R} \\ n &\mapsto x_{2,n} \end{aligned}$$

$$\vdots$$

$$\begin{aligned} x_m\colon \mathbb{N} &\to \mathbb{R} \\ n &\mapsto x_{m,n} \end{aligned}$$

Como em Cálculo I, podemos definir o conceito de convergência: $\mathbf{x}_n$ converge para $\mathbf{p}=(p_1,p_2,\ldots,p_n)$ se, e somente se, cada seqüência numérica $x_{i,n}$ converge para p_i, isto é, $x_{1,n}\to p_1$, $x_{2,n}\to p_2$, $\ldots$, $x_{m,n}\to p_m$. Uma outra maneira de se caracterizar seqüências convergentes em $\mathbb{R}^m$ é através da função distância e das bolas fechadas.

> Dizemos que $\mathbf{x}_n$ converge para $\mathbf{p}$ se, dado qualquer $K \in \mathbb{N}$, existe um índice $n_0 \in \mathbb{N}$ tal que para todo $n \geq n_0$ tem-se que a distância entre $\mathbf{x}_n$ e $\mathbf{p}$ é menor ou igual a $1/10^K$.

Ora, dizer que a distância entre $\mathbf{x}_n$ e $\mathbf{p}$ é menor ou igual a $1/10^K$ é justamente dizer que $\mathbf{x}_n$ está na bola fechada $\overline{B_{1/10^K}(\mathbf{p})}$ de centro em $\mathbf{p}$ e raio $1/10^K$. Assim:

> Dizemos que $\mathbf{x}_n$ converge para $\mathbf{p}$ se, dado qualquer $K \in \mathbb{N}$, existe um índice $n_0 \in \mathbb{N}$ tal que para todo $n \geq n_0$ tem-se que $\mathbf{x}_n \in \overline{B_{1/10^K}(\mathbf{p})}$, isto é, o elemento $\mathbf{x}_n$ da seqüência está na bola fechada de centro em $\mathbf{p}$ e raio $1/10^K$.

Intuitivamente falando, uma seqüência x_n converge para $\mathbf{p}$ se, dada qualquer bola fechada de centro em $\mathbf{p}$ (e por menor que seja o raio desta bola), existe um índice n_0 a partir do qual todos os elementos da seqüência ficam dentro da bola.

*[20] Mostre que $\mathbf{x}_n \to \mathbf{p}$ segundo a definição que vimos agora se, e somente se, cada seqüência $x_{i,n} \to p_i$, $i = 1, \ldots, N$.

*[21] Mostre que se $f\colon \mathbb{R}^n \to \mathbb{R}$ é contínua em $\mathbf{p} \in \mathbb{R}^n$ e $f(\mathbf{p}) > 0$, então existe uma bola fechada $\overline{B_r(\mathbf{p})}$ de centro em $\mathbf{p}$ e raio $r > 0$ tal que $f(\mathbf{q}) > 0$ para todo $\mathbf{q} \in \overline{B_r(\mathbf{p})}$.

Vamos agora generalizar o conceito de distância introduzido na seção (4.3).

Definição 4.14 (DISTÂNCIA EM $\mathbb{R}^n$) Uma *distância* em $\mathbb{R}^n$ é uma função $d\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ que satisfaz três propriedades:

(1) $d(\mathbf{x}, \mathbf{y}) \geq 0$ e $d(\mathbf{x}, \mathbf{x}) = 0$, para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

(2) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$, para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

(3) $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$, para todo $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$.

A última condição é conhecida como *desigualdade triangular*.

[22] Mostre que $d\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definida por

$$d(\mathrm{x}, \mathrm{y}) = \max\{|x_1 - y_1|, |x_2 - y_2|\},$$

com $\mathrm{x} = (x_1, x_2)$ e $\mathrm{y} = (y_1, y_2)$, é uma função distância em $\mathbb{R}^2$. Faça um desenho da bola de centro em $(0, 0)$ e raio 1 construída com esta distância. Mais geralmente, mostre que $d\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definida por

$$d(\mathrm{x}, \mathrm{y}) = \max\{|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n|\},$$

com $\mathrm{x} = (x_1, x_2, \ldots, x_n)$ e $\mathrm{y} = (y_1, y_2, \ldots, y_n)$ é uma distância em $\mathbb{R}^n$.

[23] Mostre que $d\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definida por

$$d(\mathrm{x}, \mathrm{y}) = |x_1 - y_1| + |x_2 - y_2|,$$

com $\mathrm{x} = (x_1, x_2)$ e $\mathrm{y} = (y_1, y_2)$, é uma função distância em $\mathbb{R}^2$. Faça um desenho da bola de centro em $(0, 0)$ e raio 1 construída com esta distância. Mais geralmente, mostre que $d\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definida por

$$d(\mathrm{x}, \mathrm{y}) = |x_1 - y_1| + |x_2 - y_2| + \cdots + |x_n - y_n|,$$

com $\mathrm{x} = (x_1, x_2, \ldots, x_n)$ e $\mathrm{y} = (y_1, y_2, \ldots, y_n)$ é uma distância em $\mathbb{R}^n$.

*[24] Mostre que $d\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definida por

$$d(\mathrm{x}, \mathrm{y}) = \begin{cases} 0 & \text{se } \mathrm{x} = \mathrm{y}, \\ 1 & \text{se } \mathrm{x} \neq \mathrm{y}. \end{cases}$$

é uma função distância em $\mathbb{R}^2$. Faça um desenho da bola de centro em $(0, 0)$ e raio 1 construída com esta distância.

*[25] O objetivo deste exercício é demonstrar o teorema (4.3), isto é, provar que a distância euclidiana definida por

$$d(\mathrm{x}, \mathrm{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2},$$

com $\mathrm{x} = (x_1, x_2, \ldots, x_n)$ e $\mathrm{y} = (y_1, y_2, \ldots, y_n)$ atende as três propriedades da definição de função distância.

(a) Mostre que $d(\mathrm{x}, \mathrm{x}) = 0$ para todo $\mathrm{x} \in \mathbb{R}^n$.

(b) Mostre que $d(\mathrm{x}, \mathrm{y}) = d(\mathrm{y}, \mathrm{x})$ para todo $\mathrm{x}, \mathrm{y} \in \mathbb{R}^n$.

(A desigualdade de Cauchy-Schwarz) Para demonstrarmos que a distância euclidiana satisfaz a desigualdade triangular, vamos precisar de uma desigualdade auxiliar, conhecida como desigualdade de Cauchy-Schwarz:

$$\sum_{i=1}^{n} |x_i \cdot y_i| \leq \sqrt{\sum_{i=1}^{n} x_i^2} \cdot \sqrt{\sum_{i=1}^{n} y_i^2}.$$

Para simplificar, vamos usar a notação:

$$||\mathrm{x}|| = \sqrt{\sum_{i=1}^{n} x_i^2} \qquad \text{e} \qquad ||\mathrm{y}|| = \sqrt{\sum_{i=1}^{n} y_i^2}.$$

(c) Mostre que $||\mathrm{x}|| = 0$ se, e somente se, $x_i = 0$ para todo $i = 1, \ldots, n$.

(d) Mostre que a desigualdade de Cauchy-Schwarz é imediata se $||\mathrm{x}|| = 0$ ou $||\mathrm{y}|| = 0$.

Assim, resta demonstrarmos a desigualdade de Cauchy-Schwarz para pontos $\mathbf{x}$ e $\mathbf{y}$ tais que $||\mathrm{x}|| \neq 0$ e $||\mathrm{y}|| \neq 0$. Ora, para quaisquer números reais a e b, $0 \leq (a-b)^2 = a^2 - 2 \cdot a \cdot b + b^2$ ou, equivalentemente,

$$2 \cdot a \cdot b \leq a^2 + b^2. \tag{$*$}$$

Como a e b são números arbitrários, podemos fazer $a = |x_i|/||\mathrm{x}||$ e $b = |y_i|/||\mathrm{y}||$ em ($*$). Assim, para qualquer i

$$2 \cdot \frac{|x_i|}{||\mathrm{x}||} \cdot \frac{|y_i|}{||\mathrm{y}||} \leq \frac{|x_i|^2}{||\mathrm{x}||^2} + \frac{|y_i|^2}{||\mathrm{y}||^2}. \tag{$**$}$$

Então, somando ($**$) com relação a i e usando $|x_i \cdot y_i| = |x_i| \cdot |y_i|$, temos

$$2 \cdot \frac{\sum_{i=1}^{n} |x_i \cdot y_i|}{||\mathrm{x}|| \cdot ||\mathrm{y}||} \leq \frac{\sum_{i=1}^{n} |x_i|}{||\mathrm{x}||} + \frac{\sum_{i=1}^{n} |y_i|}{||\mathrm{y}||} = \frac{||\mathrm{x}||}{||\mathrm{x}||} + \frac{||\mathrm{y}||}{||\mathrm{y}||},$$

isto é,

$$2 \cdot \frac{\sum_{i=1}^{n} |x_i \cdot y_i|}{||\mathrm{x}|| \cdot ||\mathrm{y}||} \leq 1.$$

Multiplicando-se ambos os membros por $||\mathrm{x}|| \cdot ||\mathrm{y}||$, temos a desigualdade de Cauchy-Schwarz.

Com a desigualdade de Cauchy-Schwarz podemos facilmente demonstrar a desigualdade triangular para a distância euclidiana:

$$d(\mathrm{x}, \mathrm{y}) \leq d(\mathrm{x}, \mathrm{z}) + d(\mathrm{z}, \mathrm{y}).$$

Observe que se $d(\mathrm{x}, \mathrm{y}) = 0$, então a desigualdade é imediata. Suponha então que $d(\mathrm{x}, \mathrm{y}) > 0$. Agora,

$$d(\mathrm{x}, \mathrm{y})^2 = \sum_{i=1}^{n}(x_i - y_i)^2 = \sum_{i=1}^{n}((x_i - z_i) + (z_i - y_i))^2 =$$

$$\sum_{i=1}^{n} |(x_i - z_i) + (z_i - y_i)| \cdot |(x_i - z_i) + (z_i - y_i)|$$

Pela desigualdade triangular para números reais temos que

$$|(x_i - z_i) + (z_i - y_i)| \leq |x_i - z_i| + |z_i - y_i|.$$

Assim,

$$d(\mathrm{x}, \mathrm{y})^2 \leq \sum_{i=1}^{n} |x_i - y_i| \cdot (|x_i - z_i| + |z_i - y_i|) =$$

$$\sum_{i=1}^{n} |x_i - y_i| \cdot |x_i - z_i| + \sum_{i=1}^{n} |x_i - y_i| \cdot |z_i - y_i|.$$

Mas, pela desigualdade de Cauchy-Schwarz,

$$\sum_{i=1}^{n} |x_i - y_i| \cdot |x_i - z_i| \leq \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \cdot \sqrt{\sum_{i=1}^{n}(x_i - z_i)^2} \quad \text{e}$$

$$\sum_{i=1}^{n} |x_i - y_i| \cdot |z_i - y_i| \leq \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \cdot \sqrt{\sum_{i=1}^{n}(z_i - y_i)^2},$$

isto é,

$$\sum_{i=1}^{n} |x_i - y_i| \cdot |x_i - z_i| \leq d(\mathrm{x}, \mathrm{y}) \cdot d(\mathrm{x}, \mathrm{z}) \qquad \text{e}$$

$$\sum_{i=1}^{n} |x_i - y_i| \cdot |z_i - y_i| \leq d(\mathrm{x}, \mathrm{y}) \cdot d(\mathrm{z}, \mathrm{y}).$$

Então $d(\mathrm{x}, \mathrm{y})^2 \leq d(\mathrm{x}, \mathrm{y}) \cdot d(\mathrm{x}, \mathrm{z}) + d(\mathrm{x}, \mathrm{y}) \cdot d(\mathrm{z}, \mathrm{y})$. Como estamos supondo $d(\mathrm{x}, \mathrm{y}) > 0$, podemos dividir por $d(\mathrm{x}, \mathrm{y})$ e obter a desigualdade triangular $d(\mathrm{x}, \mathrm{y}) \leq d(\mathrm{x}, \mathrm{z}) + d(\mathrm{z}, \mathrm{y})$ para a distância euclidiana d.

[26] O toro do exercício [42] da página 119, definido como o nível $w = 0$ da função

$$w = f(x, y, z) = (x^2 + y^2 + z^2 + R^2 - r^2)^2 - 4\,R^2\,(x^2 + y^2),$$

é um conjunto compacto em $\mathbb{R}^3$?

Secção 1 - 2

Capítulo 5

Derivadas parciais

Suponha que a função $z = f(x_1, x_2, \ldots, x_n)$ modele o lucro z de uma determinada empresa em termos dos insumos $x_1, x_2, \ldots, x_n$. Suponha ainda que se conheça o lucro z^* para uma escolha $(x_1^*, x_2^*, \ldots, x_n^*)$ dos insumos: $z^* = f(x_1^*, x_2^*, \ldots, x_n^*)$. Uma pergunta que surge imediatamente é como variações nos valores dos insumos $x_1^*, x_2^*, \ldots, x_n^*$ afetam o lucro da empresa. Por exemplo, aumentando-se *apenas* a quantidade do primeiro insumo e mantendo-se constantes os demais valores, o lucro da empresa estará aumentando ou diminuindo? E se fizéssemos isto para os outros tipos de insumo? O que seria mais lucrativo? Por que não variar todos os insumos ao mesmo tempo? Neste capítulo vamos introduzir os objetos matemáticos que permitem encontrar as respostas para estas perguntas.

5.1 Lembrando Cálculo I

Você aprendeu em Cálculo I que a derivada de uma função $f\colon \mathbb{R} \to \mathbb{R}$ mede o quanto a função está variando: se $(df/dx)(x_0) > 0$ e df/dx é contínua, então f é crescente em um intervalo aberto contendo x_0. Do mesmo modo, se $(df/dx)(x_0) < 0$ e df/dx é contínua, então f é decrescente em um intervalo aberto contendo x_0. Quanto maior o valor da derivada, maior é a taxa de variação. Mais ainda, você viu que o valor da derivada $\frac{df}{dx}(x_0)$ no ponto x_0 é o coeficiente angular da reta tangente ao gráfico de f neste ponto e este valor pode ser obtido através do limite do coeficiente angular da reta secante ao gráfico de f pelos pontos $(x_0, f(x_0))$ e $(x, f(x))$ quando $x \to x_0$:

$$\frac{df}{dx}(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

Podemos ainda reescrever este limite com a mudança de variável $x = x_0 + h$:

$$\frac{df}{dx}(x_0) = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h}.$$

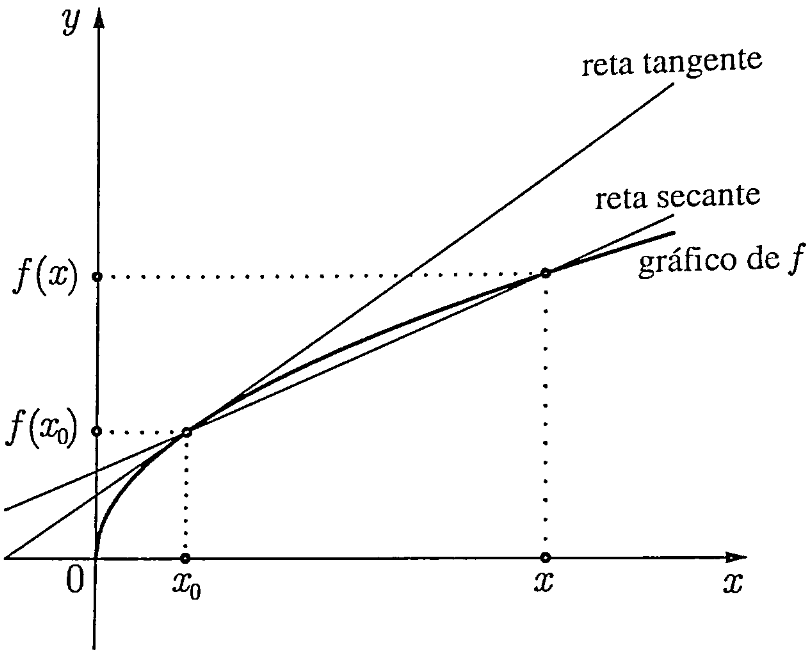

Figura 5.1: Interpretação geométrica da derivada de uma função de uma variável.

5.2 Definições e exemplos

Para começar, considere uma função

$$\begin{array}{rccl} f\colon & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x,y) & \mapsto & z = f(x,y) \end{array}$$

e $\mathbf{p} = (a,b)$. O que acontece com o valor $z = f(x,y)$ da função quando variamos x nas proximidades de a e mantemos y constante e igual a b? Ora, se $y = b$, então podemos construir uma nova função de uma única variável a partir de nossa função de duas variáveis:

$$\begin{array}{rccl} g\colon & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & z = g(x) = f(x,b) \end{array}.$$

Para saber a taxa de variação de f no ponto (a,b) mantendo-se $y = b$ constante, basta derivar g no ponto $x = a$:

$$\frac{dg}{dx}(a) = \lim_{x \to a} \frac{g(x) - g(a)}{x - a} = \lim_{x \to a} \frac{f(x,b) - f(a,b)}{x - a}.$$

Este limite é tão importante que vamos usar um símbolo especial para ele:

$$\frac{\partial f}{\partial x}(a,b) = \lim_{x \to a} \frac{f(x,b) - f(a,b)}{x - a}.$$

$(\partial f/\partial x)(a,b)$ é denominada *derivada parcial* de f *com relação* a x no ponto (a,b). Geometricamente, estamos fazendo a *restrição* de f sobre a reta $y = b$ e olhando para a curva correspondente sobre o gráfico de f. Desta maneira, o número $(\partial f/\partial x)(a,b) = (dg/dx)(a)$ é o coeficiente angular da reta tangente a esta curva no ponto (a,b) (veja a figura (5.2)).

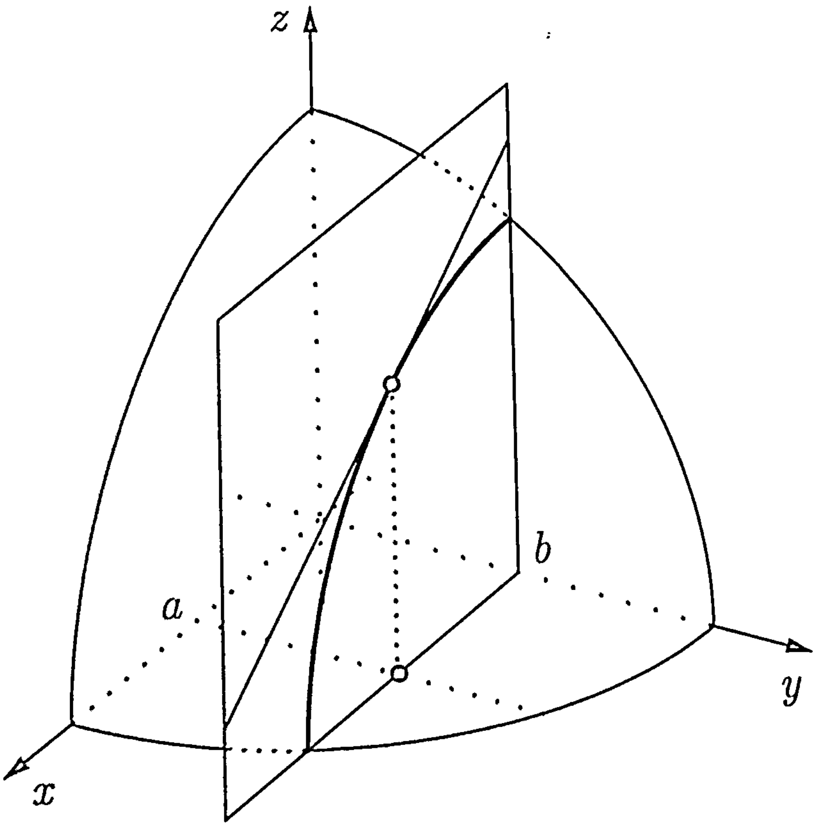

Figura 5.2: Interpretação geométrica da derivada parcial de $z = f(x,y)$ com relação a x.

Analogamente, podemos perguntar o que acontece com o valor da função $z = f(x,y)$ quando variamos y nas proximidades de b e mantemos o valor de x constante e igual a a. Novamente, se $x = a$, então podemos construir uma nova função de uma única variável a partir de nossa função de duas variáveis:

$$\begin{array}{rcl} h\colon \mathbb{R} & \to & \mathbb{R} \\ y & \mapsto & z = h(y) = f(a,y) \end{array}.$$

Para saber a taxa de variação de f no ponto (a,b), mantendo-se $x = a$ constante, basta derivar h no ponto $y = b$:

$$\frac{dh}{dy}(b) = \lim_{y\to b}\frac{h(y)-h(b)}{y-b} = \lim_{y\to b}\frac{f(a,y)-f(a,b)}{y-b}.$$

Como antes, vamos usar um símbolo especial para denotar este limite:

$$\frac{\partial f}{\partial y}(a,b) = \lim_{y\to b}\frac{f(a,y)-f(a,b)}{y-b}.$$

$(\partial f/\partial y)(a,b)$ é denominada *derivada parcial* de f *com relação* a y no ponto (a,b). Geometricamente, estamos fazendo a restrição de f sobre a reta $x = a$ e olhando para a curva correspondente sobre o gráfico de f. Assim, $(\partial f/\partial y)(a,b) = (dh/dy)(b)$ é o coeficiente angular da reta tangente a esta curva no ponto (a,b) (veja a figura (5.3)).

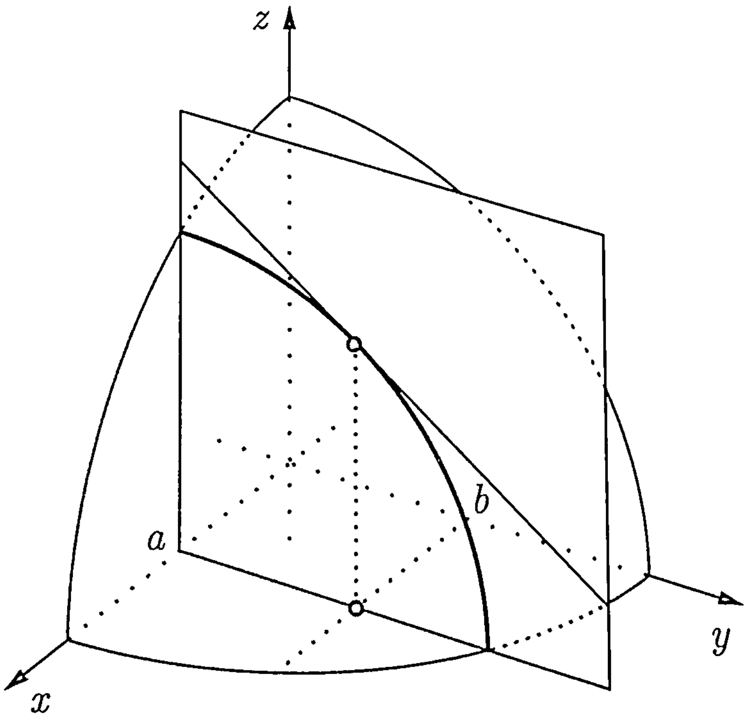

Figura 5.3: Interpretação geométrica da derivada parcial de $z = f(x,y)$ com relação a y.

Como $(\partial f/\partial x)(a,b)$ e $(\partial f/\partial y)(a,b)$ são derivadas, você pode aplicar tudo o que aprendeu em Cálculo I para concluir propriedades sobre o comportamento de f. Por exemplo, para funções f tais que suas derivadas parciais $\partial f/\partial x$ e $\partial f/\partial y$ são contínuas, vale que:

- Se $(\partial f/\partial x)(a,b) > 0$, então f cresce quando mantemos $y = b$ constante e aumentamos localmente o valor de x com relação a a.
- Se $(\partial f/\partial y)(a,b) > 0$, então f cresce quando mantemos $x = a$ constante e aumentamos localmente o valor de y com relação a b.
- Se $(\partial f/\partial x)(a,b) < 0$, então f decresce quando mantemos $y = b$ constante e aumentamos localmente o valor de x com relação a a.
- Se $(\partial f/\partial y)(a,b) < 0$, então f decresce quando mantemos $x = a$ constante e aumentamos localmente o valor de y com relação a b.

Vamos ver como tudo isto funciona com um exemplo.

Exemplo 5.1 A função de produção *Cobb-Douglas*

$$z = f(x,y) = 4 \cdot x^{3/4} \cdot y^{1/4}$$

modela a produção de uma certa empresa em função do *capital* x e do *trabalho* y. Assim, por exemplo, com capital $x^* = 10000$ e trabalho $y^* = 625$, a produção desta firma é de

$$z^* = f(x^*, y^*) = f(10000, 625) = 4 \cdot \left(10^4\right)^{3/4} \cdot \left(5^4\right)^{1/4} = 20000$$

unidades. Vamos agora calcular a taxa de variação da produção no ponto $(10000, 625)$ quando mantemos a quantidade de trabalho constante e variamos a quantidade de capital [o que significa calcular $(\partial f/\partial x)(10000, 625)$] e, de maneira recíproca, calcular a taxa de variação da produção no ponto $(10000, 625)$ quando mantemos a quantidade de capital constante e variamos a quantidade de trabalho [o que significa calcular $(\partial f/\partial y)(10000, 625)$]. Para isto, vamos construir as funções auxiliares

$$\begin{aligned} z &= g(x) &= f(x, 625) &= 4 \cdot x^{3/4} \cdot (5^4)^{1/4} &= 20 \cdot x^{3/4}, \\ z &= h(y) &= f(10000, y) &= 4 \cdot (10^4)^{3/4} \cdot y^{1/4} &= 4000 \cdot y^{1/4}. \end{aligned}$$

Desta maneira,

$$\begin{aligned} \frac{\partial f}{\partial x}(10000, 625) &= \frac{dg}{dx}(10000) &= 15 \cdot x^{-1/4}\Big|_{x=10000} &= 1.5, \\ \frac{\partial f}{\partial y}(10000, 625) &= \frac{dh}{dy}(625) &= 1000 \cdot y^{-3/4}\Big|_{y=625} &= 8. \end{aligned}$$

Uma vez que $(\partial f/\partial x)(10000, 625) = 1.5$, se a quantidade de trabalho y^* é mantida constante e aumentamos a quantidade de capital x^* em h unidades, então a produção sofrerá um aumento de aproximadamente $1.5 \cdot h$ unidades. Por exemplo, para um *acréscimo* de 10 unidades no capital x^*, a produção aumentará em aproximadamente 15 unidades. Analogamente, uma vez que $(\partial f/\partial y)(10000, 625) = 8$, para um *decréscimo* de 2 unidades no trabalho y^*, a produção diminuirá em aproximadamente 16 unidades.

Para esta empresa, no ponto $(x^*, y^*) = (10000, 625)$, entre aumentar o capital ou aumentar o trabalho, é mais produtivo aumentar o trabalho. Contudo, como veremos, a estratégia mais produtiva de todas é aumentar o trabalho e o capital ao mesmo tempo na proporção $1.5/8 = 0.1875$ (veja o capítulo 8 sobre derivadas direcionais e o vetor gradiente). □

Dada uma função f e um ponto (a, b), podemos calcular as derivadas parciais

$$\frac{\partial f}{\partial x}(a,b) = \lim_{x\to a}\frac{f(x,b)-f(a,b)}{x-a} \quad \text{e}$$

$$\frac{\partial f}{\partial y}(a,b) = \lim_{y\to b}\frac{f(a,y)-f(a,b)}{y-b}$$

(caso os limites existam) associadas ao ponto (a, b):

$$(a,b)\mapsto \frac{\partial f}{\partial x}(a,b) \qquad \text{e} \qquad (a,b)\mapsto \frac{\partial f}{\partial y}(a,b).$$

Se fizermos esta associação para cada ponto (a, b) de $\mathbb{R}^2$ (supondo que os limites sempre existam) estaremos criando duas novas funções de *duas variáveis*:

$$\begin{aligned}\frac{\partial f}{\partial x}:\ \mathbb{R}^2 &\to \mathbb{R}\\ (x,y) &\mapsto \frac{\partial f}{\partial x}(x,y)\end{aligned} \qquad \text{e} \qquad \begin{aligned}\frac{\partial f}{\partial y}:\ \mathbb{R}^2 &\to \mathbb{R}\\ (x,y) &\mapsto \frac{\partial f}{\partial y}(x,y)\end{aligned},$$

que a cada ponto (x, y) em $\mathbb{R}^2$ associam, respectivamente, as derivadas parciais com relação a x e a y no ponto (x, y).

Mais ainda, a própria definição de derivadas parciais sugere um método prático para calcular estas funções de derivada parcial: *quando queremos derivar f com relação a x, consideramos y como uma constante e efetuamos a derivada como se estivéssemos em Cálculo I e, analogamente, quando* *queremos derivar f com relação a y, consideramos x como uma constante e procedemos normalmente com o cálculo da derivada.*

Exemplo 5.2 Considere a função $z = f(x,y) = 3\,x^2y^2 + 4\,xy^3 + 7\,y$. Vamos calcular $\partial f/\partial x$ considerando y como uma constante. A primeira parcela é x^2 vezes uma "constante" $(3\,y^2)$ de modo que sua derivada é $2\,x$ vezes a constante, isto é

$$\frac{\partial}{\partial x}\left(3\,x^2y^2\right) = 2\,x \cdot 3y^2 = 6\,xy^2.$$

A segunda parcela é uma "constante" vezes x, assim, sua derivada é a própria constante

$$\frac{\partial}{\partial x}\left(4\,xy^3\right) = 4\,y^3.$$

Finalmente, desde que consideramos $7\,y$ como uma constante no cálculo da derivada parcial com relação a x, segue-se que sua derivada é zero:

$$\frac{\partial}{\partial x}(7\,y) = 0.$$

Colocando tudo isto junto e usando o fato de que a derivada da soma é a soma das derivadas, obtemos que

$$\frac{\partial}{\partial x}\left(3\,x^2y^2 + 4\,xy^3 + 7\,y\right) = 6\,xy^2 + 4\,y^3.$$

Para calcular a derivada parcial com relação a y, tratamos x como uma constante. A primeira parcela de f é y^2 vezes uma constante de modo que sua derivada com relação a y é $2\,y$ vezes a constante:

$$\frac{\partial}{\partial y}\left(3\,x^2y^2\right) = (2\,y)\cdot\left(3\,x^2\right) = 6\,x^2y.$$

A segunda parcela é y^3 vezes uma constante, assim, sua derivada com relação a y é $3\,y^2$ vezes a constante:

$$\frac{\partial}{\partial y}\left(4\,xy^3\right) = \left(3\,y^2\right)\cdot(4\,x) = 12\,xy^2.$$

Finalmente, a derivada com relação a y de $7\,y$ é, naturalmente, 7. Colocando tudo isto junto, encontramos que

$$\frac{\partial}{\partial y}\left(3\,x^2y^2 + 4\,xy^3 + 7\,y\right) = 6\,x^2y + 12\,xy^2 + 7.$$

□

Para o caso de uma função f de três variáveis x, y e z, temos três derivadas parciais, uma para cada variável:

$$\begin{aligned}
\frac{\partial f}{\partial x}(a,b,c) &= \lim_{x\to a}\frac{f(x,b,c)-f(a,b,c)}{x-a},\\
\frac{\partial f}{\partial y}(a,b,c) &= \lim_{y\to a}\frac{f(a,y,c)-f(a,b,c)}{y-b},\\
\frac{\partial f}{\partial z}(a,b,c) &= \lim_{z\to c}\frac{f(a,b,z)-f(a,b,c)}{z-c}.
\end{aligned}$$

A regra prática continua valendo: quando derivamos f com relação a x consideramos y e z constantes, quando derivamos f com relação a y consideramos x e z constantes e quando derivamos f com relação a z consideramos x e y constantes. Mais geralmente, para uma função de n variáveis. temos a definição a seguir.

Definição 5.1 (Derivada parcial) Seja $f : \mathbb{R}^n \to \mathbb{R}$. Então para cada variável x_i em cada ponto $\mathrm{p} = (p_1, \ldots, p_n)$ no domínio de f, definimos a *derivada parcial de f com relação a x_i no ponto* p como o número

$$\frac{\partial f}{\partial x_i}(p_1,\ldots,p_n) = \lim_{x_i\to x_i^*}\frac{f(p_1,\ldots,x_i,\ldots,p_n)-f(p_1,\ldots,p_i,\ldots,p_n)}{x_i-p_i},$$

caso o limite exista. Somente a i-ésima coordenada sofre variação. As demais coordenadas permanecem constantes.

Com a mudança de variável $x_i = p_i + h$, podemos escrever ainda

$$\frac{\partial f}{\partial x_i}(p_1,\ldots,p_n) = \lim_{h\to 0}\frac{f(p_1,\ldots,p_i+h,\ldots,p_n)-f(p_1,\ldots,p_i,\ldots,p_n)}{h}.$$

Existem outras duas notações para a derivada parcial de uma função $z = f(x_1, \ldots, x_n)$ com relação a x_i:

$$\frac{\partial f}{\partial x_i}(p_1,\ldots,p_n) = f_{x_i}(p_1,\ldots,p_n) = D_i f(p_1,\ldots,p_n).$$

CUIDADO! CUIDADO! CUIDADO!

O "x" que aparece na notação $\partial f/\partial x$ serve apenas para indicar que estamos derivando a função f com relação à *primeira* variável! Por exemplo, as duas funções abaixo

$$z = f(x, y) = 4 \cdot x^{3/4} \cdot y^{1/4} \qquad \text{e} \qquad u = g(s, t) = 4 \cdot s^{3/4} \cdot t^{1/4}$$

são *iguais* pois elas possuem o mesmo domínio, o mesmo contradomínio e a mesma regra de associação, apesar de aparecerem "letras" diferentes na definição de cada uma. Assim, tanto $\partial f/\partial x$ quanto $\partial g/\partial s$ querem dizer a mesma coisa: a derivada parcial da (mesma) função com relação à *primeira* variável. Neste sentido, a notação $D_1 f$ é a mais clara, uma vez que ela não se compromete com o uso de "variáveis auxiliares".

A definição (5.1) de derivada parcial pressupõe que a função f esteja definida em *todo* $\mathbb{R}^n$. Contudo, é possível fazer uma extensão desta definição para funções mais gerais: se $f\colon D_f \subset \mathbb{R}^n \to \mathbb{R}$ e $\mathbf{p} = (p_1, \ldots, p_n) \in D_f$, então basta garantirmos que seja possível avaliar o *quociente de Newton*

$$\frac{f(p_1, \ldots, x_i, \ldots, p_n) - f(p_1, \ldots, p_i, \ldots, p_n)}{x_i - p_i}$$

em pontos $(p_1, \ldots, x_i, \ldots, p_n)$ para todo x_i suficientemente próximo de p_i e, sendo assim, podemos indagar se o limite

$$\lim_{x_i \to x_i^*} \frac{f(p_1, \ldots, x_i, \ldots, p_n) - f(p_1, \ldots, p_i, \ldots, p_n)}{x_i - p_i}$$

existe ou não. Caso ele exista, definimos o seu valor como sendo a derivada parcial de f com relação a x_i no ponto $\mathbf{p}$.

Geometricamente, dizer que é possível avaliar o quociente de Newton para todos os pontos $(p_1, \ldots, x_i, \ldots, p_n)$ próximos de $(p_1, \ldots, p_i, \ldots, p_n)$ significa dizer que existe um segmento de reta paralelo ao eixo x_i inteiramente contido em D_f, onde $\mathbf{p}$ é o ponto médio deste segmento. Mais informalmente, podemos "caminhar" um pouco para frente e para trás, a partir do ponto $\mathbf{p}$, paralelamente ao eixo x_i, sem sairmos do domínio D_f de f.

Por exemplo, considere uma função $f\colon D_f \subset \mathbb{R}^2 \to \mathbb{R}$ definida em um conjunto D_f dado na figura (5.4). Podemos definir as duas derivadas parciais de f no ponto $\mathbf{p}$. Agora, para o ponto $\mathbf{q}$, podemos definir apenas a derivada parcial de f com relação a x. No ponto $\mathbf{r}$ não podemos definir nem a derivada parcial com relação a x e nem a derivada parcial com relação a y.

Observe que sempre podemos definir todas as derivadas parciais em um ponto interior do domínio da função.

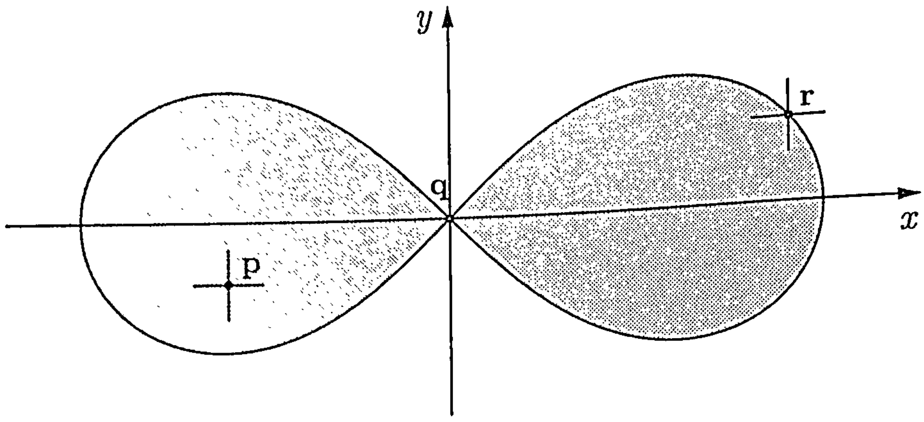

Figura 5.4: A relação entre a topologia do domínio de uma função e o conceito de derivada parcial.

CUIDADO! CUIDADO! CUIDADO!

Um abuso de notação cometido por muitos autores é o de se omitir os pontos onde as derivadas parciais são calculadas, com a justificativa de economia de espaço nas fórmulas. Assim, por exemplo, as derivadas parciais de $z = f(x,y) = 4\,x^{3/4}y^{1/4}$ são indicadas por

$$\frac{\partial f}{\partial x} = 3\,x^{-1/4}y^{1/4} \quad \text{e} \quad \frac{\partial f}{\partial y} = x^{-1/4}y^{1/4}$$

ao invés de

$$\frac{\partial f}{\partial x}(x,y) = 3\,x^{-1/4}y^{1/4} \quad \text{e} \quad \frac{\partial f}{\partial y}(x,y) = x^{-1/4}y^{1/4}.$$

Mais ainda, alguns autores trocam o nome da função pela variável que representa o valor da função em um ponto do domínio:

$$\frac{\partial z}{\partial x} = 3\,x^{-1/4}y^{1/4} \quad \text{e} \quad \frac{\partial z}{\partial y} = x^{-1/4}y^{1/4}.$$

Estes abusos de notação podem causar confusão (veja, por exemplo, as observações da página 278 sobre as complicações deste abuso de notação com relação à regra da cadeia). Em nosso texto, estes abusos foram evitados ao máximo e eles aparecem apenas em alguns exercícios.

Uma vez que a derivada parcial de uma função de várias variáveis nada mais é do que a derivada de uma função adequada de uma única variável, as propriedades da derivada com relação à soma, multiplicação e divisão transferem-se imediatamente para o caso de funções de várias variáveis.

Teorema 5.1 (Propriedades da derivada parcial) Considere $f\colon D_f \subset \mathbb{R}^n \to \mathbb{R}$, $g\colon D_g \subset \mathbb{R}^n \to \mathbb{R}$ e $\mathrm{p} \in D_f \cap D_g$. Suponha que as derivadas parciais

$$\frac{\partial f}{\partial x_i}(\mathrm{p}) \quad \text{e} \quad \frac{\partial g}{\partial x_i}(\mathrm{p})$$

de f e g com relação a x_i no ponto p existam. Então:

(a) A derivada parcial de $f+g$ com relação a x_i no ponto p existe e

$$\frac{\partial}{\partial x_i}(f+g)(\mathrm{p}) = \frac{\partial f}{\partial x_i}(\mathrm{p}) + \frac{\partial g}{\partial x_i}(\mathrm{p}).$$

(b) A derivada parcial de $f\cdot g$ com relação a x_i no ponto p existe e

$$\frac{\partial}{\partial x_i}(f\cdot g)(\mathrm{p}) = \frac{\partial f}{\partial x_i}(\mathrm{p})\cdot g(\mathrm{p}) + f(\mathrm{p})\cdot\frac{\partial g}{\partial x_i}(\mathrm{p}).$$

(c) Se $g(\mathrm{p}) \neq 0$, então a derivada parcial de f/g com relação a x_i no ponto p existe e

$$\frac{\partial}{\partial x_i}\left(\frac{f}{g}\right)(\mathrm{p}) = \frac{\dfrac{\partial f}{\partial x_i}(\mathrm{p})\cdot g(\mathrm{p}) - f(\mathrm{p})\cdot\dfrac{\partial g}{\partial x_i}(\mathrm{p})}{(g(\mathrm{p}))^2}.$$

5.3 Derivadas parciais de ordem superior

Dada uma função $z = f(x, y)$ de duas variáveis, suas derivadas parciais

$$\frac{\partial f}{\partial x} \quad \text{e} \quad \frac{\partial f}{\partial y}$$

também são funções de duas variáveis! Podemos então derivar parcialmente $\partial f/\partial x$ e $\partial f/\partial y$ gerando as derivadas parciais de *segunda ordem* da função f:

$$\begin{tikzcd}
 & & \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} \\
 & \frac{\partial f}{\partial x} \arrow[ur] \arrow[r] & \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} \\
f \arrow[ur] \arrow[dr] & & \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} \\
 & \frac{\partial f}{\partial y} \arrow[ur] \arrow[r] & \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}
\end{tikzcd}$$

Exemplo 5.3 No exemplo (5.2) calculamos as derivadas parciais da função $z = f(x, y) = 3\,x^2y^2 + 4\,xy^3 + 7\,y$:

$$\frac{\partial f}{\partial x}(x, y) = 6\,xy^2 + 4\,y^3 \qquad \text{e} \qquad \frac{\partial f}{\partial y}(x, y) = 6\,x^2y + 12\,xy^2 + 7.$$

Desta maneira,

$$\begin{aligned}
\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)(x, y) &= \frac{\partial}{\partial x}\left(6\,xy^2 + 4\,y^3\right) &&= 6\,y^2, \\
\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(x, y) &= \frac{\partial}{\partial y}\left(6\,xy^2 + 4\,y^3\right) &&= 12\,xy + 12\,y^2, \\
\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(x, y) &= \frac{\partial}{\partial x}\left(6\,x^2y + 12\,xy^2 + 7\right) &&= 12\,xy + 12\,y^2, \\
\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)(x, y) &= \frac{\partial}{\partial y}\left(6\,x^2y + 12\,xy^2 + 7\right) &&= 6\,x^2 + 24\,xy.
\end{aligned}$$

Portanto,

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 6\,y^2, \quad \frac{\partial^2 f}{\partial y \partial x}(x,y) = 12\,xy + 12\,y^2,$$

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = 12\,xy + 12\,y^2 \quad \text{e} \quad \frac{\partial^2 f}{\partial y^2}(x,y) = 6\,x^2 + 24\,xy. \qquad \square$$

Uma função de três variáveis possui 9 derivadas parciais de segunda ordem:

$$\frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial x \partial z}, \quad \frac{\partial^2 f}{\partial y \partial x}, \quad \frac{\partial^2 f}{\partial y^2}, \quad \frac{\partial^2 f}{\partial y \partial z}, \quad \frac{\partial^2 f}{\partial z \partial x}, \quad \frac{\partial^2 f}{\partial z \partial y}, \quad \frac{\partial^2 f}{\partial z^2}.$$

Uma função de n variáveis possui n^2 derivadas parciais de segunda ordem:

$$\frac{\partial^2 f}{\partial x_i \partial x_j},$$

para $i = 1, \ldots, n$ e $j = 1, \ldots, n$. A derivada de segunda ordem com relação a $x_i x_i$ é usualmente denotada por

$$\frac{\partial^2 f}{\partial x_i^2}$$

ao invés de $\partial^2 f/\partial x_i \partial x_i$. As derivadas parciais de segunda ordem da forma $\partial^2 f/\partial x_i \partial x_j$, com $i \neq j$, são denominadas *derivadas parciais mistas*. Como as derivadas parciais de segunda ordem de uma função de n variáveis também são funções de n variáveis, podemos derivá-las parcialmente e obter as *derivadas parciais de terceira ordem* de f. Este processo pode ser continuado a fim de obter as derivadas parciais de ordem superior, enquanto os limites que definem as derivadas parciais existirem. Observe que uma função de n variáveis possui n^3 derivadas de terceira ordem, n^4 derivadas parciais de quarta ordem, etc.

A definição a seguir estabelece um critério de suavidade da função em termos da continuidade das derivadas parciais de ordem superior.

Definição 5.2 (Função de classe C^k) Dizemos que uma função $f \colon D_f \subset \mathbb{R}^n \to \mathbb{R}$ é de *classe* C^k se todas as derivadas parciais de ordem $\leq k$ existem e são contínuas em D_f. Uma função é de *classe* C^0 se ela é contínua e é de *classe* C^∞ se possui todas as derivadas parciais de todas as ordens contínuas.

5.4 O teorema de Young

No exemplo (5.2), calculamos as derivadas parciais de segunda ordem de $z = f(x,y) = 3x^2y^2 + 4xy^3 + 7y$. Observe que

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = 12xy + 12y^2 = \frac{\partial^2 f}{\partial y \partial x}(x,y)$$

isto é, as duas derivadas parciais mistas de f são iguais. Isto não é uma coincidência!

Teorema 5.2 (Young) *Suponha que $z = f(x_1, \ldots, x_n)$ seja de classe C^2 em um conjunto aberto U de $\mathbb{R}^n$. Então, para todo $x \in U$ e para cada par de índices i, j, vale que*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x).$$

A demonstração do teorema de Young requer o uso do teorema do valor médio do Cálculo I. Você pode encontrá-la na referência [illegible], página [illegible].

Exemplo 5.4 Considere a função de produção Cobb-Douglas geral $z = f(x,y) = kx^a y^b$. Então

$$\frac{\partial f}{\partial x}(x,y) = akx^{a-1}y^b, \qquad \frac{\partial f}{\partial y}(x,y) = bkx^a y^{b-1},$$

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = abkx^{a-1}y^{b-1} \quad \text{e} \quad \frac{\partial^2 f}{\partial y \partial x}(x,y) = abkx^{a-1}y^{b-1},$$

com as duas últimas expressões iguais, como afirma o teorema de Young. ■

O teorema de Young pode ser generalizado para derivadas parciais de ordem superior. Por exemplo, se $z = f(x_1, x_2, x_3, x_4)$ é de classe C^3 então

$$\frac{\partial^3 f}{\partial x_1 \partial x_3 \partial x_4} = \frac{\partial^3 f}{\partial x_1 \partial x_4 \partial x_3} = \frac{\partial^3 f}{\partial x_3 \partial x_1 \partial x_4} =$$

$$\frac{\partial^3 f}{\partial x_3 \partial x_4 \partial x_1} = \frac{\partial^3 f}{\partial x_4 \partial x_1 \partial x_3} = \frac{\partial^3 f}{\partial x_4 \partial x_3 \partial x_1}.$$

Notação. Outra notação freqüentemente usada para derivadas parciais de segunda ordem inclui

$$\frac{\partial^2 f}{\partial x \partial y} = f_{yx} = D_{2,1} f \qquad \text{e} \qquad \frac{\partial^2 f}{\partial x_i \partial x_j} = f_{x_j x_i} = D_{ji} f.$$

Estes mesmos truques de notação também são usados para indicar derivadas parciais de ordem superior, tais como f_{ijk} e assim por diante.

5.5 Exercícios

[01] Determine as derivadas parciais de primeira ordem das funções de duas variáveis abaixo.

(a) $z = f(r,s) = \sqrt{r^2 + s^2}$.

(b) $z = f(s,t) = t/s - s/t$.

(c) $z = f(x,y) = 2\,x^4 y^3 - xy^2 + 3\,y + 1$.

(d) $z = f(t,v) = \ln \sqrt{(t+v)/(t-v)}$.

(e) $z = f(x,y) = (x^3 - y^2)^2$.

(f) $z = f(x,y) = xe^y + y\,\mathrm{sen}(x)$.

(g) $z = f(x,y) = e^x \ln(xy)$.

(h) $z = f(x,y) = x\cos(x/y)$.

[02] Determine as derivadas parciais de primeira ordem das funções de três variáveis abaixo.

(a) $w = f(x,y,t) = (x^2 - t^2)/(1 + \mathrm{sen}(3\,y))$.

(b) $w = f(x,y,z) = (y^2 + z^2)^x$.

(c) $w = f(r,s,v) = (2\,r + 3\,s)^{\cos(v)}$.

(d) $w = f(r,s,t) = r^2 e^{2\,s}\cos(t)$.

(e) $w = f(x,y,z) = xe^z - ye^x + ze^{-y}$.

(f) $w = f(x,y,z) = xyze^{xyz}$.

[03] Na figura (5.5) encontram-se os gráficos de três funções de duas variáveis:

$$f, \quad \partial f/\partial x \quad \text{e} \quad \partial f/\partial y.$$

Faça uma identificação entre os gráficos e as funções.

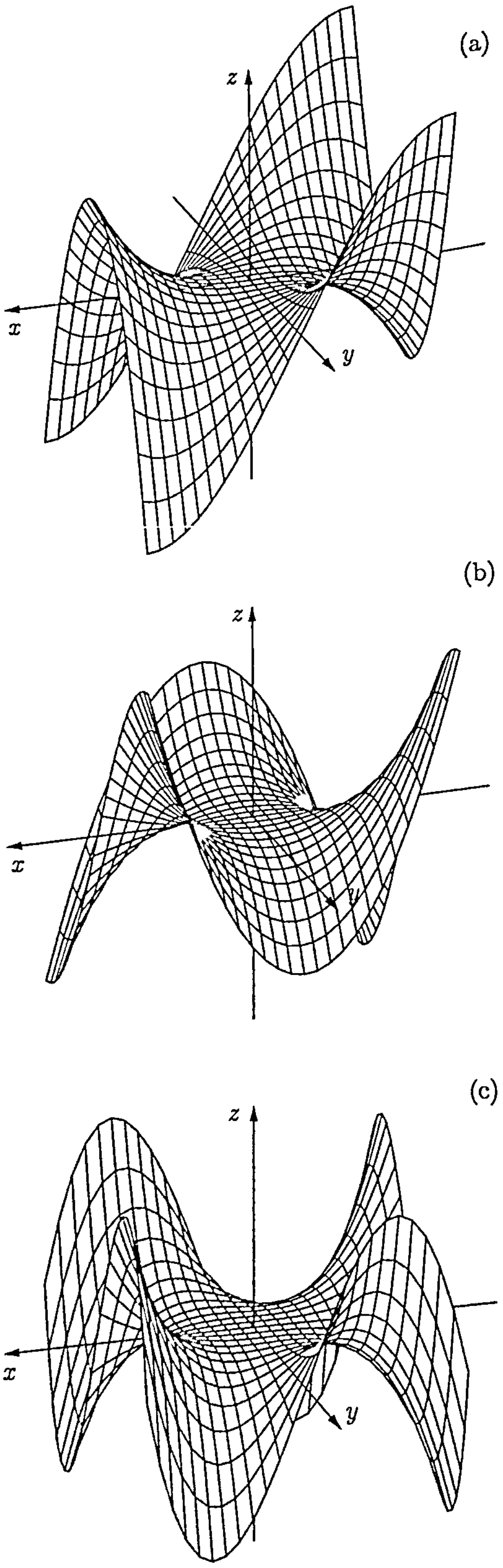

Figura 5.5: Os gráficos de f, $\partial f/\partial x$ e $\partial f/\partial y$.

[04] Para cada uma das funções abaixo verifique que $w_{xy} = w_{yx}$.

(a) $w = xyz$.

(b) $w = xy^4 - 2x^2y^3 + 4x^2 - 3y$.

(c) $w = x^2/(x+y)$.

(d) $w = x^3e^{-2y} + y^{-2}\cos(x)$.

(e) $w = y^2e^{x^2} + 1/(x^2y^3)$.

(f) $w = \sqrt{x^2+y^2+z^2}$.

[05] Resolva as questões abaixo.

(a) Calcule w_{xyz} para $w = 3x^2y^3z + xy^4z^2 - yz$.

(b) Calcule w_{tut} para $w = u^4vt^2 - 3uv^2t^3$.

(c) Calcule w_{zzy} para $w = y\ln(x^2+z^4)$.

[06] Resolva as questões abaixo.

(a) Calcule $\partial^3 w/\partial z\partial y^2$ para $w = x^2/(y^2+z^2)$.

(b) Calcule $\partial^3 w/\partial z\partial y\partial x$ para $w = \text{sen}(xyz)$.

[07] (**Função harmônica**) Uma função $z = f(x_1, \ldots, x_n)$ é *harmônica* se ela satisfaz a *equação de Laplace*

$$\frac{\partial^2 z}{\partial x_1^2} + \frac{\partial^2 z}{\partial x_2^2} + \ldots + \frac{\partial^2 z}{\partial x_n^2} = 0.$$

Mostre que $z = f(x,y) = x^3 - 3xy^2$ e $z = g(x,y) = \ln\sqrt{x^2+y^2}$ são funções harmônicas.

[08] Quais destas funções satisfazem a equação de Laplace (veja a questão anterior)?

(a) $z = f(x,y) = x^2 - y^2$.

(b) $z = f(x,y) = x^2 + y^2$.

(c) $z = f(x,y) = xy$.

(d) $z = f(x,y) = y^3 + 3x^2y$.

(e) $z = f(x,y) = e^x \,\text{sen}\, y$.

[09] Se $w = e^{-c^2t}\,\text{sen}(cx)$, mostre que $w_{xx} = w_t$ para todo real c.

[10] Mostre que $v = \operatorname{sen}(k\,x)\cdot\operatorname{sen}(ak\,t)$ satisfaz a *equação da onda*

$$\frac{\partial^2 v}{\partial t^2} = a^2\frac{\partial^2 v}{\partial x^2}.$$

[11] Seja C a curva resultante da interseção do parabolóide $z = 9 - x^2 - y^2$ com o plano $x = 1$. Determine a equação da tangente l a C no ponto $P = (1, 2, 4)$. Esboce o gráfico do parabolóide, de C e de l.

[12] Calcule

$$\lim_{h\to 0}\frac{3 + (x+y+h)^2 z - (3 + (x+y)^2 z)}{h}.$$

[13] Utilizando diretamente a definição de derivadas parciais, calcule $\partial f/\partial x$ e $\partial f/\partial y$ para as funções f dadas abaixo.

(a) $z = f(x,y) = x + 2\,y$.

(b) $z = f(x,y) = x^2 + 3\,y^2$.

[14] Encontre $f_x(0,0)$ e $f_y(0,0)$ para

$$z = f(x,y) = \begin{cases} \dfrac{x^3 + y^3}{x^2 + y^2} & \text{se } (x,y) \neq (0,0),\\ 0 & \text{se } (x,y) = (0,0), \end{cases}$$

e $g_x(0,y)$, com $y \in \mathbb{R}$, para

$$z = g(x,y) = \begin{cases} \dfrac{x^2 - xy}{x + y} & \text{se } (x,y) \neq (0,0),\\ 0 & \text{se } (x,y) = (0,0). \end{cases}$$

[15] Encontre $f_x(x,y)$ e $f_y(x,y)$ para as funções abaixo.

(a) $z = f(x,y) = \displaystyle\int_x^y \ln(\operatorname{sen}(t))\,dt$.

(b) $z = f(x,y) = \displaystyle\int_x^y e^{\cos(t)}\,dt$.

[16] Calcule todas as derivadas parciais de primeira ordem da função

$$u = g(x,y,z) = xz + e^z\left(\int_0^x t^2 e^t\,dt\right).$$

[17] Calcule todas as derivadas de terceira ordem da função de produção $Q = 4K^{3/4}L^{1/4}$. Use o teorema de Young para acelerar o processo.

[18] Seja $z = f(x, y) = x^4y^3 - x^8 + y^4$.

(a) Calcule $\partial^3 z/\partial y \partial x \partial x$, $\partial^3 z/\partial x \partial y \partial x$ e $\partial^3 z/\partial x \partial x \partial y$.

(b) Calcule $\partial^3 z/\partial x \partial y \partial y$, $\partial^3 z/\partial y \partial x \partial y$ e $\partial^3 z/\partial y \partial y \partial x$.

*[19] O objetivo deste exercício é examinar uma função de classe C^1 para a qual a tese do teorema de Young falha: as derivadas parciais mistas não são iguais. Seja

$$z = f(x, y) = \begin{cases} \dfrac{x^3y - xy^3}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0), \\ 0 & \text{se } (x, y) = (0, 0). \end{cases}$$

(a) Prove que f se anula sobre os eixos x e y. Conclua então que $(\partial f/\partial x)(0, 0)$ e $(\partial f/\partial y)(0, 0)$ são iguais a zero.

(b) Calcule $\partial f/\partial x$ e $\partial f/\partial y$ para pontos $(x, y) \neq (0, 0)$ e conclua que $(\partial f/\partial x)(0, y) = -y$ e $(\partial f/\partial y)(x, 0) = x$.

(c) Mostre que

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0, 0) = \lim_{y \to 0} \frac{\dfrac{\partial f}{\partial x}(0, y) - \dfrac{\partial f}{\partial x}(0, 0)}{y - 0} = -1.$$

(d) Mostre que

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0, 0) = \lim_{x \to 0} \frac{\dfrac{\partial f}{\partial y}(x, 0) - \dfrac{\partial f}{\partial y}(0, 0)}{x - 0} = +1$$

e conclua que as derivadas parciais mistas de f não são iguais em $(0, 0)$.

(e) Calcule $(\partial^2 f/\partial x \partial y)(x, y)$ para $(x, y) \neq (0, 0)$.

(f) Use o item anterior para mostrar que $(\partial f^2/\partial x \partial y)(x, x) = 0$ para $x > 0$.

(g) Conclua que $(\partial^2 f/\partial x \partial y)(x, y)$ é descontínua na origem $(0, 0)$ (assim, f não é uma função de classe C^2) que é a hipótese do teorema de Young não se verifica.

[20] Seja $z = f(x, y) = g(x) \cdot h(y)$ onde g e h são funções diferenciáveis de uma única variável. Encontre uma expressão para $\partial f/\partial x$ e $\partial f/\partial y$. *Dica: lembre-se como calcular uma derivada parcial de maneira prática!*

[21] Estabeleça os mesmos resultados da questão anterior com o uso direto da definição de derivadas parciais.

[22] Qual é a diferença (se é que existe) entre $\partial^2 f/\partial x^2$ e $(\partial f/\partial x)^2$?

[23] Calcule todas as derivadas parciais de ordem 1 da função

$$z = f(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2.$$

[24] Diga se cada uma das sentenças a seguir é verdadeira ou falsa, apresentando uma justificativa caso ela seja verdadeira ou um contra-exemplo caso ela seja falsa.

(a) Se $f\colon \mathbb{R}^3 \to \mathbb{R}$ é uma função de classe C^1 satisfazendo $f_x(0,0,0) = 0$, $f_y(0,0,0) = 0$ e $f_z(0,0,0) = 0$, então $f(x,y,z) = 0$ para todo $(x,y,z) \in \mathbb{R}^3$.

(b) Se $f\colon \mathbb{R}^3 \to \mathbb{R}$ é uma função de classe C^1 satisfazendo $f_x(x,y,z) = 0$, $f_y(x,y,z) = 0$ e $f_z(x,y,z) = 0$ para todo $(x,y,z) \in \mathbb{R}^3$, então $f(x,y,z) = 0$ para todo $(x,y,z) \in \mathbb{R}^3$.

[25] Diga se cada uma das sentenças a seguir é verdadeira ou falsa, apresentando uma justificativa caso ela seja verdadeira ou um contra-exemplo caso ela seja falsa.

(a) Se $f\colon D \subset \mathbb{R}^3 \to \mathbb{R}$ é uma função de classe C^1 definida no conjunto aberto

$$D = \{(x,y,z) \in \mathbb{R}^3 \mid x < 1\} \cup \{(x,y,z) \in \mathbb{R}^3 \mid x > 2\}$$

satisfazendo $f_x(x,y,z) = 0$, $f_y(x,y,z) = 0$ e $f_z(x,y,z) = 0$ para todo $(x,y,z) \in D$, então $f(x,y,z) = 0$ para todo $(x,y,z) \in D$.

(b) Se $f\colon D \subset \mathbb{R}^3 \to \mathbb{R}$ é uma função de classe C^1 definida no conjunto aberto

$$D = \{(x,y,z) \in \mathbb{R}^3 \mid x < 1\} \cup \{(x,y,z) \in \mathbb{R}^3 \mid x > 2\}$$

satisfazendo $f_x(x,y,z) = 0$, $f_y(x,y,z) = 0$ e $f_z(x,y,z) = 0$ para todo $(x,y,z) \in D$ e $f(0,0,0) = 0$, então $f(x,y,z) = 0$ para todo $(x,y,z) \in D$.

[26] A Companhia ACME produz três tipos de *roller skates*. O custo em reais para se produzir x, y e z unidades de cada tipo de *skate* é

$$c(x,y,z) = 30000 + 27\,x + 36\,y + 47\,z.$$

(a) O valor $\partial c/\partial x$ representa a taxa variação no custo devido ao acréscimo de uma unidade na produção do *skate* mais barato, sendo que os níveis de produção das outras unidades mais caras permanecem constantes. Encontre esta taxa.

(b) Encontre $\partial c/\partial z$ e forneça uma interpretação.

[27] Quando um poluente tal como óxido nítrico é emitido por uma chaminé de h metros de altura, a concentração $C(x,y)$ (em μ/m^3) do poluente em um ponto P situado a y metros acima do chão e cuja projeção ortogonal sobre o chão está a x quilômetros da base da chaminé pode ser representada por

$$C(x,y) = \frac{a}{x^2}\left(e^{-b\,(y-h)^2/x^2} + e^{-b\,(y+h)^2/x^2}\right)$$

em que a e b são constantes positivas que dependem das condições atmosféricas e da taxa de emissão do poluente (veja a figura (5.6)). Suponha que

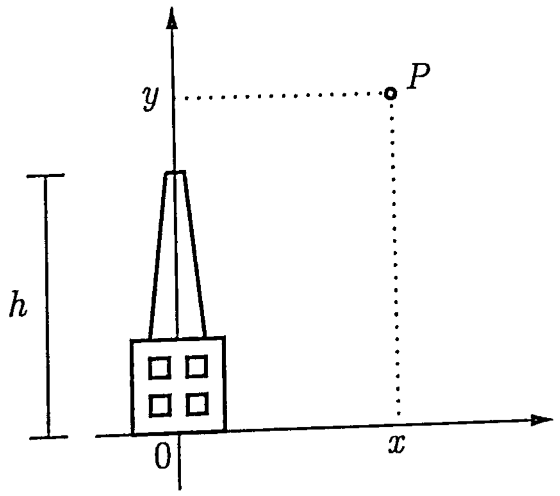

Figura 5.6: Concentração de um poluente emitido por uma chaminé.

$$C(x,y) = \frac{200}{x^2}\cdot\left(e^{-0.02\,(y-10)^2/x^2} + e^{-0.02\,(y+10)^2/x^2}\right).$$

Calcule e interprete $\partial C/\partial x$ e $\partial C/\partial y$ no ponto $P = (2,5)$.

[28] A análise de certos circuitos eletrônicos envolve a fórmula

$$I = \frac{V}{\sqrt{R^2 + L^2 \cdot \omega}},$$

em que I é a corrente, V a voltagem, R a resistência, L a indutância e ω uma constante positiva. Calcule e interprete $\partial I/\partial R$ e $\partial I/\partial L$.

[29] A maioria dos computadores tem apenas um processador que pode ser utilizado para cálculos. Os supercomputadores modernos, entretanto, têm entre dois a vários milhares de processadores. Um supercomputador multiprocessador é comparado a um computador uniprocessador em termos de *speed up*. A *speed up* S é o número de vezes mais rápido que um cálculo pode ser feito com um multiprocessador do que com um uniprocessador. A *lei de Amdahl* é uma fórmula para determinar S:

$$S(p,q) = \frac{p}{q + p \cdot (1-q)},$$

em que p é o número de processadores e q é a fração de cálculo que pode ser realizada utilizando todos os processadores disponíveis em paralelo, isto é, usando-os de maneira que os dados sejam processados concomitantemente por unidades separadas. A situação ideal, *paralelismo completo*, ocorre quando $q = 1$.

(a) Se $q = 0.8$, ache o *speed up* quando p é igual a 10, 100 e 1000. Mostre que o *speed up* S não pode exceder 5, independentemente do número de processadores disponíveis.

(b) Ache a taxa de variação de S em relação a q.

(c) Qual a taxa de variação no item anterior se há paralelismo completo? Como o número de processadores afeta esta taxa de variação?

A *eficiência* E de um cálculo por multiprocessador pode ser calculada pela equação

$$E = \frac{S(p,q)}{p} = \frac{1}{q + p \cdot (1-q)}.$$

(d) Mostre que se $0 \leq q < 1$, então E é uma função decrescente de p. Conclua que sem paralelismo completo, o aumento do número de processadores não aumenta a eficiência do cálculo.

[30] No estudo da penetração da geada em uma rodovia, a temperatura T no instante t e à profundidade x pode ser dada aproximadamente por

$$T = T_0\, e^{-\lambda \cdot x}\ \mathrm{sen}(\omega\, t - \lambda\, x),$$

em que T_0, ω e λ são constantes.

(a) Calcule e interprete as derivadas parciais $\partial T/\partial t$ e $\partial T/\partial x$.

(b) Mostre que T satisfaz a *equação unidimensional do calor*:

$$\frac{\partial T}{\partial t} = k\,\frac{\partial^2 T}{\partial x^2},$$

com $k = 2\,\lambda^2/w$.

[31] A *capacidade vital* V dos pulmões é o maior volume de ar que pode ser exalado após uma inalação de ar. Para um indivíduo do sexo masculino com x anos de idade e y centímetros de altura, V pode ser aproximada pela fórmula

$$V = 27.63\,x - 0.112\,y.$$

Calcule e interprete as derivadas parciais $\partial V/\partial x$ e $\partial V/\partial y$.

[32] Em um dia claro, a intensidade da luz solar às t horas após o nascente e à profundidade oceânica de x metros, pode ser aproximada por

$$I(x, t) = I_0\, e^{-k\,x}\ \mathrm{sen}^3(\pi\, t/D),$$

com I_0 a intensidade da luz solar ao meio-dia, D a quantidade horas do dia com luz solar e k uma constante positiva. Se $I_0 = 1000$, $D = 12$ e $k = 0.10$, calcule e interprete as derivadas parciais $\partial I/\partial t$ e $\partial I/\partial x$ quando $t = 6$ horas e $x = 5$ metros.

[33] Em Economia, a *elasticidade de preço de procura* de um artigo indica a reação dos consumidores a uma alteração no preço de mercado do artigo. Suponhamos que n artigos A_1, A_2, ..., A_n tenham preços p_1, p_2, ..., p_n, respectivamente, e que a demanda pelo artigo A_k seja uma função q_k dos preços p_1, p_2, ..., p_n:

$$q_k = f_k(p_1, p_2, \ldots, p_n).$$

A *elasticidade* de preço do artigo A_k é a função e_k (que depende dos preços p_1, p_2, ..., p_n) definida por

$$e_k = \frac{p_k}{q_k} \cdot \frac{\partial q_k}{\partial p_k}.$$

Mostre que se modelarmos a demanda q_k com uma Cobb-Douglas

$$q_k = b_k \cdot p_1^{-a_{k,1}} \cdot p_2^{-a_{k,2}} \cdot \cdots \cdot p_n^{-a_{k,n}}, \qquad (*)$$

onde $a_{k,1}$, $a_{k,2}$, ..., $a_{k,n}$ são constantes não-negativas, então a elasticidade e_k será uma função constante.

Diz-se que o artigo A_k é *independente* do artigo A_j se uma variação no preço p_k de A_k não afeta a demanda q_j de A_j. Isto equivale à condição $\partial q_j/\partial p_k = 0$. Mostre que se a demanda q_k é modelada pela Cobb-Douglas $(*)$, então A_k é independente de A_j se, e somente se, $a_{j,k} = 0$.

[34] Mostre que para toda constante positiva c, a função

$$u(x,t) = \frac{12\,c\,e^{(x-ct)\sqrt{c}}}{\left(e^{(x-ct)\sqrt{c}} + 1\right)^2}$$

satisfaz a equação de Korteweg-de Vries (KdV):

$$u_t(x,t) + u_{xxx}(x,t) + u(x,t) \cdot u_x(x,t) = 0$$

Esta solução (denominada *sóliton*) representa o perfil de uma onda de água navegando por um canal estreito.

5.6 Nota histórica

A notação $\partial u/\partial x$ foi usada pela primeira vez por Adrien Marie Legendre em 1786 em sua obra *"Memoire sur la manière de distinguer les maxima des minima dans le Calcul des Variations"*. Na página 8 está escrito:

> *Pour éviter toute ambiguité, je représenterai par $\partial u/\partial x$ le coefficient de x dans la différence de u, & par du/dx la différence complète de u divisée par dx.*

Legendre abandonou o símbolo e ele foi reintroduzido por Carl Gustav Jacob Jacobi, em 1841, em seu artigo *"De determinantibus Functionalibus"*:

> *Sed quia uncorum accumulatio et legenti et scribenti molestior fieri solet, praetuli characteristica d differentialia vulgaria, differentialia autem partialia characteristica ∂ denotare.*

O símbolo ∂ é algumas vezes chamado de delta de Jacobi e ele corresponde à letra "d" do alfabeto russo.