Teorema da Função Implícita

1 Gráficos de Funções

Dadas k funções de classe C^r $(r \geq 1)$, $f_1, \ldots, f_k : D \to \mathbb{R}$, onde $D \subset \mathbb{R}^n$ é um domínio aberto, consideremos o conjunto das soluções do sistema não linear

$$X := \{x \in D\colon f_i(x) = b_i,\ \forall\, i = 1, \ldots, k\,\}.$$

O Teorema da Função Implícita dá uma condição em $a \in X$ que é suficiente para que numa vizinhança aberta $U \subset \mathbb{R}^n$ de a, $X \cap U$ seja o gráfico de uma função de classe C^r.

Dizemos que $X \subset \mathbb{R}^n$ é o gráfico de uma função quando a relação

$$(x_1, \ldots, x_n) \in X$$

define implicitamente algumas das variáveis x_i em função das restantes. Para precisar este conceito introduzimos alguma notação. Dado um conjunto de índices $I = \{i_1, \ldots, i_k\}$, onde $1 \leq i_1 < \ldots < i_k \leq n$, e um vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, designamos por $x_I = (x_{i_1}, \ldots, x_{i_k}) \in \mathbb{R}^k$ o vector com as componentes associadas a índices em I, e por I^c o complementar de I em $\{1, \ldots, n\}$.

Dado $I \subset \{1, \ldots, n\}$, um conjunto $X \subseteq \mathbb{R}^n$ diz-se o gráfico de uma *função do tipo I* se definindo $U := \{x_{I^c}\colon x \in X\}$, existir uma função $\varphi\colon U \to \mathbb{R}^k$ tal que

$$X = \{x \in \mathbb{R}^n\colon\ x_{I^c} \in U \ \text{ e } \ x_I = \varphi(x_{I^c})\}.$$

Quando U for aberto e φ for de classe C^r dizemos que X é o gráfico duma função de tipo I e classe C^r.

Exercício 1. *Para que conjuntos $I \subset \{1, 2, 3\}$, $X := \{(x, y, z) \in \mathbb{R}^3\colon x^2 + y^2 + z^2 = 1,\ z = 1/2,\ y > 0\}$ é o gráfico duma função de tipo I?*

2 Sistemas Lineares

Dada uma matriz $A \in \mathrm{Mat}_{k\times n}(\mathbb{R})$, e um conjunto de índices $I \subset \{1, \ldots, n\}$ com $\#I = k$, designamos por $A_I \in \mathrm{Mat}_{k\times k}(\mathbb{R})$ a matriz formada pelas k colunas da matriz A com índices em I. Com esta notação qualquer que seja $x \in \mathbb{R}^n$,

$$A\,x = A_I\,x_I + A_{I^c}\,x_{I^c}\ .$$

Por exemplo, se $I = \{1, 3\}$ e $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$, temos

$$\underbrace{\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}}_{A_I} \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} + \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{A_{I^c}} [x_2]$$

Em qualquer matriz $A \in \mathrm{Mat}_{k\times n}(\mathbb{R})$ o número máximo de linhas linearmente independentes é igual ao número máximo de colunas linearmente independentes. Este número diz-se a *característica* da matriz A, e será aqui denotado por $\mathrm{rank}(A)$. É claro que $\mathrm{rank}(A) \leq \min\{k, n\}$. Uma matriz, quadrada ou rectangular, diz-se *singular* se $\mathrm{rank}(A) < \min\{k, n\}$. Caso contrário, se $\mathrm{rank}(A) = \min\{k, n\}$, A diz-se *não singular*. Uma matriz $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, é não singular se e somente se $\mathrm{rank}(A) = n$, o que equivale a dizer que $\det(A) \neq 0$. Qualquer sistema de $k \leq n$ equações em n incógnitas pode escrever-se na forma $A\,x = b$ para certas matrizes $A \in \mathrm{Mat}_{k\times n}(\mathbb{R})$ e $b \in \mathrm{Mat}_{k\times 1}(\mathbb{R})$. Eliminando as equações redundantes podemos sempre supor que todas as equações são linearmente independentes, o que equivale a dizer que $\mathrm{rank}(A) = k$, ou seja que A é não singular.

Proposição 1. *Dados $A \in \mathrm{Mat}_{k\times n}(A)$ com $\mathrm{rank}(A) = k \leq n$ e $I \subseteq \{1, \ldots, n\}$ com $\#I = k$, o sub-espaço afim $X = \{x \in \mathbb{R}^n \colon A\,x = b\}$ é o gráfico duma função (afim) de tipo I se e somente se $\det(A_I) \neq 0$.*

Demonstração. Seja $I \subseteq \{1, \ldots, n\}$ com com $|I| = k$ tal que $\det(A_I) \neq 0$. Como

$$\begin{aligned} A\,x = b \quad &\Leftrightarrow \quad A_I\,x_I + A_{I^c}\,x_{I^c} = b \\ &\Leftrightarrow \quad A_I\,x_I = b - A_{I^c}\,x_{I^c} \\ &\Leftrightarrow \quad x_I = (A_I)^{-1}\,(b - A_{I^c}\,x_{I^c})\,, \end{aligned}$$

X é o gráfico duma função de tipo I, em concreto de $\varphi(x) := (A_I)^{-1}\,(b - A_{I^c}\,x)$.

Reciprocamente, se $\det(A_I) = 0$, fixados valores para as variáveis x_{I^c} o sistema $A_I\,x_I = b - A_{I^c}\,x_{I^c}$ é impossível ou então admite uma infinidade de soluções nas variáveis x_I. Isto mostra que X não é o gráfico duma função de tipo I. □

Exemplo 1. *O sub-espaço afim definido pelo sistema de equações*

$$\begin{cases} x_1 + x_2 + 2\,x_3 = 0 \\ \qquad\;\; x_2 + 2\,x_3 = 1 \end{cases} \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

é o gráfico duma função afim de tipo $\{1, 3\}$, porque $\begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} = 2 \neq 0$, mas não de tipo $\{2, 3\}$, porque $\begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = 0$.

Consideremos agora o subespaço linear $S := \mathrm{Nuc}(A) = \{x \in \mathbb{R}^n \colon A\,x = 0\}$. O subespaço afim X é paralelo a S no sentido em que $X = p + S$, qualquer que seja $p \in X$. Seja $S^\perp := \{v \in \mathbb{R}^n \colon v \cdot x = 0, \, \forall\, x \in S\}$ o complemento ortogonal de S.

Exercício 2. *O complemento ortogonal $S^\perp$ do núcleo da matriz A é o espaço das linhas de A, i.e., $S^\perp$ é gerado pelas linhas de A.*

Exercício 3. *Mostre que para quaisquer subespaços lineares $S, S' \subset \mathbb{R}^n$,*

(a) $\mathbb{R}^n = S \oplus S^\perp$,

(b) $S = (S^\perp)^\perp$,

(c) $(S \cap S')^\perp = S^\perp + (S')^\perp$,

(d) $(S + S')^\perp = S^\perp \cap (S')^\perp$.

Sejam $\pi_{I^c} : \mathbb{R}^n \to \mathbb{R}^{n-k}$, $\pi_I : \mathbb{R}^n \to \mathbb{R}^k$, as projectções canónicas definidas por $\pi_{I^c}(x) := x_{I^c}$, $\pi_I(x) := x_I$.

Proposição 2. *Sob as hipóteses da Proposição 1 são equivalentes:*

(1) *X é o gráfico duma função do tipo I,*

(2) *S é o gráfico duma função do tipo I,*

(3) $\pi_{I^c}|_S \colon S \to \mathbb{R}^{n-k}$ *é um isomorfismo,*

(4) $\pi_I|_{S^\perp} \colon S^\perp \to \mathbb{R}^k$ *é um isomorfismo,*

(5) $\det(A_I) \neq 0$.

Demonstração. A equivalência (1)⇔(2) segue de X e S serem subespaços afins paralelos, i.e., de se ter $X = p + S$.

Suponhamos agora que S é o gráfico duma função de tipo I,

$$S = \{x \in \mathbb{R}^n \colon x_{I^c} \in \pi_{I^c}(S), \; x_I = \varphi(x_{I^c})\}.$$

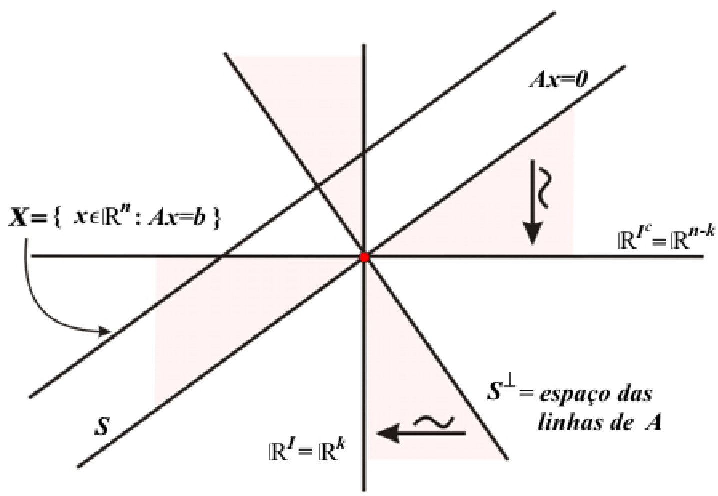
Dados $x, y \in S$ tais que $\pi_{I^c}(x) = \pi_{I^c}(y)$, temos $x_{I^c} = y_{I^c}$ o que implica que $x_I = \varphi(x_{I^c}) = \varphi(y_{I^c}) = y_I$. Logo $x_I = y_I$ e $x_{I^c} = y_{I^c}$, donde $x = y$. Mostrámos que a projecção $\pi_{I^c}|_S \colon S \to \mathbb{R}^{n-k}$ é injectiva, mas como da hipótese $\mathrm{rank}(A) = k$ segue que $\dim(S) = \dim(\mathrm{Nuc}(A)) = n - k$, esta projecção tem de ser um isomorfismo. Reciprocamente, se $\pi_{I^c}|_S \colon S \to \mathbb{R}^{n-k}$ for um isomorfismo, definindo $\varphi := (\pi_I|_S) \circ (\pi_{I^c}|_S)^{-1}$ temos $\pi_{I^c}(S) = \mathbb{R}^{n-k}$ e $S = \{x \in \mathbb{R}^n \colon x_{I^c} \in \mathbb{R}^{n-k}, \; x_I = \varphi(x_{I^c})\}$, o que prova que S é o gráfico duma função de tipo I. Logo (2)⇔(3).

Para a equivalência seguinte seja $\mathbb{R}^n = \mathbb{R}^I \oplus \mathbb{R}^{I^c}$ a decomposição em soma directa onde $\mathbb{R}^I := \{x \in \mathbb{R}^n \colon x_i = 0 \text{ se } i \notin I\}$ e $\mathbb{R}^{I^c} := \{x \in \mathbb{R}^n \colon x_i = 0 \text{ se } i \in I\}$. Fazendo as identificações $\mathbb{R}^I \equiv \mathbb{R}^k$ e $\mathbb{R}^{I^c} \equiv \mathbb{R}^{n-k}$ as aplicações π_I e π_{I^c} são respectivamente as projecções ortogonais sobre of subespaços $\mathbb{R}^I$ e $\mathbb{R}^{I^c}$ de $\mathbb{R}^n$, com $\mathbb{R}^I = \mathrm{Nuc}(\pi_{I^c})$ e $\mathbb{R}^{I^c} = \mathrm{Nuc}(\pi_I)$. Logo usando as propriedades dos complementos ortogonais expressas no Exercício 3 vemos que são sucessivamente equivalentes as afirmações:

- $\pi_{I^c}|_S : S \to \mathbb{R}^{I^c}$ é isomorfismo,
- $\pi_{I^c}|_S : S \to \mathbb{R}^{I^c}$ é injectiva, porque $\dim(S) = \dim(\mathbb{R}^{I^c})$,
- $S \cap \mathbb{R}^I = \{0\}$, porque $\mathrm{Nuc}(\pi_{I^c}|_S) = S \cap \mathbb{R}^I$,
- $S^\perp + R^{I^c} = \mathbb{R}^n$, porque $S^\perp + R^{I^c} = (S \cap \mathbb{R}^I)^\perp$,
- $S^\perp \cap \mathbb{R}^{I^c} = \{0\}$, porque $\dim(S^\perp) + \dim(\mathbb{R}^{I^c}) = n$,
- $\pi_I|_{S^\perp} : S^\perp \to \mathbb{R}^I$ é injectiva, porque $\mathrm{Nuc}(\pi_I|_{S^\perp}) = S^\perp \cap \mathbb{R}^{I^c}$,
- $\pi_I|_{S^\perp} : S^\perp \to \mathbb{R}^I$ é um isomorfismo, porque $\dim(S^\perp) = \dim(\mathbb{R}^I)$.

Vemos assim que (3)⇔(4)

Finalmente como $\mathrm{rank}(A) = k$, as linhas de A formam uma base de $S^\perp$. Por outro lado as linhas da matriz A_I são as imagens pela projecção $\pi_I|_{S^\perp}$ das linhas de A, donde segue que (4)⇔(5). $\square$



3 Enunciado e Demonstração

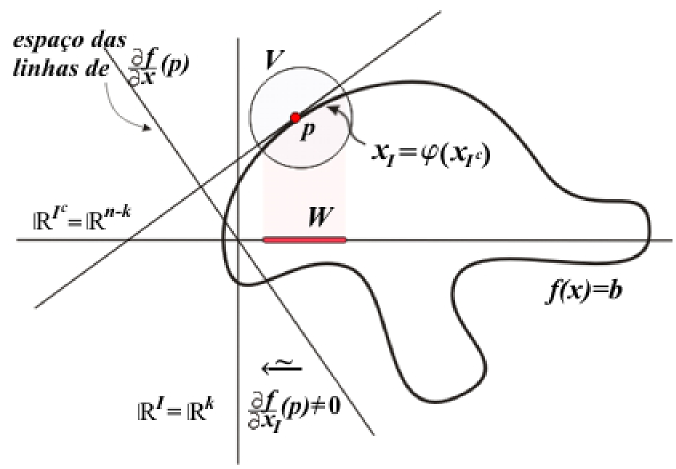
Dada uma aplicação diferenciável $f : D \to \mathbb{R}^k$, onde $D \subset \mathbb{R}^n$ é um aberto, designamos por $\frac{\partial f}{\partial x}(p) = J_f(p)$ a matriz Jacobiana de f num ponto $p \in D$. Dado um conjunto de índices $I \subset \{1, \ldots, n\}$, designamos por $\frac{\partial f}{\partial x_I}(p) = J_f(p)_I$ a submatriz de $J_f(p)$ formada pelas colunas indexadas em I.

Teorema 1. *Sejam $1 \le k \le n$ inteiros, $D \subseteq \mathbb{R}^n$ um domínio aberto e $f : D \to \mathbb{R}^k$ uma função de classe C^r, $r \ge 1$. Dado $I \subset \{1, \ldots, n\}$ com $|I| = k$, e $p \in D$, se*

1. $f(p) = b$,

2. $\det\left(\frac{\partial f}{\partial x_I}(p)\right) \neq 0$,

então existe $V \subset D$ aberto com $p \in V$ tal que o conjunto $X = \{x \in V : f(x) = b\}$ é o gráfico de uma função do tipo I e de classe C^r.



O subespaço afim

$$P := \{x \in \mathbb{R}^n : J_f(p)\,(x-p) = 0\,\}$$

é o plano (afim de dimensão $n-k$) tangente a X no ponto p. A justificação deste facto e a definição de plano tangente serão dadas adiante. A segunda hipótese do Teorema da Função Implícita é exactamente a condição necessária e suficiente para que o plano tangente P seja o gráfico duma função (afim) de tipo I.

Demonstração. Sem perda de generalidade supomos que $I^c = \{1, \ldots, n-k\}$, $I = \{n-k+1, \ldots, n\}$ e identificamos $\mathbb{R}^n = \mathbb{R}^{n-k} \times \mathbb{R}^k$. Vamos denotar os elementos de $\mathbb{R}^n$ como pares ordenados (x, y), onde $x \in \mathbb{R}^{n-k} = \mathbb{R}^{I^c}$ e $y \in \mathbb{R}^k = \mathbb{R}^I$.

Seja $p = (x_0, y_0) \in D$ um ponto tal que $f(x_0, y_0) = b$ e a matriz $A = \frac{\partial f}{\partial y}(x_0, y_0) = \frac{\partial f}{\partial x_I}(p)$ é invertível. Definindo a aplicação $g : D \to \mathbb{R}^k$,

$$g(x, y) := y - A^{-1}\,(f(x, y) - b)$$

resulta que

$$f(x, y) = b \quad \Leftrightarrow \quad g(x, y) = y. \tag{1}$$

Assim, um ponto $(x, y) \in D$ é solução da equação $f(x, y) = b$ sse y for um ponto fixo do mapa $g(x, \cdot) : y \mapsto g(x, y)$. Por construção o mapa $g(x_0, \cdot)$ tem derivada nula no ponto y_0.

$$\frac{\partial g}{\partial y}(x_0, y_0) = I - A^{-1}\,\frac{\partial f}{\partial y}(x_0, y_0) = I - A^{-1}\,A = 0.$$

Seja $W = \overline{B}_r(x_0) \times \overline{B}_r(y_0)$ onde $\overline{B}_r(x_0) \subset \mathbb{R}^{n-k}$ e $\overline{B}_r(y_0) \subset \mathbb{R}^k$ representam as bolas fechadas, respectivamente centradas em x_0 e y_0, com raios $r > 0$. Tomando r suficientemente pequeno podemos supor que para todo $(x, y) \in W$, $\|\frac{\partial g}{\partial y}(x, y)\| \leq 1/2$. Pelo Teorema do valor médio segue que

$$\|g(x, y) - g(x, y')\| \leq \frac{1}{2}\,\|y - y'\|$$

quaisquer que sejam $x \in \overline{B}_r(x_0)$ e $y, y' \in \overline{B}_r(y_0)$. Em particular, para todo $x \in \overline{B}_r(x_0)$, o mapa $g(x, \cdot)$ é uma contracção Lipschitz com constante de Lipschitz $k = \frac{1}{2}$.

Seja agora $L = \sup_{(x,y)\in W} \|\frac{\partial g}{\partial x}(x, y)\|$. Pelo Teorema de Weierstrass, como W é fechado e limitado, $L < \infty$ e $r' := r/(2L) > 0$. Dados $x \in B_{r'}(x_0)$ e $y \in \overline{B}_r(y_0)$, pelo Teorema do Valor Médio

$$\begin{aligned}
\|g(x, y) - b\| &= \|g(x, y) - g(x_0, y_0)\| \\
&\leq \|g(x, y) - g(x, y_0)\| + \|g(x, y_0) - g(x_0, y_0)\| \\
&\leq \frac{1}{2}\,\|y - y_0\| + L\,\|x - x_0\| \leq \frac{r}{2} + L\,r' = \frac{r}{2} + \frac{r}{2} = r
\end{aligned}$$

o que mostra que $g(x, \cdot)(\overline{B}_r(y_0)) \subseteq \overline{B}_r(y_0)$. Logo, pelo Teorema do Ponto Fixo de Banach para cada $x \in B_{r'}(x_0)$, o mapa $g(x, \cdot)$ admite um único ponto fixo $\varphi(x) \in \overline{B}_r(y_0)$. Segue que $g(x, \varphi(x)) = \varphi(x)$, o que por (1) implica $f(x, \varphi(x)) = b$ para todo $x \in B_{r'}(x_0)$.

Vejamos agora a continuidade da aplicação $\varphi\colon B_{r'}(x_0) \to \overline{B}_r(y_0)$. Pelo Teorema do valor médio tem-se

$$\begin{aligned}
\|\varphi(x) - \varphi(x')\| &= \|g(x, \varphi(x)) - g(x', \varphi(x'))\| \\
&\leq \|g(x, \varphi(x)) - g(x', \varphi(x))\| + \|g(x', \varphi(x)) - g(x', \varphi(x'))\| \\
&\leq \max_W \|\frac{\partial g}{\partial x}\|\,\|x - x'\| + \frac{1}{2}\,\|\varphi(x) - \varphi(x')\|
\end{aligned}$$

donde resulta que φ é Lipschitz contínua

$$\|\varphi(x) - \varphi(x')\| \leq 2\,L\,\|x - x'\|.$$

Para vermos que φ é diferenciável em $a \in B_{r'}(x_0)$ mostramos primeiro que a aplicação $h\colon B_{r'}(x_0) \to \mathbb{R}^k$, $h(x) = f(a, \varphi(x))$, é diferenciável em a. Consideremos

$$\Delta_h(x) := -h(x) + h(a) - \frac{\partial f}{\partial x}(a, \varphi(a))\,(x - a).$$

Iremos mostrar que

$$\lim_{x\to a} \frac{\|\Delta_h(x)\|}{\|x - a\|} = 0 \tag{2}$$

justificando assim a diferenciabilidade de h no ponto a. Pelo Teorema fundamental do

Cálculo temos

$$\begin{aligned}\Delta_h(x) &= f(a,\varphi(a)) - f(a,\varphi(x)) - \frac{\partial f}{\partial x}(a,\varphi(a))\,(x-a)\\ &= f(x,\varphi(x)) - f(a,\varphi(x)) - \frac{\partial f}{\partial x}(a,\varphi(a))\,(x-a)\\ &= \int_0^1 \left[\frac{\partial f}{\partial x}(a+t\,(x-a),\varphi(x)) - \frac{\partial f}{\partial x}(a,\varphi(a))\right](x-a)\,dt\end{aligned}$$

Usando a continuidade de φ e de $\frac{\partial f}{\partial x}$ o limite (2) é uma consequência desta fórmula integral para $\Delta_h(x)$.

Consideremos em seguida o erro na fórmula de Taylor de primeira ordem da função $y \mapsto f(a,y)$ no ponto $y=\varphi(a)$,

$$\Delta_f(y) := f(a,y) - f(a,\varphi(a)) - \frac{\partial f}{\partial y}(a,\varphi(a))\,(y-\varphi(a)).$$

Como φ é Lipschitziana e $f(a,\cdot)$ é uma função diferenciável temos

$$\lim_{x\to a}\frac{\|\Delta_f(\varphi(x))\|}{\|x-a\|} = \lim_{x\to a}\frac{\|\Delta_f(\varphi(x))\|}{\|\varphi(x)-\varphi(a)\|}\,\underbrace{\frac{\|\varphi(x)-\varphi(a)\|}{\|x-a\|}}_{\leq 2L} = 0. \tag{3}$$

Das definições de $\Delta_f(\varphi(x))$ e $\Delta_h(x)$ obtemos

$$\begin{aligned}\varphi(x)-\varphi(a) &= \left[\frac{\partial f}{\partial y}(a,\varphi(a))\right]^{-1}[f(a,\varphi(x)) - f(a,\varphi(a)) - \Delta_f(\varphi(x))]\\ &= \left[\frac{\partial f}{\partial y}(a,\varphi(a))\right]^{-1}[h(x)-h(a)-\Delta_f(\varphi(x))]\\ &= \left[\frac{\partial f}{\partial y}(a,\varphi(a))\right]^{-1}\left[-\frac{\partial f}{\partial x}(a,\varphi(a))\,(x-a) - \Delta_h(x) - \Delta_f(\varphi(x))\right]\\ &= -\left(\frac{\partial f}{\partial y}\right)^{-1}\frac{\partial f}{\partial x}\,(x-a) - \left(\frac{\partial f}{\partial y}\right)^{-1}(\Delta_h(x)+\Delta_f(\varphi(x)))\end{aligned}$$

onde $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x}(a,\varphi(a))$ e $\frac{\partial f}{\partial y} = \frac{\partial f}{\partial y}(a,\varphi(a))$. Logo os limites (2) e (3) implicam a diferenciabilidade de $\varphi(x)$ em $x=a$ com derivada

$$\frac{\partial \varphi}{\partial x}(a) = -\left(\frac{\partial f}{\partial y}(a,\varphi(a))\right)^{-1}\frac{\partial f}{\partial x}(a,\varphi(a)). \tag{4}$$

Chamando $R_\varphi(a)$ à expressão do lado direito em (4) obtemos que a função R_φ é de classe C^0, o que por (4) implica que $\frac{\partial\varphi}{\partial x} = R_\varphi$ seja de classe C^0. Logo φ é de classe C^1 o que implica que $\frac{\partial\varphi}{\partial x} = R_\varphi$ seja de classe C^1. Prosseguindo indutivamente, enquanto tivermos que φ é de classe C^j com $j<r$ podemos concluir que R_φ é também de classe C^j, o que implica que $\frac{\partial\varphi}{\partial x}$ seja de classe C^j e portanto que φ seja de classe C^{j+1}. Logo por indução podemos concluir que, tal como f, φ é de classe C^r. $\square$