Existência e dependência contínua de soluções de EDOs através do teorema da função implícita

Ronaldo B. Assunção* Paulo C. Carrião

Equações diferenciais são fundamentais para a compreensão de muitos problemas de matemática e de suas aplicações às outras ciências. Por exemplo, já no século XVII Newton e Leibnitz utilizaram a teoria do Cálculo para resolver problemas sobre o movimento de partículas e também de planetas. E a longa história desse assunto continua a se desenvolver até a atualidade, com importantes aplicações na engenharia de controle e automação. Dois dos resultados mais importantes da teoria de equações diferenciais ordinárias (EDOs) são o teorema de existência de solução para o problema de Cauchy e o teorema sobre a diferenciabilidade contínua da solução em relação às condições iniciais. Em muitos livros sobre esse assunto as demonstrações são apresentadas separadamente. Em primeiro lugar, o resultado de existência é obtido pelo Teorema de Picard-Lindelöf, que usa o método das aproximações sucessivas e o teorema de pontos fixos para contrações; em seguida, o resultado de dependência contínua e diferenciável é demonstrado usando o lema de Gronwall. O objetivo principal destas notas é demonstrar os dois resultados de uma só vez, através de uma aplicação bem estruturada do teorema da função implícita.

A disposição do assunto pode ser percebida através dos títulos das seções; entretanto, salientamos alguns aspectos importantes. No que diz respeito ao teorema da função implícita, necessitamos de uma versão para espaços de Banach. Dessa forma, seguimos a sugestão de Lima em [7] e apresentamos uma demonstração baseada no texto de Kolmogorov e Fomin em [4] que segue, sem modificações essenciais, do correspondente teorema para espaços vetoriais de dimensões finitas. Além disso, para que o texto seja relativamente completo em vista do objetivo, enunciamos muitos resultados que se encontram disponíveis de forma detalhada na literatura, como por exemplo o teorema do ponto fixo para contrações, inspirado no artigo de Palais em [8]. Por fim, é importante ressaltar que a demonstração do resultado principal, Teorema 11.1, baseia-se no interessante artigo de Robbin em [9].

Sumário

1 Definições e resultados preliminares

Nesta seção apresentamos as definições e notações básicas que serão usadas em todo o texto, bem como alguns resultados importantes e ilustrativos.

*R. B. Assunção foi parcialmente apoiado pela Fapemig (Projeto CEX APQ-00748-08).

1.1 Definição Seja $X \equiv (X, \|\cdot\|)$ um espaço métrico. Uma sequência $(x_n)_{n\in\mathbb{N}} \subset X$ é chamada de sequência de Cauchy se, para cada $\varepsilon > 0$, existe $N = N(\varepsilon) \in \mathbb{N}$ tal que $\|x_m - x_n\| < \varepsilon$ para quaisquer $m, n \geqslant N$.

Observamos que sequências de Cauchy em um espaço métrico X são invariantes por translações. Mais especificamente, se $(x_n)_{n\in\mathbb{N}}$ é uma sequência de Cauchy e $a \in X$ é um elemento arbitrário, então a sequência $(x_n + a)_{n\in\mathbb{N}}$ também é uma sequência de Cauchy. Isto mostra, em particular, que sequências de Cauchy não podem ser definidas usando o conceito de vizinhança. Outro fato importante é que ao definir o conceito de convergência usando a idéia de vizinhança, o limite da sequência aparece explicitamente na definição; dessa forma, em princípio, para mostrar que uma sequência converge é necessário conhecer previamente o candidato a limite da sequência. Por outro lado, com a noção de sequência de Cauchy é possível reconhecer a convergência de uma sequência sem determinar o limite. Essa é uma das principais propriedades das sequências de Cauchy. A seguir enunciamos, sem demonstrar, outros importantes resultados sobre esse tipo de sequência. Deixamos as demonstrações para o leitor interessado.

1.2 Proposição
1. *Toda sequência convergente é uma sequência de Cauchy.*
2. *Toda sequência de Cauchy é limitada.*
3. *Se uma sequência de Cauchy possui uma subsequência convergente, então a sequência original é convergente.*

A recíproca do item 1 da Proposição 1.2 não é verdadeira em geral; em outras palavras, existem espaços métricos nos quais nem toda sequência de Cauchy é convergente.

1.3 Exemplo Definimos recursivamente, no espaço $\mathbb{Q} \equiv (\mathbb{Q}, |\cdot|)$, a sequência $(x_n)_{n\in\mathbb{N}}$ por $x_0 = 2$ e $x_{n+1} = (x_n + 2/x_n)/2$ para todo $n \in \mathbb{N}$. Claramente temos que $x_n \in \mathbb{Q}$ para todo $n \in \mathbb{N}$. Além disso, é fácil verificar que esta sequência converge para $\sqrt{2}$ em $\mathbb{R}$. Assim, pela Proposição 1.2 temos que $(x_n)_{n\in\mathbb{N}}$ é uma sequência de Cauchy em $\mathbb{R}$ e, portanto, também é uma sequência de Cauchy em $\mathbb{Q}$.

Por outro lado, a sequência $(x_n)_{n\in\mathbb{N}}$ não pode convergir em $\mathbb{Q}$. De fato, se $x_n \to a$ para algum $a \in \mathbb{Q}$, então $x_n \to a$ em $\mathbb{R}$ também. Pela unicidade do limite, obtemos finalmente que $a = \sqrt{2} \in \mathbb{R}\backslash\mathbb{Q}$, o que é uma contradição. $\triangle$

Esse exemplo motiva a definição dos espaços de Banach, que apresentamos a seguir.

1.4 Definição Um espaço métrico X é chamado de espaço métrico completo se toda sequência de Cauchy em X é convergente. Um espaço vetorial normado completo é chamado de espaço de Banach.

1.5 Exemplos
1. A reta real $\mathbb{R}$ é um espaço de Banach ao ser munida da norma $\|x\| = |x|$ para qualquer $x \in \mathbb{R}$.
2. O espaço real $\mathbb{R}^n$ com elementos denotados por $x = (x_1, x_2, \ldots, x_n)$ é um espaço de Banach ao ser munido da norma
$$\|x\| = \Big[\sum_{i=1}^{i=n} x_i^2\Big]^{1/2};$$
a métrica associada a essa norma é dada por $\rho(x, y) = \|x - y\|$.

3. Se $(X, \|\cdot\|_1)$ é um espaço de Banach e se as normas $\|\cdot\|_1$ e $\|\cdot\|_2$ são equivalentes entre si, então $(X, \|\cdot\|_2)$ também é um espaço de Banach.

4. O espaço $C([a,b];\mathbb{R})$ das funções contínuas $f\colon [a,b] \to \mathbb{R}$ munido da norma

$$\|f\|_\infty = \sup\{|f(x)|\colon x \in [a,b]\}$$

é um espaço de Banach.

5. Seja $K \subset \mathbb{R}^n$ um subconjunto compacto. O conjunto das funções reais e contínuas definidas em K, isto é, $C(K;\mathbb{R}) = \{f\colon K \to \mathbb{R}\colon f \text{ é contínua}\}$ munido da norma

$$\|f\|_\infty = \sup_{x\in K} |f(x)|$$

para $f \in C(K;\mathbb{R})$ define um espaço de Banach.

Já o conjunto $C^m(K;\mathbb{R}) = \{f\colon K \to \mathbb{R}; f \text{ é } m \text{ vezes continuamente diferenciável}\}$ em que $m \in \mathbb{N}$, munido da norma $\|\cdot\|_\infty$ *não* é um espaço de Banach, visto que não é um espaço completo. Especificamente, a convergência em relação à norma $\|\cdot\|_\infty$ é a convergência uniforme e enquanto o limite uniforme de uma sequência de funções contínuas é uma função contínua, em geral o limite uniforme de uma sequência de funções diferenciáveis não é uma função diferenciável. △

1.6 Proposição *Seja X um espaço de Banach e seja $F \subset X$ um subconjunto qualquer. Então F é um subespaço de Banach se, e somente se, F é fechado.*

DEMONSTRAÇÃO. Suponhamos que $F \subset X$ seja um subespaço fechado e seja $(x_n)_{n\in\mathbb{N}}$ uma sequência de Cauchy em F. Então esta sequência também é de Cauchy em X; logo, existe o limite $\lim_{n\to\infty} x_n \equiv x$ em X. Como F é fechado, resulta que $x \in F$.

A recíproca é imediata. □

No que segue, o conjunto das aplicações lineares contínuas $S\colon X \to Y$ é denotado por $L(X;Y)$ e o conjunto das aplicações limitadas $T\colon X \to Y$ é denotado por $B(X;Y)$.

1.7 Proposição *Seja V um conjunto qualquer e seja X um espaço de Banach. Então $B(V;X)$ é um espaço de Banach.*

DEMONSTRAÇÃO. Seja $(f_n)_{n\in\mathbb{N}}$ uma sequência de Cauchy em $B(V;X)$. Dado $\varepsilon > 0$ existe $N = N(\varepsilon) \in \mathbb{N}$ tal que, se $m, n \geqslant N$, então $\|f_m - f_n\| < \varepsilon$. Dessa forma, para cada $x_0 \in X$ fixado, se $m, n \geqslant N$, então

$$\|f_m(x_0) - f_n(x_0)\| < \varepsilon. \tag{1}$$

Dessa forma, a sequência $(f_n(x_0))_{n\in\mathbb{N}}$ é de Cauchy em X; por hipótese, existe o limite dessa sequência. Assim, definimos uma função $f\colon V \to X$ através da fórmula $f(x_0) \equiv \lim_{n\to\infty} f_n(x_0)$. Para demonstrar que $f \in B(V;X)$, basta verificar que f é o limite uniforme da sequência e que é limitada.

Passando ao limite na expressão (1), obtemos

$$\lim_{n\to\infty} \|f_m(x_0) - f_n(x_0)\| = \|f_m(x_0) - f(x_0)\| < \varepsilon$$

para todo $x_0 \in X$. Se $m \geqslant N$, então $f_n \to f$ uniformemente. Em particular, para $\varepsilon = 1$, escolhendo $m > N(\varepsilon)$ resulta $\|f_m - f\| < 1$; portanto, $\|f(x)\| < \|f_m\| + 1$. Assim, obtemos a limitação de $f \in B(V;X)$ e isto conclui a demonstração da proposição. □

1.8 Proposição *Sejam X e Y espaços normados. Suponhamos que Y seja espaço de Banach e que $V \subset X$ seja um subconjunto qualquer. Então o conjunto $C(V;Y)$ é um subespaço fechado de $B(V;Y)$. Em particular, $C(V;Y)$ é um espaço de Banach.*

DEMONSTRAÇÃO. Para verificar que $C(V;Y)$ é um subconjunto fechado, mostraremos que o limite uniforme f de uma sequência $(f_n)_{n\in\mathbb{N}}$ de funções contínuas é sempre uma função contínua em cada ponto $x_0 \in X$. De fato, suponhamos que $\lim_{n\to\infty} f_n = f$. Dado $\varepsilon > 0$, existe $N(\varepsilon) \in \mathbb{N}$ tal que, se $m \geqslant N(\varepsilon)$, então $\|f_m - f\| < \varepsilon/3$. Como f_n é contínua para todo $n \in \mathbb{N}$, escolhendo $n \geqslant N(\varepsilon)$, existe $\delta(\varepsilon)$ tal que $\|f_n(x) - f_n(x_0)\| < \varepsilon/3$ sempre que $\|x - x_0\|$. Assim,

$$\begin{aligned}\|f(x) - f(a)\| &\leqslant \|f(x) - f_n(x)\| + \|f_n(x) - f_n(a)\| + \|f_n(a) - f(a)\| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon\end{aligned}$$

sempre que $\|x - x_0\| \leqslant \delta(\varepsilon)$. Portanto, f é contínua e o conjunto $C(V;Y)$ é fechado em $B(V;Y)$. Finalmente, usando a Proposição 1.6 concluímos que $C(V;Y)$ é um espaço de Banach. Isto conclui a demonstração da proposição. □

2 Derivadas de Fréchet e de Gâteaux

Tendo em vista a aplicação que apresentamos para o teorema da função implícita, necessitamos generalizar os conceitos de derivada e de derivada direcional do cálculo de várias variáveis. As derivadas de Fréchet e de Gâteaux servem a esse propósito. Nesta seção apresentamos as definições e na sequência estabelecemos as relações entre os dois conceitos.

2.1 Definição Sejam X e Y espaços de Banach, e seja $V \subset X$ um subconjunto aberto. Dizemos que a função $F\colon V \to Y$ é Fréchet diferenciável no ponto $x_0 \in V$ se existe uma aplicação linear $DF(x_0)\colon X \to Y$ tal que

$$\lim_{h\to 0} \frac{|F(x_0 + h) - F(x_0) - DF(x_0)h|}{|h|} = 0. \tag{2}$$

Se F é diferenciável em cada ponto $x \in V$ e a aplicação $DF\colon \Omega \to L(X,Y)$ é contínua, então dizemos que F é continuamente diferenciável em V e escrevemos $F \in C^1(V;Y)$. A expressão $DF(x_0)h$, que denota, para qualquer $h \in X$, um elemento do espaço Y, é denominada diferencial de Fréchet (ou diferencial forte) da aplicação F no ponto x_0. O próprio operador $DF(x_0)$ é denominado derivada forte da aplicação F no ponto x_0 e frequentemente é denotada também por $F'(x_0)$.

A seguir listamos algumas das propriedades fundamentais da derivada de Fréchet.

2.2 Exemplos

1. Se F é Fréchet diferenciável no ponto x_0, então $DF(x_0)$ é univocamente determinada.
2. Se F é Fréchet diferenciável no ponto x_0, então F é contínua nesse ponto.
3. Se $F(x) = y_0$ é uma função constante, então $DF(x) \equiv 0$, ou seja, a derivada de Fréchet de F é o operador nulo.
4. A derivada de Fréchet da aplicação linear $L\colon X \to Y$ coincide com a própria aplicação, ou seja, $DL(x) = L$.

5. Sejam $F, G\colon X \to Y$ aplicações Fréchet difenciáveis no ponto x_0. Então as aplicações $F + G$ e αF em que $\alpha \in \mathbb{R}$ também são Fréchet diferenciáveis nesse ponto e

$$\begin{aligned} D(F+G)(x_0) &= DF(x_0) + DG(x_0) \\ D(\alpha F)(x_0) &= \alpha DF(x_0). \end{aligned}$$

6. Sejam X, Y, Z espaços normados; seja $U(x_0)$ uma vizinhança de $x_0 \in X$ e $F\colon U(x_0) \to Y$ uma aplicação dessa vizinhança em Y com $F(x_0) = y_0$; seja $V(y_0)$ uma vizinhança de $y_0 \in Y$ e $G\colon V(y_0) \to Z$ uma aplicação dessa vizinhança em Z. Nessas condições, se a aplicação F é Fréchet diferenciável no ponto x_0 e se a aplicação G é Fréchet diferenciável no ponto y_0, então a aplicação $H = G \circ F$, definida em uma certa vizinhança de x_0, é Fréchet diferenciável nesse ponto e vale a relação

$$DH(x_0) = DG(y_0) \circ DF(x_0). \tag{3}$$

Se F, G e H são funções numéricas, a fórmula (3) corresponde à regra de derivação da função composta. △

2.3 Definição Um espaço vetorial normado completo X e munido de um produto interno $\langle \cdot, \cdot \rangle\colon X \times X \to \mathbb{R}$ é chamado de espaço de Hilbert.

2.4 Exemplo Seja X um espaço de Hilbert com produto interno $\langle \cdot, \cdot \rangle$. Então a derivada de Fréchet da função $F\colon X \to \mathbb{R}$ definida por $F(x) = \|x\| = [\langle x, x \rangle]^{1/2}$ é dada por $DF(x)h = \langle \frac{x}{\|x\|}, h \rangle$ para $x \neq 0$.

De fato, definindo a função $G(x) = \|x\|^2$, temos que

$$\begin{aligned} \frac{\|x+th\|^2 - \|x\|^2}{t} &= \frac{\langle x+th, x+th \rangle - \langle x, x \rangle}{t} \\ &= 2\langle x, h \rangle + t\langle h, h \rangle. \end{aligned}$$

Portanto, $DG(x,h) = 2\langle x, h \rangle$. Esta aplicação é contínua para todos os pontos $x \in X$; logo, G possui derivada de Fréchet e

$$DG(x)h = 2\langle x, h \rangle.$$

Como $F = G^{1/2}$, usando a regra da cadeia obtemos $DG(x) = 2\|x\|DF(x)$, e como $x \neq 0$, resulta que $DF(x)h = \langle x/\|x\|, h \rangle$. △

2.5 Definição Sejam X e Y espaços de Banach e seja $V \subset X$ um subconjunto aberto. Dizemos que a função $F\colon X \to Y$ é Gâteaux diferenciável no ponto x_0 em relação ao acréscimo $h \in X$ se existe o limite

$$DF(x_0, h) = \lim_{t \to 0} \frac{F(x_0 + th) - F(x_0)}{t}, \tag{4}$$

em que este limite é avaliado relativamente à norma de Y. A expressão $DF(x_0, h)$ é chamada de diferencial de Gâteaux; esta pode não ser linear em relação ao acréscimo h. No entanto, se tal linearidade se cumprir, isto é, se

$$DF(x, h) = DF_G(x_0)h,$$

em que DF_G é um operador linear limitado, então este último é denominado derivada de Gâteaux (ou derivada fraca) de F no ponto x.

2.6 Exemplos 1. Se F é Gâteaux diferenciável no ponto x_0, então $D_GF(x_0, h)$ é univocamente determinada.

2. Se F é Gâteaux diferenciável no ponto x_0, então $D_GF(x_0, th) = tD_GF(x_0, h)$ para todo $t \in \mathbb{R}$.

3. Se F é Gâteaux diferenciável no ponto x_0, então para todo $h \in X$ e para qualquer funcional linear $y^* \in Y^*$, a função real $\phi\colon \mathbb{R} \to \mathbb{R}$ definida por $\phi(t) = \langle y^*, f(x_0 + th)\rangle$, em que o símbolo $\langle\cdot,\cdot\rangle$ denota o produto de dualidade, isto é, $\langle y^*, f(x_0 + th)\rangle = y^*(f(x_0 + th)$, é diferenciável no ponto $t = 0$ e $\phi'(t) = \langle y^*, D_GF(x_0, h)\rangle$. △

2.7 Definição Sejam X_1, X_2 e Y espaços de Banach. Sejam $V_1 \subset X_1$ e $V_2 \subset X_2$ subconjuntos abertos e consideremos a função $F\colon V_1 \times V_2 \to Y$. Se $(x_1, x_2) \in V_1 \times V_2$ e se mantivermos $x_2 \in V_2$ fixo, então como função da primeira variável definimos a derivada de Fréchet conforme o fizemos anteriormente. Se existir, esta derivada será denotada por $D_1F(x_1, x_2)$ e, nesse caso, temos a aplicação $D_1F\colon V_1 \times V_2 \to L(X_1; Y)$, denominada derivada parcial em relação à primeira variável. Analogamente, podemos definir a derivada parcial em relação à segunda variável e denotá-la por $D_2F(x_1, x_2)$, ou seja, $D_2F\colon V_1 \times V_2 \to L(X_2; Y)$.

De forma geral, podemos considerar uma função $F\colon V_1 \times V_2 \times \cdots \times V_n \to Y$ definida em subconjuntos abertos $V_k \subset X_k$ de espaços de Banach X_k para $1 \leqslant k \leqslant n$ e, dado $x = (x_1, x_2, \ldots, x_n) \in V_1 \times V_2 \times \cdots \times V_n$, se mantivermos todas as coordenadas fixas com exceção da k-ésima, definimos a derivada parcial $D_kF((x_1, x_2, \ldots, x_k, \ldots, x_n))$ em relação a essa variável, ou seja, $D_kF\colon V_1 \times V_2 \times \cdots \times V_n \to L(X_k; Y)$.

Derivadas de ordens superiores podem ser definidas da maneira análoga; para mais detalhes, confira o Apêndice. Na próxima seção apresentamos a fórmula dos acréscimos finitos e uma de suas aplicações para relacionar os conceitos de derivadas de Fréchet e de Gâteaux.

3 Fórmula dos acréscimos finitos

Nesta seção apresentamos a fórmula dos acréscimos finitos, também chamada de desigualdade do valor médio. Seja $U \subset X$ um subconjunto aberto e seja $[x_0, x]$ um segmento contido em U. Seja $F\colon U \to Y$ uma função possuindo derivada fraca DF_G em todos os pontos do segmento $[x_0, x]$. Dado um funcional arbitrário $\phi \in Y^*$, denotamos por $\Delta x = x - x_0$ e consideramos a função numérica $f\colon [0, 1] \to \mathbb{R}$ definida por

$$f(t) = \phi(F(x_0 + t\Delta x)). \tag{5}$$

Esta função é diferenciável como função da variável t. De fato, dada a expressão

$$\frac{f(t + \Delta t) - f(t)}{\Delta t} = \phi\Big(\frac{F(x_0 + t\Delta x + \Delta t\Delta x) - F(x_0 + t\Delta x)}{\Delta t}\Big)$$

podemos passar ao limite sob o sinal do funcional linear contínuo ϕ, e assim obtemos

$$f'(t) = \phi(DF_G(x_0 + t\Delta x)\Delta x).$$

Aplicando à função f a fórmula dos acréscimos finitos no intervalo $[0, 1]$, obtemos

$$f(1) = f(0) + f'(\theta) \qquad \theta \in [0, 1],$$

ou seja,

$$\phi(F(x) - F(x_0)) = \phi(DF_G(x_0 + \theta\Delta x)\Delta x). \tag{6}$$

Esta relação é válida para qualquer funcional $\phi \in Y^*$, em que θ depende de ϕ. A partir da equação (6), obtemos

$$|\phi(F(x) - F(x_0))| \leqslant \|\phi\| \sup_{\theta\in[0,1]} \|DF_G(x_0 + \theta\Delta x)\| \|x\|. \tag{7}$$

Agora escolhemos um funcional não nulo ϕ verificando a condição

$$\phi(F(x) - F(x_0)) = \|\phi\| \|F(x) - F(x_0)\| \tag{8}$$

que existe em decorrência do Teorema de Hahn-Banach. Da relação (7), resulta que

$$\|F(x) - F(x_0)\| \leqslant \sup_{\theta\in[0,1]} \|DF_G(x_0 + \theta\Delta x)\| \|\Delta x\|. \tag{9}$$

Esta desigualdade pode ser compreendida como o análogo para espaços de Banach da fórmula dos acréscimos finitos para funções numéricas (também chamada de teorema do valor médio no caso de funções reais ou de desigualdade do valor médio no caso de funções vetoriais).

No caso particular em que a aplicação linear tem a forma $x \mapsto F(x) - DF_G(x_0)\Delta x$, da fórmula (9) dos acréscimos finitos obtemos

$$\|F(x) - F(x_0) - DF_G(x_0)\Delta x\| \leqslant \sup_{\theta\in[0,1]} \|DF_G(x_0 + \theta\Delta x) - DF_G(x_0)\| \|\Delta x\|. \tag{10}$$

4 Relação entre as derivadas de Fréchet e de Gâteaux

Os conceitos de derivada forte de Fréchet e de derivada fraca de Gâteaux não coincidem nem mesmo no caso de espaços de dimensão finita. Sabemos que, dada a função numérica $f\colon \mathbb{R}^n \to \mathbb{R}$ em que $n \in \mathbb{N}$ é tal que $n \geqslant 2$, a existência da derivada direcional $D_h f(x)$ para qualquer $h \in \mathbb{R}^n$ fixo não é condição suficiente para que esta função seja diferenciável, isto é, para que o acréscimo finito $f(x+h) - f(x)$ possa ser representado como a soma de uma função linear de h e de um termo que tende a zero mais rapidamente do que h.

4.1 Exemplo Consideremos a função de duas variáveis $f\colon \mathbb{R}^2 \to \mathbb{R}$ definida por

$$f(x,y) \equiv \begin{cases} \dfrac{x^3 y}{x^4 + y^2} & \text{se } (x,y) \neq (0,0), \\ 0 & \text{se } (x,y) = (0,0). \end{cases} \tag{11}$$

Esta função é contínua em todo o plano $\mathbb{R}^2$, inclusive na origem. Além disso, existe a derivada fraca de f na origem, que é a função nula, pois dado $h = (h_1, h_2) \in \mathbb{R}^2$, temos

$$\lim_{t\to 0} \frac{f(0 + th_1, 0 + th_2) + f(0,0)}{t} = \lim_{t\to 0} \frac{t^4 h_1^2 h_2}{t^4 h_1^4 + t^2 h_2^2} = 0.$$

Por outro lado, esta derivada fraca não é a parte linear da variação da função f em uma vizinhança da origem. De fato, se usarmos $h_2 = h_1^2$, obtemos

$$\lim_{\|h\|\to 0} \frac{f(h_1, h_2)}{\|h\|} = \lim_{\|h\|\to 0} \frac{h_1^5}{2h_1^4\,(h_1^2 + h_1^4)^{1/2}} = \frac{1}{2} \neq 0.$$

Isto significa que a função f não possui derivada de Fréchet na origem. △

Entretanto, se uma aplicação F possui derivada forte, então também possui derivada fraca e ambas estas derivadas coincidem. De fato, seja F uma aplicação Fréchet diferenciável; então

$$\lim_{t \to 0} \frac{F(x+th) - F(x)}{t} = \lim_{t \to 0} F'(x)h.$$

O resultado a seguir estabelece condições adicionais para que uma função Gâteaux diferenciável também seja Fréchet diferenciável.

4.2 Teorema *Seja $F\colon X \to Y$ uma função fracamente diferenciável no ponto $x_0 \in X$. Se a derivada de Gâteaux $DF_G(x)$ existe em cada ponto de uma vizinhança U de x_0 e se a função que a cada $x \in U$ associa o operador $DF_G(x)$ for contínua em x_0, então a derivada de Fréchet $DF(x_0)$ existe e coincide com a derivada de Gâteaux.*

DEMONSTRAÇÃO. Dado $\varepsilon > 0$, existe $\delta > 0$ tal que, para qualquer $\|h\| < \delta$, vale a desigualdade

$$\|DF_G(x_0 + h) - DF_G(x_0)\| \leqslant \varepsilon.$$

Aplicando a fórmula (10) à função F, obtemos

$$\|F(x_0 + h) - F(x_0) - DF_G(x_0)h\| \leqslant \sup_{\theta \in [0,1]} \|DF_G(x_0 + \theta h) - DF_G(x_0)\| \|h\| \leqslant \varepsilon \|h\|.$$

Dessa forma, é válido o limite (2); portanto, obtemos a existência da derivada de Fréchet e a sua igualdade com a derivada de Gâteaux. $\square$

A importância do Teorema 4.2 reside no fato de que, em geral, não é fácil determinar diretamente a derivada de Fréchet de uma aplicação, mas o cálculo da derivada de Gâteaux dessa aplicação reduz-se ao cálculo diferencial de uma única variável. Podemos fazer uma analogia com o estudo de funções de várias variáveis, em que o cálculo do gradiente reduz-se ao cálculo das derivadas parciais.

Encerramos esta seção apresentando um resultado similar ao anterior, mas referente às derivadas parciais. Omitimos a demonstração omitimos, já que possui o mesmo esquema geral daquela que apresentamos.

4.3 Proposição *Sejam $X_1, X_2, \ldots, X_n$ e Y espaços de Banach. Sejam $V_k \subset X_k$ subconjuntos abertos para $1 \leqslant k \leqslant n$ e consideremos a função $F\colon V_1 \times V_2 \times \cdots \times V_n \to Y$. Se as derivadas parciais $D_kF\colon V_1 \times V_2 \times \cdots \times V_n \to L(X_k; Y)$ existem e são contínuas, então para $x = (x_1, x_2, \ldots, x_n) \in V_1 \times V_2 \times \cdots \times V_n$ e $h = (h_1, h_2, \ldots, h_n) \in X_1 \times X_2 \times \cdots \times X_n$, temos a relação*

$$DF(x)(h_1, h_2, \ldots, h_n) = \sum_{k=1}^{k=n} D_kF(x)h_k.$$

5 O teorema do ponto fixo para contrações

Muitos problemas que tratam de existência e de unicidade de soluções de equações algébricas ou de equações diferenciais podem ser reduzidos ao problema de existência de pontos fixos por certas aplicações de um espaço métrico em si mesmo. O princípio das aplicações contractantes trata exatamente da existência de um único ponto fixo para certas aplicações.

5.1 Definição Seja M um espaço métrico e consideremos a aplicação $A\colon M \to M$. Dizemos que A é uma contração, se existe um número real $\alpha < 1$ tal que

$$\rho(A(x), A(y)) \leqslant \alpha\rho(x, y) \qquad \text{para quaisquer } x, y \in M. \tag{12}$$

É imediato verificar que toda contração é uma aplicação contínua. De fato, seja A uma contração e seja $(x_n)_{n\in\mathbb{N}} \subset M$ uma sequência tal que $x_n \to x$ em M quando $n \to +\infty$. Então $\rho(A(x_n), A(x)) \leqslant \alpha\rho(x_n, x) \to 0$ quando $n \to +\infty$, ou seja, $A(x_n) \to A(x)$ quando $n \to +\infty$.

Além disso, usando a desigualdade triangular, obtemos

$$\begin{aligned}\rho(x_1, x_2) &\leqslant \rho(x_1, A(x_1)) + \rho(A(x_1, A(x_2)) + \rho(A(x_2, x_2)) \\ &\leqslant \rho(x_1, A(x_1)) + \alpha\rho(x_1, x_2) + \rho(A(x_2, x_2))\end{aligned}$$

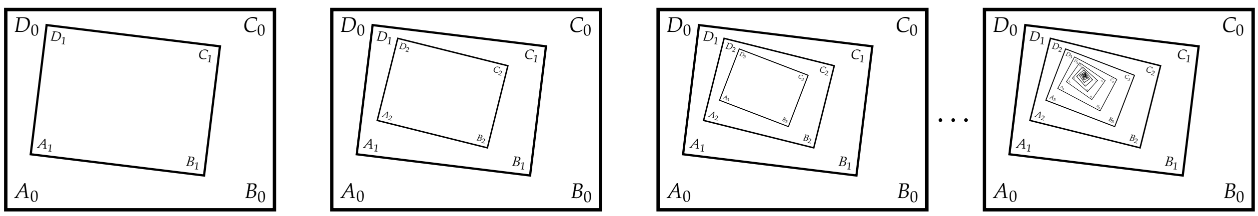
e disso resulta o que Palais em [8] denomina de desigualdade fundamental para contrações, a saber,

$$\rho(x_1, x_2) \leqslant \frac{1}{1-\alpha}[\rho(x_1, A(x_1)) + \rho(x_2, A(x_2))]. \tag{13}$$

5.2 Definição Seja $F\colon M \to M$ uma aplicação definida em um espaço métrico M. Dizemos que $x \in M$ é um ponto fixo por F se $F(x) = x$.

5.3 Teorema *Sejam M um espaço métrico completo e $A\colon M \to M$ uma contração. Então existe um único ponto fixo por A.*

DEMONSTRAÇÃO. Inicialmente mostraremos a unicidade do ponto fixo. Sejam $x_1, x_2 \in M$ pontos fixos para a aplicação A. Então pela desigualdade (13) temos que $\rho(x_1, x_2) = 0$, ou seja, $x_1 = x_2$. Assim, a contração A pode ter no máximo um ponto fixo.



Usando novamente a desigualdade (13), mostraremos agora que a sequência das iteradas de um ponto $x \in M$ arbitrário, isto é, a sequência $(A^n(x))_{n\in\mathbb{N}}$ é uma sequência de Cauchy. De fato, dado um ponto arbitrário $x \in M$, definimos $x_m = A^m(x)$ e $x_n = A^n(x)$ para números naturais $m, n \in \mathbb{N}$, e assim temos

$$\begin{aligned}\rho(x_m, x_n) &= \rho(A^m(x), A^n(x)) \\ &\leqslant \frac{1}{1-\alpha}\big[\rho(A^m(x), A(A^m(x))) + \rho(A^n(x), A(A^n(x)))\big] \\ &= \frac{1}{1-\alpha}\big[\rho(A^m(x), A^m(A(x))) + \rho(A^n(x), A^n(A(x)))\big] \\ &\leqslant \frac{1}{1-\alpha}\big[\alpha^m\rho(x, A(x)) + \alpha^n\rho(x, A(x))\big] \\ &= \frac{\alpha^m + \alpha^n}{1-\alpha}\,\rho(x, A(x)).\end{aligned}$$

Como $\alpha < 1$, resulta que $\alpha^n \to 0$ quando $n \to +\infty$. Portanto, quando $m \to +\infty$ e $n \to +\infty$, resulta que também $\rho(A^m(x), A^n(x)) \to 0$. Dessa forma, a sequência das iteradas do ponto arbitrário $x \in M$ é uma sequência de Cauchy.

Finalmente, como M é um espaço métrico completo, a sequência das iteradas converge para um ponto $p \in M$, que é claramente um ponto fixo para a aplicação A, já que, por definição de p, temos

$$p = \lim_{m \to +\infty} A^m(x) = \lim_{m \to +\infty} A^{m+1}(x) = A(p)$$

e isto conclui a demonstração do teorema. □

Como consequência da demonstração do teorema, para o ponto fixo $p \in M$ temos a desigualdade seguinte,

$$\rho(f^n(x), p) \leqslant \frac{\alpha^n}{1-\alpha} \rho(x, f(x)) \tag{14}$$

válida para qualquer ponto $x \in M$ e para qualquer $n \in \mathbb{N}$.

A relevância da desigualdade é que permite-nos estimar o número N de iterações que são necessárias para o erro na aproximação do ponto fixo seja menor do que um valor previamente fixado. De fato, suponhamos que desejamos determinar, no lugar do ponto fixo $p \in M$ para a aplicação f, um outro ponto denotado por $\tilde{p} \in M$ de modo que $\rho(p, \tilde{p}) < \varepsilon$ em que o valor $\varepsilon > 0$ indica o erro máximo da aproximação. Denotando por $N \in \mathbb{N}$ o número tal que $\tilde{p} = f^N(x)$, devemos ter $\rho(p, f^N(x)) < \varepsilon$; dessa forma, basta usar um valor de N de modo que

$$\frac{\alpha^N}{1-\alpha} \rho(x, f(x)) < \varepsilon.$$

A partir da quantidade $\rho(x, f(x)) = d$, que pode ser obtida a partir da primeira iteração em que usamos um ponto arbitrário $x \in M$, podemos usar a função logarítmica e calcular uma cota inferior para N, a saber,

$$N > \frac{\log(\varepsilon) + \log(1-\alpha) - \log(d)}{\log(\alpha)}. \tag{15}$$

Assim sendo, para um valor de N verificando a desigualdade (15), temos necessariamente $\rho(x, f^N(x)) \leqslant \varepsilon$.

Suponhamos, que $\varepsilon = 10^{-m}$; pela desigualdade (15) vemos que o crescimento de N com o valor de m é uma constante mais o valor de $m/|\log(\alpha)|$, ou seja, para obtermos uma casa decimal adicional na aproximação do ponto fixo devemos efeturar aproximadamente $1/|\log(\alpha)|$ iterações suplementares. Equivalentemente, se necessitamos de N iterações para obtermos uma aproximação com m casas decimais de precisão, então necessitaremos aproximadamente de outras N iterações para obtermos uma precisão de $2m$ casas decimais.

6 Aplicações do teorema do ponto fixo para contrações

O princípio do ponto fixo para contrações é não apenas um dos mais antigos e simples dos teoremas de pontos fixos, como também tem a vantagem de que os métodos de suas demonstrações incluem instruções de como determinar os pontos fixos. Como vimos acima, algumas demonstrações indicam também o erro associado às sequências de aproximações sucessivas.

Como aplicação principal do princípio do ponto fixo para contrações citamos o teorema da função implícita e uma de suas consequências, a saber, o teorema de existência de soluções para equações diferenciais ordinárias e de dependência contínua das soluções em relação aos dados iniciais. Essas aplicações são apresentadas nas seções seguintes. Por ora, mostramos um método de cálculo aproximado de soluções de equações da forma $A(x) = x$, comumente denominado de método das aproximações sucessivas.

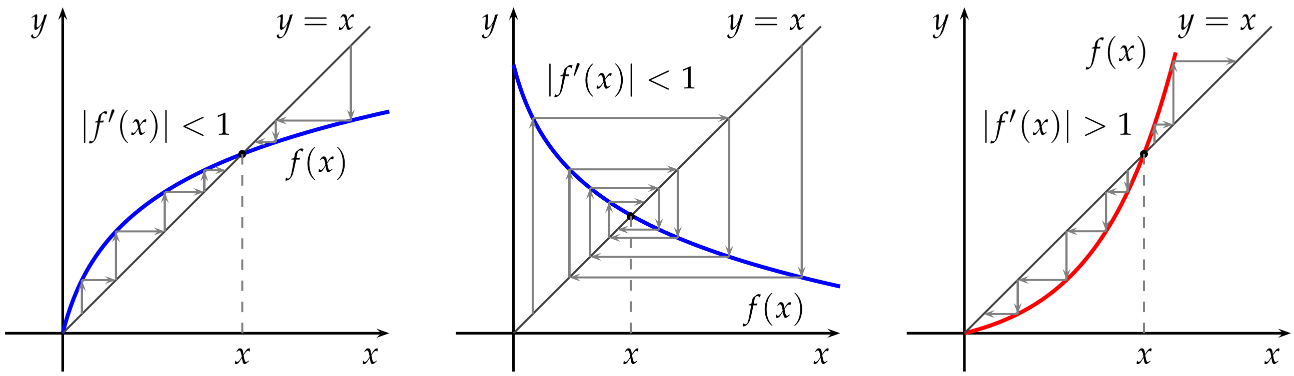
6.1 Exemplo Seja $f\colon [a,b] \to [a,b]$ uma função lipschtiziana, isto é, uma função para a qual existe uma constante $K \in \mathbb{R}$ tal que

$$|f(x) - f(y)| \leqslant K|x - y| \qquad \text{para quaisquer } x, y \in [a,b]. \tag{16}$$

Suponhamos que a constante de Lipschitz K seja tal que $K < 1$. Então a aplicação f é uma contração e, portanto, possui um único ponto fixo que é solução da equação $f(x) = x$. Pela demonstração do teorema, a sequência $(x_n)_{n\in\mathbb{N}}$ definida por $x_0 \in [a,b]$, $x_1 = f(x_0)$, $x_2 = f(x_1) = f^2(x_0)$, e em geral, $x_n = f(x_{n-1}) = \cdots = f^n(x_0)$ para $n \in \mathbb{N}$ converge para o ponto fixo.

Uma situação particular em que f é uma contração é a seguinte. Suponhamos que $f\colon [a,b] \to [a,b]$ seja derivável em (a,b) e que a derivada verifique a desigualdade $|f'(x)| \leqslant K < 1$. Então f é uma contração e a sequência acima converge para a solução da equação $f(x) = x$.

Consideremos agora a função diferenciável $F\colon [a,b] \to \mathbb{R}$ tal que $F(a) < 0$, $F(b) > 0$ e existem constantes $K_1, K_2 \in \mathbb{R}$ tais que $0 < K_1 \leqslant F'(x) \leqslant K_2$ para todo $x \in [a,b]$. Definimos a função $f(x) \equiv x - \lambda F(x)$ e procuramos a solução da equação $f(x) = x$, que é equivalente a $F(x) = 0$ para $\lambda \neq 0$. Como $f'(x) = 1 - \lambda F'(x)$ resulta que $1 - \lambda K_2 \leqslant f'(x) \leqslant 1 - \lambda K_1$, podemos escolher um número $\lambda \neq 0$ que permite aplicar o método das aproximações sucessivas e assim, podemos determinar aproximações das raízes de $F(x) = 0$.



A hipótese $f'(x) < K < 1$ é fundamental para que o método das aproximações sucessivas convirja para o ponto fixo. Na verdade, se $f'(x) > 1$ então a sequência de pontos afasta-se do ponto fixo. △

7 Teorema da função implícita

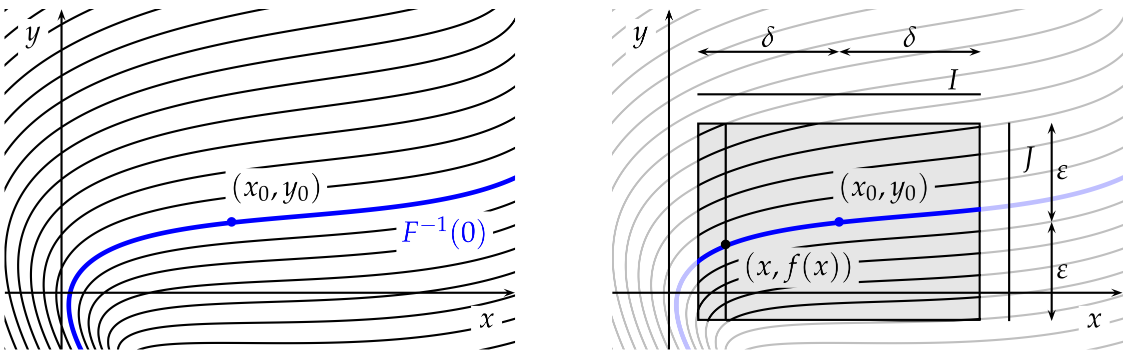
De acordo com Krantz e Parks em [5], o germe da idéia para o teorema da função implícita aparece nos trabalhos de Newton e de Leibniz, que usaram o que chamamos de diferenciação implícita; além disso, Lagrange descobriu uma versão do teorema da função inversa, que é equivalente ao teorema da função implícita. Entretanto, foi Cauchy que abordou o teorema da função implícita com o devido rigor matemático e é geralmente reconhecido como o descobridor do teorema. Nesta seção enunciamos o teorema da função implícita e na seção seguinte apresentamos o teorema sobre a diferenciabilidade da função implícita; as demonstrações seguem estreitamente as idéias de Kolmogorov e Fomin em [4].

7.1 Teorema *Sejam X, Y e Z espaços de Banach. Seja U uma vizinhança de $(x_0, y_0) \in X \times Y$ e seja $F\colon U \to Z$ uma aplicação verificando as propriedades seguintes.*

1. *F é contínua em (x_0, y_0).*
2. *$F(x_0, y_0) = 0$.*
3. *Existe a derivada parcial $D_2F(x, y)$ para todo $(x, y) \in U$; além disso, esta derivada parcial é contínua em (x_0, y_0) e o operador $D_2F(x_0, y_0)$ possui um inverso limitado.*

Então a equação $F(x, y) = 0$ possui uma única solução em uma vizinhança de (x_0, y_0). Em outros termos, dado $\varepsilon > 0$, existem $\delta > 0$ e uma aplicação $f\colon B \to Y$, em que $B = B_\delta(x_0)$ é a bola de centro x_0 e raio δ, tal que $y = f(x)$ é contínua em x_0 e tal que quaisquer elementos (x, y) com $x \in B_\delta(x_0)$ e $y = f(x)$ verificam também a equação $F(x, y) = 0$. Reciprocamente, quaisquer elementos (x, y) tais que $x \in B_\delta(x_0)$ e $y \in B_\varepsilon(y_0)$ e verificando a igualdade $F(x, y) = 0$ também se relacionam pela fórmula $y = f(x)$.

7.2 Observação Gostaríamos de tornar o teorema da função implícita tão simples como o seguinte caso em dimensão dois. Consideremos uma função $F\colon V \subset \mathbb{R}^2 \to \mathbb{R}$ definida em uma vizinhança aberta $V \subset \mathbb{R}^2$ e verificando as hipóteses do teorema. Dizemos que a equação $F(x, y) = 0$ define y implicitamente como função da variável x se existe uma função $f\colon I \subset \mathbb{R} \to \mathbb{R}$ definida em um intervalo aberto $I \subset \mathbb{R}$ e tal que $F(x, y) = 0$ se, e somente se, $y = f(x)$. Nesse caso, o conjunto $F^{-1}(0)$ é o gráfico da função f.



Para fixar as idéias, suponhamos que $D_2F(x_0, y_0) > 0$. Pela continuidade da função F, existem vizinhanças abertas $I = (x_0 - \delta, x_0 + \delta)$ e $J = (y_0 - \varepsilon, y_0 + \epsilon)$ de x_0 e y_0, respectivamente, tais que $I \times \overline{J} \subset V$ e com $D_2F(x, y) > 0$ para todo ponto $(x, y) \in I \times \overline{J}$. Mais ainda, para cada $x \in I$ fixado, valem as desigualdades $F(x, y_0 - \varepsilon) < 0$ e $F(x, y_0 + \varepsilon) > 0$.

Aplicando o teorema do valor intermediário, para cada elemento $x \in I$ existe um único elemento $y = f(x) \in \overline{J}$ tal que $F(x, y) = 0$. Assim temos $y \in J$ e, portanto, $F^{-1}(0) \cap I \times \overline{J} = F^{-1}(0) I \times J$ é o gráfico de uma função $f\colon I \times J$. Isto garante a existência da função implícita.

No caso mais geral de espaços de Banach, e até mesmo nos casos de dimensões maiores do que a situação que analisamos, não temos a noção de ordem e, portanto, não podemos aplicar o teorema do valor intermediário. Uma possibilidade para contornar essa dificuldade é usar o teorema da perturbação da aplicação identidade; outra alternativa é aplicar o Teorema 5.3 do ponto fixo para contrações. É essa última idéia que utilizamos a seguir para demonstar o teorema. $\triangle$

DEMONSTRAÇÃO DO TEOREMA 7.1. Seja $x \in X$; denotamos por $U_x \subset Y$ o conjunto dos elementos $y \in Y$ tais que $(x, y) \in U$. Suponhamos que x esteja suficientemente próximo de x_0 de modo que $y_0 \in U_x$.

Seja a aplicação $A_x \colon U_x \to Y$ definida por

$$A_x(y) \equiv y - [D_2F(x, y_0)]^{-1}F(x, y). \tag{17}$$

Dessa forma, são equivalentes as afirmativas: $A_x(y) = y$ e $F(x, y) = 0$.

Para demonstrar que a equação (17) possui um ponto fixo $y \in U_x$ para cada x fixado, usaremos o Teorema 5.3 do ponto fixo para contrações. Para isso, devemos mostrar que, dado $\varepsilon > 0$ suficientemente pequeno, existe $\delta = \delta(\varepsilon)$ tal que, se $x \in B_\delta(x_0)$, então a aplicação A_x é uma contração e aplica a bola $B_\varepsilon(y_0)$ em si mesma.

Começamos calculando a derivada primeira da aplicação A_x. Usando as propriedades mais elementares da derivada, temos

$$\begin{aligned} DA_x(y) &= I - [D_2F(x_0, y_0)]^{-1}D_2F(x, y) \\ &= [D_2F(x_0, y_0)]^{-1}\{D_2F(x_0, y_0) - D_2F(x, y)\}. \end{aligned}$$

Agora fazemos uma estimativa da norma da derivada da aplicação A_x. Como por hipótese a derivada D_2F é contínua em (x_0, y_0), existe um número real r tal que

$$\|DA_x(y)\| \leqslant r < 1. \tag{18}$$

Usando a fórmula dos acréscimos finitos, a desigualdade (18) garante que a aplicação A_x definida em Y é, para qualquer elemento $x \in X$ tal que $x \in B_\delta(x_0)$, uma contração da bola fechada $\overline{B_\varepsilon(y_0)}$.

Além disso, temos a estimativa

$$\begin{aligned} \|A_x(y_0) - y_0\| &\leqslant \|[D_2F(x_0, y_0)]^{-1}\| \cdot \|F(x, y_0)\| \\ &= \|[D_2F(x_0, y_0)]^{-1}\| \cdot \|F(x, y_0) - F(x_0, y_0)\| \end{aligned}$$

já que $F(x_0, y_0) = 0$. Como a aplicação F é contínua em (x_0, y_0), a norma da expressão acima torna-se arbitrariamente pequena, desde que δ seja escolhido adequadamente. Seja então $\delta > 0$ suficientemente pequeno de modo que

$$\|A_x(y_0) - y_0\| < \varepsilon(1 - r)\text{sempre que } \|x - x_0\| < \delta.$$

Escolhendo δ dessa forma a aplicação A_x é tal que $A_x(B_\varepsilon(y_0)) \subset B_\varepsilon(y_0)$. De fato, das desigualdades $\|x - x_0\| < \delta$ e $\|y - y_0\| \leqslant \varepsilon$ e da fórmula dos acréscimos finitos resulta

$$\begin{aligned} \|A_x(y) - y_0\| &\leqslant \|A_x(y_0) - y_0\| + \|A_x(y) - A_x(y_0)\| \\ &\leqslant \varepsilon(1 - r) + \sup_{\theta \in [0,1]} \|DA_x(y_0 + \theta(y - y_0))\| \cdot \|y - y_0\| \\ &\leqslant \varepsilon(1 - r) + \varepsilon r = \varepsilon. \end{aligned}$$

Portanto, para $\|x - x_0\| < \delta$, a aplicação A_x é uma contração da bola $B_\varepsilon(y_0)$ em si mesma. Nessa bola existe um único ponto fixo $y = f(x)$, isto é, um ponto verificando a condição

$$y = y - [D_2F(x_0, y_0)]^{-1}F(x, y).$$

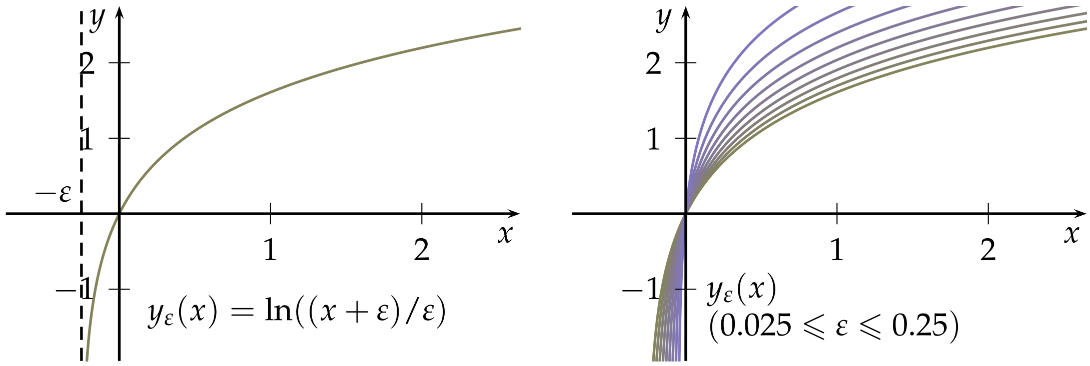
A aplicação f definida acima é a aplicação requerida. De fato, já temos $F(x, y) = 0$ pela hipótese 3 do teorema. Da unicidade do ponto fixo garantida pelo princípio das aplicações contractantes, resulta que $f(x_0) = y_0$. Finalmente, a continuidade de f é consequência do fato de que podemos escolher o valor de $\varepsilon > 0$ arbitrariamente pequeno. Isto conclui a demonstração do teorema. □

7.3 Observação Pela demonstração acima, concluímos que o Teorema 7.1 da função implícita é, essencialmente, uma consequência do Teorema 5.3 do ponto fixo para contrações. △

Conforme observa Chilov em [2], o Teorema 7.1 da função implícita possui um caráter local; isto significa que, dado o ponto (x_0, y_0) tal que $F(x_0, y_0) = 0$, a existência da função $y = f(x)$ representando uma solução da equação $F(x, y) = 0$ é garantida pelo teorema apenas em uma vizinhança do ponto x_0. Para determinar de forma precisa o domínio de existência da função f, consideremos a seguinte situação. Suponhamos que a função $F(x, y)$ seja definida, contínua e derivável em relação à variável y para quaisquer $x \in X$ e $y \in Y$, e que $D_2F(x, y)$ seja definida, contínua e inversível em todo o domínio. Dada a equação $F(x, y) = 0$, a função implícita $f(x)$ (cuja existência é garantida na bola $|x - x_0| < \delta$ pelo teorema da função implicita), pode ser definida não para todo $x \in X$, mas para todos os pontos de uma bola $|x - x_0| < r$ em que $r > 0$. Entretanto, o número r não é fixo para todo x, mas depende da função $F(x, y)$.

7.4 Exemplo Para todo $\varepsilon > 0$ consideremos a função $F_\varepsilon \colon \mathbb{R}^2 \to \mathbb{R}$ definida por

$$F_\varepsilon(x, y) = x + \varepsilon - \varepsilon e^y.$$



Essa função é contínua e derivável em relação à variável y para todos os pontos $(x, y) \in \mathbb{R}^2$; além disso, essa derivada é contínua e inversível para todos os pontos $(x, y) \in \mathbb{R}^2$; também vale a relação $F_\varepsilon(0, 0) = 0$. Entretanto, o intervalo de existência da função implícita é tal que $x > -\varepsilon$, pois nesse caso temos que $y = y_\varepsilon(x) = \ln((x + \varepsilon)/\varepsilon)$. △

No que diz respeito à unicidade da função implícita, a situação é bem mais satisfatória. De fato, suponhamos que X seja um espaço métrico conexo, isto é, X não admite subconjuntos próprios que sejam simultaneamente abertos e fechados. Sejam $f_1, f_2 \colon X \to Y$ duas funções contínuas tais que $f_1(a) = f_2(a)$ e com $F(x, f_1(x)) = 0$ e $F(x, f_2(x)) = 0$. Se as hipóteses do teorema da função implícita são verificadas em todos os pontos $(x, f_2(x)) \in X \times Y$, então temos $f_1(x) = f_2(x)$ para todo $x \in X$.

Com efeito, seja $B = \{x \in X \colon f_1(x) = f_2(x)\}$. Claramente temos que o conjunto B é fechado. Também temos que o conjunto B é aberto, já que pelo teorema da função implícita, B contém um vizinhança de todos os seus pontos. Como $f_1(x_0) = f_2(x_0)$, resulta que B é não vazio; e como X é um espaço métrico conexo, obtemos finalmente que $B = X$, e isto implica que $f_1 = f_2$.

Observamos que se supusermos que a aplicação F é contínua não apenas em (x_0, y_0) mas em toda a vizinhança U, então podemos demonstrar que a aplicação f, definida implicitamente pela fórmula $F(x, y) = 0$, é contínua em uma vizinhança de x.

8 Diferenciabilidade da função implícita

Para garantir a diferenciabilidade da aplicação f de que fala o teorema da função implícita devemos impor novas restrições à função F. Este é o assunto do próximo teorema.

8.1 Teorema *Suponhamos que sejam válidas as hipóteses do Teorema 7.1. Suponhamos ainda que no subconjunto $U \subset X \times Y$ existe a derivada parcial $D_1F(x,y)$ e que esta seja contínua em $x_0, y_0)$. Então a aplicação f definida implicitamente pela fórmula $F(x,y) = 0$ é diferenciável em x_0 e vale a relação*

$$Df(x_0) = -[D_2F(x_0,y_0)]^{-1}DF_1(x_0,y_0). \tag{19}$$

DEMONSTRAÇÃO. Seja $L\colon X \to Y$ o operador definido pelo lado direito da fórmula (19), isto é,

$$L \equiv -[D_2F(x_0,y_0)]^{-1}DF_1(x_0,y_0).$$

Para demonstrar que o operador L é a derivada da aplicação f em x_0 é suficiente demonstrar que, dado qualquer $\varepsilon > 0$, existe $\delta > 0$ tal que, para todo $x \in B_\delta(x_0)$ vale a desigualdade

$$\|f(x) - f(x_0) - L(x - x_0)\| < \varepsilon\|x - x_0\|. \tag{20}$$

Denotando $y = f(x)$, substituindo $y_0 = f(x_0)$ e usando a definição do operador L, temos

$$\begin{aligned} f(x) - f(x_0) - L(x - x_0) &= y - y_0 + [D_2F(x_0,y_0)]^{-1}D_1F(x_0,y_0)(x - x_0) \\ &= [D_2F(x_0,y_0)]^{-1} \\ &\qquad \times \{D_1F(x_0,y_0)(x - x_0) + D_2F(x_0,y_0)(y - y_0)\}. \end{aligned}$$

Como $F(x,y) = F(x_0,y_0) = 0$, pela fórmula dos acréscimos finitos obtemos a estimativa

$$\begin{aligned} &\|f(x) - f(x_0) - L(x - x_0)\| \\ &\quad \leqslant \|[D_2F(x_0,y_0)]^{-1}\| \\ &\qquad \times \|\{F(x,y) - F(x_0,y_0) - D_1F(x_0,y_0)(x - x_0) - D_2F(x_0,y_0)(y - y_0)\}\| \\ &\quad \leqslant \|[D_2F(x_0,y_0)]^{-1}\| \\ &\qquad \times \Big[\sup_{\theta_1,\theta_2 \in [1,0]} \|D_1F(x_0 + \theta_1(x - x_0), y_0 + \theta_2(y - y_0))\| \cdot \|x - x_0\| \\ &\qquad + \sup_{\theta_1,\theta_2 \in [1,0]} \|D_1F(x_0 + \theta_1(x - x_0), y_0 + \theta_2(y - y_0)) - D_2F(x_0,y_0)\| \cdot \|y - y_0\| \Big] \\ &\quad \leqslant \eta\,[\|x - x_0\| + \|y - y_0\|], \end{aligned}$$

em que η torna-se arbitrariamente pequena se escolhermos δ convenientemente pequeno (em decorrência da continuidade das derivadas parciais $D_1F(x_0,y_0)$ e $D_2F(x_0,y_0)$).

Dessa forma, temos

$$\begin{aligned} \|f(x) - f(x_0) - L(x - x_0)\| &\leqslant \eta\,[\|x - x_0\| + \|f(x) - f(x_0)\|] \\ &\leqslant \eta\,[\|x - x_0\| + \|L(x - x_0)\|] \\ &\qquad + \eta\|f(x) - f(x_0) - L(x - x_0)\|. \end{aligned}$$

Portanto, se η for suficientemente pequeno, obtemos

$$\|f(x) - f(x_0) - L(x - x_0)\| \leqslant \eta(1 - \eta)^{-1}(1 + \|L\|)\|x - x_0\|. \tag{21}$$

Para demonstrar a desigualdade (20), basta escolher η de modo que $\eta(1 - \eta)^{-1}(1 + \|L\|) \leqslant \varepsilon$, e isto conclui a demonstração do teorema. $\square$

9 Aplicações do teorema da função implícita

Como primeira aplicação do Teorema 8.1 sobre a diferenciabilidade da função implícita, apresentamos o teorema da função inversa.

9.1 Teorema *Sejam X e Y espaços de Banach e seja $f\colon Y \to X$ uma função derivável em uma vizinhança de $y_0 \in Y$, com $f(y_0) = x_0$, e tal que o operador $f'(y)$ seja contínuo e inversível em y_0. Então existem uma vizinhança $V_\delta = \{x \in X\colon |x - x_0| < \delta\}$ e uma função derivável $g\colon V_\delta \to Y$ denotada por $y = g(x)$ tal que $g(f(y)) = y$ para todo $y \in V_\delta$. Além disso, o operador $g'\colon X \to Y$ é o inverso do operador $f'\colon Y \to X$, isto é,*

$$g'(x) = [f'(y)]^{-1}.$$

DEMONSTRAÇÃO. Definimos a aplicação $F\colon X \times Y \to Y$ pela fórmula

$$F(x,y) = x - f(y)$$

e observamos que $D_1F(x,y)$ é o operador identidade e que $D_2F(x,y) = -f'(y)$. Dessa forma, podemos aplicar o Teorema 7.1 da função implícita para garantir a existência da função implicita g bem como oTeorema 8.1 sobre a diferenciabilidade da função implícita para obtermos as correspondentes propriedades de diferenciabilidade de g. □

No caso de funções numéricas $F\colon \mathbb{R}^2 \to \mathbb{R}$, o valor da derivada da função implícita $f\colon \mathbb{R} \to \mathbb{R}$ pode ser calculado a partir da fórmula

$$f'(x) = -[D_2F(x,y(x))]^{-1}D_1F(x,y(x)). \tag{22}$$

Além disso, aplicando aos dois lados da equação (22) o operador $D_2F(x,y(x))$, obtemos a fórmula

$$D_1F(x,y) + D_2F(x,y)y'(x) = 0. \tag{23}$$

Dessa forma, sob as hipóteses do Teorema 8.1, podemos determinar $y'(x)$ *derivando implicitamente* a igualdade $F(x,y) = 0$.

9.2 Exemplo Consideremos a curva definida implicitamente pela equação

$$x^3 + y^3 - 2xy = 0, \tag{24}$$

denominada folium de Descartes. Esta curva contém o ponto $A = (1,1)$ e para determinar a equação da reta tangente ao gráfico do folium nesse ponto definimos a função $F\colon \mathbb{R}^2 \to \mathbb{R}$ por $F(x,y) = x^3 + y^3 - 2xy$. Claramente, temos $F(1,1) = 0$ e F é diferenciável, com derivadas parciais dadas por

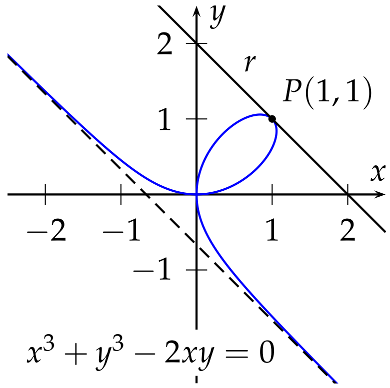
$$\frac{\partial F}{\partial x}(x,y) = 3x^2 - 2y \qquad \frac{\partial F}{\partial y}(x,y) = 3y^2 - 2x$$
$$\frac{\partial F}{\partial x}(1,1) = 1 \qquad \frac{\partial F}{\partial y}(1,1) = 1 \neq 0.$$

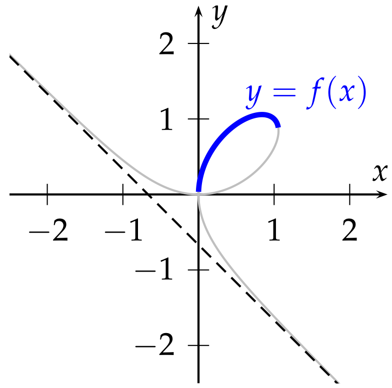
Dessa forma, podemos aplicar o Teorema 8.1 sobre a diferenciabilidade da função implícita que garante a existência de uma função $y = f(x)$, solução da equação $F(x,y(x)) = 0$ em uma vizinhança do ponto $A_x = 1$. Além disso, essa função é diferenciável e temos

$$f'(x) = -\Big[\frac{\partial F(x,y)}{\partial y}\Big]^{-1}\frac{\partial F(x,y)}{\partial y}; \qquad f'(1) = -1.$$

Consequentemente, a equação da reta r tangente ao folium de Descartes no ponto A tem a forma

$$y - 1 = -1(x - 1).$$





Uma vez justificada a fórmula de diferenciação implícita, podemos determinar esse resultado derivando diretamente a equação (24) em relação à variável x e obter

$$3x^2 + 3y^2\frac{\mathrm{d}y}{\mathrm{d}x} - 2y - 2x\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

e substituindo as coordenadas do ponto A, obtemos $\dfrac{\mathrm{d}y}{\mathrm{d}x} = -1$.

Esse raciocínio deixa de ser válido na origem do sistema de coordenadas, já que no ponto $O = (0,0)$ temos $\dfrac{\partial F(0,0)}{\partial y} = 0$. Com efeito, este é um ponto comum aos dois ramos do folium. △

9.3 Exemplo Consideremos a superfície definida implicitamente pela equação

$$x^3 + y^3 + z^3 - 6xyz = 0. \tag{25}$$

Esta superfície contém o ponto $P = (1,2,3)$e para determinar a equação do plano tangente ao gráfico da superfície nesse ponto definimos a função $F\colon \mathbb{R}^3 \to \mathbb{R}$ por $F(x,y,z) = x^3 + y^3 + z^3 - 6xyz$. É evidente que $F(1,2,3) = 0$ e que

$$\frac{\partial F}{\partial x}(x,y,z) = 3x^2 - 6yz \qquad \frac{\partial F}{\partial y}(x,y,z) = 3y^2 - 6xz \qquad \frac{\partial F}{\partial z}(x,y,z) = 3z^2 - 6xy$$
$$\frac{\partial F}{\partial x}(1,2,3) = -33 \qquad \frac{\partial F}{\partial y}(1,2,3) = -6 \qquad \frac{\partial F}{\partial y}(1,2,3) = 15 \neq 0.$$

Dessa forma, podemos aplicar o Teorema 7.1 da função implícita que garante a existência de uma função $z = f(x,y)$, solução da equação $F(x,y,z(x,y)) = 0$ em uma vizinhança do ponto $P_{xy} = (1,2)$ e tal que $f(1,2) = 3$. Além disso, a função $f\colon \mathbb{R}^2 \to \mathbb{R}$ é diferenciável e temos

$$\begin{aligned}
\operatorname{grad} f(x,y) &= \Big(\frac{\partial f}{\partial x}(x,y,z), \frac{\partial f}{\partial y}(x,y,z)\Big) \\
&= \Big(-\Big[\frac{\partial F(x,y,z)}{\partial x}\Big]^{-1}\frac{\partial F(x,y,z)}{\partial z}, -\Big[\frac{\partial F(x,y,z)}{\partial y}\Big]^{-1}\frac{\partial F(x,y,z)}{\partial z}\Big) \\
\operatorname{grad} f(1,2) &= \Big(\frac{33}{15}, \frac{6}{15}\Big)
\end{aligned}$$

Consequentemente, o plano de equação

$$z - 3 = \frac{33}{15}(x-1) + \frac{6}{15}(y-2).$$

é tangente ao gráfico da superfície no ponto $P = (1, 2, 3)$. $\triangle$

9.4 Exemplo Consideremos a curva definida pelo sistema de equações implícitas

$$x^2 - yz = 0; \qquad 3x^3 - y - 2z = 0. \tag{26}$$

Esta curva contém o ponto $C = (1, 1, 1)$ e para determinar a equação da reta tangente ao gráfico da curva nesse ponto definimos a função $F\colon \mathbb{R}^3 \to \mathbb{R}^2$ por $F(x, y, z) = (u, v)$ em que

$$u = x^2 - yz; \qquad v = 3x^2 - y - 2z.$$

Claramente temos que $F(1, 1, 1) = (0, 0)$ e que

$$\frac{\partial F(x,y,z)}{\partial(y,z)} = \begin{bmatrix} \dfrac{\partial u(x,y,z)}{\partial y} & \dfrac{\partial u(x,y,z)}{\partial z} \\ \dfrac{\partial v(x,y,z)}{\partial y} & \dfrac{\partial v(x,y,z)}{\partial z} \end{bmatrix} = \begin{bmatrix} -z & -y \\ -1 & -2 \end{bmatrix}.$$

Dessa forma, obtemos

$$\frac{\partial F(1,1,1)}{\partial(y,z)} = \begin{bmatrix} -1 & -1 \\ -1 & -2 \end{bmatrix} \qquad \text{e} \qquad \Big[\frac{\partial F(1,1,1)}{\partial(y,z)}\Big]^{-1} = \begin{bmatrix} -2 & +1 \\ +1 & -1 \end{bmatrix}$$

Como a matriz $\partial F(1, 1, 1)/\partial(y, z)$ é inversível, podemos aplicar o teorema 7.1 da função implícita e garantir que existe uma solução $y = y(x)$, $z = z(x)$ do sistema $u = 0$, $v = 0$, para a qual $y(1) = 1$ e $z|(1) = 1$. Além disso,

$$\frac{\partial F(x,y,z)}{\partial x} = \begin{bmatrix} 2x \\ 9x^2 \end{bmatrix} \qquad \frac{\partial F(1,1,1)}{\partial x} = \begin{bmatrix} 2 \\ 9 \end{bmatrix}$$

de modo que

$$\begin{aligned} (y'(1), z'(1)) &= -\Big[\frac{\partial F(1,1,1)}{\partial(y,z)}\Big]^{-1} \cdot \frac{\partial F(1,1,1)}{\partial x} \\ &= -\begin{bmatrix} -2 & +1 \\ +1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 9 \end{bmatrix} = \begin{bmatrix} -5 \\ 7 \end{bmatrix}. \end{aligned}$$

Consequentemente, a reta definida pela interseção dos planos

$$y - 1 = -5(x-1) \qquad \text{e} \qquad z - 1 = 7(x-1)$$

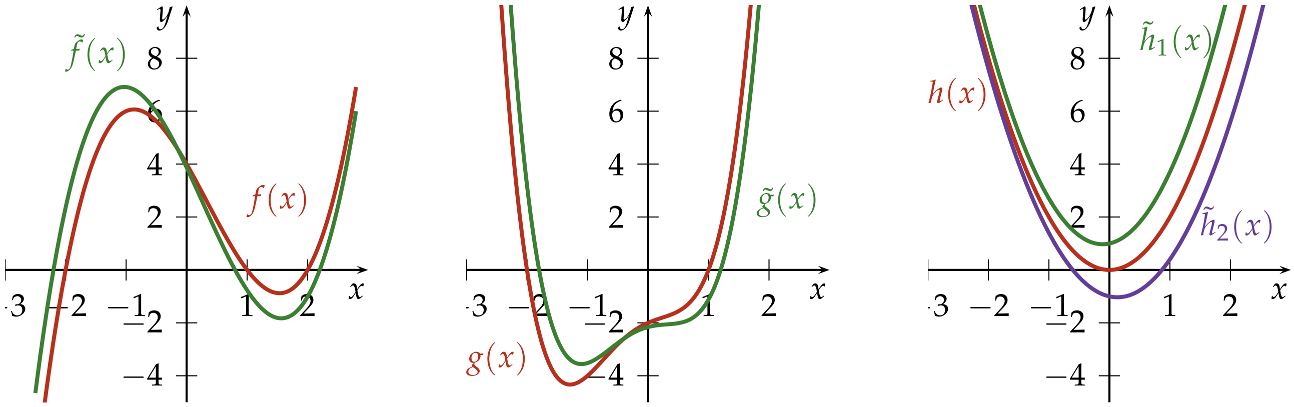
é a reta tangente à curva no ponto C. $\triangle$

9.5 Exemplo Consideremos um polinômio algébrico dado por

$$p(z) = z^n - a_1 z^{n-1} + a_2 z^{n-2} - \cdots + (-1)^n a_n \tag{27}$$

com coeficientes complexos $a_1, a_2, \ldots, a_n \in \mathbb{C}$. O teorema fundamental da álgebra afirma que $p(z)$ possui n raízes $\lambda_1, \lambda_2, \ldots, \lambda_n$, contadas as multiplicidades. E essas raízes são funções contínuas e deriváveis de seus coeficientes. De fato, a todo conjunto de coeficientes $a = (a_1, a_2, \ldots, a_n) \in \mathbb{C}^n$ corresponde um único conjunto de raízes $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{C}^n$, podemos definir uma função $\lambda \colon \mathbb{C}^n \to \mathbb{C}^n$ por $\lambda = \lambda(a)$. Mostraremos que esta função é contínua e derivável para todo elemento $a \in \mathbb{C}^n$ para o qual as raízes são duas a duas distintas.

As figuras a seguir ilustram os casos dos polinômios $f(x) = x^3 - x^2 - 4x + 4$ e $\tilde{f}(x) = x^3 - 0.8x^2 - 4.84x + 3.872$ bem como $g(x) = x^4 + x^3 - 3x^2 - x + 2$ e $\tilde{g}(x) = x^4 + 0.6x^3 - 1.16x^2 + 0.6x - 2.16$. A hipótese de que as raízes sejam distintas é importante, como mostra o caso em que $h(x) = 2x^2$, $\tilde{h}_1(x) = 2.2x^2 + 0.5x + 1$ e $\tilde{h}_2(x) = 1.9x^2 - 0.5x - 1$. Nessa situação $h(x)$ possui uma raiz dupla, $h_1(x)$ não possui raízes reais e $h_2(x)$ possui duas raízes distintas.



Os coeficientes do polinômio (27) são relacionados às suas raízes pelas relações de Viète-Girard, a saber,

$$
\begin{aligned}
a_1 &= \lambda_1 + \lambda_2 + \cdots + \lambda_n \\
a_2 &= \lambda_1\lambda_2 + \lambda_1\lambda_3 + \cdots + \lambda_{n-1}\lambda_n \\
a_3 &= \lambda_1\lambda_2\lambda_3 + \cdots + \lambda_{n-2}\lambda_{n-1}\lambda_n \\
\vdots &= \vdots \\
a_n &= \lambda_1\lambda_2 \ldots \lambda_n.
\end{aligned} \tag{28}
$$

Além disso, a função obtida $a = a(\lambda)$ é contínua e derivável em todos os pontos $\lambda \in \mathbb{C}^n$. Denotando a matriz jacobiana complexa do sistema (30) por $\dfrac{\partial(a_1, a_2, \ldots, a_n)}{\partial(\lambda_1, \lambda_2, \ldots, \lambda_n)}$, então temos que o operador diferencial

$$
\mathrm{d}a(\lambda) = \frac{\partial(a_1, a_2, \ldots, a_n)}{\partial(\lambda_1, \lambda_2, \ldots, \lambda_n)} \, \mathrm{d}\lambda \tag{29}
$$

é inversível desde que

$$
\det \frac{\partial(a_1, a_2, \ldots, a_n)}{\partial(\lambda_1, \lambda_2, \ldots, \lambda_n)} \neq 0.
$$

Nesse caso, o Teorema 9.1 pode ser aplicado para garantir a existência, a continuidade e a diferenciabilidade da função inversa $\lambda = \lambda(a)$. Portanto, devemos mostrar apenas que o determinante da matriz jacobiana é não nulo, desde que os números $\lambda_1, \lambda_2, \ldots, \lambda_n$ sejam dois a dois distintos.

A demonstração desse fato se faz usando o princípio de indução finita no grau do polinômio. A afirmativa é verdadeira no caso em que o grau do polinômio é $m = 1$, pois $\dfrac{\partial a_1}{\partial \lambda_1} = 1$. Suponhamos que a afirmativa seja verdadeira no caso em que o grau do polinômio é $m = n - 1$. Agora formamos as relações

$$\begin{aligned}
b_1 &= \lambda_1 + \lambda_2 + \cdots + \lambda_{n-1} \\
b_2 &= \lambda_1\lambda_2 + \lambda_1\lambda_3 + \cdots + \lambda_{n-2}\lambda_{n-1} \\
b_3 &= \lambda_1\lambda_2\lambda_3 + \ldots \lambda_{n-3}\lambda_{n-2}\lambda_{n-1} \\
\vdots &= \vdots \\
b_n &= \lambda_1\lambda_2 \ldots \lambda_{n-1}.
\end{aligned} \tag{30}$$

e definimos o polinômio

$$q(z) = z^{n-1} - b_1 z^{n-2} + b_2 z^{n-3} - \cdots + (-1)^{n-1} b_{n-1}. \tag{31}$$

Usando novamente as relações de Viète-Girard, temos que os números $\lambda_1, \lambda_2, \ldots, \lambda_{n-1}$ são as raízes do polinômio (31); pela hipótese de indução temos que

$$\det \frac{\partial(b_1, b_2, \ldots, b_{n-1})}{\partial(\lambda_1, \lambda_2, \ldots, \lambda_{n-1})} \neq 0.$$

Portanto, completanto as fórmulas do sistema (30) com a equação adicional $\lambda_n = \lambda_n$, obtemos a relação

$$\det \frac{\partial(b_1, b_2, \ldots, b_{n-1}, \lambda_n)}{\partial(\lambda_1, \lambda_2, \ldots, \lambda_{n-1}, \lambda_n)} \neq 0. \tag{32}$$

Usando as relações do sistema (30), as fórmulas do sistema (28) podem ser reescritas como

$$\begin{aligned}
a_1 &= b_1 + \lambda_1 \\
a_2 &= b_2 + b_1\lambda_n \\
a_3 &= b_3 + b_2\lambda_n \\
\vdots &= \vdots \\
a_n &= b_{n-1}\lambda_n.
\end{aligned}$$

Dessa forma, temos que a matriz jacobiana $\dfrac{\partial(a_1, a_2, \ldots, a_n)}{\partial(b_1, b_2, \ldots, b_{n-1}, \lambda_n)}$ tem a forma

$$\frac{\partial(a_1, a_2, \ldots, a_n)}{\partial(b_1, b_2, \ldots, b_{n-1}, \lambda_n)} = \begin{bmatrix}
1 & & & \ldots & & & 1 \\
\lambda_n & 1 & & \ldots & & & b_1 \\
& \lambda_n & 1 & \ldots & & & b_2 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
& & & \ldots & \lambda_n & 1 & b_{n-2} \\
& & & \ldots & & \lambda_n & b_{n-1}
\end{bmatrix}$$

Para calcular o determinante dessa matriz, faremos a seguinte sequência de transformações: a partir da segunda linha, substituímos cada linha pela sua soma com a linha precedente multiplicada por λ_n. Dessa forma,

$$\det\frac{\partial(a_1,a_2,\dots,a_n)}{\partial(b_1,b_2,\dots,b_{n-1},\lambda_n)} = \begin{vmatrix} 1 & & & \dots & & & 1 \\ 0 & 1 & & \dots & & & b_1-\lambda_n \\ & 0 & 1 & \dots & & & b_2-b_1\lambda_n+\lambda_n^2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ & & & \dots & 0 & 1 & \vdots \\ & & & \dots & & 0 & b_{n-1}-b_{n-2}\lambda_n+\dots+(-1)^{n-1}\lambda_n^{n-1} \end{vmatrix}$$
$$= (-1)^{n-1}[\lambda_n^{n-1}-b_1\lambda_n^{n-2}+\dots+(-1)^{n-1}b_{n-1}]$$
$$= q(\lambda_n)\neq 0$$

já que, por hipótese, λ_n é distinto de cada um dos números λ_1, λ_2, ..., λ_{n-1} e, consequentemente, não é raiz do polinômio (31). Assim, em vista da desigualdade (32), obtemos a desigualdade

$$\det\frac{\partial(a_1,a_2,\dots,a_n)}{\partial(\lambda_1,\lambda_2,\dots\lambda_n)} = \det\frac{\partial(a_1,a_2,\dots,a_n)}{\partial(b_1,b_2,\dots b_{n-1},\lambda_n)}\det\frac{\partial(b_1,b_2,\dots,b_{n-1},\lambda_n)}{\partial(\lambda_1,\lambda_2,\dots\lambda_{n-1},\lambda_n)}\neq 0$$

que é a relação requerida. △

10 Equações diferenciais ordinárias de primeira ordem

O estudo de equações diferenciais originou-se da tentativa de descrever matematicamente uma ampla classe de problemas em várias ciências. Por exemplo, o estudo do crescimento de populações, ou do movimento de partículas, conduzem a certas equações diferenciais juntamente com variadas condições iniciais. O problema matemático, comumente denominado problema de Cauchy, é o de determinar uma função diferenciável verificando a equação e a condição inicial.

Mais especificamente, sejam X um espaço de Banach e $U\subset X$ um subconjunto aberto; sejam $I=(a,b)\subset\mathbb{R}$ e $f\in C(I\times U;X)$ uma função que pode ser interpretada como um campo vetorial dependente do tempo e definido em U. O problema de Cauchy é

$$\frac{\mathrm{d}\phi}{\mathrm{d}t}(x,t)=f(t,\phi(t,x))\qquad \phi(x,t_0)=x_0, \tag{33}$$

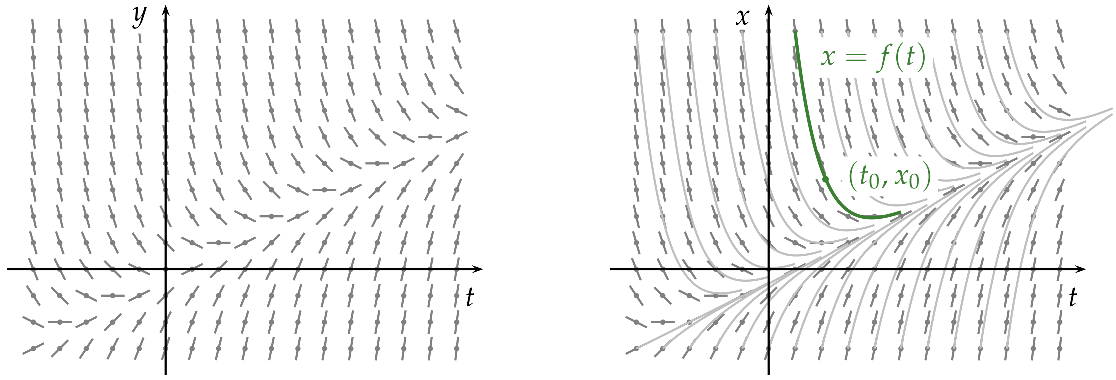
em que $P_0=(t_0,x_0)\in I\times U$ e f é uma função contínua.

O leitor pode verificar que qualquer solução $\phi\in C^1(I;\mathbb{R})$ do problema de Cauchy (33) é uma solução $\phi\in C(I;\mathbb{R})$ da equação integral

$$\phi(t)=x_0+\int_{t_0}^{t} f(t,\phi(\tau,x(\tau)))\,\mathrm{d}\tau. \tag{34}$$

Reciprocamente, se $\phi\in C(I;\mathbb{R})$ é solução da equação (34), então $\phi\in C^1(I/;\mathbb{R})$ e ϕ é solução do problema (33).

A interpretação geométrica da equação diferencial do problema de Cauchy é a seguinte. Por cada ponto $(t,x)\in I\times U$ desenhamos um pequeno segmento de reta cuja inclinação vale $f(t,x)$. A coleção desses segmentos é uma representação do campo de direções correspondente à equação diferencial $[\mathrm{d}\phi/\mathrm{d}t](x,t)=f(t,\phi(t,x))$. Uma solução da equação diferencial do problema de Cauchy é, portanto, uma curva que, em cada um de seus pontos, tem como reta tangente o correspondente segmento do campo de direções. Uma solução do problema de Cauchy é uma dessas curvas que contém o ponto inicial $P_0(t_0, x_0)$.



10.1 Observação A hipótese sobre a continuidade da função f muitas vezes é substituída pela condição de Lipschitz na variável x, isto é, a de que existe uma constante K tal que

$$|f(t, x_1) - f(t, x_2)| \leqslant K|x_1 - x_2|$$

para quaisquer $x_1, x_2 \in \mathbb{R}$. Com essa hipótese, podemos garantir não apenas a existência mas também a unicidade da solução para o problema de Cauchy. △

10.2 Observação Por simplicidade, normalmente supomos que a condição inicial para o problema de Cauchy (33) é considerada em $t_0 = 0$. De fato, não há perda de generalidade ao utilizarmos esta suposição, já que se $\tilde{\phi}$ é solução do problema de Cauchy

$$\frac{\mathrm{d}\tilde{\phi}}{\mathrm{d}t}(x, t) = \tilde{f}(t, \tilde{\phi}(t, x)) \qquad \tilde{\phi}(x, 0) = x_0,$$

então $\phi(t) \equiv \tilde{\phi}(t - t_0)$ é solução do problema (33) em que $f(t, x) \equiv \tilde{f}(t - t_0, x)$. △

10.3 Exemplo Sejam $X = \mathbb{R}$ e $f(t, x) = x^n$ em que $n \in \mathbb{N}$ e consideremos o problema de Cauchy

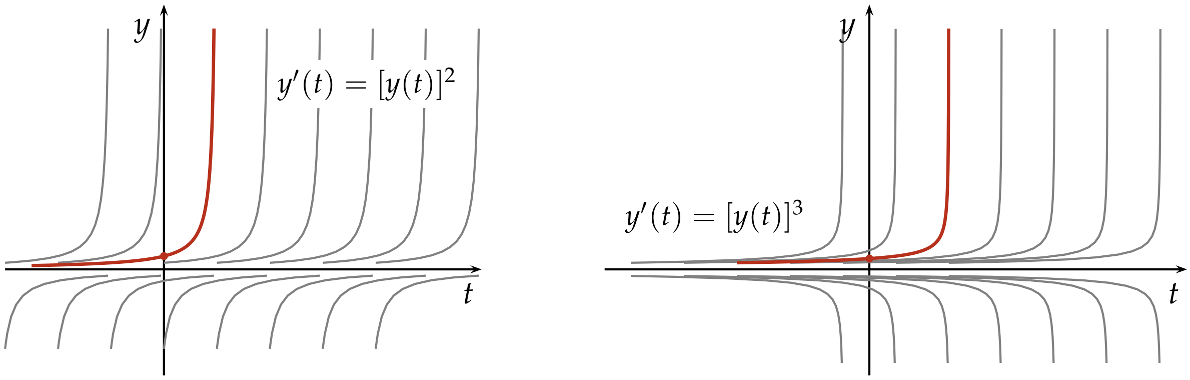
$$\frac{\mathrm{d}y}{\mathrm{d}t}(t) = [y(t)]^n \qquad y(0) = y_0 \in \mathbb{R}. \tag{35}$$

Se a função y é solução do problema (35) com $y_0 \neq 0$, então $y(t)$ é não nula para t próximo de zero. Portanto, pelo menos até o ponto em que possivelmente $y(t)$ se anula, devemos ter

$$\begin{aligned} t &= \int_0^t \frac{1}{[y(\tau)]^n} \frac{\mathrm{d}y(\tau)}{\mathrm{d}\tau} \,\mathrm{d}\tau = \int_{y(0)}^{y(t)} u^{-n} \,\mathrm{d}u \\ &= \begin{cases} \ln\left|\dfrac{y(t)}{y_0}\right| & \text{se } n = 1 \\ \dfrac{[y(t)]^{n-1} - [y_0]^{n-1}}{1 - n} & \text{se } n > 1. \end{cases} \end{aligned}$$

Resolvendo para $y(t)$, obtemos

$$y(t) = y(t, y_0) = \begin{cases} y_0 \,\mathrm{e}^t & \text{se } n = 1 \\ \dfrac{y_0}{[1 - (n-1)t y_0^{n-1}]^{n-1}} & \text{se } n > 1. \end{cases} \tag{36}$$



Efetuando os cálculos diretamente, que $y(t, y_0)$ definida acima de fato é solução do problema de Cauchy (35). Os argumentos garantem que essas são as únicas soluções possíveis. Observamos ainda que no caso $n = 1$ não há restrição para os valores de t; entretanto, para $n > 1$ devemos impor a restrição $1 - (n-1)ty_0^{n-1} > 0$, ou equivalentemente,

$$
\begin{aligned}
t &< +\frac{1}{(n-1)y_0^{n-1}} \qquad \text{se } y_0^{n-1} > 0,\\
t &> -\frac{1}{(n-1)|y_0|^{n-1}} \qquad \text{se } y_0^{n-1} < 0.
\end{aligned}
$$

Além disso, para $n > 1$ a solução tem uma assíntota vertical em $t_a = y_0^{1-n}/(n-1)$.

Finalmente, para s e t suficientemente próximos de zero, vale a relação

$$y(t, y(s, y_0)) = y(t + s, x).$$

Em particular, no caso $n = 1$ esta propriedade é equivalente à identidade $\mathrm{e}^t\mathrm{e}^s = \mathrm{e}^{t+s}$. △

A condição de Lipschitz é importante para garantirmos a unicidade da solução do problema de Cauchy, conforme ilustra o exemplo a seguir.

10.4 Exemplo Sejam $X = \mathbb{R}$ e $f(t, x) = x^\alpha$ em que $\alpha \in \mathbb{R}$ é tal que $0 < \alpha < 1$ e consideremos o problema de Cauchy

$$\frac{\mathrm{d}y}{\mathrm{d}t}(t) = |y(t)|^\alpha \qquad y(0) = y_0. \tag{37}$$

Realizando cálculos semelhantes aos do Exemplo 10.3, para $y_0 \neq 0$ temos

$$
\begin{aligned}
t &= \int_0^t \frac{1}{[y(\tau)]^\alpha} \frac{\mathrm{d}y(\tau)}{\mathrm{d}\tau} \,\mathrm{d}\tau = \int_{y(0)}^{y(t)} u^{-\alpha} \,\mathrm{d}u\\
&= \frac{[y(t)]^{1-\alpha} - [y_0]^{1-\alpha}}{1-\alpha}
\end{aligned}
$$

em que, por simplicidade, usamos a notação $u^{1-\alpha} \equiv u|u|^{-\alpha}$. Como $y(t)$ e y_0 têm o mesmo sinal (que denotaremos por $\sigma(y(t)) = \sigma(y_0)$), a equação anterior implica que

$$
\begin{aligned}
\sigma(y_0)(1-\alpha)t &= \sigma(y_0)\left[\sigma(y(t))|y(t)|^{1-\alpha} - \sigma(y_0)|y_0|^{1-\alpha}\right]\\
&= |y(t)|^{1-\alpha} - |y_0|^{1-\alpha}
\end{aligned}
$$

e, portanto,

$$y(t, y_0) = \sigma(x)\left[|x|^{1-\alpha} + \sigma(y_0)(1-\alpha)t\right]^{1/(1-\alpha)}. \tag{38}$$

Essa solução é determinada univocamente até o primeiro valor de t para o qual $|y_0|^{1-\alpha} + \sigma(y_0)(1-\alpha)t = 0$.

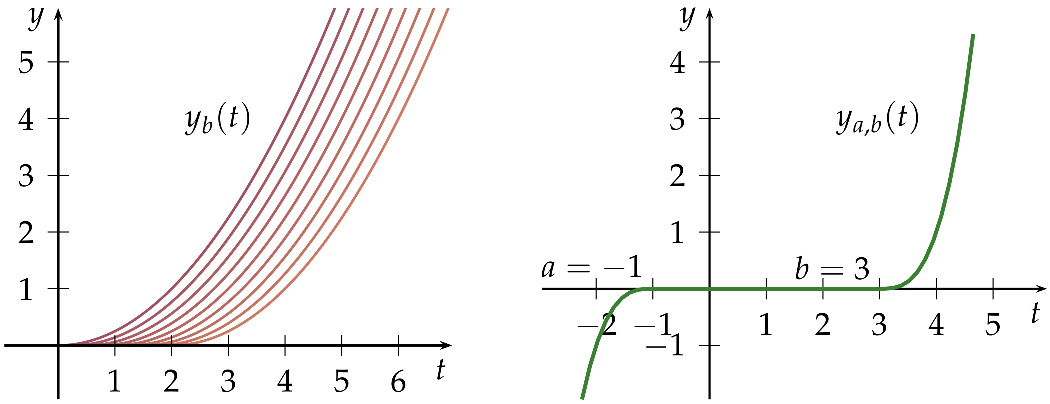
Claramente, $y(t) \equiv 0$ é uma solução do problema de Cauchy (37) quando $y_0 = 0$; entretanto, tal solução não é única, por fazendo y_0 tender a zero por valores positivos na equação (38) obtemos a função

$$y(t, 0^+) = [(1-\alpha)t]^{1/(1-\alpha)}$$

que resolve o problema (37) para $t > 0$. Além disso, se definimos a função

$$y(t) \equiv \begin{cases} 0 & \text{se } t \leqslant 0, \\ [(1-\alpha)t]^{1/(1-\alpha)} & \text{se } t > 0, \end{cases} \tag{39}$$

então podemos verificar prontamente que $y(t)$ é solução do problema de Cauchy (37). É fácil verificar que a família de funções definidas por $y_b(t) \equiv y(t-b)$ também é solução do problema de Cauchy (37) para $b \geqslant 0$.



Para o problema $\mathrm{d}y(t)/\mathrm{d}t = 3y^{2/3}$, obtemos a família de soluções $y_{a,b}(t) = (t-a)^3$ para $t \leqslant a$, $y_{a,b}(t) = 0$ para $a \leqslant t \leqslant b$ e $y_{a,b}(t) = (t-b)^3$ para $b \leqslant t$ em que $a, b \in \mathbb{R}$ são arbitrários. O caso em que $a = -1$ e $b = 3$ é ilustrado acima. △

11 Teorema de existência e dependência contínua da solução do problema de Cauchy

Para que certos problemas de equações diferenciais possam ter algum significado importante para as aplicações é fundamental estabelecer de que forma uma solução depende das condições iniciais. Em outros termos, um resultado de diferenciabilidade das soluções do problema de Cauchy em relação às condições iniciais fornece um método eficiente de se estudar a influência exercida na solução por uma pequena perturbação da condição inicial. Estabelecer um resultado dessa natureza é o principal objetivo destas notas. Para uma demonstração simples do resultado citamos o trabalho de Dai em [3], que utiliza a sequência de aproximações de Picard-Lindelöf e o lema de Ascoli; entretanto, nessa abordagem é utilizado também o resultado de existência de solução para o problema de Cauchy. Por outro lado, Sotomayor em [10] demonstra de uma só vez tanto a existência quanto a diferenciabilidade da solução do problema de Cauchy em relação às condições iniciais usando o teorema de contração nas fibras. Robbin em [9] também demonstra o resultado usando o Teorema 7.1 da função implícita e é nesse último trabalho que nos baseamos para o restante desta seção.

11.1 Teorema *Seja E um espaço de Banach e seja $U \subset E$ um subconjunto aberto; seja $f\colon \mathbb{R} \times U \to E$ uma aplicação de classe $f \in C^r(\mathbb{R} \times U; E)$, em que $r \geqslant 1$. Então para cada elemento $x_0 \in U$ existem uma vizinhança aberta V de x_0, um intervalo aberto $(-\varepsilon, +\varepsilon) \subset \mathbb{R}$ e uma aplicação $\phi\colon (-\varepsilon, +\varepsilon) \times V \to U$ tais que valem as afirmativas seguintes.*

1. $\phi \in C^r((-\varepsilon, +\varepsilon) \times V; U)$.
2. $\phi(0, x) = x$ *para todo* $x \in V$.
3. $\dfrac{\mathrm{d}\phi}{\mathrm{d}t}(t, x) = f(t, \phi(t, x))$ *para* $(t, x) \in (-\varepsilon, +\varepsilon) \times V$.

DEMONSTRAÇÃO. Podemos supor, sem perda de generalidade, que o ponto x_0 é a origem do espaço de Banach E e que a vizinhança U de x_0 é uma bola de raio r centrada nesse ponto, isto é, $U = B_r(x_0)$ e seja $U_0 = B_{r/2}(x_0)$ a bola centrada em x_0 e cujo raio é a metade do raio de U. Seja $I \equiv [-1, +1] \subset \mathbb{R}$. Dado o número natural $p \in \mathbb{N}$, seja $C^p(I; E)$ o espaço de Banach das aplicações de classe C^p da forma $\gamma\colon I \to E$ com a topologia de C^p e seja $C^p_0(I; E)$ o subespaço fechado de $C^p(I; E)$ consistindo das aplicações $\gamma \in C^p(I; E)$ tais que $\gamma(0) = 0$; finalmente, seja $C^p_0(I; U_0)$ o conjunto das aplicações $\gamma \in C^p_0(I; E)$ tais que $\gamma(I) \subseteq U_0$. Notamos que $C^p_0(I; U_0)$ é um subconjunto aberto no espaço de Banach $C^p_0(I; E)$.

Seja $F\colon \mathbb{R} \times U_0 \times C^1_0(I; U_0) \to C^0(I; E)$ a aplicação definida por

$$F(a, x, \gamma)(t) \equiv \dot{\gamma}(t) - af(at, x + \gamma(t)), \tag{40}$$

em que $a \in \mathbb{R}$, $x \in U_0$ e $\gamma \in C^1_0(I; U_0)$ para $t \in I$. Assim, temos que $F \in C^1$ entre os espaços de Banach envolvidos. Esta afirmativa é consequência do Lema A.8; notamos que a aplicação $\gamma \mapsto \dot{\gamma}$ é uma aplicação linear contínua. A derivada parcial em relação a γ no ponto $a = 0$, $x = x_0$ e $\gamma = 0$ avaliada no vetor tangente $\delta \in C^1_0(I; E)$ é dada por

$$D_3F(0, x_0, 0)\delta(t) = \frac{\mathrm{d}}{\mathrm{d}t}\delta(t). \tag{41}$$

Claramente esta aplicação é um isomorfismo linear. Como $F(0, x_0, 0) = 0$, podemos aplicar o Teorema 7.1 da função implícita e dessa forma obtemos uma vizinhança $(-2\varepsilon, +2\varepsilon) \times V$ de $(0, x_0) \in \mathbb{R} \times U_0$ e uma aplicação $H\colon (-2\varepsilon, +2\varepsilon) \times V \to C^1_0(I; U_0)$ de classe C^1 tal que

$$F(a, x, H(a, x)) = 0 \tag{42}$$

para $(a, x) \in (-2\varepsilon, +2\varepsilon) \times V$. Podemos definir $\phi\colon (-\varepsilon, +\varepsilon) \times V \to U$ pela fórmula

$$\phi(t, x) \equiv H(\varepsilon, x)(t/\varepsilon) + x. \tag{43}$$

A aplicação ϕ é de classe C^1 pois a aplicação avaliação $A\colon C^1_0(I; U_0) \times I \to U_0$ é de classe C^1. Além disso, $\phi(0, x) = x$ pois $H(\varepsilon, x) \in C^1_0(I; U_0)$. Finalmente, como

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(t, x) - f(t, \phi(t, x)) = (1/\varepsilon)F(\varepsilon, x, H(\varepsilon, x))(t/\varepsilon) = 0, \tag{44}$$

resulta que ϕ é a curva solução. Isto demonstra o teorema no caso em que $r = 1$.

O caso geral segue do caso demonstrado através de um argumento de indução padrão. De fato, consideremos o espaço de Banach $B = C^{r-1}(\overline{U}; E)$ e seja a aplicação $\omega_f\colon B \to B$ definida por $\omega_f(\eta) \equiv f \circ \eta$. Esta aplicação é de classe C^1 pelo Lema A.8. Considerando ω_f como uma aplicação sobre B, existe uma única solução η_t da equação diferencial do item 3 que é de classe C^1 tal que $\eta_0 =$ identidade pelo que já foi demonstrado. Essa curva integral é o fluxo local de F e, portanto, é de classe C^{r-1} pois pertence a B. Como $r \geqslant 2$, resulta que η é pelo menos de classe C^1 e assim vemos que $D\eta_t = u_t$ verifica a equação diferencial do item 3, isto é,

$$\frac{\mathrm{d}}{\mathrm{d}t}u_t = DF(\eta_t)u_t. \tag{45}$$

Novamente pelo que já foi demonstrado, vemos que $u_t \in C^{r-1}((-\varepsilon, +\varepsilon) \times V; U)$, ou seja, $\eta_t \in C^r((-\varepsilon, +\varepsilon) \times V; U)$. Isto conclui a demonstração do teorema. $\square$

A Resultados auxiliares

Neste apêndice apresentamos os resultados utilizados na demonstração do Teorema 11.1.

A.1 Derivadas de ordens superiores

Sejam X e Y espaços de Banach e seja $F\colon X \to Y$ uma aplicação diferenciável. Para cada $x \in X$, a derivada $DF(x)$ é um elemento do espaço $L(X;Y)$ dos operadores lineares. Se esta aplicação for diferenciável, então sua derivada é denomidada derivada segunda da aplicação F e é denotada por D^2F. Assim, $D^2F(x)$ é um elemento do espaço $L(X;L(X;Y))$ dos operadores que aplicam X em $L(X;Y)$.

Os elementos desse espaço têm uma interpretação bastante conveniente que se realiza através das aplicações bilineares, que definimos a seguir.

A.1 Definição Sejam X e Y espaços de Banach. Uma aplicação $B\colon X \to Y$ é denominada aplicação bilinear se, a cada par ordenado $(u_1, v_1) \in X \times X$, associa um elemento $y \in B(u_1, v_1) \in Y$ de modo que valem as propriedades seguintes.

1. A aplicação B depende linearmente de cada uma das componentes, isto é,

$$\begin{aligned} &B(\alpha_1 u_1 + \alpha_2 u_2, v_1) = \alpha_1 B(u_1, v_1) + \alpha_2 B(u_2, v_1), \\ &B(u_1, \beta_1 v_1 + \beta_2 v_2) = \beta_1 B(u_1, v_2) + \beta_2 B(u_1, v_2). \end{aligned}$$

2. A aplicação B é contínua como função de duas variáveis, isto é, existe um número positivo $M \in \mathbb{R}^+$ tal que, para quaisquer $(u_1, v_1) \in X \times X$ vale a desigualdade

$$\|B(u_1, v_1) \leqslant M\|u_1\|\|v_1\|. \tag{46}$$

O menor dos números M verificando a desigualdade (46) é denominado norma da aplicação bilinear B e é denotado por $\|B\|$.

Usando as propriedades acima podemos verificar que as aplicações bilineares de X em Y constituem um espaço vetorial normado, denotado por $B(X \times X; Y)$. Em particular, se Y for um espaço completo, então $B(X \times X; Y)$ também é completo, conforme a Proposição 1.8.

Definimos a aplicação $T\colon L(X;L(X;Y)) \to B(X \times X; Y)$ associando um elemento $A \in L(X;L(X;Y))$ a um elemento de $B(X \times X; Y)$, através da equação

$$B(x_1, x_2) = (A(x_1))x_2. \tag{47}$$

A aplicação T é linear; mais ainda, é uma isometria entre os espaços envolvidos. De fato, dado $w = B(u,v) = (A(u))v$, temos

$$\|w\| \leqslant \|A(u)\|\|v\| \leqslant \|A\|\|u\|\|v\|,$$

e, portanto, $\|B\| \leqslant \|A\|$.

Por outro lado, dada uma aplicação bilinear B, a aplicação $x_2 \mapsto (A(x_1))x_2 = B(x_1, x_2)$ é, para cada elemento fixo $x_1 \in X$, uma aplicação linear. Assim, a cada $x_1 \in X$ associamos um elemento $A(x_1) \in L(X;Y)$. Esta correspodência é linear e, assim, a aplicação bilinear B define um elemento $A \in L(X, L(X;Y))$. Nessas condições, a aplicação B pode ser obtida de A a partir da relação (47); dessa forma,

$$\|A(x_1)\| = \sup_{\|x_2\| \leqslant 1} \|(A(x_1))x_2\| = \sup_{\|x_2\| \leqslant 1} \|B(x_1, x_2)\| \leqslant \|B\|\|x_1\|,$$

e, portanto, $\|A\| \leqslant \|B\|$.

Combinando as desigualdades dos finais dos dois parágrafos anteriores, obtemos $\|A\| = \|B\|$. Assim, a correspondência entre os espaços $L(X;L(X;Y))$ e $B(X \times X;Y)$ definida pela equação (47) é de fato uma isometria linear. A importância dessa isometria está em que a derivada segunda $D^2F(x)$ foi definida como um elemento do espaço $L(X;L(X;Y))$; e pelo que acabamos de verificar, esta derivada segunda pode ser interpretada como um elemento do espaço $B(X \times X;Y)$.

A.2 Exemplo Sejam $X = \mathbb{R}^m$ e $Y = \mathbb{R}^n$ espaços euclideanos de dimensões m e n, respectivamente. Qualquer aplicação linear $T\colon X \to Y$ representa-se por uma matriz de dimensões $n \times m$. Logo, dada a aplicação diferenciável $F\colon X \to Y$, a derivada $DF(x)$ é uma matriz. Consideremos as bases $\{e_1, e_2, \ldots, e_m\}$ de X e $\{f_1, f_2, \ldots, f_n\}$ de Y, e sejam os elementos $x \in X$ e $y \in Y$, representados respectivamente por

$$x = x_1e_1 + x_2e_2 + \cdots + x_me_m,$$
$$y = y_1f_1 + y_2f_2 + \cdots + y_nf_n.$$

Podemos escrever a aplicação $y = F(x)$ e sua derivada $DF(x)$, respectivamente, nas formas

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} F_1(x_1, x_2, \ldots, x_m), \\ F_2(x_1, x_2, \ldots, x_m), \\ \vdots \\ F_n(x_1, x_2, \ldots, x_m), \end{bmatrix} \qquad \text{e} \qquad DF(x) = \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_x} & \cdots & \dfrac{\partial y_1}{\partial x_m} \\ \dfrac{\partial y_2}{\partial x_1} & \dfrac{\partial y_2}{\partial x_x} & \cdots & \dfrac{\partial y_2}{\partial x_m} \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial y_n}{\partial x_1} & \dfrac{\partial y_n}{\partial x_x} & \cdots & \dfrac{\partial y_n}{\partial x_m} \end{bmatrix}.$$

A matriz $DF(x)$ é chamada de matriz jacobiana e frequentemente é denotada também por $\dfrac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_m)}$.

Para um elemento $x \in X$, a derivada segunda $D^2F(x)$ é definida por $a_{k,ij} = \dfrac{\partial^2 y_k}{\partial x_i \partial x_j}(x)$, em que $1 \leqslant k \leqslant n$ e $1 \leqslant i, j \leqslant m$. Esses números definem uma aplicação linear $S\colon X \to L(X;Y)$ mediante a fórmula

$$b_{k,j} = \sum_{i=1}^{i=m} a_{k,ij}x_i$$

ou uma aplicaçao bilinear $B\colon X \times X \to Y$ mediante a fórmula

$$y_k = \sum_{i=1}^{i=m} \sum_{i=1}^{i=m} a_{k,ij} u_i v_j.$$

Os conceitos de derivada terceira, derivada quarta e, de modo geral, o conceito de derivada n-ésima de uma aplicação $F\colon X \to Y$ entre espaços de Banach, podem ser desenvolvidos da seguinte forma. A n-ésima derivada de F é definida como a derivada da derivada de ordem $n-1$; será, portanto, um elemento do espaço $L(X; L(X; L(X, \ldots, L(X; Y)) \ldots))$, em que o símbolo L aparece n vezes na expressão. Argumentando de forma análoga ao caso da derivada segunda, podemos mostrar que a cada elemento desse espaço corresponde um elemento do espaço $N(X \times X \times \cdots \times X; Y) = N(X^n; Y)$ das aplicações n-lineares entre os espaços X^n e Y, que verificam propriedades semelhantes àquelas listadas para aplicações bilineares. Especificamente, dizemos que uma função $A\colon X^n \to Y$ é uma aplicação n-linear se aplica qualquer elemento $(u_1, u_2, \ldots, u_n) \in X^n$ em um elemento $y = A(u_1, u_2, \ldots, u_n) \in Y$ de modo que A é linear em relação a cada uma das componentes $u_i \in X$, os demais elementos permanecendo fixados, e que verifica, para algum número real positivo $M \in \mathbb{R}^+$ a desigualdade

$$\|A(u_1, u_2, \ldots, u_n)\| \leqslant M \|u_1\| \|u_2\| \ldots \|u_n\|,$$

para qualquer elemento $(u_1, u_2, \ldots, u_n) \in X^n$. Dessa maneira, podemos interpretar a n-ésima derivada da aplicação F, denotada por DF^n, como um elemento do espaço $N(X^n; Y)$, isto é, $D^n F(x) \in N(X^n; Y)$.

Definimos a diferencial de Fréchet de uma aplicação $F\colon X \to Y$ como o resultado da aplicação do operador derivada $DF(x)$ no ponto $x \in X$ a um elemento $h \in X$, isto é, $DF = DF(x)h$. A diferencial de segunda ordem é definida como $D^2F = D^2F(x)(h,h)$, ou seja, é a função quadrática associada à aplicação $D^2F \in B(X^2; Y)$. Analogamente, a diferencial de orden n é o elemento $D^nF(x)(h, h, \ldots, h) \in N(Z^n; Y) \in Y$, ou seja, é a imagem por $D^nF(x)$ do elemento $(h, h, \ldots, h) \in X^n$.

A.2 Fórmula de Taylor

A partir da definição das derivadas de ordens superiores, podemos generalizar a conhecida fórmula de Taylor para espaços de Banach. Sabemos que se uma aplicação $F\colon X \to Y$ possui derivada de Fréchet, então a diferença $F(x+h) - F(x)$ pode ser representada como a soma de um termo linear relativamente a h e de um termo negligenciável em relação a $\|h\|$, que é frequentemente representado pela notação de Bachmann-Landau, a saber,

$$F(x+h) = F(x) + DF(x)h + o(\|h\|),$$

em que

$$\lim_{h \to 0} \frac{o(\|h\|)}{\|h\|} = 0.$$

Generalizando a idéia da fórmula de Taylor para funções numéricas, obtemos a expressão seguinte, denominada de fórmula de Taylor para funções aplicações.

A.3 Teorema *Sejam X e Y espaços de Banach e seja $F\colon U \subset X \to Y$ uma aplicação definida em um subconjunto aberto $U \subset X$, de modo que $DF^n(x)$ está definida e é uniformemente continua para todo $x \in U$. Então é válida a relação*

$$\begin{aligned} F(x+h) &= F(x) + DF(x)h + \frac{1}{2!}D^2F(x)(h,h) + \frac{1}{3!}D^3F(x)(h,h,h) \\ &\qquad + \cdots + \frac{1}{n!}D^nF(x)(h,h,\ldots,h) + \omega(x,h), \end{aligned} \tag{48}$$

em que $\|\omega(x,h)\| = o(\|h\|^n)$.

DEMONSTRAÇÃO. Argumentamos por indução em $k \in \mathbb{N}$. Para $k = 1$ a identidade (48) é imediata, pela definição da derivada. Suponhamos que seja válida a igualdade (48) para $k = n-1$, ou seja, cumpre-se a fórmula para todas as aplicações que verificam as hipóteses do teorema, enunciadas para $n-1$. Assim, para a aplicação DF, obtemos

$$\begin{aligned} DF(x+h) &= DF(x) + D^2F(x)h + \frac{1}{2}D^3(x)(h,h) \\ &\qquad + \cdots + \frac{1}{(n-1)!}D^nF(x)(h,h,\ldots,h) + \omega_1(x,h), \end{aligned} \tag{49}$$

em que $\omega_1(x,h) = o(\|h\|^{n-1})$. Integrando ambos os membros da equação (49) obtemos

$$\begin{aligned} F(x+h) &= F(x) + \int_0^1 DF(x+th)h\,\mathrm{d}t \\ &= \int_0^1 \Big[DF(x) + tD^2F(x)h + \frac{1}{2!}t^2D^3F(x)(h,h) \\ &\qquad + \cdots + \frac{1}{(n-1)!}t^{n-1}D^nF(x)(h,h,\ldots,h)\Big]h\,\mathrm{d}t + R_n(x,h), \end{aligned} \tag{50}$$

em que $R_n(x,h) \equiv \displaystyle\int_0^1 \omega_1(x,th)h\,\mathrm{d}t$. Da igualdade (50) obtemos

$$F(x+h) = F(x) + DF(x)h + \frac{1}{2}D^2F(x)(h,h) + \cdots + \frac{1}{n!}D^nF(x)(h,h,\ldots,h) + R_n(x,h),$$

em que

$$\|R_n(x,h)\| \leqslant \int_0^1 \|\omega_1(x,th)\|\|h\|\,\mathrm{d}t = o(\|h\|^n).$$

Assim, se a fórmula (48) vale para $k = n-1$, então também vale para $k = n$ e isso conclui a demonstração do teorema. □

A.4 Observação Na demonstração do Teorema A.3 utilizamos o conceito de integral; entretanto, visando tornar o texto mais sucinto, omitimos completamente a definição de integral de uma função abstrata, cujo domínio é um subconjunto dos números reais e que tem valores em um espaço de Banach. A teoria de integração nesse caso segue de perto a teoria de integração de funções numéricas; ao leitor interessado, recomendamos os textos de Kolmogorov e Fomin [4, Cap. 10, seção 1] e de Lang [6, Cap. 1, seção 4]. △

O próximo resultado trata da recíproca do teorema sobre a fórmula de Taylor.

A.5 Teorema *Seja $F\colon U \subset X \to Y$ uma função r vezes diferenciável e para $1 \leqslant k \leqslant r$, sejam $L_k \in N(X^k;Y)$ aplicações k-lineares simétricas tais que*

$$\begin{aligned} F(x+h) &= F(x) + L_1(x)h + \frac{1}{2!}L_2F(x)(h,h) + \frac{1}{3!}L_3F(x)(h,h,h) \\ &\quad + \cdots + \frac{1}{n!}L_rF(x)(h,h,\ldots,h) + \rho(h)\|h\|^r, \end{aligned} \tag{51}$$

em que $\|\rho(h)\| = o(1)$. Então $L_k(x) = \dfrac{1}{k!}D^kF(x)$.

DEMONSTRAÇÃO. Consulte o livro de Abraham, Marsden e Ratiu [1, Box 2.4B]. Veja também o livro de Lima [7, Cap. 9, exercício 2]. □

A.3 Teorema de Hahn-Banach

A seguir enunciamos, sem demonstrar, o teorema de Hahn-Banach para espaços normados. Este importante teorema garante que qualquer funcional linear $f_0\colon L \to \mathbb{R}$, definido em um subespaço L de um espaço vetorial E e verificando a desigualdade

$$|f_0(x)| \leqslant p(x), \tag{52}$$

em que $p\colon E \to \mathbb{R}$ é um funcional convexo e homogêneo, pode ser prolongado sobre todo o espaço E de modo que ainda se verifica a desigualdade (52). Para a aplicação desse resultado em espaços normados, o teorema pode ser enunciado da forma seguinte.

A.6 Teorema *Seja $L \subset E$ um subespaço do espaço normado E e seja $f_0\colon L \to \mathbb{R}$ um funcional limitado em L. Então o funcional f_0 pode ser prolongado em um funcional linear f definido sobre todo o espaço E, de modo que seja válida a igualdade*

$$\|f_0\|_L = \|f\|_E.$$

DEMONSTRAÇÃO. Consulte o livro de Kolmogorov e Fomin [4, Cap. 3, seção 2]. □

A.4 Aplicação avaliação

Seja $I = [0,1]$ e sejam E e F espaços de Banach. O espaço vetorial $C^r(I;E)$ das aplicações definidas em I e com valores em E e que possuem derivadas de ordem r continuamente diferenciáveis é um espaço de Banach quando munido da norma

$$\|f\|_k = \max_{1\leqslant i\leqslant k} \sup_{t\in I} \|D^i f(t)\|.$$

Além disso, se $U \subset E$ é aberto, então o conjunto $C^r(I;U) = \{f \in C^r(I;E)\colon f(I) \subset U\}$ é aberto em $C^r(I;E)$.

A.7 Proposição *A aplicação avaliação $A\colon C^1(I;U) \times I \to U$ definida por $A(f,t) = f(t)$ é Fréchet diferenciável e para $(g,s) \in C^1(I;U) \times \mathbb{R}$, vale a relação*

$$DA(f,t)(g,s) = Df(t)s + g(t).$$

A.5 Lema ômega

O Lema A.8 é conhecido na literatura como Lema ômega. Seja M um espaço topológico compacto e sejam X e Y espaços de Banach. Sabemos que o espaço vetorial $C(M;X)$ das funções contínuas definidas em M e com valores em E é um espaço de Banach quando munido da norma

$$\|f\| = \sup_{m \in M} \|f(m)\|.$$

Se $V \subset X$ é um subconjunto aberto, então o subconjunto $C(M;V)$ das funções $f \in C(M;E)$ tais que $f(M) \subset V$ também é aberto.

A.8 Lema *Sejam X e Y espaços de Banach e seja $V \subset X$ subconjunto aberto. Sejam $f\colon M \to X$ em que $f \in C(M;X)$ e $g\colon V \to Y$ em que $g \in C^1(V;Y)$. A aplicação $\Omega_g\colon C(M;V) \to C(M;Y)$ definida por $\Omega_g(f) \equiv g \circ f$ é tal que $\Omega_g \in C^r(C(M;V);C(M;Y))$. A derivada de Fréchet de Ωg é*

$$[D\Omega_g(f)h](x) = Dg(f(x))h(x).$$

DEMONSTRAÇÃO. Seja $f \in C(M;V)$. Pela continuidade da aplicação g e pela compacidade do espaço topológico M, dado $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\|\Omega_g(f_1) - \Omega_g(f_2)\| = \sup_{m \in M} \|g(f_1(m)) - g(f_2(m))\| < \varepsilon$$

sempre que $\|f_1 - f_2\| < \delta$. Dessa forma, a aplicação Ω_g é contínua em f.

Seja $A\colon C(M;L(X;Y)) \to L(C(M;X);C(M;Y))$ uma aplicação definida por

$$A(H)(h)(m) = H(m)h(m)$$

para $H \in C(M;L(X;Y))$, $v \in C(M;X)$ e $m \in M$. A aplicação A é linear e contínua; além disso, $\|A\| \leqslant 1$. Como $Dg\colon V \to L(X;Y)$ é contínua, o argumento precedente mostra que a aplicação $\Omega_{Dg}\colon C(M;V) \to C(M;L(X;Y))$ é contínua e, portanto, a aplicação $A \cdot \Omega_{Dg}\colon C(M;V) \to L(C(M;X);C(M;Y))$ também é contínua. Aplicando o Teorema A.5 à função g obtemos

$$g(f(m) + v(m)) = g(f(m)) + D_g(f(m))h(m) + R(f(m), v(m))h(m).$$

Definindo

$$[(Dg \circ f)h](m) = Dg(f(m))h(m)$$

e

$$[R(f,v)(h)](m) = R(f(m), v(m))h(m),$$

temos que R é uma aplicação contínua, $R(f,0) = 0$ e

$$\begin{aligned}\Omega_g(f+v) &= g \circ (f+h)\\ &= g \circ f + (Dg \circ f)h + R(f,v)h\\ &= \Omega_g(f) + (A \circ \Omega_{Dg})(f)h_R(f,v)h.\end{aligned}$$

Pela recíproca do Teorema de Taylor, $D\Omega g = A \circ \Omega_{Dg}$ e, portanto, Ω_g é continuamente Fréchet diferenciável, isto é, $\Omega_g \in C^1(C(M;V);C(M;Y))$. Isto conclui a demonstração do teorema. $\square$

Referências

[1] R. Abraham, J. E. Marsden, T. Ratiu, *Manifolds, Tensor Analysis, and Applications*, Addison-Wesley, Reading, 1983.

[2] G. Chilov, *Analyse Mathématique, Fonctions de Plusieurs Variables Réeles*, Editions Mir, Moscou, 1975.

[3] X. Dai, *Continuous Differentiability of Solutions of ODEs with respect to Initial Conditions*, Amer. Math. Monthly, 113 (2006) 66–70.

[4] A. N. Kolmogorov, S. V. Fomin, *Elementos da Teoria das Funções e de Análise Funcional*, Editora Mir, Moscou, 1982.

[5] S. G. Krantz, H. R. Parks, *The Implicit Function Theorem, History, Theory, And Applications*, Birkhäuser, Boston, 2002.

[6] S. Lang, *Introduction to Differentiable Manifolds*, Springer-Verlag, New York, 2002.

[7] E. L. Lima, *Análise no espaço* $\mathbb{R}^n$, Editora Edgard Blücher, São Paulo, 1970.

[8] R. S. Palais, *A simple proof of the Banach contraction principle*, J. Fixed Point Theory Appl. 2 (2007) 221–223.

[9] J. W. Robbin, *On the existence theorem for differential equations*, Proc. Amer. Math. Soc. 19 (1968), 1005–1006.

[10] J. Sotomayor, *Smooth dependence of solutions of differential equations on initial data: A simple proof*, Bol. Soc. Brasil. Mat. 04 (1973), n. 1, 55–59.

Ronaldo B. Assunção
Departamento de Matemática - UFMG
ronaldo@mat.ufmg.br

Paulo C. Carrião
Departamento de Matemática - UFMG
carrion@mat.ufmg.br