

Nosso objetivo na primeira parte destas notas é provar que para uma função suficientemente regular as derivadas parciais mistas f_{xy} e f_{yx} coincidem.

1 Derivação sob o sinal de integral

Sejam $D \subset \mathbb{R}^2$ aberto, $f : D \rightarrow \mathbb{R}$ uma função de classe C^1 em D e $[a, b] \times [c, d] \subset D$. Considere a função $\varphi : [a, b] \rightarrow \mathbb{R}$ dada por

$$\varphi(x) = \int_a^b f(x, y) dy.$$

A função φ é chamada de *integral dependendo do parâmetro x* . Gostaríamos de calcular a derivada de φ .

Por exemplo, se $f(x, y)$ é soma de funções do tipo $g(x)h(y)$, então, evidentemente, temos $\varphi(x) = g(x) \int_a^b h(y) dy$, portanto

$$\frac{d\varphi}{dx} = \frac{dg}{dx} \int_a^b h(y) dy = \int_a^b \frac{dg}{dx} h(y) dy = \int_a^b \frac{\partial f}{\partial x}(x, y) dy. \quad (1)$$

Isso mostra que as operações de derivação em relação a x e integração em relação a y *comutam* no caso em que f é uma soma de funções do tipo anterior.

Uma abordagem semelhante à anterior pode ser usada no caso mais geral. Nesta situação, podemos aproximar a integral por uma soma de Riemann

$$\varphi(x) = \int_a^b f(x, y) dy \approx \sum_j f(x, y_j^*) \Delta y_j,$$

onde $a = y_0 < y_1 < \dots < y_n = b$ é uma partição de $[a, b]$, $y_j^* \in [y_{j-1}, y_j]$ e $\Delta y_j = y_j - y_{j-1}$, $j = 1, \dots, n$. Abusando da fé, concluímos que

$$\frac{d\varphi}{dx} \approx \sum_j \frac{\partial f}{\partial x}(x, y_j^*) \Delta y_j.$$

Como a última soma é uma soma de Riemann para a função $\frac{\partial f}{\partial x}$, fazendo $\Delta y_j \rightarrow 0$, concluímos que a fórmula (2) deve valer também neste caso mais geral. Nas condições anteriores, vale o resultado abaixo.

Teorema 1 (Derivação sob o sinal de integral) Derivação em relação a x e integração em relação a y *comutam*, i.e.,

$$\frac{d}{dx} \int_a^b f(x, y) dy = \int_a^b \frac{\partial f}{\partial x}(x, y) dy.$$

Dada uma função $f : D \rightarrow \mathbb{R}$ de classe C^2 sejam $(x_0, y_0) \in D$, B um disco fechado centrado em (x_0, y_0) contido em D e $(x, y) \in D$. Pelo teorema fundamental do cálculo em uma variável, temos que

$$f(x, y) = f(x_0, y) + \int_{x_0}^x f_x(t, y) dt.$$

Derivando a expressão acima em relação à y , obtemos

$$f_y(x, y) = f_y(x_0, y) + \int_{x_0}^x f_{xy}(t, y) dt;$$

derivando em relação a x e usando o teorema fundamental do cálculo, obtemos $f_{yx}(x, y) = f_{xy}(x, y)$. Como (x, y) é arbitrário em D , concluímos que $f_{xy} = f_{yx}$.

Teorema 2 (Schwarz/Clairaut) Se $f : D \rightarrow \mathbb{R}$ é de classe C^2 então $f_{xy} = f_{yx}$.

Como consequência do teorema anterior, podemos observar que para uma função de classe C^3 , temos $f_{xxy} = f_{xyx} = f_{yxx}$, $f_{xyy} = f_{yxy} = f_{yyx}$, para uma função de classe C^4 temos $f_{xyxy} = f_{xyyx} = f_{xxyy}$, e assim por diante.

2 A fórmula de Taylor em duas variáveis

Nesta seção, veremos como as derivadas de ordem superior de uma função podem ser utilizadas para estudar seu comportamento. Embora os resultados aqui trabalhados possam ser feitos utilizando derivadas de ordens maiores, vamos estudar a fórmula de Taylor e suas consequências apenas em ordem 2.

Inicialmente, lembramos que se φ é uma função real de uma variável de classe C^2 definida em um intervalo aberto I contendo t_0 , então

$$\begin{aligned} \varphi(t_0 + h) &= \varphi(t_0) + \varphi'(t_0)h + r_1(h) \text{ e} \\ \varphi(t_0 + h) &= \varphi(t_0) + \varphi'(t_0)h + \varphi''(t_0)h^2 + r_2(h) \end{aligned}$$

onde r_1, r_2 são funções tais que $\lim_{h \rightarrow 0} r_1(h)/h = \lim_{h \rightarrow 0} r_2(h)/h^2 = 0$ quando $t \rightarrow t_0$. A segunda fórmula acima é chamada de *fórmula de Taylor com resto infinitesimal (de grau 2)*. Uma aplicação muito importante desta fórmula acima ocorre quando t_0 é um *ponto crítico não-degenerado* de φ , i.e., $\varphi'(t_0) = 0$ e $\varphi''(t_0) \neq 0$. Neste caso, é fácil ver que $\varphi(t_0 + h) - \varphi(t_0)$ é essencialmente uma função *quadrática* com concavidade dependendo do sinal de $\varphi''(t_0)$. Em particular, podemos determinar se t_0 é um ponto de máximo ou mínimo local.

A questão central é saber se existe uma fórmula análoga às fórmulas acima para funções de duas variáveis. Vamos estudar esta questão mais de perto. Sejam $f : D \rightarrow \mathbb{R}$ uma função de classe C^2 definida no aberto $D \subset \mathbb{R}^2$ e $(x_0, y_0) \in D$. Como f é diferenciável, vale

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + \{f_x(x_0, y_0)h + f_y(x_0, y_0)k\} + r(h, k),$$

onde $r(h, k)/\sqrt{h^2 + k^2} \rightarrow 0$ quando $(h, k) \rightarrow (0, 0)$. Como comentamos anteriormente, a expressão que aparece acima entre chaves é chamada de *derivada de f no ponto (x_0, y_0) na direção (h, k)* e é denotada por $df(x_0, y_0) \cdot (h, k)$. Assim, a fórmula acima pode ser reescrita como

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + df(x_0, y_0) \cdot (h, k) + r(h, k).$$

Vejam como podemos utilizar as derivadas parciais segundas para estudar f . Considere, para (h, k) fixado suficientemente pequeno, a função $\varphi : t \mapsto f(x_0 + th, y_0 + tk)$, definida para $|t|$ pequeno. Pela regra da cadeia, temos $\varphi(0) = f(x_0, y_0)$ e $\varphi'(t) = f_x(x_0 + th, y_0 + tk)h + f_y(x_0 + th, y_0 + tk)k$; em particular, $\varphi'(0) = f_x(x_0, y_0)h + f_y(x_0, y_0)k$. Além disso, pelo teorema de Schwarz,

$$\varphi''(0) = f_{xx}(x_0, y_0)h^2 + 2f_{xy}(x_0, y_0)hk + f_{yy}(x_0, y_0)k^2.$$

A expressão acima é uma *função quadrática* de duas variáveis e pode ser escrita em forma matricial como

$$\begin{pmatrix} h & k \end{pmatrix} \begin{pmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{xy}(x_0, y_0) & f_{yy}(x_0, y_0) \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}. \quad (2)$$

A matriz 2×2 que aparece acima é muito importante no estudo de pontos críticos e é chamada de *matriz hessiana de f* no ponto (x_0, y_0) e denotada por $Hf(x_0, y_0)$. A expressão acima é denotada por $d^2 f(x_0, y_0) \cdot (h, k)^2$, por conveniência.

Aplicando a fórmula de Taylor à função φ no ponto $t = 0$, com um pouco mais de esforço, obtemos¹

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + df(x_0, y_0) \cdot (h, k) + d^2 f(x_0, y_0) \cdot (h, k)^2 + r(h, k),$$

onde $r(h, k)/|(h, k)|^2 \rightarrow 0$ quando $(h, k) \rightarrow 0$. Esta fórmula é chamada de *fórmula de Taylor com resto infinitesimal (de grau 2) em duas variáveis*.

3 Máximos e mínimos em conjuntos abertos

Como no caso de funções de uma variável, a fórmula de Taylor de ordem 2 pode ser usada para estudar pontos críticos de uma função de duas variáveis.

Antes de mais nada, fixemos a nomenclatura. Dada $f : D \rightarrow \mathbb{R}$ de classe C^2 definida em um *aberto* $D \subset \mathbb{R}^2$ e $(x_0, y_0) \in D$, dizemos que (x_0, y_0) é um ponto de *máximo local* (respectivamente, *mínimo local*) para f se existe um disco fechado $B \subset D$ centrado em (x_0, y_0) tal que $f(x, y) \leq f(x_0, y_0)$ (respectivamente, $f(x, y) \geq f(x_0, y_0)$) para todo $(x, y) \in B \cap D$. Se alguma das desigualdades anteriores valer para *todo* $(x, y) \in D$, substituímos o adjetivo *local* por *global*. Os pontos de máximo ou mínimo local (respectivamente, global) são chamados genericamente de *extremantes locais* (respectivamente, *extremantes globais*).

Exemplos simples como $f(x, y) = x$, $(x, y) \in \mathbb{R}^2$, mostram que uma função pode não admitir extremantes locais nem globais.

A próxima proposição dá uma boa dica sobre onde começar a procurar.

Proposição 3 Seja $f : D \rightarrow \mathbb{R}$ uma função definida em um *aberto* $D \subset \mathbb{R}^2$ e $(x_0, y_0) \in D$ um extremante local para f . Se existem $f_x(x_0, y_0)$ e $f_y(x_0, y_0)$ então $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$.

Sendo assim, os extremantes locais de uma função de classe C^1 em um *aberto* são pontos onde ambas as derivadas parciais se anulam. Tais pontos recebem um nome especial.

Definição 4 Se $f : D \rightarrow \mathbb{R}$ é uma função de classe C^1 definida em um *aberto* $D \subset \mathbb{R}^2$. Um ponto $(x_0, y_0) \in D$ é dito *ponto crítico* para f se $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$.

¹O resultado aqui utilizado é o seguinte: Se r é uma função de classe C^2 definida em um *aberto* contendo a origem, então r e todas as suas derivadas parciais de ordem ≤ 2 se anulam na origem se e só se $\lim_{(h,k) \rightarrow (0,0)} \frac{r(h,k)}{|(h,k)|^2} = 0$.

Sendo assim, os extremantes locais (e globais) de uma função de classe C^1 definida em um aberto são pontos críticos de f . O exemplo clássico de ponto crítico que não é extremante local é descrito abaixo.

Exemplo 5 Sejam $f(x, y) = y^2 - x^2$ e $(x, y) \in \mathbb{R}^2$. Evidentemente, o único ponto crítico de f é $(0, 0)$, mas este não é um extremante local, pois $(0, 0)$ é ponto de mínimo global para a restrição de f ao eixo x e máximo global para a restrição de f ao eixo y .

Seja (x_0, y_0) um ponto crítico para a função f de classe C^2 . Aplicando a fórmula de Taylor de grau 2 para f , obtemos para todo (h, k) suficientemente pequeno,

$$\begin{aligned} f(x_0 + h, y_0 + k) - f(x_0, y_0) &= f_{xx}(x_0, y_0)h^2 + 2f_{xy}(x_0, y_0)hk + f_{yy}(x_0, y_0)k^2 + r(h, k) \\ &= d^2 f(x_0, y_0) \cdot (h, k)^2 + r(h, k), \end{aligned} \tag{3}$$

onde $r(h, k)/|(h, k)|^2 \rightarrow 0$ quando $|(h, k)| \rightarrow 0$. Se a expressão quadrática acima tiver um sinal definido para $|(h, k)|$ pequeno, temos a indicação de que este será o sinal de $f(x_0 + h, y_0 + k) - f(x_0, y_0)$, uma vez que $r(h, k)$ é muito pequeno quando comparado com $|(h, k)|^2$. Isso nos leva a pensar em um critério que permita decidir se uma função quadrática de duas variáveis tem sinal definido perto da origem.

Lema 6 Seja $F(h, k) = ah^2 + 2bhk + ck^2$, onde $a, b, c \in \mathbb{R}$ são fixados.

1. O único ponto crítico de F é $(0, 0)$;
2. Se $a > 0$ e $ac - b^2 > 0$ então $F(h, k) > 0$ para todo $(h, k) \neq (0, 0)$. Neste caso, existe uma constante positiva α (dependendo somente de a, b, c) tal que $F(h, k) \geq \alpha(h^2 + k^2)$ para todo $(h, k) \in \mathbb{R}^2$;
3. Se $a < 0$ e $ac - b^2 > 0$ então $F(h, k) < 0$ para todo $(h, k) \neq (0, 0)$. Neste caso, existe uma constante positiva β (dependendo somente de a, b, c) tal que $F(h, k) \leq -\beta(h^2 + k^2)$ para todo $(h, k) \in \mathbb{R}^2$;
4. Se $ac - b^2 < 0$ então o sinal de F sobre a reta $h = -\frac{b}{a}k$ é contrário ao sinal de F sobre a reta $k = 0$.

Prova. O caso $a = 0$ pode ser analisado diretamente a partir da expressão da função F . Se $a \neq 0$, completando quadrados obtemos

$$F(h, k) = a \left(h + \frac{b}{a}k \right)^2 + \frac{ac - b^2}{a}k^2.$$

As afirmações feitas a respeito de positividade decorrem da expressão acima. Para obter as constantes α e β , basta observar que a função $F(h, k)$ é homogênea de grau 2, portanto, podemos tomar $\alpha = \min_{h^2+k^2=1} F(h, k)$ e $\beta = \max_{h^2+k^2=1} F(h, k)$. ■

No caso em que estamos interessados, os coeficientes a, b, c são as derivadas parciais segundas de f em (x_0, y_0) . Se (x_0, y_0) é de fato um extremante local para f então considerando as restrições $x \mapsto f(x, y_0)$ e $y \mapsto f(x_0, y)$, temos, do cálculo de uma variável que os números $f_{xx}(x_0, y_0)$ e $f_{yy}(x_0, y_0)$ podem ser um deles ou ambos nulos, mas nunca podem ter sinais contrários. Dito de outra forma, temos que

$$f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) \geq 0.$$

Se, ao contrário disso, tivéssemos que $f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) < 0$, então, evidentemente o determinante da matriz Hessiana $Hf(x_0, y_0)$ seria negativo. Portanto (x_0, y_0) seria máximo local em uma direção e mínimo local em outra direção, e portanto, não poderia ser extremante local de f . Esta situação é bastante comum e inspira a definição abaixo (conforme o exemplo (5).)

Definição 7 Dizemos que (x_0, y_0) é *ponto de sela* para f se $\det Hf(x_0, y_0) < 0$.

Todas estas observações nos permitem concluir o teorema abaixo.

Teorema 8 Seja $f : D \rightarrow \mathbb{R}$ de classe C^2 definida em um aberto $D \subset \mathbb{R}^2$ e $(x_0, y_0) \in D$ um ponto crítico de f . São verdadeiras as seguintes afirmações:

1. Se $f_{xx}(x_0, y_0) > 0$ e $\det Hf(x_0, y_0) > 0$ então (x_0, y_0) é um ponto de mínimo local;
2. Se $f_{xx}(x_0, y_0) < 0$ e $\det Hf(x_0, y_0) > 0$ então (x_0, y_0) é um ponto de máximo local;
3. Se $\det Hf(x_0, y_0) < 0$ então existe uma direção ao longo da qual (x_0, y_0) é ponto de máximo local e outra direção ao longo da qual (x_0, y_0) é ponto de mínimo local
4. Se $\det Hf(x_0, y_0) = 0$, nada se pode afirmar.

Prova. Observando a equação (3) e a propriedade que caracteriza a função r , podemos escrever $g(h, k) = \frac{r(h, k)}{h^2 + k^2}$ e observar que $r(h, k) = F(h, k)(h^2 + k^2)$ e $g(h, k) \rightarrow 0$ quando $|(h, k)| \rightarrow 0$. Se $f_{xx}(x_0, y_0) > 0$, então pelo primeiro item do lema (6), obtemos $\alpha > 0$ tal que $d^2 f(x_0, y_0) \cdot (h, k)^2 \geq \alpha(h^2 + k^2)$ para todo $(h, k) \in \mathbb{R}^2$. Logo,

$$\begin{aligned} f(x_0 + h, y_0 + k) - f(x_0, y_0) &= d^2 f(x_0, y_0) \cdot (h, k)^2 + r(h, k) \\ &\geq (\alpha + g(h, k))(h^2 + k^2). \end{aligned}$$

Como $g(h, k) \rightarrow 0$ quando $|(h, k)| \rightarrow 0$, segue que existe $\delta > 0$ tal que para todo (h, k) com $|(h, k)| < \delta$, tem-se $\alpha + g(h, k) > 0$, portanto, $f(x_0 + h, y_0 + k) \geq f(x_0, y_0)$ se $|(h, k)| < \delta$. Isso significa que (x_0, y_0) é um mínimo local. A situação é completamente análoga para o caso de máximo local. ■

Observação 9 (Para quem entende de álgebra linear) O teorema anterior mostra que o comportamento de f numa vizinhança de um ponto crítico satisfazendo as condições do enunciado é essencialmente o comportamento da função quadrática

$$d^2 f(x_0, y_0) \cdot (h, k)^2 \doteq f_{xx}(x_0, y_0)h^2 + 2f_{xy}(x_0, y_0)hk + f_{yy}(x_0, y_0)k^2.$$

Como observamos anteriormente (veja a expressão (2)), esta função pode ser escrita como

$$G(u) = \langle u, Hu \rangle,$$

onde $u = (h, k)$, $\langle \cdot, \cdot \rangle$ denota o produto interno usual e H é operador linear em \mathbb{R}^2 cuja matriz na base canônica é a matriz hessiana de f no ponto (x_0, y_0) . Podemos estudar o sinal da função G usando ferramentas de álgebra linear. Lembramos que, como a matriz hessiana é simétrica, então seus autovalores são ambos reais, digamos λ_1, λ_2 . Os autovalores são raízes da equação polinomial $\det(H - \lambda I) = 0$. É bem sabido que esta equação polinomial pode ser escrita como

$$\lambda^2 - (\text{tr } H)\lambda + \det H = 0,$$

onde $\text{tr } H$ denota o *traço* de H . Assim, $\lambda_1 \lambda_2 = \det H$ e $\lambda_1 + \lambda_2 = \text{tr } H$. No caso $\det H > 0$, segue que λ_1, λ_2 são não-nulos de mesmo sinal. Se assumirmos que $f_{xx}(x_0, y_0) > 0$ então necessariamente $f_{yy}(x_0, y_0) > 0$ (senão, não teríamos $\det H > 0$!), e portanto $\text{tr } H > 0$. A recíproca também é verdadeira, pois se assumirmos que $\text{tr } H > 0$, então, como $\det H > 0$, os números $f_{xx}(x_0, y_0)$ e $f_{yy}(x_0, y_0) > 0$ devem ter

o mesmo sinal. Como $0 < \text{tr } H = f_{xx}(x_0, y_0) + f_{yy}(x_0, y_0)$, segue que $f_{xx}(x_0, y_0) > 0$ e $f_{yy}(x_0, y_0) > 0$. Portanto, a hipótese $f_{xx}(x_0, y_0) > 0$ ($f_{xx}(x_0, y_0) < 0$) no primeiro (segundo) ítem do teorema pode ser substituída por $\text{tr } H > 0$ ($\text{tr } H < 0$). Se $\lambda_1 = \lambda_2 = \lambda$, então $T = \lambda I$, e, portanto, a positividade da função G depende somente do sinal de λ . Caso $\lambda_1 \neq \lambda_2$, podemos tomar vetores unitários $u_1, u_2 \in \mathbb{R}^2$ tais que $Tu_1 = \lambda_1 u_1$ e $Tu_2 = \lambda_2 u_2$. Como H é simétrica, segue que u_1 e u_2 são ortogonais, e portanto, $\{u_1, u_2\}$ é uma base ortonormal de \mathbb{R}^2 . Dado $u \in \mathbb{R}^2$, escrevendo $u = x_1 u_1 + x_2 u_2$, temos então

$$G(u) = \lambda_1 x_1^2 + \lambda_2 x_2^2.$$

No caso $\det H > 0$, temos que λ_1, λ_2 têm o mesmo sinal, e portanto, segue imediatamente da expressão acima que $(0, 0)$ é um ponto de máximo local se $\lambda_1 < 0$ e mínimo local se $\lambda_1 > 0$. Se $\det H < 0$ então λ_1 e λ_2 têm sinais distintos, digamos $\lambda_1 > 0$ e $\lambda_2 < 0$, portanto, G admite um mínimo em $(0, 0)$ quando restrita à reta $x_2 = 0$ e um máximo em $(0, 0)$ quando restrita à reta $x_1 = 0$.

Se $\det H = 0$ então $\lambda_1 = 0$ ou $\lambda_2 = 0$. Exemplos distintos desta situação são $G(u) = x_1^2$ (mínimo) e $G(u) = -x_2^2$ (máximo); evidentemente, nada se pode afirmar sobre a natureza do ponto crítico nesta condição.

4 Máximos e mínimos em conjuntos compactos

Como vimos anteriormente, nada garante que uma função definida em um aberto admita extremantes locais ou globais, porém, em diversas situações práticas, seria muito conveniente *adicionar pontos ao domínio* para que a função admita extremantes. Por exemplo, a função $f(x, y) = x$ não admite extremantes no domínio $D = \{(x, y) : 0 < x < 1, 0 < y < 1\}$, mas se adicionarmos as *bordas* a D , observamos que os pontos da forma $(0, y)$, $0 \leq y \leq 1$, são mínimos globais e os pontos da forma $(1, y)$, $0 \leq y \leq 1$, são máximos globais.

Evidentemente, não se pode esperar muito no caso em que f seja descontínua; por exemplo, $f(x, y) = (x^2 + y^2)^{-1}$, $f(0, 0) = 0$, definida no disco de centro $(0, 0)$ e raio 1 não admite máximos locais. A função $f(x, y) = (x^2 + y^2)^{-1}$ definida no domínio $\mathbb{R}^2 \setminus \{(0, 0)\}$ não admite nem máximos nem mínimos locais, mas é contínua.

Assim, vemos que a continuidade de f , a limitação do domínio e a inclusão das *bordas* no conjunto são condições que, sozinhas, não bastam para garantir a existência de extremantes para f . Por isso, para encontrar condições que garantam existência de extremantes, devemos abrir mão de considerar apenas funções definidas em abertos e passar a considerar funções definidas em domínios genéricos de \mathbb{R}^2 . Fixemos algumas notações.

Dada $f : D \rightarrow \mathbb{R}$ de classe C^2 definida em um domínio $D \subset \mathbb{R}^2$ e $(x_0, y_0) \in D$, dizemos que (x_0, y_0) é um ponto de *máximo local* (respectivamente, *mínimo local*) para f se existe um disco fechado $B \subset D$ centrado em (x_0, y_0) tal que $f(x, y) \leq f(x_0, y_0)$ (respectivamente, $f(x, y) \geq f(x_0, y_0)$) para todo $(x, y) \in B \cap D$. Se alguma das desigualdades anteriores valer para *todo* $(x, y) \in D$, substituímos o adjetivo *local* por *global*. Os pontos de máximo ou mínimo local (respectivamente, global) são chamados genericamente de *extremantes locais* (respectivamente, *extremantes globais*).

O teorema abaixo estabelece uma classe de funções e domínios para os quais podemos garantir, em geral, a existência de extremantes.

Teorema 10 (Weierstrass) Se D é um conjunto fechado e limitado e $f : D \rightarrow \mathbb{R}$ é contínua então existem $(x_0, y_0), (x_1, y_1) \in D$ tais que

$$f(x_0, y_0) \leq f(x, y) \leq f(x_1, y_1)$$

para todo $(x, y) \in D$. Em particular, (x_0, y_0) é um ponto de mínimo global para f e (x_1, y_1) é um ponto de máximo global para f .

O enunciado e a demonstração do resultado acima são muito relevantes do ponto de vista do desenvolvimento conceitual do Cálculo e da Análise Matemática. Este resultado foi enunciado e provado somente no século XIX pelo matemático alemão Karl Weierstrass e ilustra bem o método famoso de *caça ao leão*. Um conjunto fechado e limitado em \mathbb{R}^2 é dito *compacto*.

Um tipo de problema muito comum é o de determinar os extremantes locais/globais de uma função de classe C^2 em um disco fechado ou em um retângulo fechado. Pelo teorema de Weierstrass, sabemos que estes extremantes sempre existem, por isso, separamos a busca em duas etapas: em uma delas, procuramos os máximos/mínimos no *interior* do conjunto utilizando o teorema (8) e na outra, estudamos a função na *fronteira* do conjunto. A próxima seção fornece um método para a segunda etapa.

5 Exercícios

Exercício 1 Considere $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ contínua e defina

$$g(x) = \int_c^d f(x, y) dy,$$

para $x \in [a, b]$.

1. Mostre que g é contínua. Em particular, podemos considerar $\int_a^b g(x) dx = \int_a^b \left\{ \int_c^d f(x, y) dy \right\} dx$.
2. Considere $h(y) = \int_a^b f(x, y) dx$; o mesmo argumento usado para g prova que h é contínua, e portanto, também podemos considerar $\int_c^d h(y) dy = \int_c^d \left\{ \int_a^b f(x, y) dx \right\} dy$.
3. Vamos provar que as integrais dos itens (1) e (2) coincidem. Para isto, considere

$$\varphi(y) = \int_a^b \left\{ \int_c^y f(x, t) dt \right\} dx,$$

para $y \in [c, d]$. Mostre que φ é diferenciável e $\varphi'(y) = \int_a^b f(x, y) dx$.

4. Use o teorema fundamental do cálculo para φ na forma $\varphi(d) = \varphi(c) + \int_c^d \varphi'(t) dt$ para concluir que

$$\int_a^b \left\{ \int_c^d f(x, y) dy \right\} dx = \int_c^d \left\{ \int_a^b f(x, y) dx \right\} dy.$$