CAPíTULO 5

Cálculo em várias variáveis

[1]

Neste capítulo estudamos as propriedades de funções de várias variáveis. Nos concentramos principalmente nas questões relativas a diferenciabilidade destas funções e aplicações.

5.1. Introdução

Consideramos aqui funções $f : \mathbb{R}^m \to \mathbb{R}^n$. Como na imensa maioria das questões $m = 2$ e $n = 1$, suporemos aqui este caso, i.e., que f é uma função real definida no plano. Mas os resultados apresentados podem ser trivialmente estendidos para o caso mais geral. Assim como no caso unidimensional, dizemos que uma função f é contínua em $\mathbf{x}_0$ se dado $\epsilon > 0$ existe $\delta > 0$ tal que

$$\|\mathbf{x} - \mathbf{x}_0\| < \delta \implies \|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \epsilon.$$

Analogamente ao caso unidimensional, podemos definir o conceito de limite de funções num determinado ponto. Entretanto, a noção de *limites laterais* não faz mais sentido.

Outro conceito importante é o de curvas (em duas dimensões) ou superfícies (em três dimensões) de nível, que chamaremos sempre de curvas de nível. Uma curva de nível é dada pelo conjunto de pontos que têm imagem constante, i.e., dado $c \in \mathbb{R}$, definimos

$$\{(x, y) \in \mathbb{R}^2 : f(x, y) = c\}.$$

ANPEC 2005, Questão 09. *Avalie as afirmativas:*

[1] Última Atualização: 24/06/2025

1 *Se uma função $f(x,y)$ é contínua em um ponto (x_0,y_0), então as funções $\phi(x) = f(x,y_0)$ e $\varphi(y) = f(x_0,y)$ são contínuas em x_0 e y_0, respectivamente.*

Resolução: Como f é contínua em (x_0,y_0), então $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0,y_0)$. A função ϕ será contínua em x_0 se $\lim_{x\to x_0}\phi(x) = \phi(x_0)$. Dado que $\phi(x) = f(x,y_0)$ e $\varphi(y) = f(x_0,y)$, então

$$\lim_{x\to x_0}\phi(x) = \lim_{x\to x_0} f(x,y_0) = f(x_0,y_0) = \phi(x_0).$$

De forma análoga,

$$\lim_{y\to y_0}\varphi(y) = \lim_{y\to y_0} f(x_0,y) = f(x_0,y_0) = \varphi(y_0).$$

Portanto ambas funções são contínuas. Afirmativa verdadeira. ✓

2 *A função $f(x,y) = \frac{x^2y^2}{x^4+y^4}$ é contínua em $(0,0)$.*

Resolução: Observe primeiro que a pergunta não faz muito sentido, pois $f(0,0)$ não foi definida. De qualquer forma, nos eixos vertical e horizontal tem-se $f(x,0) = f(0,y) = 0$. Portanto, para termos continuidade, teríamos que ter $f(0,0) = 0$. Entretanto, em cima da reta $y = x$ temos que $f(x,x) = 1/2$. Logo,

$$\lim_{(x,x)\to(0,0)} f(x,y) = \frac{1}{2}.$$

e portanto não é possível definir f na origem de forma que seja contínua. Afirmativa falsa. ✓

4 *Seja $h(x) = f(x)g(x)$. Se $h(x)$ é contínua, então f e g também o são.*

Resolução: Considere como contra-exemplo $f(x) = 0$ para todo $x \in \mathbb{R}$ e g uma função descontínua qualquer. Neste caso, $h(x) = f(x)g(x) = 0$ para todo $x \in \mathbb{R}$ é contínua. Mas g não o é. Falso. ✓

5.2. Derivadas parciais e planos tangentes

Seja $f : \mathbb{R}^2 \to \mathbb{R}$. Então definimos as *derivadas parciais* de f num ponto (x,y) por

$$\frac{\partial f}{\partial x}(x,y) = \lim_{h\to 0}\frac{f(x+h,y)-f(x,y)}{h}, \qquad \frac{\partial f}{\partial y}(x,y) = \lim_{h\to 0}\frac{f(x,y+h)-f(x,y)}{h}$$

quando o limite acima existir.

Destaque 5.1: Gradiente

Se cada uma das derivadas parciais existir num ponto (x_0, y_0), definimos o *vetor gradiente* (também chamado de *gradiente* pelos íntimos) dado por

$$\nabla f(x_0, y_0) = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0)\right).$$

Dizemos que um ponto (x_0, y_0) é crítico se $\nabla f(x_0, y_0) = (0, 0)$.

Uma propriedade importante do vetor gradiente é que ele aponta na direção de maior crescimente da função. Considere um vetor $\vec{u}$ de norma um. Dado que a derivada direcional $\partial f(x)/\partial \vec{u}$ indica o crescimento da f na direção $\vec{u}$, temos que

$$\frac{\partial f(x)}{\partial \vec{u}} = \nabla f(x) \cdot \vec{u} = \| \nabla f(x)\| \cdot \|\vec{u}\| \cos\theta = \| \nabla f(x)\| \cos\theta,$$

onde θ é o ângulo entre os dois vetores. Para que $\partial f(x)/\partial \vec{u}$ seja máximo, devemos ter $\cos(\theta) = 1$. Então

$$\frac{\partial f(x)}{\partial \vec{u}} = \nabla f(x) \cdot \vec{u} = \| \nabla f(x)\|.$$

Logo, a direção de máximo crescimento, no ponto x, é

$$\vec{u} = \frac{\nabla f(x)}{\| \nabla f(x)\|}.$$

Outra propriedade importante do vetor gradiente é que ele é ortogonal às curvas de nível. De fato, seja a curva $\vec{\gamma} : \mathbb{R} \to \mathbb{R}^2$ tal que $f(\vec{\gamma}(t)) =$ constante. Derivando em função de t e usando a regra da cadeia, temos

$$0 = \frac{\partial f}{\partial x}\frac{\partial \gamma_1}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial \gamma_2}{\partial t} = \nabla f(\gamma(t)) \cdot (\gamma_1'(t), \gamma_2'(t)).$$

Como $\vec{\gamma}' = (\gamma_1', \gamma_2')$ é vetor tangente à curva de nível, então ∇f é ortogonal à curva de nível.

Definimos a seguir planos tangentes. No caso unidimensional, a reta tangente $r(x)$ a uma função $f(x)$ num determinado ponto x_0 é definida por

$$r(x) = f(x_0) + f'(x_0)(x - x_0) = 0.$$

Note que $r(x_0) = f(x_0)$ e $r'(x_0) = f'(x_0)$.

O plano tangente ao gráfico de $f : \mathbb{R}^2 \to \mathbb{R}$ num determinado ponto (x_0, y_0) é definido por $p(x, y) = ax + by + d$ tal que

$$p(x_0, y_0) = f(x_0, y_0), \quad a = \frac{\partial p}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0), \quad b = \frac{\partial p}{\partial y}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0).$$

Destaque 5.2: Vetor gradiente

Dada uma função $f : \mathbb{R}^2 \to \mathbb{R}$, o vetor gradiente é dado por

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right).$$

O vetor $\nabla f(x, y)$ aponta na direção de maior crescimento da f, e é ortogonal às curvas de nível de f.

O plano tangente à função $f(x, y)$ no ponto (x_0, y_0) é dado por

$$p(x, y) = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + f(x_0, y_0).$$

Note pela fórmula acima que $\nabla f(\mathbf{x}_0)$ é o vetor ortogonal ao plano.

Uma interessante forma de analisarmos uma função em várias variáveis é restringindo esta função numa direção e usando propriedades de funções de apenas uma variável. Obtemos assim o conceito de derivada direcional.

Destaque 5.3: Derivada Direcional

Sejam $\mathbf{u} \in \mathbb{R}^2$ com $\|\mathbf{u}\| = 1$, e $f : \mathbb{R}^2 \to \mathbb{R}$. Dado $\mathbf{x} \in \mathbb{R}^2$, definimos a *derivada direcional de f em $\mathbf{x}$ na direção $\mathbf{u}$* por

$$\mathbf{D_u} f(\mathbf{x}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{u}) - f(\mathbf{x})}{t}.$$

quando o limite acima existir. Se $\mathbf{D_u}f(\mathbf{x}) > 0$ então a direção $\mathbf{u}$ é de crescimento, e se $\mathbf{D_u}f(\mathbf{x}) < 0$, f decresce na direção $\mathbf{u}$.

Quando $\mathbf{u} = \mathbf{e}_i$ é vetor da base canônica, temos as *derivadas parciais*

$$\mathbf{D}_{\mathbf{e}_1}f(\mathbf{x}) = \frac{\partial f}{\partial x}(\mathbf{x}), \qquad \mathbf{D}_{\mathbf{e}_2}f(\mathbf{x}) = \frac{\partial f}{\partial y}(\mathbf{x}).$$

Se f for diferenciável em $\mathbf{x}$, então existe a derivada direcional, que é dada por

$$\mathbf{D_u}f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{u} = \frac{\partial f}{\partial x}(\mathbf{x})u_1 + \frac{\partial f}{\partial y}(\mathbf{x})u_2.$$

ANPEC 2020, Questão 12. *Julgue a veracidade das afirmações abaixo.*

(2) *Dada a função* $f(x,y) = \frac{2y\ln(1+x)}{y^2+\ln^2(1+x)}$, *a curva de nível* $C(1) = \{(x,y) \in D_f : f(x,y) = 1\}$ *coincide com o gráfico da função* $h(x) = \ln(1+x)$, $x > -1$.

Resolução: Note que $h(0) = 0$, e que f não pode ser definida de forma contínua em $(0,0)$. De fato, $f(x, \ln(1+x)) = 1$ para todo $x \neq 0$, portanto $\lim_{x\to 0} f(x, \ln(1+x)) = 1$. Entrento, $\lim_{x\to 0} f(x,0) = 0$. Portanto, $(0,0)$ pertence ao gráfico de h, mas $(0,0) \notin C(1)$.

Veja que esta é uma tentativa de "pegadinha", pois $f(x,y) = 1$ implica em $y = \ln(1+x)$. A afirmativa é falsa. ✓

(3) *Dada* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função de classe* C^1 *e um ponto* $x \in \mathbb{R}^n$ *tal que o gradiente* $\nabla f(x)$ *é não nulo, o vetor* $\nabla f(x)$ *indica a direção de maior crescimento da função* f *a partir do ponto* x.

Resolução: Basta ver o Destaque 5.2. A afirmativa é verdadeira. ✓

(4) *Considere a curva derivável* $\gamma : (a,b) \subseteq \mathbb{R} \to \mathbb{R}^2, \gamma(t) = (\gamma_1(t), \gamma_2(t))$, *a função* $f : \mathbb{R}^2 \to \mathbb{R}$ *de classe* C^1 *e* $f(\gamma(t)) = c \in \mathbb{R}$, *para todo* $t \in (a,b)$. *Então os vetores* $\nabla f(\gamma(t))$ *e* $(\gamma_1'(t), \gamma_2'(t))$ *são ortogonais para todo* $t \in (a,b)$.

Resolução: Novamente, pelo Destaque 5.2, a afirmativa é verdadeira. ✓

ANPEC 2017, Questão 8. *Considere a função* $f(x,y) = x^2\ln y + y^3e^x + 3x + 2y$.

(0) *No ponto $(x, y) = (1, 1)$ a direção $(-3, 1)$ é uma direção de crescimento da função f;*

Resolução: Note que

$$\nabla f(x, y) = (2x \ln y + y^3 e^x + 3, (x^2/y) + 3y^2 e^x + 2),$$

e portanto $\nabla f(1, 1) = (e + 3, 3 + 3e)$. Na direção $(-3, 1)$ temos que a derivada direcional é dada por

$$\nabla f(1, 1) \cdot (-3, 1) = -6 < 0.$$

Assim, a direção não é de crescimento. Logo, a afirmativa é falsa. ✓

(1) *No ponto $(x, y) = (0, 1)$ a direção $(4, 5)$ é a direção de máximo incremento da função f;*

Resolução: No ponto $(x, y) = (0, 1)$ temos $\nabla f(0, 1) = (4, 5)$, e a direção do gradiente é exatamente a de máximo crescimento. Logo, a afirmativa é verdadeira. ✓

(2) *A função f tem um máximo relativo no interior no seu domínio;*

Resolução: Como em uma dimensão, um ponto interior de máximo relativo é um ponto crítico (o gradiente se anula). Note que a função só está definida para $y > 0$, e portanto

$$\frac{\partial f}{\partial y} = (x^2/y) + 3y^2 e^x + 2$$

nunca se anula. Portanto não existe ponto crítico. Logo, a afirmativa é falsa. ✓

(3) *Ao longo do eixo vertical (quando $x = 0$) a direção horizontal à direita (ou seja, $(1, 0)$) é a direção de máximo incremento de f ;*

Resolução: Quando $x = 0$, temos

$$\nabla f(x, y) = (y^3 + 3, 3y^2 + 2),$$

que nunca está alinhado com a direção $(1, 0)$. Logo, a afirmativa é falsa. ✓

(4) *Em todo ponto do domínio a função f é crescente em ambas variáveis.*

Resolução: Note que para $x = 1$, por exemplo,

$$\frac{\partial f}{\partial x}(0, y) = 2 \ln y + y^3 e + 3 < 0$$

se y estiver próximo do zero. Portanto a função é decrescente na direção x. Logo, a afirmativa é falsa. ✓

5.3. Diferenciabilidade

Assim como em uma dimensão, $f : \mathbb{R}^2 \to \mathbb{R}$ é diferenciável em $\mathbf{x}$ se e somente se existir uma função $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$\mathbf{r}(\mathbf{h}) = f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - \nabla f'(\mathbf{x}) \cdot \mathbf{h} \quad \text{com} \lim_{\mathbf{h} \to \mathbf{0}} \frac{\|\mathbf{r}(\mathbf{h})\|}{\|\mathbf{h}\|} = 0. \tag{5.1}$$

Note que pela identidade acima, diferenciabilidade implica em continuidade.

Observação 5.1. No caso de uma função $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$, a definição é semelhante, mas o gradiente é substituido pela *matriz jacobiana*

$$[\mathbf{f}'(\mathbf{x})] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_m} \end{bmatrix}.$$

Chamamos $\mathbf{f}'(\mathbf{x})$ de derivada de $\mathbf{f}$ em $\mathbf{x}$, e que também denotamos por $\boldsymbol{D}f(\mathbf{x})$. A matriz é $n \times m$, que presenta uma transformação linear $\mathbb{R}^m \to \mathbb{R}^n$.

Os casos mais importantes e comuns são quando $m = 1$ ou $n = 1$. No primeiro caso, para $\gamma(t) = (\gamma_1(t), \ldots, \gamma_n(t))$, obtemos

$$\gamma'(t) = \begin{bmatrix} \gamma_1'(t) \\ \gamma_2'(t) \\ \vdots \\ \gamma_n'(t) \end{bmatrix}.$$

Para o caso $f : \mathbb{R} \to \mathbb{R}$, obtemos a gradiente:

$$[f'(\mathbf{x})] = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}) & \frac{\partial f}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial f}{\partial x_m}(\mathbf{x}) \end{bmatrix}.$$

É importante ressaltar que a *existência de derivadas parciais em relação às coordenadas não implica na existência de derivadas direcionais em geral.* Considere o simples exemplo abaixo.

Exemplo 5.2. Seja $f : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x,y) = \begin{cases} \frac{x}{y} & \text{se } y \neq 0, \\ 0 & \text{se } y = 0. \end{cases}$$

Então

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

mas a derivada direcional na direção (a,b) não existe se a e b são não nulos, pois não existe o limite quando $t \to 0$ de

$$\frac{f(ta,tb) - f(0,0)}{t} = \frac{1}{t}\frac{a}{b}.$$

A existência de derivadas direcionais *não implica em diferenciabilidade.* Para ilustrar tal fato, considere a função

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

Então

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

mas dado (a,b) com $\|(a,b)\|^2 = a^2 + b^2 = 1$ e $b \neq 0$, temos

$$\lim_{t \to 0} \frac{f(ta,tb) - f(0,0)}{t} = \frac{a^2}{b},$$

e a derivada direcional é dada por

$$D_{(a,b)} f(0,0) = \frac{a^2}{b}. \tag{5.2}$$

Entretanto, se f fosse diferenciável, teríamos

$$D_{(a,b)}f(0,0) = \mathbf{f}'(0,0)(a,b) = \frac{\partial f}{\partial x}(0,0)a + \frac{\partial f}{\partial y}(0,0)b = 0,$$

uma contradição com (5.2). Logo f não é diferenciável em $(0,0)$ apesar de ter todas as derivadas direcionais neste ponto. Note que $f(x,x^2) = 1$ para $x \neq 0$, e portanto f é descontínua em $(0,0)$.

Apesar da existência de derivadas direcionais num determinado ponto não garantir a diferenciabilidade neste ponto, a existência e continuidade das derivadas parciais numa *vizinhança* dum ponto garante a diferenciabilidade, como podemos ver no resultado a seguir.

Teorema 5.1 (Diferenciabilidade)**.** *Seja* $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}$. *Se* $\partial \mathbf{f}/\partial x_i$ *existir e for contínua numa vizinhança de* $\mathbf{x}$ *para* $i = 1, \ldots, m$, *então* $\mathbf{f}$ *é diferenciável em* $\mathbf{x}$.

ANPEC 2008, Questão 11. *Considere a função:*

$$f(x,y) = \begin{cases} 0 & \textit{se } (x,y) = (0,0), \\ \frac{x^3y - xy^3}{x^2+y^2} & \textit{em caso contrário.} \end{cases}$$

Com relação à função acima, julgue as afirmativas abaixo.

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

Resolução: Note que f é identicamente nula nos eixos x e y. Logo, a afirmativa é verdadeira. ✓

(1) *Se* $g(x,y) = \frac{\partial f(0,y)}{\partial x} + \frac{\partial f(x,0)}{\partial y}$, *então* $g(2,2) = 0$.

Resolução: Para $(x,y) \neq (0,0)$

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{(3x^2y - y^3)(x^2+y^2) - 2x(x^3y - xy^3)}{(x^2+y^2)^2}, \\ \frac{\partial f}{\partial y} &= \frac{(x^3 - 3xy^2)(x^2+y^2) - 2y(x^3y - xy^3)}{(x^2+y^2)^2}. \end{aligned}$$

Logo,

$$g(x,y) = \frac{\partial f(0,y)}{\partial x} + \frac{\partial f(x,0)}{\partial y} = -y + x \implies g(2,2) = 0.$$

Logo, a afirmativa é verdadeira. ✓

(2) $\frac{\partial^2 f}{\partial x \partial y}(0,0) = \frac{\partial^2 f}{\partial y \partial x}(0,0)$.

Resolução: Note que

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

onde usamos as contas do item (1). Logo, a afirmativa é falsa. ✓

(3) $\frac{\partial^2 f(x,y)}{\partial x \partial y}$ *é contínua na origem.*

Resolução: Pelo Teorema de Schwarz ou Clairaut ou Young, se a segunda derivada fosse contínua, a ordem de derivação não importaria. Mas o item (2) mostra que este não é o caso. Logo, a afirmativa é falsa. ✓

(4) $\frac{\partial^2 f(x,x)}{\partial x \partial y} = 0$ *para* $x > 0$.

Resolução: Do item anterior e para $x > 0$, obtemos

$$\frac{\partial^2 f(x,y)}{\partial x \partial y} = \frac{(x^2 - y^2)(x^2 + 9y^2)}{(x^2 + y^2)^2}.$$

Se $x = y$, então $\partial^2 f / \partial x \partial y = 0$.
Logo, a afirmativa é verdadeira, apesar do gabarito oficial ser falso. ✓

Outro resultado de grande importância diz respeito à diferenciabilidade de composições de funções, garantindo que se duas funções são diferenciáveis, então a composição também o é.

DESTAQUE 5.4: REGRA DA CADEIA

Sejam $\mathbf{g} : \mathbb{R}^l \to \mathbb{R}^m$ e $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$. Se $\mathbf{g}$ é diferenciável em $\mathbf{x} \in \mathbb{R}^l$ e $\mathbf{f}$ é diferenciável em $\mathbf{g}(\mathbf{x})$, então $\mathbf{f} \circ \mathbf{g} : \mathbb{R}^l \to \mathbb{R}^n$ é diferenciável em $\mathbf{x}$ e

$$(\mathbf{f} \circ \mathbf{g})'(\mathbf{x}) = \mathbf{f}'(\mathbf{g}(\mathbf{x})) \circ \mathbf{g}'(\mathbf{x}).$$

Considerando $f : \mathbb{R}^2 \to \mathbb{R}$, os casos mais comuns são os seguintes:

→ Para $f(x(t), y(t))$ temos

$$\frac{d}{dt} f(x(t), y(t)) = \nabla f(x(t), y(t)) \cdot (x'(t), y'(t)) = \frac{\partial f}{\partial x}\frac{dx}{dt}(t) + \frac{\partial f}{\partial y}\frac{dy}{dt}(t)$$

→ Para $f(x(s,t), y(s,t))$ temos

$$\begin{bmatrix} \frac{\partial}{\partial s} & \frac{\partial}{\partial t} \end{bmatrix} f(x(s,t), y(s,t)) = f'g' = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{bmatrix}.$$

Logo,

$$\frac{\partial}{\partial s} f(x(s,t), y(s,t)) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s},$$
$$\frac{\partial}{\partial t} f(x(s,t), y(s,t)) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t},$$

Veremos agora várias aplicações da regra da cadeia.

- **Aplicação 1:** A derivada da função inversa é a inversa da derivada (matriz jacobiana). Seja $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$, e seja a função $\mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^2$ inversa de $\mathbf{f}$, isto é,

$$\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x}, \qquad \mathbf{f}(\mathbf{g}(\mathbf{y})) = \mathbf{y},$$

para todo $\mathbf{x}$, $\mathbf{y}$ em $\mathbb{R}^2$. Suponha $\mathbf{f}$ e $\mathbf{g}$ diferenciáveis. Então as matrizes jacobianas $\mathbf{f}'(\mathbf{x})$ e $\mathbf{g}'(\mathbf{y})$ são inversas uma da outra, isto é,

$$\mathbf{f}'(\mathbf{x})\mathbf{g}'(\mathbf{y}) = \mathbf{g}'(\mathbf{y})\mathbf{f}'(\mathbf{x}) = \mathbf{I},$$

onde $\mathbf{I}$ é a matriz identidade.

De fato, seja $\mathbf{h}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x}$. Derivando $\mathbf{h}(\mathbf{x}) = \mathbf{x}$, temos $\mathbf{h}'(\mathbf{x}) = \mathbf{I}$. Usando a regra da cadeia para $\mathbf{h}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$, temos $\mathbf{h}'(\mathbf{x}) = \mathbf{g}'(\mathbf{y})\mathbf{f}'(\mathbf{x})$. Logo, $\mathbf{g}'(\mathbf{y})\mathbf{f}'(\mathbf{x}) = \mathbf{I}$. De forma análoga segue-se que $\mathbf{f}'(\mathbf{x})\mathbf{g}'(\mathbf{y}) = \mathbf{I}$.

- **Aplicação 2:** Ver Seção 5.2.
- **Aplicação 3:** Uma outra aplicação da regra da cadeia tem a ver com funções homogêneas. Dizemos que $f : \mathbb{R}^m \to \mathbb{R}$ é *homogênea de grau* k se $f(t\mathbf{x}) = t^k f(\mathbf{x})$ para todo $t \in \mathbb{R}$. Por exemplo, $x^2/(yz)$ é homogênea de grau zero, e $\sqrt{x^5}$ é homogênea de grau $5/2$. O Teorema de Euler para funções

homogêneas diz que $\nabla f(\mathbf{x}) \cdot \mathbf{x} = kf(\mathbf{x})$ é equivalente a f ser homogênea. De fato, derivando $f(t\mathbf{x}) = t^k f(\mathbf{x})$ em relação a t, temos

$$\nabla f(t\mathbf{x}) \cdot \mathbf{x} = kt^{k-1} f(\mathbf{x}).$$

Obtemos a identidade tomando $t = 1$.

DESTAQUE 5.5: FUNÇÃO HOMOGÊNEA

O Teorema de Euler para funções homogêneas afirma que f é homogênea de grau k se e somente se

$$\nabla f(\mathbf{x}) \cdot \mathbf{x} = kf(\mathbf{x}).$$

- **Aplicação 5:** É possível obter a regra de Leibniz através da regra da cadeia. Seja
$$F(t_1, t_2, t_3) = \int_{t_1}^{t_2} f(t_3, y)\, dy.$$
Então
$$\frac{d}{dx} F(u(x), v(x), x) = \frac{\partial F}{\partial t_1}\frac{du}{dx} + \frac{\partial F}{\partial t_2}\frac{dv}{dx} + \frac{\partial F}{\partial t_3},$$
implica em
$$\frac{d}{dx}\int_{u(x)}^{v(x)} f(x, y)\, dy = -f(x, u(x))\frac{du}{dx} + f(x, v(x))\frac{dv}{dx} + \int_{u(x)}^{v(x)} \frac{\partial f(x, y)}{\partial x}\, dy.$$

ANPEC 2021, Questão 8. *Para um número natural $N \geq 1$ denotamos por $\mathbb{R}^N_{++}$ o conjunto dos vetores $x = (x_1, x_2, \dots, x_N)$ em $\mathbb{R}^N$ com $x_1 > 0, x_2 > 0, \dots, x_N > 0$. Uma função $f : \mathbb{R}^N_{++} \to \mathbb{R}$ é chamada de positivamente homogênea de grau p, sendo $p \geq 0$ um número inteiro, se para todo número real $\alpha > 0$ tivermos $f(\alpha x) = \alpha^p f(x)$. Classifique as seguintes afirmações como verdadeiras ou falsas:*

0 *No caso $N = 1$ a função definida por $f(x) = x|x|$ é um exemplo de função positivamente homogênea de grau 2.*

Resolução: Basta ver que

$$f(\alpha x) = (\alpha x)|\alpha x| = \alpha|\alpha|x|x| = \alpha^2 f(x), \quad \text{para } \alpha > 0.$$

Logo, a afirmativa é verdadeira. ✓

(1) *Se $g : \mathbb{R}_{++} \to \mathbb{R}$ é uma função qualquer, então $f(x_1, x_2) = x_2 g\left(\frac{x_1}{x_2}\right)$ define uma função positivamente homogênea de grau 1 sobre $\mathbb{R}^2_{++}$.*

Resolução: Veja que

$$f(\alpha x_1, \alpha x_2) = \alpha x_2 g\left(\frac{x_1}{x_2}\right) = \alpha f(x_1, x_2).$$

Logo, a afirmativa é verdadeira. ✓

(2) *Se $g : \mathbb{R}_{++} \to \mathbb{R}$ é uma função positivamente homogênea de grau 1, então a função $f : \mathbb{R}^2_{++} \to \mathbb{R}$ definida por $f(x_1, x_2) = x_2 g\left(\frac{x_1}{x_2}\right)$ é côncava.*

Resolução: Note que

$$f(x_1, x_2) = x_2 g\left(\frac{x_1}{x_2}\right).$$

Se g é homogênea de grau 1, então $g\left(\frac{x_1}{x_2}\right) = \frac{x_1}{x_2} g(1)$, e assim

$$f(x_1, x_2) = x_2 \cdot \frac{x_1}{x_2} \cdot g(1) = x_1 g(1),$$

ou seja, f é linear, portanto côncava. Logo, a afirmativa é verdadeira. ✓

(3) *Qualquer função $g : \mathbb{R}_{++} \to \mathbb{R}$ positivamente homogênea de grau p satisfaz*

$$\lim_{x \to 0^+} g(x) > 0.$$

Resolução: Considere como contraexemplo a função $g(x) = x$, que é homogênea de grau 1, mas

$$\lim_{x \to 0^+} g(x) = 0.$$

Logo, a afirmativa é falsa. ✓

(4) *Sejam $f : \mathbb{R}^N_{++} \to \mathbb{R}$ e $g : \mathbb{R}^N_{++} \to \mathbb{R}$ funções positivamente homogêneas de grau p. Defina, sobre o mesmo domínio, a soma $f + g$ por $(f + g)(x) = f(x) + g(x)$ e o produto fg por $(fg)(x) = f(x)g(x)$. Portanto, $f + g$ e fg são positivamente homogêneas de grau p.*

Resolução: Tome $f(x) = g(x) = x$, que são homogêneas de grau 1, mas

$$(fg)(\alpha x) = f(\alpha x)g(\alpha x) = \alpha x \cdot \alpha x = \alpha^2 x^2 = \alpha^2 f(x)g(x),$$

ou seja, fg é homogênea de grau 2. Logo a afirmativa é falsa. ✓

ANPEC 2020, Questão 12. *Julgue a veracidade das afirmações abaixo;*

(0) *Considere a função* $f(x,y) = e^{\frac{x}{y}+1} + \frac{y}{x}$, *em que* $x > 0$ *e* $y > 0$. *Existem* $a > 0$ *e* $b > 0$ *tais que* $\big\langle (a,b), \nabla f(a,b) \big\rangle > 0$.

Resolução: Note que f é homogênea de grau zero, e portanto

$$\nabla f(\mathbf{x}) \cdot \mathbf{x} = kf(\mathbf{x}) = 0,$$

pois $k = 0$. Logo, a afirmativa é falsa.

Outra forma de se ver tal propriedade sem fazer contas é notando que f é constante em cima das retas $y = \alpha x$, com $\alpha \neq 0$. Logo, a derivada direcional na direção (a,b) calculada no ponto (a,b) tem que ser zero.

Outra forma, trabalhosa, é fazendo as contas. O gradiente

$$\nabla f(x,y) = \Big(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\Big) = \Big(\frac{1}{y}e^{\frac{x}{y}+1} - \frac{y}{x^2}, -\frac{x}{y^2}e^{\frac{x}{y}+1} + \frac{1}{x}\Big).$$

Assim, mais uma vez, $\langle (a,b), \nabla f(a,b) \rangle = 0$. Para todo $a > 0$ e $b > 0$, e a afirmativa é falsa. ✓

(1) *Dada a função* $f(x,y) = xe^{\frac{x}{y}}$, *todo plano tangente ao gráfico de* f *contém a origem.*

Resolução: Usando que o plano tangente é da forma

$$z = \frac{\partial f}{\partial x}(x_0,y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y - y_0) + f(x_0,y_0).$$

basta checar que o ponto $(0,0,0)$ pertence ao plano, i.e.,

$$0 = -\frac{\partial f}{\partial x}(x_0,y_0)x_0 - \frac{\partial f}{\partial y}(x_0,y_0)y_0 + f(x_0,y_0) \tag{5.3}$$

para todo (x_0,y_0). Mas esta relação (5.3) é verdadeira, pois f é homogênea de grau um, e então

$$\nabla f(\mathbf{x}_0) \cdot \mathbf{x}_0 = f(\mathbf{x}_0).$$

Outra forma de checar (5.3) é fazendo as contas. Como

$$\frac{\partial f(x_0,y_0)}{\partial x} = e^{\frac{x_0}{y_0}} + \frac{x_0}{y_0}e^{\frac{x_0}{y_0}}, \qquad \frac{\partial f(x_0,y_0)}{\partial y} = -\frac{x_0^2}{y_0^2}e^{\frac{x_0}{y_0}},$$

então

$$-\frac{\partial f}{\partial x}(x_0,y_0)x_0 - \frac{\partial f}{\partial y}(x_0,y_0)y_0 + f(x_0,y_0) = -[e^{\frac{x_0}{y_0}} + \frac{x_0}{y_0}e^{\frac{x_0}{y_0}}]x_0 - [-\frac{x_0^2}{y_0^2}e^{\frac{x_0}{y_0}}]y_0 + x_0e^{\frac{x_0}{y_0}}$$
$$= e^{\frac{x_0}{y_0}}[-x_0 - \frac{x_0^2}{y_0} + \frac{x_0^2}{y_0}] + x_0e^{\frac{x_0}{y_0}} = 0.$$

Logo, a afirmativa é verdadeira. ✓

ANPEC 2018, Questão 10. *Considere a função*

$$f(x,y,z) = \frac{xy^2z + x^2z^2 + \frac{y^6}{x^2}}{yz + xz + y^2}$$

e a direção $\vec{u} = (\sqrt{3}, \sqrt{3}, \sqrt{3})$*. Calcule a derivada direcional de* f *na direção* $\vec{u}$ *no ponto* $(\sqrt{3}, \sqrt{3}, \sqrt{3})$*.*

Resolução: Note que

$$f(x,y,z) = \frac{x^3y^2z + x^4z^2 + y^6}{x^2yz + x^3z + x^2y^2},$$

e portanto a função f é homogênea de grau 2, dado que $f(tx,ty,tz) = t^2f(x,y,z)$. Pela identidade de Euler, obtemos

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

A derivada direcional de f na direção $\vec{u}$ é dado por

$$D_{\vec{u}}f(\vec{u}) = \nabla f(\vec{u}) \cdot \frac{\vec{u}}{\|\vec{u}\|} = 2f(\vec{u})\frac{1}{\|\vec{u}\|} = 2.\frac{3}{3} = 2,$$

já que $f(\vec{u}) = f(\sqrt{3}(1,1,1)) = 3f(1,1,1) = 3.$ ✓

5.4. Matriz hessiana, fórmula de Taylor e pontos críticos

Depois de obtida uma derivada parcial, pode-se derivar novamente para se obter segundas derivadas parciais. Por exemplo, derivando-se na direção i e depois na direção j obtemos a função $\partial^2 f/\partial x_j \partial x_i$. Para funções "suaves", a ordem em que se deriva não importa. Em particular, se $\partial^2 f/\partial x_j \partial x_i(\mathbf{x}_0)$ e $\partial^2 f/\partial x_i \partial x_j(\mathbf{x}_0)$ existem e são contínuas numa vizinhança de $\mathbf{x}_0$, então

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}_0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0).$$

Este resultado é atribuído a Schwarz, a Clairaut ou a Young; ver [**3**, **19**, **35**].

Note que a derivada de uma função de uma função de $f : \mathbb{R}^m \to \mathbb{R}$ num determinado ponto $\mathbf{x}$ pode ser vista como uma aplicação linear de $\mathbb{R}^m$ em $\mathbb{R}$. No caso, para $\mathbf{x}$ fixo, teríamos $\mathbf{f}'(\mathbf{x}) : \mathbb{R}^m \to \mathbb{R}$ dada por

$$\mathbf{f}'(\mathbf{x})(\mathbf{y}) = \frac{\partial f}{\partial x_1}(\mathbf{x}) y_1 + \frac{\partial f}{\partial x_2}(\mathbf{x}) y_2 + \cdots + \frac{\partial f}{\partial x_m}(\mathbf{x}) y_m,$$

onde $\mathbf{y} \in \mathbb{R}^m$.

De forma análoga, definimos a segunda derivada de f num ponto $\mathbf{x}$ fixado como sendo a função bilinear $f''(\mathbf{x}) : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ tal que

$$f''(\mathbf{x})(\mathbf{y}, \mathbf{z}) = \sum_{i,j=1}^{m} \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} y_i z_j, \quad \text{onde } \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right),$$

e $\mathbf{y}, \mathbf{z} \in \mathbb{R}^m$. Uma forma mais compacta de escrever a definição acima é usando-se a *matriz hessiana* H dada por $H_{ij}(\mathbf{x}) = \partial^2 f(\mathbf{x})/\partial x_i \partial x_j$. Logo

$$f''(\mathbf{x})(\mathbf{y}, \mathbf{z}) = (\vec{\mathbf{y}})^t H(\mathbf{x}) \vec{\mathbf{z}}.$$

Note que se f for suave o suficiente, então a hessiana é simétrica.

Definições para derivadas de ordem mais alta seguem o mesmo formato, sendo estas *aplicações multilineares*. Entretanto para os nossos propósitos, a matriz hessiana basta.

Apresentamos no teorema a seguir a fórmula de Taylor, e nos restringimos ao caso particular de polinômios quadráticos. Este teorema será de fundamental importância para caracterizarmos pontos extremos.

Destaque 5.6: Teorema de Taylor

Seja $f : \mathbb{R}^m \to \mathbb{R}$ duas vezes diferenciável em Ω, com derivadas contínuas. Para $\mathbf{x} \in \mathbb{R}^m$ e $\mathbf{h} \in \mathbb{R}^m$, existe $\hat{t} \in (0,1)$ tal que para $\boldsymbol{\xi} = \mathbf{x} + \hat{t}\mathbf{h}$ tem-se

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{h}) + \frac{1}{2} f''(\boldsymbol{\xi})(\mathbf{h},\mathbf{h}). \tag{5.4}$$

Observação 5.3. Exigindo que as segundas derivadas sejam contínuas, podemos usar o fato de que a "ordem" das segundas derivadas não importam.

Assim como em uma dimensão, usaremos o Teorema de Taylor para estudarmos pontos extremos de uma função. Dizemos que $f : \mathbb{R}^m \to \mathbb{R}$ tem um *máximo local* em $\mathbf{x} \in \mathbb{R}^m$ se existe $\delta > 0$ tal que

$$\|\mathbf{y}-\mathbf{x}\| < \delta \implies f(\mathbf{y}) \le f(\mathbf{x}). \tag{5.5}$$

Dizemos que $\mathbf{x}$ é *máximo estrito local* se valer a desigualdade estrita em (5.5). Definição análoga serve para *mínimo local* e *mínimo estrito local*. Chamamos um ponto de máximo ou mínimo local de *ponto extremo local*, e um ponto de máximo ou mínimo estrito local de *ponto extremo estrito local*.

O resultado que obtemos a seguir, relativo a pontos extremos interiores, é análogo ao caso unidimensional, ver o Destaque 4.5, e diz primeiro que pontos extremos interiores são *pontos críticos*, i.e., pontos em que a derivada se anula. O resultado mostra também que se um ponto $\mathbf{x}$ é de mínimo local, então a forma bilinear $f''(\mathbf{x})$ é *semi-definida positiva*, i.e, $f''(\mathbf{x})(\mathbf{h},\mathbf{h}) \ge 0$ para todo $\mathbf{h} \in \mathbb{R}^m$. De forma análoga se um ponto é de máximo local, então $f''(\mathbf{x})$ é *semi-definida negativa*, i.e, $f''(\mathbf{x})(\mathbf{h},\mathbf{h}) \le 0$ para todo $\mathbf{h} \in \mathbb{R}^m$.

Em termos matriciais, $f''(\mathbf{x})$ é semi-definida positiva se a matriz hessiana $H(\mathbf{x})$ o for, i.e., se $(\vec{\mathbf{h}})^t H(\mathbf{x})\vec{\mathbf{h}} \ge 0$ para todo $\mathbf{h} \in \mathbb{R}^m$, e semi-definida negativa se valer $(\vec{\mathbf{h}})^t H(\mathbf{x})\vec{\mathbf{h}} \le 0$ para todo $\mathbf{h} \in \mathbb{R}^m$.

Destaque 5.7: Ponto extremo interior

Seja $f : \mathbb{R}^m \to \mathbb{R}$ e $\mathbf{x} \in \mathbb{R}^m$ ponto extremo local. Se f é diferenciável em $\mathbf{x}$, então $\mathbf{x}$ é ponto crítico, i.e., $\nabla f(\mathbf{x}) = 0$. Se além disto, f for duas vezes diferenciável com derivadas segundas contínuas, então temos que

(1) se $\mathbf{x}$ for ponto de mínimo local, então a hessiana é semi-definida positiva: $f''(\mathbf{x})(\mathbf{h}, \mathbf{h}) \geq 0$ para todo $\mathbf{h} \in \mathbb{R}^m$,

(2) se $\mathbf{x}$ for ponto de máximo local, então a hessiana é semi-definida negativa: $f''(\mathbf{x})(\mathbf{h}, \mathbf{h}) \leq 0$ para todo $\mathbf{h} \in \mathbb{R}^m$.

Os resultados acima nos dão condições necessárias para um ponto interior ser extremo local, porém estas não são suficientes (vide exemplo $f(x) = x^3$). Dizemos que um ponto é *de sela* quando a derivada se anula mas este não é extremo local. Um caso interessante é quando a função é localmente crescente na direção de uma coordenada e decrescente na direção de outra. Por exemplo, as funções dadas por $f(x, y) = x^2 - y^2$ e por $f(x, y) = x^3 - 3xy^2$; ver Figura 1.

---

arrumar figuras pretoe branco sem direitos autorais

O resultado a seguir nos fornece algumas condições suficientes para um ponto ser de máximo, mínimo ou de sela. Mais precisamente, temos que se um ponto crítico $\mathbf{x}$ de uma função suave tem $f''(\mathbf{x})$ *positiva definida*, i.e, $f''(\mathbf{x})(\mathbf{h}, \mathbf{h}) > 0$ para todo $\mathbf{h} \in \mathbb{R}^m \backslash \{\mathbf{0}\}$, então ele é mínimo estrito local. De forma análoga, se $f''(\mathbf{x})$ é *negativa definida*, i.e, $f''(\mathbf{x})(\mathbf{h}, \mathbf{h}) < 0$ para todo $\mathbf{h} \in \mathbb{R}^m \backslash \{\mathbf{0}\}$, então ele é máximo estrito local. O último caso é quando $f''(\mathbf{x})$ é *indefinida* i.e, existem $\mathbf{h}$, $\boldsymbol{\xi}$ em $\mathbb{R}^m$ tais que $[f''(\mathbf{x})(\mathbf{h}, \mathbf{h})][f''(\mathbf{x})(\boldsymbol{\xi}, \boldsymbol{\xi})] < 0$. Aí então $\mathbf{x}$ é ponto de sela.

Destaque 5.8: Pontos Críticos

Seja $f : \mathbb{R}^m \to \mathbb{R}$ duas vezes diferenciável, com derivadas contínuas, e $\mathbf{x} \in \mathbb{R}^m$ ponto crítico. Temos então que

(1) se $f''(\mathbf{x})$ for positiva definida então $\mathbf{x}$ é mínimo estrito local,

(2) se $f''(\mathbf{x})$ for negativa definida então $\mathbf{x}$ é máximo estrito local,

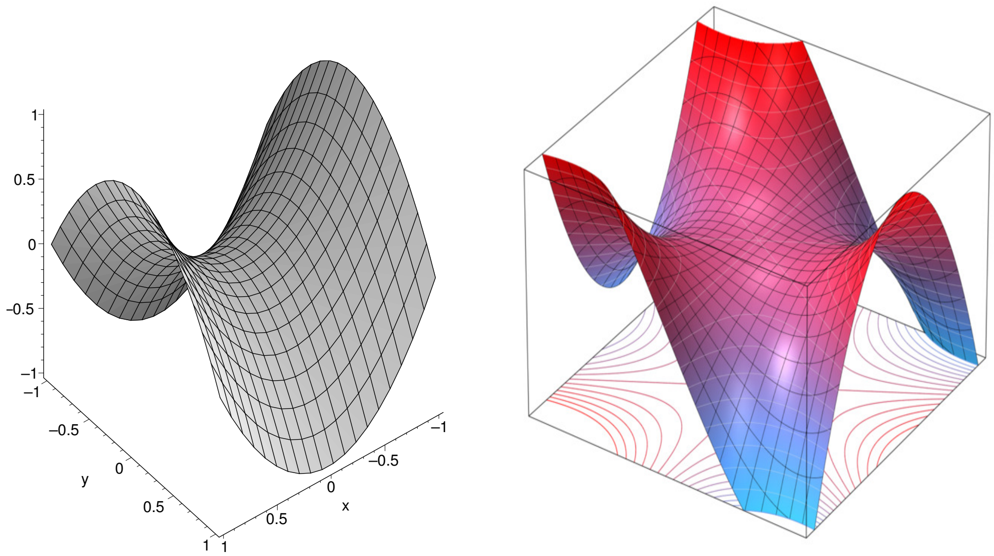

FIGURA 1. Gráfico de $x^2 - y^2$ (esquerda) e gráfico da sela de macaco dada por $x^3 - 3xy^2$ (direita), que tem ponto de sela em $(0, 0)$.

(3) se $f''(\mathbf{x})$ for indefinida então $\mathbf{x}$ é ponto de sela.

Note que apesar do teorema anterior dar condições suficientes para determinar se um ponto crítico é ou não extremo local, ainda é preciso descobrir se a f'' é positiva ou negativa definida ou indeterminada. Esta dificuldade é contornável, pois existem vários resultados de álgebra linear que dizem, por exemplo, quando uma matriz é ou não positiva definida. Por exemplo, uma matriz simétrica é positiva definida se e somente se seus autovalores são positivos. A referência [**13**] apresenta este e vários outros resultados relacionados ao tema.

DESTAQUE 5.9: MATRIZES SIMÉTRICAS POSITIVA DEFINIDAS

Seja M matriz simétrica $n \times n$. Temos então que

→ M é positiva definida se e somente se todos seus autovalores são positivos

- → M é positiva semi-definida se e somente se todos seus autovalores são não negativos
- → M é negativa definida se e somente se todos seus autovalores são negativos
- → M é positiva semi-definida se e somente se todos seus autovalores são não positivos
- → M é indefinida se e somente se possui autovalores positivos e negativos

Como para $n = 2$ podemos determinar os sinais dos autovalores via o determinante e o traço da matriz, alguns resultados valem para matrizes 2×2:

- → M é positiva definida se e somente se $\det M > 0$ e $\operatorname{tr} M > 0$
- → M é positiva semi-definida se e somente se $\det M \geq 0$ e $\operatorname{tr} M \geq 0$
- → M é negativa definida se e somente se $\det M > 0$ e $\operatorname{tr} M < 0$
- → M é negativa semi-definida implica em $\det M \geq 0$ e $\operatorname{tr} M \leq 0$
- → M é indefinida se e somente se $\det M < 0$

Uma segunda aplicação do Destaque 5.6 diz respeito à *funções convexas* definidas em *conjuntos convexos*. Dizemos que um conjunto $\Omega \subseteq \mathbb{R}^m$ é convexo se $\mathbf{x}$, $\mathbf{y} \in \Omega$ implica em $(1-t)\mathbf{x} + t\mathbf{y} \in \Omega$ para todo $t \in [0, 1]$. Dizemos que $f : \Omega \to \mathbb{R}$ é convexa em Ω se

$$f\big((1-t)\mathbf{x} + t\mathbf{y}\big) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y}).$$

para todo $t \in [0, 1]$ (f é estritamente convexa se a desigualdade for estrita). Graficamente, uma função é convexa se o gráfico de f entre $\mathbf{x}$ e $\mathbf{y}$ está abaixo da reta que une os pontos $(\mathbf{x}, f(\mathbf{x}))$ e $(\mathbf{y}, f(\mathbf{y}))$, como ilustra a Figura 2.

Existem vários resultados relacionados a convexidade. Em particular, um mínimo local é também global, e se o mínimo local é estrito, segue-se a unicidade de mínimo global [**23**].

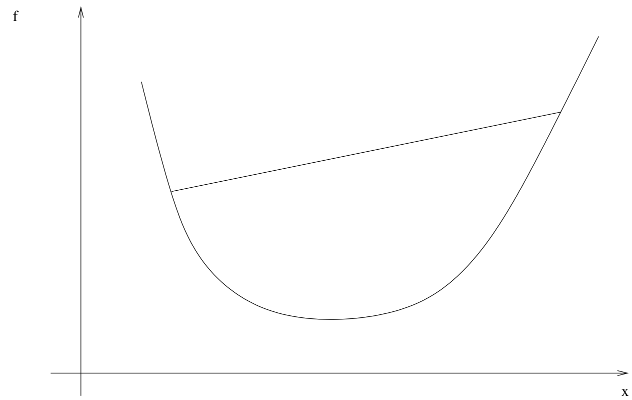

FIGURA 2. Função convexa.

DESTAQUE 5.10: CONVEXIDADE

Seja $f : \mathbb{R}^m \to \mathbb{R}$ duas vezes diferenciável, com derivadas contínuas. Então as afirmativas abaixo são equivalentes:

- → f é convexa
- → $f''(\mathbf{x})$ (matriz hessiana) é semi-definida positiva para todo $\mathbf{x} \in \mathbb{R}^m$
- → $\nabla f(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y}) - f(\mathbf{x})$ para todo $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^m$

Se a hessiana for positiva definida em todos os pontos, então f é estritamente convexa. Se uma função convexa tem um mínimo local, este mínimo é global.

ANPEC 2022, Questão 10. *Considere a função $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x, y) = x^2 + (1 + x)^3 y^2$ para todo $(x, y) \in \mathbb{R}^2$, e classifique as afirmações abaixo.*

1 *A função f tem mais do que um ponto crítico.*

Resolução: Note que $\nabla f(x, y) = (2x + 3y^2(1 + x)^2, 2y(1 + x)^3)$. Um ponto crítico satisfaz $y(1+x)^3 = 0$ e portanto $y = 0$ ou $x = -1$. Se $y = 0$ então $x = 0$. Se $x = -1$ então $\partial f/\partial x$ não se anula. Logo $(0, 0)$ é único ponto crítico. Logo, a afirmativa é falsa. ✓

2 *Existe um ponto de mínimo local para f.*

Resolução: Basta ver que a hessiana é dada por

$$H_f(x,y) = \begin{pmatrix} 2+6y^2(1+x) & 6y(1+x)^2 \\ 6y(1+x)^2 & 2(1+x)^3 \end{pmatrix}.$$

e então

$$H_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Como a hessiana é positiva definida no ponto crítico, então este ponto é mínimo local. Logo, a afirmativa é verdadeira. ✓

3 *Existe um ponto de máximo local para f.*

Resolução: Pelos itens **1** e **2**, o único ponto crítico é de mínimo. Logo, a afirmativa é falsa. ✓

4 *A função f assume um valor mínimo em seu domínio.*

Resolução: Sabemos que para $y \neq 0$,

$$\lim_{x \to -\infty} f(x,y) = -\infty.$$

Logo,a afirmativa é falsa. ✓

ANPEC 2022, Questão 13. *Julgue a veracidade da seguinte afirmativa.*

4 *A matriz hessiana associada à função $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x_1, x_2) = 2x_1 - 2x_1x_2 - 4x_2^2$ é semi-definida positiva em qualquer ponto do domínio, e logo a função f é convexa.*

Resolução: Note que

$$\nabla f = (2 - 2x_2, -2x_1 - 8x_2), \qquad H = \begin{pmatrix} 0 & -2 \\ -2 & -4 \end{pmatrix}.$$

Então H não é semidefinida positiva pois $\mathbf{e}_2^t H \mathbf{e}_2 = -4 < 0$, onde $\mathbf{e}_2 = (0,1)^t$. Logo, a afirmativa é falsa. ✓

ANPEC 2019, Questão 7. *Considere a função* $f : \mathbb{R}^4 \to \mathbb{R}$, *definida por* $f(x_1, x_2, x_3, x_4) = -(x_1)^2 + \sum_{k=1}^{4}(-x_k)^k$, *e verifique a veracidade das seguintes afirmações:*

0 *Seja* $\mathbf{x}^* = (x_1^*, x_2^*, x_3^*, x_4^*)$ *um ponto no* $\mathbb{R}^4$. *Para que* $\mathbf{x}^* \in \mathbb{R}^4$ *seja um ponto crítico é necessário que* $-2x_1^* - 1 = 2x_2^* = -3(x_3^*)^2 = 4(x_4^*)^3 = 0$.

Resolução: A função

$$f(x_1, x_2, x_3, x_4) = -x_1^2 - x_1 + x_2^2 - x_3^3 + x_4^4.$$

Se x^* é um ponto crítico então

$$\begin{aligned}\nabla f(x^*) &= \left(\frac{\partial f}{\partial x_1}(x^*), \frac{\partial f}{\partial x_2}(x^*), \frac{\partial f}{\partial x_3}(x^*), \frac{\partial f}{\partial x_4}(x^*)\right) \\ &= (-2x_1^* - 1, 2x_2^*, -3x_3^{*^2}, 4x_4^{*^3}) = (0,0,0,0).\end{aligned}$$

Logo, a afirmativa é verdadeira. ✓

1 *A matriz Hessiana* H *de* f *no ponto* $\mathbf{x} = (x_1, x_2, x_3, x_4)$ *é*

$$\begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -3x_3 & 0 \\ 0 & 0 & 0 & 4x_4 \end{pmatrix}.$$

Resolução: A matriz Hessiana

$$H(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_3}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_4}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_3}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_4}(x) \\ \frac{\partial^2 f}{\partial x_3 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_3 \partial x_2}(x) & \frac{\partial^2 f}{\partial x_3^3}(x) & \frac{\partial^2 f}{\partial x^3 \partial x_4}(x) \\ \frac{\partial^2 f}{\partial x_4 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_4 \partial x_2}(x) & \frac{\partial^2 f}{\partial x_4 \partial x_3}(x) & \frac{\partial^2 f}{\partial x_4^2}(x) \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & -6x_3 & 0 \\ 0 & 0 & 0 & 12x_4^2 \end{pmatrix}.$$

Logo, a afirmativa é falsa. ✓

2 *A matriz Hessiana* H *de* f *é indefinida em* R^4.

Resolução: Seja $y = (y_1, y_2, y_3, y_4)$. Então

$$yH(x)y^T = -2y_1^2 + 2y_2^2 - 6x_3y_3^2 + 12x_4^2y_2^2.$$

$$\text{para} \quad y = (1,0,0,0) \text{ temos } yH(x)y^T = -2 < 0,$$
$$\text{para} \quad y = (0,1,0,0) \text{ temos } yH(x)y^T = 2 > 0.$$

Portanto, a matriz Hessiana H é indefinida em todo o R^4. Logo, a afirmativa é verdadeira. ✓

3 *f possui um máximo local em* $\mathbf{x}^* = (1,0,0,0)$.
Resolução: Do item **2**, temos que a hessiana é indefinida em todos o R^4. Portanto, nenhum ponto crítico é extremo local.

De forma alternativa, do item **0**, temos

$$-2x_1^* - 1 = 2x_2^* = -3(x_3^*)^2 = 4(x_4^*)^2 = 0.$$

Logo $x_1^* = -1/2$, $x_2^* = x_3^* = x_4^* = 0$. Portanto o único candidato a extremo local é $(-1/2, 0, 0, 0)$. Logo, a alternativa é falsa. ✓

4 *O determinante da matriz Hessiana de f é positivo para todo* $\mathbf{x} \in \mathbb{R}^4$*, isto é,* $\det H_f(\mathbf{x}) > 0$.
Resolução: Do item **1**, temos $\det(H(x)) = 288x_3x_4^2$. Para $x = (0,0,0,0)$ tem-se $\det(H(x)) = 0$. Logo, a alternativa é falsa.

Para uma solução alternativa que dispensa o cálculo de determinantes, basta notar que $H_f(\mathbf{0})[0,0,1,1]^t = \mathbf{0}$, e portanto o núcleo de $H_f(\mathbf{0})$ é não trivial. Logo, $\det H_f(\mathbf{x}) = 0$ na origem. ✓

ANPEC 2007, Questão 5. *As seguintes informações são fornecidas: Se $f : U \to R$ uma função duas vezes diferenciável, definida em $U = \{(x,y) : x, y > 0\}$ e $H_f(x,y)$ a matriz hessiana de f no ponto $(x,y) \in U$. Avalie as afirmativas abaixo.*

0 *A função f é convexa se e somente se $H_f(x,y)$ é semidefinida positiva em todos os pontos de U .*

Resolução: O gabarito tem como falsa a afirmação, que tem toda a aparência de verdadeira. Possivelmente foi devido ao detalhe da f não ser duas vezes *continuamente* diferenciável em U, i.e., $C^2(U)$.

Olhar esta questão

✓

(1) *Se $f(x,y) = -x^{1/3}y^{1/4}$, então f é convexa.*

Resolução: Veja que

$$\nabla f = \begin{pmatrix} -\frac{1}{3}x^{-2/3}y^{1/4} & -\frac{1}{4}x^{1/3}y^{-3/4} \end{pmatrix}, \qquad H_f = \begin{pmatrix} \frac{2}{9}x^{-5/3}y^{1/4} & -\frac{1}{12}x^{-2/3}y^{-3/4} \\ -\frac{1}{12}x^{-2/3}y^{-3/4} & \frac{3}{16}x^{1/3}y^{-7/4} \end{pmatrix}.$$

Para determinar se H_f 'positiva definida, usamos o Destaque 5.9. Como para $(x,y) \in U$, o traço de H_f é positivo, precisamos apenas conferir o sinal do determinante:

$$\begin{aligned} \det H_f = \frac{2}{9}x^{-5/3}y^{1/4}\frac{3}{16}x^{1/3}y^{-7/4} - \frac{1}{12}x^{-2/3}y^{-3/4}\frac{1}{12}x^{-2/3}y^{-3/4} \\ = \left(\frac{1}{24} - \frac{1}{144}\right)x^{-4/3}y^{-6/4} > 0. \end{aligned}$$

Então H_f é positiva definida e a afirmativa é verdadeira. ✓

(2) *Se f é convexa, então $H_f(x,y)$ é positiva definida em todos os pontos de U.*

Resolução: Tome $f(x,y) = (x-1)^4 + (y-1)^4$. Então

$$\nabla f = (4(x-1)^3, 4(y-1)^3), \qquad H_f(x,y) = 12\begin{pmatrix} (x-1)^2 & 0 \\ 0 & (y-1)^2 \end{pmatrix}.$$

Note que f não é positiva definida no ponto $(1,1) \in U$. A afirmativa é falsa. ✓

(3) *Se $f(x,y) = x^2y^2$, então f é convexa.*

Resolução: Sem fazer contas, veja que a função se anula nos eixos e é positiva fora deles. Não pode mesmo ser convexa. Fazendo as contas, $\nabla f = (2xy^2, 2x^2y)$ e

$$H_f(x,y) = 2\begin{pmatrix} y^2 & 2xy \\ 2xy & x^2 \end{pmatrix}.$$

Como $\det H_f(x,y) = -6x^2y^2 < 0$ em $(x,y) = (1,1)$ por exemplo, a hessiana é indeterminada, e a função f não pode ser convexa. Logo, a afirmativa é falsa. ✓

4 *Se f é convexa e (x_0, y_0) é um ponto crítico de f, então $f(x,y) \geq f(x_0, y_0)$, para todo $(x,y) \in U$.*

Resolução: Um ponto crítico de uma função convexa é mínimo global. Logo, a afirmativa é verdadeira. ✓

5.5. Teorema da Função Implícita

O Teorema da função implícita determina condições em que pode-se "escrever" uma variável em função de outras. Por exemplo, seja $F : \mathbb{R}^2 \to \mathbb{R}$ suave e $(x_0, y_0) \in \mathbb{R}^2$ tal que $F(x_0, y_0) = c$ para alguma constante $c \in \mathbb{R}$, e $\partial F/\partial y(x_0, y_0) \neq 0$. Então existe função ϕ definida numa vizinhança de x_0 e tal que $\phi(x_0) = y_0$ e $F(x, \phi(x)) = c$. Note que

$$0 = \frac{d}{dx}\big[F(x, \phi(x))\big] = \frac{\partial F}{\partial x}(x, \phi(x)) + \frac{\partial F}{\partial y}(x, \phi(x))\phi'(x) \implies \phi'(x) = -\frac{\partial F/\partial x}{\partial F/\partial y}.$$

O Teorema de função implícita trata da importante questão de solvabilidade de equações dadas de forma implícita. A pergunta é simples: dados os pontos (x, y), soluções de uma equação $F(x, y) = 0$, será que é possível escrever y em função de x, numa vizinhança de x?

Como uma primeira motivação, considere $F(x, y) = x^2 + y^2 - 1$. Então a curva de nível determinada por $F(x, y) = 0$ é dada pelo círculo de raio unitário, como nos mostra a Figura 3 (esquerda). Seja $(a, b) \in \mathbb{R}^2$ tal que $F(a, b) = 0$. Por exemplo $(0, 1)$ e $(0, -1)$ satisfazem esta condição. Numa vizinhaça de $(0, 1)$, podemos escrever y como função de x, e.g., $y = \sqrt{1 - x^2}$. E em torno de $(0, -1)$ podemos escrever $y = -\sqrt{1 - x^2}$. Uma pergunta natural é se existe uma função ϕ tal que $F(x, \phi(x)) = 0$, e $\phi(a) = b$. A resposta é *globalmente, não.* Mas *localmente sim*, se $\partial F/\partial y(a, b) \neq 0$.

Um segundo exemplo é dado por $F(x, y) = x - y^2$, ver Figura 3 (direita). Para se ter $F(x, \phi(x)) = 0$, pode-se escolher $\phi(x) = \sqrt{x}$ ou $\phi(x) = -\sqrt{x}$. Entretanto nenhuma das duas funções está definida em intervalos que contenham $x = 0$. Note que $\partial F/\partial y(0, 0) = 0$.

Considere então $F : \mathbb{R}^2 \to \mathbb{R}$ função suave e seja $(x_0, y_0) \in \mathbb{R}^2$ tal que $F(x_0, y_0) = c$ para alguma constante $c \in \mathbb{R}$, e $\partial F/\partial y(x_0, y_0) \neq 0$. Então, aplicando o Teorema

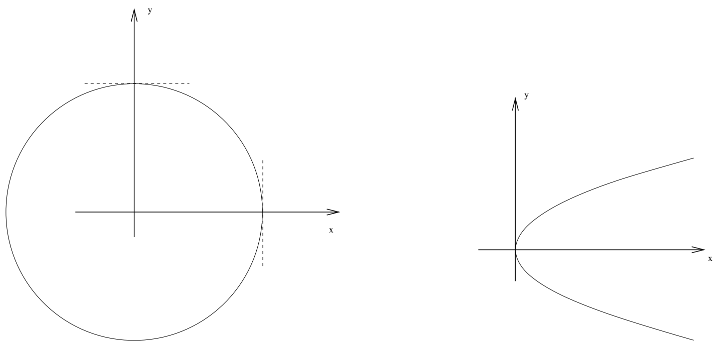

Figura 3. Conjuntos $\{(x,y) \in \mathbb{R}^2 : x^2+y^2=1\}$ (esquerda) e $\{(x,y) \in \mathbb{R}^2 : x=y^2\}$ (direita).

de função implícita para $F(x_0,y_0)-c$, existe função ϕ definida numa vizinhança de x_0 e tal que $\phi(x_0)=y_0$ e $F(x,\phi(x))=c$. Note que

$$0=\frac{d}{dx}F(x,\phi(x))=\frac{\partial F}{\partial x}(x,\phi(x))+\frac{\partial F}{\partial y}(x,\phi(x))\phi'(x),$$

e portanto $\phi'(x)=-\frac{\partial F/\partial x}{\partial F/\partial y}$ calculada no ponto $(x,\phi(x))$.

O caso tridimensional é semelhante, i.e., suponha $F:\mathbb{R}^3\to\mathbb{R}$ diferenciável com derivadas contínuas, e (x_0,y_0,z_0) tal que $F(x_0,y_0,z_0)=0$. Se $\partial f/\partial z(x_0,y_0,z_0)\neq 0$, então existe uma vizinhança W contendo (x_0,y_0), e uma única função $\phi:W\to\mathbb{R}$, que é $\mathcal{C}^1(W)$, tal que $z_0=\phi(x_0,y_0)$ e $F(x,y,\phi(x,y))=0$ para todo $(x,y)\in W$. Além disto, pela regra da cadeia

$$0=\frac{\partial}{\partial x}\big[F(x,y,\phi(x,y))\big]=\frac{\partial F}{\partial x}(x,y,\phi(x,y))+\frac{\partial F}{\partial z}(x,y,\phi(x,y))\frac{\partial}{\partial x}\phi(x,y),$$
$$\implies \frac{\partial \phi}{\partial x}(x,y)=-\frac{\frac{\partial F}{\partial x}(x,y,\phi(x,y))}{\frac{\partial F}{\partial z}(x,y,\phi(x,y))}.$$

Analogamente,

$$\frac{\partial \phi}{\partial y}(x,y)=-\frac{\frac{\partial F}{\partial y}(x,y,\phi(x,y))}{\frac{\partial F}{\partial z}(x,y,\phi(x,y))}.$$

DESTAQUE 5.11: FUNÇÃO IMPLÍCITA

Seja $F : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathcal{C}^1(\Omega)$ (diferenciável e com derivadas contínuas), e (x_0, y_0) tal que $F(x_0, y_0) = 0$. Se $\partial F/\partial y(x_0, y_0) \neq 0$, então existe um intervalo W contendo x_0, e uma única função $\phi : W \to \mathbb{R}$, que é $\mathcal{C}^1(W)$, tal que $y_0 = \phi(x_0)$ e $F(x, \phi(x)) = 0$ para todo $x \in W$. Além disto, vale que $\phi'(x) = -\frac{\partial F/\partial x}{\partial F/\partial y}$ calculada no ponto $(x, \phi(x))$.

O Teorema da função implícita *não* garante existência de um ponto (x_0, y_0) tal que $f(x_0, y_0) = 0$. Por exemplo, considere $F : \mathbb{R}^2 \to \mathbb{R}$ dada por $F(x, y) = x^2+y^2+1$. Então F não tem raiz, apesar de $\partial F/\partial y(x_0, y_0) = 2y \neq 0$ sempre que $y \neq 0$.

ANPEC 2022, Questão 12. *Dado um parâmetro $a \in \mathbb{R}$, considere a função $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x, y) = -x^2 - xy - 2y^2 + 2ax + 2ay$ e o problema de maximização:*

$$\max_{(x,y)\in\mathbb{R}^2} f(x, y).$$

Denotando por $(x^*, y^*) = (x^*(a), y^*(a))$ a solução deste problema, seja $f^* : \mathbb{R} \to \mathbb{R}$ tal que $f^*(a) = f(x^*, y^*)$ a função valor correspondente. Calcule a derivada de f^* no ponto $a = 14$, ou seja, o valor de $\frac{df^*}{da}(14)$.

Note que $\nabla f(x, y) = (-2x - y + 2a, -x - 4y + 2a)$, e

$$H_f = \begin{pmatrix} -2 & -1 \\ -1 & -4 \end{pmatrix}.$$

Como o determinante da hessiana é maior que zero, ambos autovalores têm o mesmo sinal. Logo, ela é positiva ou negativa definida. Mas como os termos da diagonal são negativos, i.e.,

$$\mathbf{e}_1^T H_f \mathbf{e}_1 = -2 < 0, \qquad \mathbf{e}_2^T H_f \mathbf{e}_2 = -4 < 0,$$

então a hessiana tem que ser negativa definida (uma forma mais rápida para se chegar à mesma conclusão é vendo que $\operatorname{tr} H_f < 0$). Em particular, a função é estritamente côncava.

Para achar pontos críticos, basta igualar o gradiente a zero. Portanto,

$$\begin{cases} x + 4y = 2a \\ 2x + y = 2a \end{cases}$$

Depois de algumas contas vê-se que a única solução é $(x^*, y^*) = (6a/7, 2a/7)$. Como função é côncava, este ponto crítico é de máximo. Substituindo, temos

$$f^*(a) = f(x^*, y^*) = a^2\left(-\frac{36}{49}-\frac{12}{49}-\frac{8}{49}+\frac{12}{7}+\frac{4}{7}\right) = \frac{a^2}{49}(-36-12-8+84+28) = \frac{56a^2}{49}.$$

Portanto,

$$\frac{df^*(a)}{da} = \frac{16a}{7} \implies \frac{df^*(14)}{da} = 32.$$

ANPEC 2022, Questão 13. *Julgue a veracidade das afirmativas abaixo.*

1 *A equação* $(x-2)^3+x(y-1)^2-\ln y = 1$ *define implicitamente* y *como função de* x *em uma vizinhança do ponto* $(3,1) \in \mathbb{R}^2$*, e denotamos para expressar isso* $y = h(x)$*. Esta função satisfaz* $\frac{dh}{dx}(3) = \frac{1}{2}$*.*

Resolução: Seja $f(x,y) = (x-2)^3 + x(y-1)^2 - \ln y$. Note que

$$\frac{\partial f}{\partial x} = 3(x-2)^2 + (y-1)^2, \qquad \frac{\partial f}{\partial y} = 2x(y-1) - \frac{1}{y}.$$

Como $\partial f/\partial y(3,1) = -1 \neq 0$, podemos aplicar o Teorema da função implícita. Logo, existe função tal que $y = h(x)$ e $f(x,h(x)) = 1$. Derivando em relação à x,

$$\frac{\partial f}{\partial x}(x,h(x)) + \frac{\partial f}{\partial y}(x,h(x))h'(x) = 0 \implies h'(x) = -\frac{\frac{\partial f}{\partial x}(x,h(x))}{\frac{\partial f}{\partial y}(x,h(x))}.$$

no ponto $(3,1)$ temos $h'(x) = 3$. Logo, a afirmativa é falsa. ✓

2 *Considerando a função* $f : \mathbb{R} \times [0,1) \to \mathbb{R}$ *definida por* $f(x,y) = (x^2+y^2)(ye^{|x|} - 1)$*, para cada* $x \in \mathbb{R}$ *existe um único* $y = h(x) \in [0,1)$ *tal que* $f(x,h(x)) = 0$*, o que define uma função contínua* $h : \mathbb{R} \to [0,1)$*.*

Resolução: Quando $x = 0$, temos que $h(0)^2(h(0)-1) = 0$. Como $h(0) = 1$ não é possível (pois 1 não pertence ao contradomínio de h), então $h(0) = 0$.

Para $x \neq 0$, tem-se $h(x) = e^{-|x|}$. Portanto a função h é dada por

$$h(x) = \begin{cases} 0 & \text{se } x = 0, \\ e^{-|x|} & \text{se } x \neq 0. \end{cases}$$

Como esta função não é contínua, a afirmativa é falsa. ✓

ANPEC 2020, Questão 5. *Seja a função $z(x, y)$ definida implicitamente por*

$$F(x, y, z) = x \ln z + z \ln y - z^2 - x \ln y + x = 0.$$

Julgue as seguintes afirmativas a seguir.

0 *O valor de F no ponto $(e^2 - e, e, e)$ é zero.*
Resolução:

$$F(e^2 - e, e, e) = (e^2 - e) \ln e + e \ln e - e^2 - (e^2 - e) \ln e + e^2 - e = 0.$$

Logo, a afirmativa é verdadeira. ✓

1 *O valor $\frac{\partial F}{\partial z}$ no ponto $(e^2 - e, e, e)$ é zero.*
Resolução: A derivada parcial de F com respeito à variável z é

$$\frac{\partial F(x, y, z)}{\partial z} = \frac{x}{z} + \ln\ y - 2z.$$

Para $(x, y, z) = (e^2 - e, e, e)$, temos

$$\frac{\partial F(e^2 - e, e, e)}{\partial z} = \frac{e^2 - e}{e} + \ln\ e - 2e = -e.$$

Logo, a afirmativa é falsa. ✓

2 *z não pode ser definida implicitamente como função de (x, y) ao redor do ponto $(e^2 - e, e, e)$.*
Resolução: Pelos itens **0** e **1**

$$F(e^2 - e, e, e) = 0, \qquad \frac{\partial F(e^2 - e, e, e)}{\partial z} \neq 0,$$

Então, pelo Teorema da função implícita, z pode ser definida implicitamente como função de (x, y) ao redor do ponto $(e^2 - e, e, e)$. Logo, a afirmativa é falsa. ✓

3 *O vetor* $\left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right)$ *no ponto* $(e^2 - e, e, e)$ *é* $\left(\frac{1}{e}, \frac{2}{e} - 1\right)$.

Resolução: Definindo z como função de (x, y), temos $F(x, y, z(x, y)) = 0$, e derivando em relação à x e à y temos

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = 0 \implies \frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}},$$

$$\frac{\partial F}{\partial y} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial y} = 0 \implies \frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}},$$

As derivadas parciais de F são

$$\frac{\partial F(x, y, z)}{\partial x} = \ln z - \ln y + 1, \qquad \frac{\partial F(x, y, z)}{\partial y} = \frac{z}{y} - \frac{x}{y},$$

$$\frac{\partial F(x, y, z)}{\partial z} = \frac{x}{z} + \ln y - 2z.$$

Para $(x, y, z) = (e^2 - e, e, e)$, temos

$$\frac{\partial F(e^2 - e, e, e)}{\partial x} = 1, \quad \frac{\partial F(e^2 - e, e, e)}{\partial y} = 2 - e, \quad \frac{\partial F(e^2 - e, e, e)}{\partial z} = -e.$$

Assim,

$$\frac{\partial z(e^2 - e, e)}{\partial x} = \frac{-\frac{\partial F}{\partial x}(e^2 - e, e, e)}{\frac{\partial F}{\partial z}(e^2 - e, e, e)} = \frac{-1}{-e} = \frac{1}{e},$$

$$\frac{\partial z(e^2 - e, e)}{\partial y} = \frac{-\frac{\partial F}{\partial y}(e^2 - e, e, e)}{\frac{\partial F}{\partial z}(e^2 - e, e, e)} = -\frac{2 - e}{-e} = \frac{2}{e} - 1.$$

Logo, a afirmativa é verdadeira. ✓

5.6. Otimização com restrições

Para problemas de otimização com restrições de igualdade, o Teorema de Lagrange nos dá condições suficientes para que um ponto seja extremo. Na prática, os problemas de otimição com restrições que têm sido objeto de questões nas provas da ANPEC podem ser resolvidos "somente" com o Teorema de Lagrange. Se a restrição é de igualdade, então este Teorema fornece as condições suficientes para se buscar os pontos candidatos a serem extremos. Se a restrição é de desigualdade ($\leq$ ou $\geq$), então basta buscar candidatos que satisfaçam a desigualdade estrita ($<$ ou $>$) e depois pontos que satisfaçam a igualdade na restrição, usando o Teorema de Lagrange.

Dadas funções reais $f, g_1, \dots, g_k$ definidas num aberto Ω de $\mathbb{R}^m$, consideramos o problema de minimizar f restrita ao conjunto de raízes de $g_1, \dots, g_k$ em Ω. O Teorema de Lagrange nos dá condições necessárias que um candidato a mínimo de tal problema tem que satisfazer.

Destaque 5.12: Teorema de Lagrange

Sejam $f, g_1, \dots, g_k$ funções reais definidas em $\mathbb{R}^m$ com derivadas contínuas. Suponha que exista $\delta > 0$ e $\mathbf{x}^* \in \mathbb{R}^m$ e tal que

$$f(\mathbf{x}^*) = \min\{f(\mathbf{x}) : \mathbf{x} \in B_\delta(\mathbf{x}^*) \text{ e } g_1(\mathbf{x}) = \cdots = g_k(\mathbf{x}) = 0\}.$$

Então existem números $\mu, \lambda_1^*, \dots, \lambda_k^*$ não todos nulos e tais que

$$\mu \nabla f(\mathbf{x}^*) = \lambda_1^* \nabla g_1(\mathbf{x}^*) + \cdots + \lambda_k^* \nabla g_k(\mathbf{x}^*). \tag{5.6}$$

Além disto, se $\{\nabla g_1(\mathbf{x}^*), \dots, \nabla g_k(\mathbf{x}^*)\}$ é linearmente independente, então pode-se tomar $\mu = 1$.

Observação 5.4. Note que uma forma compacta de descrever as condições de Lagrange vem através da definição do Lagrangeano $\mathcal{L} : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}$ dado por

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda} \cdot \mathbf{g}(\mathbf{x}), \tag{5.7}$$

onde $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$ e $\mathbf{g}(\mathbf{x}) = \big(g_1(\mathbf{x}), \dots, g_k(\mathbf{x})\big) \in \mathbb{R}^k$. Então as condições necessárias para $\mathbf{x}^*$ ser ponto de mínimo são dadas por $\nabla \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0$, i.e.,

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \nabla_{\mathbf{x}} f\big(\mathbf{x}^*\big) - \boldsymbol{\lambda}^* \cdot \nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x}^*) = \mathbf{0}, \quad \nabla_{\boldsymbol{\lambda}} \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = -\mathbf{g}(\mathbf{x}^*) = \mathbf{0}. \tag{5.8}$$

Os números $\lambda_1^*, \dots, \lambda_k^*$ acima são conhecidos por *multiplicadores de Lagrange*, e em muitas aplicações têm significado próprio.

Um caso particular importante é quando a função a ser otimizada é linear não constante restrita a um domínio poligonal. Neste caso, os pontos de ótimos *sempre* estarão na fronteira do domínio: ou todos os pontos de um lado inteiro do polígono são de ótimo, ou apenas os vértices do polígono são candidatos a pontos extremos. Ver a Seção 2.6 para maiores detalhes.

ANPEC 2021, Questão 10. *Considere a seguinte questão:*

Seja a função $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x) = -\frac{x_1^2}{4} + 4\sqrt{5}\,x_1 + x_2$, em que $x = (x_1, x_2)$. Dados dois parâmetros positivos $p, m \in (0, +\infty)$, defina o conjunto $A = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0, px_1 + x_2 = m\}$. Encontre o menor valor de m que faz com que, para qualquer parâmetro p satisfazendo $0 < p < 4\sqrt{5}$, a condição $\frac{\partial f(x)}{\partial x_1} = p$ forneça a solução para o problema de maximizar $f(x)$ sujeito a $x \in A$.

Resolução: De $x_2 = m - px_1$ e $x_2 \geq 0$, buscamos $x_1^* \in [0, m/p]$ tal que

$$g(x_1^*) = \max_{x \in [0,m/p]} g(x), \qquad \text{onde } g(x) = -\frac{x_1^2}{4} + 4\sqrt{5}x_1 - px_1 + m.$$

Quando $x_1^* \in (0, m/p)$ então

$$0 = g'(x_1^*) = -\frac{x_1^*}{2} + 4\sqrt{5} - p$$

fornece a solução para maximizar f. Neste caso,

$$2(4\sqrt{5} - p) = x_1^* \leq \frac{m}{p}. \tag{5.9}$$

Seja m^* o menor valor de m tal que (5.9) valha para todo $p \in (0, 4\sqrt{5})$. Como $2p(4\sqrt{5} - p)$ atinge seu máximo em $2\sqrt{5}$, então $m^* = 2 \times 2\sqrt{5}(4\sqrt{5} - 2\sqrt{5}) = 40$. ✓

ANPEC 2019, Questão 12. *Considere o problema do investidor que pode investir os pesos w_1 e w_2 de sua riqueza em dois instrumentos financeiros arriscados. Suas preferências implicam que ele quer maximizar a função*

$$U(w_1, w_2) = 1,15w_1 + 1,2w_2 - 0,5(0,04w_1^2 + 0,09w_2^2).$$

Sujeita às restrições:

$$w_1 + w_2 = 1,$$
$$w_1 \geq 0 \quad e \quad w_2 \geq 0.$$

4 *A solução do problema de maximização com restrição acima é* $(w_1^*, w_2^*) = (0, 1)$*, ou seja, o investidor prefere investir todo o seu dinheiro em apenas um instrumento financeiro.*

Resolução: Pelo item **1**,

$$\nabla U = \left(\frac{115}{100} - \frac{4}{100} w_1, \frac{12}{10} - \frac{9}{100} w_2 \right).$$

Ignorando por um momento as condições de positividade e impondo somente a restrição $w_1 + w_2 = 1$, temos que

$$\begin{cases} \frac{115}{100} - \frac{4}{100} w_1 = \lambda \\ \frac{12}{10} - \frac{9}{100} w_2 = \lambda \\ w_1 + w_2 = 1 \end{cases}$$

e conluímos que $w_1 = 4/13$ e $w_2 = 9/13$. Assim,

$$U(4/13, 9/13) > \max\{U(0,1), U(1,0)\}.$$

Logo, a afirmativa é falsa, apesar de ter o gabarito oficial como verdadeira. ✓

ANPEC 2015, Questão 9. *Considere as seguintes informações: Seja* $f : D \to R$ *a função definida por* $f(x, y) = x^2 - 2xy + 2y$*, em que*

$$D = \{(x, y) \in R^2 : 0 \leq x \leq 3 \ \ e \ \ 0 \leq y \leq 2\}.$$

0 *A função* f *possui um único ponto de mínimo em* D*;*

Resolução: Note que

$$\nabla f = (2x - 2y, -2x + 2), \qquad H = \begin{pmatrix} 2 & -2 \\ -2 & 0 \end{pmatrix}.$$

Como $\det H = -4$, a hessiana é indefinida (tem autovalores de sinais opostos). Logo, nenhum ponto crítico é mínimo ou máximo. Então o mínimo ocorre na fronteira. Vamos analisar cada caso:

(1) $y = 0 \implies f = x^2$. Logo a função alcança o mínimo em $(0, 0)$.
(2) $x = 0 \implies f = 2y$. Logo a função alcança o mínimo em $(0, 0)$.
(3) $y = 2 \implies f = x^2 - 4x + 4$. Logo a função alcança o mínimo em $(2, 2)$.
(4) $x = 3 \implies f = -4y + 9$. Logo a função alcança o mínimo em $(3, 2)$.

Temos portanto três candidatos a mínimo. Calculando os valores de f:

$$f(0, 0) = 0, \qquad f(2, 2) = 0, \qquad f(3, 2) = 1.$$

Assim, f tem dois mínimos em D. Logo, a afirmativa é falsa. ✓

1 *$(1, 1)$ é ponto de mínimo local de f em D;*
Resolução: Basta ver a solução do item **0**. Logo, a afirmativa é falsa. ✓

2 *O valor máximo absoluto da função em D é 9;*
Resolução: Basta repetir a argumentação do item **0** e calcular f nos pontos $(0, 1)$, $(3, 2)$ e $(3, 0)$:

$$f(0, 1) = 2, \qquad f(3, 2) = 1, \qquad f(3, 0) = 9.$$

Logo, a afirmativa é verdadeira. ✓

3 *O máximo é atingido na fronteira de D;*
Resolução: Pelo item **0**, a função não possui pontos de máximo ou mínimo no interior. Logo, a afirmativa é verdadeira. ✓

4 *A função é côncava em D.*
Resolução: Basta ver que a hessiana calculada no item **0** é indefinida. Logo, a afirmativa é falsa. ✓

ANPEC 2012, Questão 15. *Seja* (x^*, y^*) *o ponto de* R^2 *que maximiza* $f(x,y) = x^2y$ *sujeita à restrição* $2x^2 + y^2 \leq 9$. *Encontre* $a = [f(x^*, y^*)]^2$.

Resolução: Seja $g(x) = 2x^2 + y^2 - 9$. Note que

$$\nabla f = (2xy, x^2), \qquad \nabla g = (4x, 2y).$$

Portanto os pontos críticos estão concentrados em $x = 0$. Mas $f(0, y) = 0$ para todo y, e portanto estes pontos não são de máximo. Concluímos que o(s) ponto(s) de máximo se encontram na fronteira do domínio, i.e., satisfazem $2x^2 + y^2 - 9 = 0$. Podemos então resolver este problema de otimização com restrição de igualdade. Usando o Teorema de Lagrange, existe λ tal que $\nabla f = \lambda \nabla g$, i.e.,

$$\begin{cases} 2xy = 4\lambda x \\ x^2 = 2\lambda y \\ 2x^2 + y^2 = 9 \end{cases}.$$

Como sabemos que o ponto de máximo tem $x \neq 0$, podemos dividir por x e concluir da primeira equação que $y = 2\lambda$. Note que $y > 0$ no ponto de máximo, e portanto $\lambda > 0$. Da segunda equação temos $x^2 = 4\lambda^2$. Finalmente, da terceira equação temos $\lambda = 3/(2\sqrt{3})$. Logo,

$$[f(x,y)]^2 = x^4y^2 = 16\lambda^4 4\lambda^2 = 64\lambda^6 = 64\frac{3^6}{2^6 \times 3^3} = 27.$$

✓

5.6.1. Justificativas e Extensões. Justificamos a seguir o Teorema de Lagrange. Começamos com um caso mais simples, de funções definidas no plano e onde há apenas uma restrição. Sejam f e g funções do $\mathbb{R}^2$ em $\mathbb{R}$, e suponha que $\mathbf{x}^* \in \mathbb{R}^2$ seja tal que

$$f(\mathbf{x}^*) = \min\{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^2,\ g(\mathbf{x}) = 0\}.$$

Impor $\mathbf{x}^*$ como mínimo global é somente uma simplificação. A teoria não se modifica em nada se $\mathbf{x}^*$ for somente mínimo local.

Suponha agora que $\nabla g(\mathbf{x}^*) \neq 0$ e considere $\mathbf{r} : \mathbb{R} \to \mathbb{R}^2$ uma parametrização para a curva de nível determinada por $g(x,y) \equiv 0$, e tal que $\mathbf{r}(0) = \mathbf{x}^*$. A existência de tal parametrização (local) pode ser justificada através do Teorema da função implícita. Pela regra da cadeia,

$$0 = \frac{d}{dt} g(\mathbf{r}(t)) = \nabla g(\mathbf{r}(t)) \cdot \mathbf{r}'(t)$$

concluímos que $\nabla g(\mathbf{r}(t))$ é ortogonal a $\mathbf{r}'(t)$. De forma análoga, temos que $f(\mathbf{r}(t))$ atinge seu máximo em $t = 0$. Logo, $df(\mathbf{r}(0))/dt = 0$, e como

$$\frac{d}{dt} f(\mathbf{r}(t)) = \nabla f(\mathbf{r}(t)) \cdot \mathbf{r}'(t),$$

temos que $\nabla f(\mathbf{x}^*)$ é também ortogonal a $\mathbf{r}'(0)$. Em duas dimensões, se $\nabla g(\mathbf{x}^*)$ e $\nabla f(\mathbf{x}^*)$ são ortogonais a $\mathbf{r}'(0)$ isto significa que $\nabla g(\mathbf{x}^*)$ e $\nabla f(\mathbf{x}^*)$ são colineares, i.e., existe $\lambda^* \in \mathbb{R}$ tal que

$$\nabla g(\mathbf{x}^*) = \lambda^* \nabla f(\mathbf{x}^*).$$

Apresentamos agora, seguindo **[9]**, alguns argumentos que indicam o porquê do Teorema de Lagrange valer no caso geral. A argumentação é toda baseada na aproximação dada pelo Teorema do Valor Médio:

$$g_i(\mathbf{x}^* + \mathbf{s}) \approx g_i(\mathbf{x}^*) + \nabla g_i(\mathbf{x}^*) \cdot \mathbf{s}, \tag{5.10}$$

O vetor $\mathbf{s} \in \mathbb{R}^m$ é uma *direção factível* se $\mathbf{x}^* + \mathbf{s}$ satisfizer todas as restrições, i.e., $g_i(\mathbf{x}^*+\mathbf{s}) = 0$ para todo $i = 1, \dots, k$. Logo, para termos $\mathbf{s}$ factível, temos $g_i(\mathbf{x}^*+\mathbf{s}) = g_i(\mathbf{x}^*) = 0$. Logo, por (5.10), vemos que $\mathbf{s}$ é factível se e somente se

$$\nabla g_i(\mathbf{x}^*) \cdot \mathbf{s} = 0 \text{ para } i = 1, \dots, k. \tag{5.11}$$

Note que a derivada direcional da f em $\mathbf{x}^*$ na direção factível $\mathbf{s}$ não pode ser negativa, pois isto indicaria que $f(\mathbf{x}^* + \mathbf{s}) < f(\mathbf{x}^*)$, uma contradição com $\mathbf{x}^*$ ser mínimo local. Logo

$$\nabla f(\mathbf{x}^*) \cdot \mathbf{s} < 0 \tag{5.12}$$

não pode ocorrer junto com (5.11).

A condição necessária e suficiente para que (5.11) e (5.12) nunca ocorram simultaneamente é que existam $\lambda_1^*, \dots, \lambda_k^*$ tais que

$$\nabla f(\mathbf{x}^*) = \lambda_1^* \nabla g_1(\mathbf{x}^*) + \cdots + \lambda_k^* \nabla g_k(\mathbf{x}^*).$$

É imediato checar que a condição é suficiente. Para ver que é também necessária, suponha que (5.11) e (5.12) não ocorram simultaneamente para nenhuma direção factível $\mathbf{s}$, e considere P o subespaço vetorial formado por $\nabla g_1(\mathbf{x}^*), \dots, \nabla g_k(\mathbf{x}^*)$. Se $\nabla f(\mathbf{x}^*) \notin P$ então existe $\mathbf{u}$ não nulo e ortogonal a P. Por (5.11), $\mathbf{u}$ é factível. Seja $\mathbf{v} \in P$ tal que $\mathbf{u} + \mathbf{v} = \nabla f(\mathbf{x}^*)$. Então $\mathbf{s} = -\mathbf{u}$ satisfaz (5.11) e (5.12), uma contradição.

5.6.1.1. *Restrições de desigualdade.* Uma outra situação de minimização com restrições ocorre quando as restrições são dadas por desigualdades. Neste caso temos o Teorema de Kuhn–Tucker, dado abaixo.

Destaque 5.13: Teorema de Kuhn–Tucker

Sejam $f, h_1, \dots, h_k$ funções reais definidas em $\mathbb{R}^m$ com derivadas contínuas. Suponha que existam $\delta > 0$ e $\mathbf{x}^* \in \mathbb{R}^m$ tais que

$$f(\mathbf{x}^*) = \min\{f(\mathbf{x}) : \mathbf{x} \in B_\delta(\mathbf{x}^*) \text{ e } h_1(\mathbf{x}) \geq 0, \cdots, h_k(\mathbf{x}) \geq 0\}.$$

Então as seguintes afirmativas são verdadeiras:

(1) existem números $\mu, \lambda_1^*, \dots, \lambda_k^*$ não todos nulos e tais que

$$\mu \nabla f(\mathbf{x}^*) = \lambda_1^* \nabla h_1(\mathbf{x}^*) + \cdots + \lambda_k^* \nabla h_k(\mathbf{x}^*).$$

(2) seja $i \in \{1, \dots, k\}$ tal que $h_i(\mathbf{x}^*) > 0$. Então pode-se impor $\lambda_i^* = 0$.

(3) se o conjunto $V = \{\nabla h_i(\mathbf{x}^*) : h_i(\mathbf{x}^*) = 0, \text{ onde } 1 \leq i \leq k\}$ é linearmente independente, então pode-se tomar $\mu = 1$ e $\lambda_1^* \geq 0, \dots, \lambda_k^* \geq 0$.

Justificamos a seguir o Teorema de Kuhn–Tucker, seguindo novamente [**9**]. Podemos supor, sem perda de generalidade, que $h_i(x^*) = 0$ para todo $i = 1, \dots, k$. De fato, se $h_i(x^*) > 0$, então podemos tomar uma vizinhança de $\mathbf{x}^*$ tal que h_i seja estritamente positiva nesta vizinhança (pois h_i é positiva). Podemos então tomar $\lambda_i^* = 0$ nestes casos.

Se $\mathbf{s}$ é direção factível, i.e., satisfaz as desigualdades, então, $h_i(\mathbf{x}^* + \mathbf{s}) \geq 0$. Logo, como $h_i(\mathbf{x}^*) = 0$, por (5.10), temos

$$\nabla h_i(\mathbf{x}^*) \cdot \mathbf{s} \geq 0 \text{ para } i = 1, \dots, k. \tag{5.13}$$

Desde que a derivada direcional da f em $\mathbf{x}^*$ na direção factível $\mathbf{s}$ não pode ser negativa, temos que

$$\nabla f(\mathbf{x}^*) \cdot \mathbf{s} < 0 \tag{5.14}$$

não pode ocorrer junto com (5.13).

A condição necessária e suficiente para que (5.13) e (5.14) nunca ocorram simultaneamente é que existam $\lambda_1^* \geq 0, \dots, \lambda_k^* \geq 0$ tais que

$$\nabla f(\mathbf{x}^*) = \lambda_1^* \nabla h_1(\mathbf{x}^*) + \dots + \lambda_k^* \nabla h_k(\mathbf{x}^*).$$

Novamente, é fácil checar que a condição é suficiente. Para ver que $\nabla f(\mathbf{x}^*)$ é combinação linear de $\nabla h_1(\mathbf{x}^*), \dots, \nabla h_k(\mathbf{x}^*)$, basta argumentar como no caso de restrição estrita. Sem perda de generalidade suponha agora que $\lambda_1^* < 0$. Seja P o espaço vetorial formado pelas combinações lineares de $\{\nabla h_2(\mathbf{x}^*), \dots, \nabla h_k(\mathbf{x}^*)\}$. Seja agora $\mathbf{s}$ vetor ortogonal ao subespaço P, e "apontando na mesma direção que $\nabla h_1(\mathbf{x}^*)$", i.e., tal que $\nabla h_1(\mathbf{x}^*) \cdot \mathbf{s} \geq 0$. Por construção temos então que (5.13) e (5.14) ocorrem com tal $\mathbf{s}$, uma contradição. Logo temos sempre $\lambda_i^* \geq 0$. Note que no argumento acima usamos que $\nabla h_1(\mathbf{x}^*)$ é linearmente independente de $\nabla h_2(\mathbf{x}^*), \dots, \nabla h_k(\mathbf{x}^*)$ (caso contrário $\nabla h_1(\mathbf{x}^*) \in P$). Esta hipótese de independência linear é um artifício desta demonstração, e pode ser eliminada [**9**].

5.6.1.2. *Restrições de positividade.* Seja $f : \mathbb{R}^n \to \mathbb{R}$ "suave" e considere o seguinte problema de otimização com restrições:

$$\min_{\mathbf{x} \in \mathbb{R}^m} f(\mathbf{x}) \quad \text{tal que } \mathbf{x} \geq \mathbf{0}. \tag{5.15}$$

e suponha que $\mathbf{x}^*$ resolva (5.15). Seja $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ base canônica de $\mathbb{R}^m$ e $i \in \{1, \dots, m\}$. Seja $f_i : [0, \infty) \to \mathbb{R}$ dada por

$$f_i(\tau) = f(x_1^*, x_2^*, \dots, x_{i-1}^*, \tau, x_{i+1}^*, \dots, x_m^*).$$

Note então que f_i atinge mínimo em $\tau = x_i^*$. Se $x_i = 0$, então $f_i'(x_i^*) \geq 0$, caso contrário, $f_i'(x_i^*) = 0$. Podemos escrever estas condições de forma resumida por

$$f_i'(x_i^*) \geq 0, \qquad x_i^* f_i'(x_i^*) = 0, \qquad x_i^* \geq 0.$$

Como $f_i'(x_i^*) = \partial f/\partial x_i(x_i^*)$, reescrevemos as condições acima por

$$\nabla f(\mathbf{x}^*) \geq \mathbf{0}, \qquad \nabla f(\mathbf{x}^*)\cdot\mathbf{x}^* = 0, \qquad \mathbf{x}^* \geq \mathbf{0}.$$

5.6.1.3. *Restrições de positividade com desigualdade.* Com a introdução de variáveis de folga, pode-se reduzir o problema com restrições de desigualdade a problemas com restrições de positividade. Seja $\mathbf{g} : \mathbb{R}^m \to \mathbb{R}^k$ "suave" e $\mathbf{b} \in \mathbb{R}^k$, com $k \geq 1$, e

$$\min_{\mathbf{x}\in\mathbb{R}^m} f(\mathbf{x}) \quad \text{tal que } \mathbf{g}(\mathbf{x}) \geq \mathbf{b},\ \mathbf{x} \geq \mathbf{0}. \tag{5.16}$$

Introduzindo a *variável de folga* $\mathbf{s} = \mathbf{g}(\mathbf{x}) - \mathbf{b}$, obtemos um problema na forma:

$$\min_{\mathbf{x}\in\mathbb{R}^m} f(\mathbf{x}) \quad \text{tal que } \mathbf{s} + \mathbf{b} - \mathbf{g}(\mathbf{x}) = \mathbf{0}, \quad \mathbf{s} \geq \mathbf{0}, \quad \mathbf{x} \geq \mathbf{0}. \tag{5.17}$$

Para obtermos condições suficientes para pontos de ótimo, basta combinar a demonstração do Destaque 5.12 com as ideias da Seção 5.6.1.2. O lagrangeano será agora dado por

$$\mathcal{L}(\mathbf{x},\boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}\cdot\big(\mathbf{s} + \mathbf{b} - \mathbf{g}(\mathbf{x})\big),$$

e as condições necessárias de otimalidade são

$$\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x},\boldsymbol{\lambda}) = \nabla_{\mathbf{x}} f - \boldsymbol{\lambda}\cdot\nabla_{\mathbf{x}}\mathbf{g} \geq \mathbf{0}, \qquad \nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x},\boldsymbol{\lambda})\cdot\mathbf{x} = (\nabla_{\mathbf{x}} f - \boldsymbol{\lambda}\cdot\nabla_{\mathbf{x}}\mathbf{g})\cdot\mathbf{x} = 0,$$
$$\nabla_{\mathbf{s}}\mathcal{L}(\mathbf{x},\boldsymbol{\lambda}) = \boldsymbol{\lambda} \geq \mathbf{0}, \qquad \nabla_{\mathbf{s}}\mathcal{L}(\mathbf{x},\boldsymbol{\lambda})\cdot\mathbf{s} = \boldsymbol{\lambda}\cdot\mathbf{s} = 0,$$
$$\nabla_{\boldsymbol{\lambda}} L(\mathbf{x},\boldsymbol{\lambda}) = \mathbf{s} + \mathbf{b} - \mathbf{g} = \mathbf{0}, \qquad \mathbf{s} \geq \mathbf{0}, \qquad \mathbf{x} \geq \mathbf{0}.$$

computed at $\mathbf{x} = \mathbf{x}^*$, $\mathbf{s} = \mathbf{s}^*$ and $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*$. Eliminando $\mathbf{s} = \mathbf{g} - \mathbf{b}$, obtém-se finalmente

$$\nabla_{\mathbf{x}} f - \boldsymbol{\lambda}\cdot\nabla_{\mathbf{x}}\mathbf{g} \geq \mathbf{0}, \qquad (\nabla_{\mathbf{x}} f - \boldsymbol{\lambda}\cdot\nabla_{\mathbf{x}}\mathbf{g})\cdot\mathbf{x} = 0,$$
$$\boldsymbol{\lambda} \geq \mathbf{0}, \qquad \boldsymbol{\lambda}\cdot(\mathbf{g} - \mathbf{b}) = 0, \qquad \mathbf{g} - \mathbf{b} \geq \mathbf{0}, \qquad \mathbf{x} \geq \mathbf{0}.$$

5.7. Questões ANPEC

Para questões das provas ANPEC relacionadas a este capítulo, ver ano 2005 (questão 09)

---

completar esta seção
