Cálculo Infinitesimal II / Cálculo II - Apontamentos de Apoio
Capítulo 3 - Funções de n Variáveis

Neste capítulo vamos estender as noções do cálculo diferencial a funções que dependem de mais de uma variável.

1 Funções de n variáveis: domínios, curvas de nível, limites e continuidade

Consideremos a função
$$f : D \subseteq \mathbb{R}^n \to \mathbb{R}$$
que a cada $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ faz corresponder o número real $z = f(x)$ que se designa por *imagem* de x por meio de f. O conjunto D é o *domínio* da função f e é o maior conjunto onde a expressão que define f faz sentido, a não ser que se explicite uma restrição deste. O *contradomínio* de f é o conjunto de todas as imagens $f(x)$ para $x \in D$ e representa-se por $f(D)$:

$$f(D) = \Big\{ f(x_1, x_2, \ldots, x_n) : (x_1, x_2, \ldots, x_n) \in D \Big\}.$$

Definição 1.1 *Se f é uma função de duas variáveis com domínio D, o gráfico de f é o conjunto*

$$S = \Big\{ (x, y, z) \in \mathbb{R}^3 : z = f(x, y), (x, y) \in D \Big\}.$$

Para a maior parte das funções de duas variáveis é muito complicado esboçar o respectivo gráfico, nesses casos um processo que pode ajudar a visualizar a superfície $z = f(x, y)$ é considerar as respectivas curvas de nível:

Definição 1.2 *As curvas de nível de uma função de duas variáveis f são as curvas de equação $f(x, y) = k$, onde k é uma constante.*

A curva de nível de equação $f(x, y) = k$ dá-nos o conjunto dos pontos do gráfico de f que estão à mesma altura k, por outras palavras corresponde ao corte do gráfico de f pelo plano horizontal $z = k$, projectado no plano xy. Assim, se desenharmos algumas curvas de nível duma função e imaginarmos que as levantamos até à altura indicada, ficamos com uma ideia do gráfico da função. Supondo que a diferença entre os valores de k de duas curvas de nível consecutivas é constante, a superfície será inclinada onde as curvas de nível estão mais próximas e será mais plana onde estas estão mais afastadas.

Dados dois pontos $P = (x_1, x_2, \ldots, x_n)$ e $Q = (y_1, y_2, \ldots, y_n)$, recordemos que a distância entre P e Q é dada por $d(P, Q) = \|(y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n)\| = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \ldots + (y_n - x_n)^2}$.

Definição 1.3 *Dado um ponto $a \in \mathbb{R}^n$ e um número real $\delta > 0$, chama-se vizinhança de centro em a e raio δ ao conjunto*
$$V_\delta(a) = \Big\{ x \in \mathbb{R}^n : \|x - a\| < \delta \Big\}.$$

Assim, se $n = 1$ a vizinhança de centro em a e raio δ é o intervalo aberto $]a - \delta, a + \delta[$, se $n = 2$ a vizinhança de centro em a e raio δ é o disco aberto de centro em $a = (a_1, a_2)$ e raio δ dado por

$$V_\delta(a) = \Big\{ (x, y) \in \mathbb{R}^2 : (x - a_1)^2 + (y - a_2)^2 < \delta^2 \Big\}.$$

Definição 1.4 *Dado um conjunto $S \subseteq \mathbb{R}^n$, um ponto $a \in \mathbb{R}^n$ diz-se interior a S se existe um número real $\delta > 0$ tal que $V_\delta(a) \subset S$, ou seja, se existe uma vizinhança de a contida em S. Um ponto $a \in \mathbb{R}^n$ diz-se um ponto fronteiro a S se qualquer vizinhança de a contém pontos de S e do seu complementar.*

Ao conjunto dos pontos interiores a S chamamos interior de S e escrevemos int S. A fronteira de S, denotada por fr S ou ∂S, é o conjunto dos pontos fronteiros a S.

Um conjunto diz-se *aberto* se todos os seus pontos forem pontos interiores, um conjunto diz-se *fechado* se contiver todos os seus pontos fronteiros. Note-se que muitos conjuntos não são abertos nem fechados.

Definição 1.5 *Seja $S \subseteq \mathbb{R}^n$. Um ponto $a \in \mathbb{R}^n$ diz-se um ponto de acumulação de S se e só se qualquer vizinhança de a contiver infinitos pontos de S.*

Definição 1.6 *Seja $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ e seja a um ponto de acumulação de D. Dizemos que*

$$\lim_{x \to a} f(x) = b \in \mathbb{R}$$

se e só se

$$\forall \delta > 0 \; \exists \varepsilon > 0 : x \in D \wedge 0 < \|x - a\| < \varepsilon \Rightarrow |f(x) - b| < \delta.$$

Assim, dizer que $\lim_{x \to a} f(x) = b$ é equivalente a afirmar que $\lim_{x \to a} |f(x) - b| = 0$. Isto significa que a distância entre $f(x)$ e b pode ser arbitrariamente pequena desde que se tome a distância entre x e a suficientemente pequena (mas não nula). Note-se que, na definição anterior, o ponto a pode não pertencer ao domínio D de f mas tem que ser um ponto de acumulação de D para que nos possamos aproximar de a por pontos em $D \setminus \{a\}$.

Recordemos que, para funções de variável real, se $\lim_{x \to a^-} f(x) \neq \lim_{x \to a^+} f(x)$ então $\lim_{x \to a} f(x)$ não existe. Para funções de mais de uma variável a situação é mais complicada uma vez que há uma infinidade de maneiras de x tender para a e o limite, a existir, tem que ser independente do modo como $x \to a$. Portanto, se $f(x) \to L_1$ quando $x \to a$ ao longo de uma curva C_1 e $f(x) \to L_2$ quando $x \to a$ ao longo de uma curva C_2, com $L_1 \neq L_2$, então $\lim_{x \to a} f(x)$ não existe. Tornamos esta ideia mais precisa com as seguintes definições:

Definição 1.7 *Seja $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ e seja $A \subseteq D$. A restrição de f ao conjunto A é a função $f_{|A} : A \to \mathbb{R}$ definida por $f_{|A}(x) = f(x)$, $\forall x \in A$.*

Definição 1.8 *Sejam $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, $A \subseteq D$ e a um ponto de acumulação de A. Dizemos que o limite de f no ponto a, restrito ao conjunto A, é b, e escrevemos*

$$\lim_{\substack{x \to a \\ x \in A}} f(x) = b \in \mathbb{R},$$

se e só se $\lim_{x \to a} f_{|A}(x) = b$.

Note-se que quando $A = D$ as definições de limite e de limite restrito coincidem. Além disso, se f tiver limite em a, todos os limites restritos de f em a existem e são iguais. Isto mostra que no caso em que existem dois limites restritos distintos de f em a então $\lim_{x \to a} f(x)$ não existe.

Definição 1.9 *Seja $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ e seja $a \in D$. A função f diz-se contínua em a se e só se*

$$\forall \delta > 0 \; \exists \varepsilon > 0 : x \in D \wedge \|x - a\| < \varepsilon \Rightarrow |f(x) - f(a)| < \delta.$$

Se $a \in D$ for um ponto de acumulação de D então f é contínua em a se e só se

$$\lim_{x \to a} f(x) = f(a).$$

f diz-se contínua num conjunto $S \subseteq D$ se for contínua em todos os pontos de S.

Usando as propriedades algébricas dos limites, que generalizam as já conhecidas para funções reais de variável real, pode-se mostrar que somas, produtos e compostas de funções contínuas são funções contínuas. Assim, temos como exemplos de funções contínuas (nos respectivos domínios) funções polinomiais, racionais e funções que resultem de somas, produtos e compostas de funções polinomiais, racionais, trigonométricas, exponenciais, logarítmicas, etc.

2 Derivadas parciais

Definição 2.1 *Dada uma função* $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$*, definida numa vizinhança dum ponto* $(x_0, y_0) \in D$*, a derivada parcial de* f *em ordem a* x *no ponto* (x_0, y_0) *é dada por*

$$\frac{\partial f}{\partial x}(x_0, y_0) = f_x(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h},$$

se este limite existir. A derivada parcial de f *em ordem a* y *no ponto* (x_0, y_0) *é dada por*

$$\frac{\partial f}{\partial y}(x_0, y_0) = f_y(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h},$$

se este limite existir.

Note-se que $\frac{\partial f}{\partial x}(x_0, y_0) = g'(x_0)$ onde g é a função real de variável real que se obtém, a partir de f, fixando $y = y_0$ e deixando variar apenas x: $g(x) = f(x, y_0)$. Esta derivada parcial dá-nos, pois, a taxa de variação da função f em ordem à variável x, mantendo y fixo. O gráfico da função g é a curva que resulta da intersecção da superfície $z = f(x, y)$ com o plano $y = y_0$. Assim, a derivada parcial de f em ordem a x no ponto (x_0, y_0) representa o declive da recta tangente a essa curva no ponto $(x_0, y_0, f(x_0, y_0))$.

Conclusões análogas são válidas para a derivada parcial de f em ordem a y. Neste caso, fixamos o valor de x e deixamos variar apenas y tendo-se $\frac{\partial f}{\partial y}(x_0, y_0) = k'(y_0)$ onde k é dada por $k(y) = f(x_0, y)$.

Deixando variar o ponto (x_0, y_0) obtemos duas funções de duas variáveis, as funções derivadas parciais de f: f_x e f_y.

A noção de derivada parcial generaliza-se de forma natural ao caso de funções de mais de duas variáveis. Note-se ainda que, uma vez que as derivadas parciais de f se obtêm derivando a função em ordem a uma das variáveis, mantendo as outras fixas, as regras de derivação já conhecidas para funções reais de variável real permanecem válidas.

Se as funções derivadas parciais de f puderem por sua vez ser derivadas em ordem a alguma das variáveis obtemos novas funções chamadas derivadas parciais de segunda ordem de f.

Assim, uma função de duas variáveis tem, caso existam, quatro derivadas parciais de segunda ordem que se denotam por

$$f_{xx} = \frac{\partial^2 f}{\partial x^2}, \quad f_{xy} = \frac{\partial^2 f}{\partial y \partial x}, \quad f_{yx} = \frac{\partial^2 f}{\partial x \partial y} \quad \text{e} \quad f_{yy} = \frac{\partial^2 f}{\partial y^2}.$$

Observemos que a notação $f_{xy} = \frac{\partial^2 f}{\partial y \partial x}$ significa que derivamos f primeiro em ordem a x e depois em ordem a y, para $f_{yx} = \frac{\partial^2 f}{\partial x \partial y}$ a ordem de derivação é a oposta.

Analogamente se definem derivadas parciais de ordem superior à segunda. Por exemplo, a derivada parcial de terceira ordem f_{yyx} obtem-se derivando f duas vezes em ordem a y e depois uma vez em ordem a x.

Definição 2.2 *Seja* D *um conjunto aberto. Uma função* $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ *diz-se de classe* C^k *em* D*,* $k \in \mathbb{N}_0$*, e escreve-se* $f \in C^k(D)$*, se* f *e todas as suas derivadas parciais até à ordem* k *(inclusive) forem funções contínuas em* D*. Uma função de classe* C^0 *em* D *é uma função contínua em* D*.* f *diz-se de classe* C^∞ *em* D*, e escreve-se* $f \in C^\infty(D)$*, se* $f \in C^k(D)$*, para todo* $k \in \mathbb{N}_0$*.*

Uma vez que somas, produtos e compostas de funções contínuas são funções contínuas, resulta que somas, produtos e compostas de funções de classe C^k são ainda funções de classe C^k.

Teorema 2.3 (Schwarz) *Se* f *é uma função de classe* C^2 *numa vizinhança do ponto* (x_0, y_0)*, então* $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$*.*

O resultado anterior pode-se generalizar a funções de mais de duas variáveis e ao caso de derivadas parciais de ordem superior à segunda. Por exemplo, se $f \in C^4(\mathbb{R}^3)$ pode-se mostrar que $f_{xxyz} = f_{xyzx} = f_{yxzx}$. Note-se que em qualquer destes casos derivamos f duas vezes em ordem a x, uma vez em ordem a y e uma vez em ordem a z, a ordem de derivação é que é indiferente.

3 Funções diferenciáveis, noção de gradiente

Vamos ver nesta secção o que se entende por diferenciabilidade de uma função

$$f : D \subseteq \mathbb{R}^n \to \mathbb{R}.$$

Para as funções reais de variável real a diferenciabilidade num ponto implica a continuidade nesse ponto. De forma a preservarmos esta propriedade para funções de duas ou mais variáveis não podemos definir diferenciabilidade de f num ponto $x_0 \in \mathbb{R}^n$ como sendo equivalente à existência das derivadas parciais de f em x_0. Com efeito, existem exemplos de funções para as quais as derivadas parciais existem num ponto mas que são descontínuas nesse ponto. Isto acontece uma vez que a existência de derivadas parciais reflecte o comportamento da função apenas em segmentos de recta paralelos aos eixos coordenados enquanto que a continuidade de f num certo ponto está relacionada com o comportamento da função numa vizinhança desse ponto, em particular, em todas as direcções.

Definição 3.1 *Seja $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ e seja x_0 um ponto interior de D. Se f tem derivadas parciais de primeira ordem em x_0 chamamos gradiente de f no ponto x_0 ao vector*

$$\nabla f(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \frac{\partial f}{\partial x_2}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0) \right).$$

Recordemos que uma função real de variável real tem derivada $f'(x_0)$ num ponto x_0 do interior do seu domínio se e só se

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0) \Leftrightarrow \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - hf'(x_0)}{h} = 0.$$

Para uma função de n variáveis tem-se a seguinte definição:

Definição 3.2 *Seja $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ uma função real definida numa vizinhança do ponto $x_0 \in D$. A função f diz-se diferenciável no ponto x_0 se e só se existir um vector $y \in \mathbb{R}^n$ tal que*

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - y \cdot h}{\|h\|} = 0, \quad (h \in \mathbb{R}^n). \tag{3.1}$$

É fácil de ver que, quando existe um vector y nas condições anteriores, ele é único. Veremos mais adiante que se tem $y = \nabla f(x_0)$.

Teorema 3.3 *Se f é diferenciável em x_0, então f é contínua em x_0.*

Teorema 3.4 *Se a função f tem derivadas parciais contínuas numa vizinhança do ponto x_0, então f é diferenciável em x_0.*

Foi mencionado no início desta secção que o facto de existirem as derivadas parciais de f não é suficiente para garantir a diferenciabilidade da função, nem sequer a sua continuidade. O teorema anterior mostra, no entanto, que se f tem derivadas parciais contínuas numa vizinhança do ponto x_0, então f é diferenciável em x_0. Em particular, toda a função de classe C^1 é diferenciável.

Teorema 3.5 *Seja $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, seja $x_0 \in D$ e suponhamos que f é diferenciável em x_0. Então existem todas as derivadas parciais de primeira ordem de f no ponto x_0 e o vector y de (3.1) é dado por $y = \nabla f(x_0)$.*

4 Derivadas direccionais

Já sabemos que se uma função f é diferenciável, as suas derivadas parciais nos dão a taxa de variação de f nas direcções dos eixos coordenados. Definiremos nesta secção a noção de derivada direccional, que nos fornece a taxa de variação de f numa direcção arbitrária, e veremos qual a sua relação com o vector gradiente.

Definição 4.1 *Dada uma função $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, definida numa vizinhança dum ponto $a \in D$, a derivada de f no ponto a segundo o vector $u \in \mathbb{R}^n$ é dada por*

$$f'_u(a) = \lim_{h \to 0} \frac{f(a + hu) - f(a)}{h},$$

se este limite existir.

Se o vector u for unitário, isto é, se $\|u\| = 1$, a derivada de f no ponto a segundo o vector u diz-se derivada direccional ou derivada dirigida de f no ponto a na direcção e sentido de u.

Se $u = e_j$ (vector da base canónica de $\mathbb{R}^n$, $j = 1, \ldots, n$) tem-se $f'_{e_j}(a) = \dfrac{\partial f}{\partial x_j}(a)$.

Teorema 4.2 *Seja $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável no ponto $x \in \operatorname{int} D$. Então f tem derivada direccional na direcção e sentido de qualquer vector unitário $u \in \mathbb{R}^n$ e tem-se*

$$f'_u(x) = \nabla f(x) \cdot u.$$

Teorema 4.3 *Seja f uma função diferenciável em $x \in \operatorname{int} D$. Então o valor máximo da derivada direccional $f'_u(x)$ é dado por $\|\nabla f(x)\|$ e ocorre quando u tem a direcção e sentido do vector $\nabla f(x)$.*

Assim, $\|\nabla f(x)\|$ corresponde ao valor máximo da taxa de variação de f no ponto x e esse máximo ocorre na direcção e sentido do vector $\nabla f(x)$. Esta é, então, a direcção e sentido em que a função f aumenta mais rapidamente no ponto x.

5 Derivação da função composta

Recordemos a regra de derivação da função composta para funções reais de variável real. Se $y = f(x)$, onde f é diferenciável em $x_0 = g(t_0)$, e $x = g(t)$, onde g é diferenciável em t_0, então a função composta $y = f(g(t))$ é diferenciável em t_0 e tem-se

$$y'(t_0) = (f \circ g)'(t_0) = f'(g(t_0))g'(t_0) = f'(x_0)g'(t_0) \tag{5.2}$$

ou, em notação abreviada,

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt},$$

onde está subentendido que as derivadas são calculadas em pontos correspondentes.

Vamos agora generalizar este resultado ao caso em que as funções intervenientes dependem de várias variáveis.

Começamos com um caso simples: suponhamos que $z = f(u)$, onde

$$f : D \subseteq \mathbb{R} \to \mathbb{R}$$

é uma função diferenciável em $u_0 = g(x_0, y_0)$, onde

$$g : A \subseteq \mathbb{R}^2 \to \mathbb{R},$$

com $g(A) \subseteq D$, é diferenciável em (x_0, y_0). Então a função composta

$$z = f(g(x, y))$$

é diferenciável em (x_0, y_0). Atendendo a que, para calcularmos as derivadas parciais de z em ordem a x e a y deixamos variar apenas uma das variáveis, mantendo fixa a outra, resulta imediatamente de (5.2) que

$$\frac{\partial z}{\partial x}(x_0, y_0) = f'(u_0)\frac{\partial u}{\partial x}(x_0, y_0)$$

e

$$\frac{\partial z}{\partial y}(x_0, y_0) = f'(u_0)\frac{\partial u}{\partial y}(x_0, y_0).$$

Teorema 5.1 (Regra da Cadeia - Caso Particular) *Seja*

$$\begin{array}{rcl} f : D \subseteq \mathbb{R}^n & \rightarrow & \mathbb{R} \\ x = (x_1, \ldots, x_n) & \rightarrow & y = f(x) \end{array}$$

uma função diferenciável em $x_0 \in \mathrm{int}D$ e sejam

$$\begin{array}{rcl} g_i : A \subseteq \mathbb{R} & \rightarrow & \mathbb{R} \\ t & \rightarrow & x_i = g_i(t), \end{array}$$

com $i = 1, \ldots, n$, funções diferenciáveis em $t_0 \in \mathrm{int}A$ tal que $(g_1(t), g_2(t), \ldots, g_n(t)) \in D$, $\forall t \in A$, e $x_0 = (g_1(t_0), g_2(t_0), \ldots, g_n(t_0))$. Então a função composta $y(t) = f(g_1(t), g_2(t), \ldots, g_n(t))$ é diferenciável em t_0 e tem-se

$$y'(t_0) = \sum_{i=1}^{n} \frac{\partial y}{\partial x_i}(x_0) x_i'(t_0).$$

Teorema 5.2 (Regra da Cadeia - Caso Geral) *Seja*

$$\begin{array}{rcl} f : D \subseteq \mathbb{R}^n & \rightarrow & \mathbb{R} \\ x = (x_1, \ldots, x_n) & \rightarrow & y = f(x) \end{array}$$

uma função diferenciável em $x_0 \in \mathrm{int}D$ e sejam

$$\begin{array}{rcl} g_i : A \subseteq \mathbb{R}^p & \rightarrow & \mathbb{R} \\ t = (t_1, \ldots, t_p) & \rightarrow & x_i = g_i(t_1, \ldots, t_p), \end{array}$$

com $i = 1, \ldots, n$, funções diferenciáveis em $t_0 \in \mathrm{int}A$ tal que $(g_1(t), g_2(t), \ldots, g_n(t)) \in D$, $\forall t \in A$, e $x_0 = (g_1(t_0), g_2(t_0), \ldots, g_n(t_0))$. Então a função composta $y(t) = f(g_1(t), g_2(t), \ldots, g_n(t))$ é diferenciável em t_0 e tem-se

$$\frac{\partial y}{\partial t_i}(t_0) = \sum_{j=1}^{n} \frac{\partial y}{\partial x_j}(x_0)\frac{\partial x_j}{\partial t_i}(t_0), \; i = 1, \ldots, p.$$

Neste teorema intervêm três tipos de variáveis: $t_1, \ldots, t_p$ são variáveis independentes, $x_1, \ldots, x_n$ dizem-se variáveis intermédias e y é a variável dependente. Note-se que existem tantas derivadas parciais da função composta quanto o número de variáveis independentes, cada uma destas é dada por uma soma de n parcelas, sendo n o número de variáveis intermédias.

O cálculo de derivadas parciais de ordem superior à primeira para funções compostas faz-se aplicando tantas vezes quantas necessário o teorema anterior.

6 Plano tangente e recta normal a uma superfície

Teorema 6.1 *Seja $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ uma função de classe C^1 numa vizinhança do ponto $(x_0, y_0) \in D$ e suponhamos que a curva de nível da função f que passa em (x_0, y_0) é representada parametricamente pela função diferenciável $r(t) = (x(t), y(t))$ tal que $r(t_0) = (x_0, y_0)$ e $r'(t_0) \neq 0$. Então tem-se*

$$\nabla f(x_0, y_0) \cdot r'(t_0) = 0,$$

onde $\cdot$ representa produto interno.

Nas condições do teorema anterior, dizemos que o vector $\nabla f(x_0, y_0)$ é perpendicular à curva de nível $f(x, y) = k$ que passa no ponto (x_0, y_0). Este resultado pode ser generalizado ao caso de uma superfície de equação $f(x, y, z) = k$ em $\mathbb{R}^3$. Neste caso pode-se mostrar que o vector $n = \nabla f(x_0, y_0, z_0)$ é perpendicular ao vector tangente a qualquer curva diferenciável da superfície $f(x, y, z) = k$ que passa por (x_0, y_0, z_0). Por esse motivo dizemos que o vector n é um vector normal à superfície $f(x, y, z) = k$ no ponto (x_0, y_0, z_0).

Definição 6.2 *O plano tangente à superfície de equação $f(x, y, z) = k$ no ponto (x_0, y_0, z_0) dessa superfície é o plano de equação*

$$\nabla f(x_0, y_0, z_0) \cdot ((x, y, z) - (x_0, y_0, z_0)) = 0.$$

O vector $n = \nabla f(x_0, y_0, z_0)$ (ou qualquer múltiplo deste) diz-se um vector normal à superfície no mesmo ponto.

Definição 6.3 *A recta normal à superfície de equação $f(x, y, z) = k$ no ponto (x_0, y_0, z_0) dessa superfície é a recta que passa em (x_0, y_0, z_0) e que tem a direcção do vector $\nabla f(x_0, y_0, z_0)$. Assim, as equações paramétricas da recta normal são*

$$\begin{cases} x = x_0 + f_x(x_0, y_0, z_0)\, t \\ y = y_0 + f_y(x_0, y_0, z_0)\, t \\ z = z_0 + f_z(x_0, y_0, z_0)\, t, \quad t \in \mathbb{R}. \end{cases}$$

No caso em que a superfície $f(x, y, z) = k$ pode ser escrita na forma $z = g(x, y)$ o respectivo plano tangente, no ponto (x_0, y_0, z_0), é dado por

$$z - z_0 = g_x(x_0, y_0)(x - x_0) + g_y(x_0, y_0)(y - y_0),$$

sendo a recta normal dada por

$$\begin{cases} x = x_0 + g_x(x_0, y_0)\, t \\ y = y_0 + g_y(x_0, y_0)\, t \\ z = z_0 - t, \quad t \in \mathbb{R}. \end{cases}$$

7 Teorema da função implícita

Consideremos a equação $x^2+y^2 = 25$. Sabemos que o conjunto dos pontos do plano que a satisfazem estão sobre uma circunferência de centro na origem e raio 5. Note-se que nem sempre para cada x esta equação define uma função $y = g(x)$. Por exemplo, se $x = 4$ há dois valores de y que verificam $x^2 + y^2 = 25$: $y = 3$ e $y = -3$, pois $y = \pm\sqrt{25 - x^2}$. No entanto, restringindo, por exemplo, $x \in]3, 5[$ e $y \in]0, 4[$, uma vez que neste último intervalo se tem $y > 0$, podemos resolver a equação dada em ordem a y e obtemos $y = \sqrt{25 - x^2}$.

Em geral, a partir duma equação da forma

$$f(x, y) = 0$$

pode não ser possível obter uma fórmula explícita para y como função de x (ou de x como função de y). Por exemplo, não é possível a equação

$$xy + e^x \log y - x \sin y = 0$$

em ordem a y, nem em ordem a x.

Veremos nesta secção em que condições é que uma equação da forma $f(x_1, x_2, \ldots, x_n) = 0$ define implicitamente uma das variáveis como função das restantes e, apesar de em geral não ser possível explicitar tal função, veremos como calcular as suas derivadas (parciais).

Teorema 7.1 (Teorema da Função Implícita - Caso de uma Equação) *Seja*

$$\begin{aligned} f : D \subseteq \mathbb{R}^n \times \mathbb{R} &\to \mathbb{R} \\ (x, y) &\to f(x, y), \end{aligned}$$

onde $x = (x_1, x_2, \ldots, x_n)$, uma função definida num aberto $D \subseteq \mathbb{R}^n \times \mathbb{R}$ e seja $(x_0, y_0) \in D$. Suponhamos que *i)* $f \in C^1(D)$;

ii) $f(x_0, y_0) = 0$;

iii) $\dfrac{\partial f}{\partial y}(x_0, y_0) \neq 0$.

Então a equação $f(x, y) = 0$ *define implicitamente* y *como função de* x, *tendo-se* $y = g(x)$, *numa vizinhança do ponto* (x_0, y_0), *isto é, existem conjuntos abertos* $W \subseteq D$, $V(x_0) \subseteq \mathbb{R}^n$, *com* $(x_0, y_0) \in W$, $x_0 \in V(x_0)$, *e existe uma função* $g : V(x_0) \subseteq \mathbb{R}^n \to \mathbb{R}$, *tais que:*

a) $(x, g(x)) \in W$ *e* $f(x, g(x)) = 0$, $\forall x \in V(x_0)$;

b) se $(x, y) \in W$ *e* $f(x, y) = 0$ *então* $x \in V(x_0)$ *e* $y = g(x)$;

c) $g \in C^1(V(x_0))$.

Nas condições anteriores, se $f \in C^k(D)$, $k \geq 1$, então $g \in C^k(V(x_0))$.

Apesar de, na maior parte dos casos, não ser possível explicitar a função g, o teorema fornece um método que nos permite calcular as suas derivadas (parciais). No caso $n = 1$, vejamos como calcular $g'(x)$ para $x \in V(x_0)$: por *a)* tem-se

$$F(x) = f(x, g(x)) = 0, \forall x \in V(x_0)$$

pelo que $F'(x) = 0$. Mas, pela regra de derivação da função composta, vem

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

donde

$$g'(x) = -\frac{\frac{\partial f}{\partial x}(x, g(x))}{\frac{\partial f}{\partial y}(x, g(x))}$$

desde que $\dfrac{\partial f}{\partial y}(x, g(x)) \neq 0$, o que acontece numa vizinhança de (x_0, y_0) por *iii)* e por continuidade das derivadas parciais de f.

Note-se que o teorema anterior nos dá um resultado local: a função g que se afirma existir está definida numa vizinhança do ponto x_0, se alterarmos o ponto (x_0, y_0) a função também pode variar. Voltando ao exemplo da circunferência $x^2 + y^2 = 25$ vimos que se $x \in]3, 5[$ e $y \in]0, 4[$, então

$$y = g(x) = \sqrt{25 - x^2}.$$

No entanto, se $x \in]3, 5[$ e $y \in] - 4, 0[$ tem-se

$$y = h(x) = -\sqrt{25 - x^2},$$

uma vez que $y < 0$.

Notemos ainda que o teorema da função implícita é aplicável e permite definir y como função de x numa vizinhança de qualquer ponto da circunferência à excepção dos pontos $(5, 0)$ e $(-5, 0)$ onde a derivada $\dfrac{\partial f}{\partial y}$ se anula, sendo $f(x, y) = x^2 + y^2 - 25$. De facto, em qualquer vizinhança de cada um destes pontos existem pontos com $y > 0$ e outros com $y < 0$ pelo que é impossível escolher entre $\sqrt{25 - x^2}$ e $-\sqrt{25 - x^2}$ e portanto não é possível escrever univocamente y como função de x (mas pode-se escrever x como função de y...).

A condição $\dfrac{\partial f}{\partial y}(x_0, y_0) \neq 0$ no teorema anterior é apenas suficiente. Se $\dfrac{\partial f}{\partial y}(x_0, y_0) = 0$ nada se pode concluir quanto à existência de função implícita. Por exemplo, se

$$f(x, y) = (y - x)^2,$$

tem-se $\dfrac{\partial f}{\partial y}(0,0) = 0$ mas a relação $(y-x)^2 = 0$ define claramente a função $y = x$. Por outro lado, também para

$$f(x,y) = x^2 + y^2$$

se tem $\dfrac{\partial f}{\partial y}(0,0) = 0$ mas a relação $x^2 + y^2 = 0$ não define y como função de x em qualquer aberto de $\mathbb{R}$ que contenha o ponto $x = 0$ porque $(0,0)$ é a única solução da equação $x^2 + y^2 = 0$.

8 Extremos locais e absolutos

Nesta secção vamos estender ao caso das funções reais de duas variáveis as noções de máximos e mínimos locais e absolutos.

Definição 8.1 *Seja $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ e seja $(x_0, y_0) \in D$. A função f tem um máximo (respectivamente, um mínimo) local ou relativo no ponto (x_0, y_0) se existe uma vizinhança $V_\varepsilon(x_0, y_0)$ ($\varepsilon > 0$) do ponto (x_0, y_0) tal que*

$$f(x_0, y_0) \geq f(x,y), \;\; \forall (x,y) \in V_\varepsilon(x_0, y_0) \cap D$$

(respectivamente, $f(x_0, y_0) \leq f(x,y), \;\; \forall (x,y) \in V_\varepsilon(x_0, y_0) \cap D$). Em qualquer destes casos dizemos que f tem um extremo local ou relativo no ponto (x_0, y_0).

No caso das funções reais de variável real sabemos que se f tem um extremo local em x_0 pertencente ao interior do domínio então $f'(x_0) = 0$ ou $f'(x_0)$ não existe. O resultado análogo para funções de duas variáveis é dado em termos de ∇f.

Teorema 8.2 (Fermat) *Se a função $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ tem um extremo local no ponto $(x_0, y_0) \in \mathrm{int} D$ então $\nabla f(x_0, y_0) = (0,0)$ ou $\nabla f(x_0, y_0)$ não existe.*

Definição 8.3 *Chama-se ponto crítico de uma função $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ a um ponto (x_0, y_0) do interior de D para o qual $\nabla f(x_0, y_0) = (0,0)$ ou $\nabla f(x_0, y_0)$ não existe. Se $\nabla f(x_0, y_0) = (0,0)$ o ponto (x_0, y_0) diz-se um ponto de estacionaridade de f.*

O teorema anterior diz-nos que os únicos pontos interiores ao domínio onde a função f pode atingir extremos locais são os pontos críticos. Note-se, no entanto, que nem todos os pontos críticos correspondem a extremos locais.

Definição 8.4 *Chama-se ponto de sela a um ponto de estacionaridade de f que não corresponde a um extremo local.*

Se uma função real de variável real é diferenciável e atinge um extremo local no ponto x_0 interior ao domínio, então a recta tangente ao gráfico de f no ponto (x_0, y_0), onde $y_0 = f(x_0)$, é horizontal.

Analogamente, resulta do Teorema 8.2 que se f é diferenciável e atinge um extremo local no ponto (x_0, y_0) interior ao domínio de f, então o plano tangente à superfície $z = f(x,y)$ no ponto (x_0, y_0, z_0), onde $z_0 = f(x_0, y_0)$, é horizontal. Com efeito, como neste caso se tem $\nabla f(x_0, y_0) = (0,0)$, a equação do referido plano tangente é

$$z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \Leftrightarrow z = z_0.$$

Se f é uma função real de variável real duas vezes diferenciável tal que $f'(x_0) = 0$ sabemos que f tem um mínimo local em x_0 se $f''(x_0) > 0$ e que f tem um máximo local em x_0 se $f''(x_0) < 0$. Este resultado generaliza-se a funções de duas variáveis do seguinte modo:

Teorema 8.5 *Seja $f \in C^2(D)$ e seja (x_0, y_0) um ponto de estacionaridade de f. Consideremos a matriz, chamada matriz hessiana de f no ponto (x_0, y_0), dada por*

$$H_f(x_0, y_0) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{bmatrix}_{(x_0, y_0)}$$

e seja $d = det H_f(x_0, y_0)$. Então:

i) se $d < 0$ (x_0, y_0) é um ponto de sela;

ii) se $d > 0$ e $\dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0$ f tem um mínimo local em (x_0, y_0);

iii) se $d > 0$ e $\dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) < 0$ f tem um máximo local em (x_0, y_0);

iv) se $d = 0$ nada se pode concluir.

Notas 8.6 1) Uma vez que $f \in C^2(D)$ tem-se $\dfrac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \dfrac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$ pelo que a matriz hessiana é uma matriz simétrica.

2) Pela mesma razão, se $\dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) = 0$ obtemos $d \leq 0$.

3) Consideremos as funções $f(x, y) = x^4 + y^4$, $g(x, y) = -(x^4 + y^4)$ e $h(x, y) = x^4 - y^4$. É fácil de ver que $(0, 0)$ é ponto de estacionaridade para todas elas e que, neste ponto, $d = 0$ para cada uma das funções mencionadas. No entanto, $(0, 0)$ é ponto de mínimo local de f, $(0, 0)$ é ponto de máximo local de g e $(0, 0)$ é ponto de sela de h. Estes três exemplos mostram que nada se pode concluir no caso em que $d = 0$.

O facto de uma função ter um extremo **local** num ponto (x_0, y_0) depende do comportamento da função numa **vizinhança** de (x_0, y_0). Os extremos **absolutos** de f dependem do comportamento da função em **todo o seu domínio**.

Definição 8.7 *Seja $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ e seja $(x_0, y_0) \in D$. A função f tem um máximo (respectivamente, um mínimo) absoluto no ponto (x_0, y_0) se*

$$f(x_0, y_0) \geq f(x, y), \quad \forall (x, y) \in D$$

(respectivamente, $f(x_0, y_0) \leq f(x, y)$, $\forall (x, y) \in D$). Em qualquer destes casos dizemos que f tem um extremo absoluto no ponto (x_0, y_0).

Claro que se f tem um extremo absoluto em (x_0, y_0) também tem um extremo local nesse ponto, mas o recíproco é falso.

Definição 8.8 *Um subconjunto $S \subset \mathbb{R}^n$ diz-se limitado se existe $R > 0$ tal que $\|x\| \leq R$, $\forall x \in S$.*

Teorema 8.9 (Weierstrass) *Se $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ é contínua num conjunto limitado, fechado e não vazio $S \subseteq D$ então f atinge um máximo e um mínimo absolutos em S.*

Os extremos absolutos referidos no teorema anterior podem ser atingidos no interior do conjunto S ou na sua fronteira. Assim, para determinarmos os extremos absolutos de uma função contínua f num conjunto limitado e fechado S:

i) determinamos os pontos críticos de f no interior de S;

ii) determinamos os pontos da fronteira de S que podem dar origem a extremos. No caso $n = 2$, uma maneira de fazer isto é parametrizar a fronteira de S através de uma função vectorial $r(t)$ e reduzir o problema ao estudo da função de uma só variável $f(r(t))$. Veremos na próxima secção um método alternativo para resolver este passo.

iii) Calculamos o valor de f em cada um dos pontos determinados nos passos anteriores. O maior destes valores é o máximo absoluto de f em S, o menor é o mínimo absoluto.

9 Extremos condicionados

Veremos nesta secção como determinar extremos de uma função $f(x,y)$ (ou $f(x,y,z)$) no caso em que os pontos (x,y) estão sujeitos a uma condição do tipo $g(x,y)=0$ (ou $g(x,y,z)=0$). Chama-se a isto resolver um problema de *extremos condicionados*. Este problema resume-se assim a calcular os extremos da função f restrita ao conjunto, suposto não vazio,

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

(ou $C = \{(x,y,z) \in \mathbb{R}^3 : g(x,y,z) = 0\}$), notada $f_{|C}$. Vamos usar o chamado *método dos multiplicadores de Lagrange.*

Teorema 9.1 *Sejam $g : E \subseteq \mathbb{R}^2 \to \mathbb{R}$, $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ funções tais que $g \in C^1(E)$, $f \in C^1(D)$ e $E \subseteq D$ e seja $C = \{(x,y) \in E : g(x,y) = 0\}$. Se $f_{|C}$ tem um extremo local em $(x_0,y_0) \in C$ e se $\nabla g(x_0,y_0) \neq (0,0)$ então $\nabla f(x_0,y_0)$ e $\nabla g(x_0,y_0)$ são paralelos, isto é, existe $\lambda \in \mathbb{R}$ tal que*

$$\nabla f(x_0,y_0) = \lambda \nabla g(x_0,y_0).$$

Ao escalar λ referido no teorema anterior damos o nome de *multiplicador de Lagrange.* Este resultado, que também é válido em $\mathbb{R}^3$, diz-nos então que os pontos onde são atingidos os extremos locais de $f_{|C}$ estão entre as soluções do sistema

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

10 Funções vectoriais de n variáveis

Consideremos a função vectorial de variável vectorial

$$f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$$

que a cada $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ faz corresponder o vector de $\mathbb{R}^m$ dado por

$$(f_1(x_1,x_2,\ldots,x_n), f_2(x_1,x_2,\ldots,x_n), \ldots, f_m(x_1,x_2,\ldots,x_n)).$$

Nesta secção vamos estender as noções do cálculo diferencial a este tipo de funções.

Quando $m = 1$ estas funções designam-se também por *campos escalares* (ou *reais*), se $m \geq 2$ dizem-se *funções* ou *campos vectoriais.* Às funções reais

$$\begin{array}{rcl} f_i : D \subseteq \mathbb{R}^n & \to & \mathbb{R} \\ (x_1,x_2,\ldots,x_n) & \to & f_i(x_1,x_2,\ldots,x_n), \end{array}$$

para $i = 1, \ldots, m$, chamamos *funções componentes* de f.

O *domínio* da função f é a intersecção dos domínios de cada uma das suas funções componentes e é o maior conjunto onde a expressão que define f faz sentido, a não ser que se explicite uma restrição deste.

Dado um subconjunto $A \subseteq D$, chama-se *imagem* de A por meio de f ao conjunto

$$f(A) = \Big\{ f(x_1,x_2,\ldots,x_n) : (x_1,x_2,\ldots,x_n) \in A \Big\}.$$

Definição 10.1 *Seja $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ e seja a um ponto de acumulação de D. Dizemos que*

$$\lim_{x \to a} f(x) = b \in \mathbb{R}^m$$

se e só se

$$\forall \delta > 0 \ \exists \varepsilon > 0 : 0 < \|x - a\| < \varepsilon \Rightarrow \|f(x) - b\| < \delta.$$

Assim, dizer que $\lim_{x\to a} f(x) = b$ é equivalente a afirmar que $\lim_{x\to a} \|f(x) - b\| = 0$. Isto significa que a distância entre $f(x)$ e b pode ser arbitrariamente pequena desde que se tome a distância entre x e a suficientemente pequena (mas não nula).

O teorema que se segue diz-nos que, tal como para as funções vectoriais de variável real, os limites das funções vectoriais de variável vectorial se calculam componente a componente, reduzindo-se ao cálculo de m limites de funções reais.

Teorema 10.2 *Sejam $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, a um ponto de acumulação de D e $b = (b_1, \ldots, b_m)$. Então*

$$\lim_{x\to a} f(x) = b \Leftrightarrow \lim_{x\to a} f_i(x) = b_i, \forall i = 1, \ldots, m.$$

São válidas as seguintes propriedades algébricas dos limites.

Teorema 10.3 *Sejam $f, g : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ funções vectoriais definidas num subconjunto D de $\mathbb{R}^n$, seja $\lambda : D \to \mathbb{R}$ uma função real e seja a um ponto de acumulação de D. Suponhamos que $\lim_{x\to a} f(x) = b \in \mathbb{R}^m$, $\lim_{x\to a} g(x) = c \in \mathbb{R}^m$ e que $\lim_{x\to a} \lambda(x) = \alpha \in \mathbb{R}$. Então tem-se:*

i) $\lim_{x\to a} (f(x) + g(x)) = b + c$;

ii) $\lim_{x\to a} f(x) \cdot g(x) = b \cdot c$, *onde* $\cdot$ *representa produto interno;*

iii) $\lim_{x\to a} \lambda(x) f(x) = \alpha b$;

iv) $\lim_{x\to a} \|f(x)\| = \|b\|$.

Definição 10.4 *Seja $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ e seja $a \in D$. A função f diz-se contínua em a se e só se*

$$\forall \delta > 0 \; \exists \varepsilon > 0 : \|x - a\| < \varepsilon \Rightarrow \|f(x) - f(a)\| < \delta.$$

Se a for um ponto de acumulação de D, então f é contínua em a se e só se

$$\lim_{x\to a} f(x) = f(a) \in \mathbb{R}^m.$$

f diz-se contínua num conjunto $S \subseteq D$ se for contínua em todos os pontos de S.

Resulta imediatamente do Teorema 10.2 que

Teorema 10.5 *Seja $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ e seja $a \in D$. Então f é contínua em a se e só se f_i é contínua em a, $\forall i = 1, \ldots, m$.*

O próximo resultado dá-nos algumas propriedades das funções contínuas, análogos aos já conhecidos para funções escalares.

Teorema 10.6 *Sejam $f, g : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $\varphi : E \subseteq \mathbb{R}^p \to \mathbb{R}^n$ tal que $\varphi(E) \subseteq D$ e $\lambda : D \subseteq \mathbb{R}^n \to \mathbb{R}$. Então:*

i) se f, g e λ são contínuas em $a \in D$, o mesmo sucede a $\|f\|$, $f + g$, λf, $f \cdot g$, e ainda a $\dfrac{f}{\lambda}$ se $\lambda(a) \neq 0$;

ii) se φ é contínua em $a \in E$ e f é contínua em $b = \varphi(a)$, então $f \circ \varphi$ é contínua em a.

Tem-se também a seguinte generalização para funções vectoriais da noção de função de classe C^k:

Definição 10.7 *Seja $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $f = (f_1, f_2, \ldots, f_m)$. A função f diz-se de classe C^k em D, $k \in \mathbb{N}_0$, (respectivamente, de classe C^∞ em D) se $f_i \in C^k(D)$, $\forall i = 1, \ldots, m$, (respectivamente, $f_i \in C^\infty(D)$, $\forall i = 1, \ldots, m$).*

As noções de derivada parcial e de derivada segundo um vector estendem-se de modo natural às funções vectoriais $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Em vez de números reais, estas derivadas são agora vectores de $\mathbb{R}^m$ cujas componentes são exactamente as correspondentes derivadas das m funções componentes de f.

Definição 10.8 *Dada uma função $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, definida numa vizinhança dum ponto $a \in D$, a derivada de f no ponto a segundo o vector $u \in \mathbb{R}^n$ é dada por*

$$f'_u(a) = \lim_{h\to 0} \frac{f(a+hu) - f(a)}{h}$$

se este limite existir.

Se o vector u for unitário, isto é se $\|u\| = 1$, a derivada de f segundo o vector u diz-se derivada direccional ou derivada dirigida de f na direcção e sentido de u.

Se $u = e_j$ (vector da base canónica de $\mathbb{R}^n$, $j = 1, \ldots, n$) tem-se $f'_{e_j}(a) = \dfrac{\partial f}{\partial x_j}(a)$.

Note-se que $f'_u(a)$ é um vector de $\mathbb{R}^m$ que, atendendo ao Teorema 10.2, tem por componentes as derivadas das funções f_i, $i = 1, \ldots, m$, no ponto a segundo o vector u, isto é,

$$f'_u(a) = ((f_1)'_u(a), (f_2)'_u(a), \ldots, (f_m)'_u(a))$$

onde

$$(f_i)'_u(a) = \lim_{h\to 0} \frac{f_i(a+hu) - f_i(a)}{h}, \; i = 1, \ldots, m.$$

Supondo que todas as funções $f_i : D \subseteq \mathbb{R}^n \to \mathbb{R}$ são diferenciáveis no ponto a sabemos que

$$(f_i)'_u(a) = \nabla f_i(a) \cdot u$$

pelo que

$$f'_u(a) = (\nabla f_1(a) \cdot u, \nabla f_2(a) \cdot u, \ldots, \nabla f_m(a) \cdot u).$$

Usando a notação matricial podemos escrever

$$f'_u(a) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}_{(a)} \cdot \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} \nabla f_1(a) \\ \nabla f_2(a) \\ \vdots \\ \nabla f_m(a) \end{bmatrix}_{m\times n} \cdot \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}_{n\times 1}$$

identificando a matriz coluna $m \times 1$ resultante com o vector de $\mathbb{R}^m$ correspondente.

Definição 10.9 *Seja $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ e seja $a \in D$ tal que as derivadas parciais $\dfrac{\partial f_i}{\partial x_j}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, existem no ponto a. À matriz*

$$J_f(a) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}_{(a)} = \begin{bmatrix} \nabla f_1(a) \\ \nabla f_2(a) \\ \vdots \\ \nabla f_m(a) \end{bmatrix}$$

dá-se o nome de matriz jacobiana de f no ponto a.

Quando $m = n$ o determinante da matriz $J_f(a)$ diz-se o jacobiano da função f no ponto a e representa-se por

$$\det J_f(a) = \frac{\partial(f_1, f_2, \ldots, f_n)}{\partial(x_1, x_2, \ldots, x_n)}(a).$$

O teorema da função implícita pode-se generalizar ao caso em que temos um sistema de equações.

Teorema 10.10 (Teorema da Função Implícita - Caso de um Sistema de 2 Equações) *Seja*

$$\begin{array}{rcl} f: D \subseteq \mathbb{R}^{n+2} & \to & \mathbb{R}^2 \\ (x,y) & \to & f(x,y) = (f_1(x,y), f_2(x,y)), \end{array}$$

onde $x \in \mathbb{R}^n$ *e* $y \in \mathbb{R}^2$*, uma função definida num aberto* $D \subseteq \mathbb{R}^{n+2}$ *e seja* $(x_0, y_0) \in D$*, com* $x_0 \in \mathbb{R}^n$ *e* $y_0 \in \mathbb{R}^2$*. Suponhamos que*

i) $f \in C^1(D)$;

ii) $f(x_0,y_0) = 0 \Leftrightarrow \left\{ \begin{array}{l} f_1(x_0,y_0) = 0 \\ f_2(x_0,y_0) = 0 \end{array} \right.$;

iii) $\dfrac{\partial(f_1,f_2)}{\partial(y_1,y_2)}(x_0,y_0) \neq 0.$

Então o sistema de equações $f(x,y) = 0$ *define implicitamente* $y = (y_1,y_2)$ *como função de* x*, tendo-se* $y = g(x) = (g_1(x_1,\ldots,x_n), g_2(x_1,\ldots,x_n))$*, numa vizinhança do ponto* (x_0,y_0)*, isto é, existem conjuntos abertos* $W \subseteq D$*,* $V(x_0) \subseteq \mathbb{R}^n$*, com* $(x_0,y_0) \in W$*,* $x_0 \in V(x_0)$*, e existe uma função* $g : V(x_0) \subseteq \mathbb{R}^n \to \mathbb{R}^2$*, tais que:*

a) $(x, g(x)) \in W$ *e* $f(x, g(x)) = 0$, $\forall x \in V(x_0)$;

b) se $(x,y) \in W$ *e* $f(x,y) = 0$ *então* $x \in V(x_0)$ *e* $y = g(x)$;

c) $g \in C^1(V(x_0))$.

Nas condições anteriores, se $f \in C^k(D)$, $k \geq 1$, então $g \in C^k(V(x_0))$.

Apesar de, em geral, não ser possível explicitar as duas funções definidas implicitamente (como funções das restantes n variáveis) pode-se, tal como anteriormente, derivar implicitamente as equações do sistema para obter as derivadas (parciais) destas funções.