



Universidade Federal do Pará
Instituto de Ciências Exatas e Naturais
Programa de Pós-Graduação em Matemática e Estatística

Jociane dos Santos Fonseca

Método de Newton generalizado e Aplicações

Belém - PA
Junho de 2017

Jociane dos Santos Fonseca

Método de Newton generalizado e Aplicações

Trabalho apresentado ao Programa de Pós-Graduação em Matemática e Estatística da Universidade Federal do Pará, em cumprimento com as exigências legais para obtenção do título de Mestre.

Orientadora:

Prof. Dra. Cristina Vaz

Belém - PA

Junho de 2017

Dados Internacionais de Catalogação-na-Publicação (CIP)
Biblioteca Central

Fonseca, Jociane dos Santos, 1988-

Método de Newton generalizado e Aplicações / Fonseca, Jociane dos Santos. — 2017.

Orientador: Cristina Vaz

Dissertação (Mestrado) - Universidade Federal do Pará, Instituto de Ciências Exatas e Naturais, Programa de Pós-Graduação em Matemática Aplicada e Estatística, Belém, 2017.

1. Método de Newton-Raphson. 2. Problemas inversos (Equações diferenciais). 3. Análise numérica. 4. Funções (Matemática). 5. Teorema de existência. I. Título.

Universidade Federal do Pará
Instituto de Ciências Exatas e Naturais
Programa de Pós-Graduação em Matemática e Estatística

Jociane dos Santos Fonseca

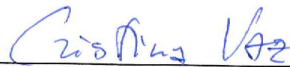
Método de Newton Generalizado e Aplicações

Trabalho apresentado ao Programa de Pós-Graduação em Matemática e Estatística da Universidade Federal do Pará, em cumprimento com as exigências legais para obtenção do título de Mestre.

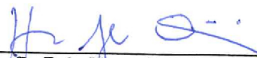
Data da defesa: 27 de Junho de 2017.

Resultado: Aprovado

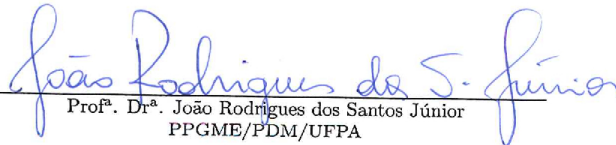
Banca Examinadora:



Prof. Dra. Cristina Lúcia Dias Vaz- UFPA
Orientadora



Prof^ª. Dr.^a. Hugo Alex Carneiro Diniz
UFOPA



Prof.^a. Dr.^a. João Rodrigues dos Santos Júnior
PPGME/PDM/UFPA

Dedicatória

Dedico esta dissertação à minha família, pelo apoio incondicional e constante incentivo, em especial à minha mãe e inspiração, Maria Messias.

Dedico também a minha orientadora Profa. Dra. Cristina Vaz, pela confiança, paciência, incentivo, amizade e excelente orientação.

Sem o apoio de ambos este trabalho não teria se realizado, a eles meu muito, muito obrigada.

Agradecimentos

Agradeço a Deus por todas as graças concedidas, pois tem me dado mais do que eu poderia imaginar, pela força e inspiração para desenvolver este trabalho.

Agradeço a minha família, meus pais José Coimbra e Maria Messias e aos meus irmãos, Josivan e Josimar pelo apoio e incentivo.

Agradeço aos meus tios Antonio, Lucia e seus filhos Allan e Artur por toda a receptividade e acolhimento em sua residência no tempo em que residir com os mesmos.

Agradeço a minha orientadora Prof. Dra. Cristina Vaz por sua orientação impecável, por seu apoio, dedicação e paciência.

Agradeço ao Prof. Dr. João Rodrigues pelo apoio, estímulo e confiança.

Agradeço aos meus amigos e colegas, em especial ao amigo Eduardo Rangel e as amigas, Jacqueline Ribeiro e Susana Santos pela solidariedade, compreensão e apoio.

Agradeço ao CAPES pelo apoio financeiro.

Resumo

Este trabalho consiste em uma análise numérica teórica do método de Newton em espaços de Banach, conhecido como *Método de Newton generalizado*. A principal ideia do método é transformar uma equação de operadores num problema de ponto fixo equivalente.

Como o método é um processo iterativo analisamos sua convergência e ordem de convergência. Além disso, aplicaremos o método de Newton generalizado na resolução de sistema não lineares, equações diferenciais ordinárias não lineares e equações integrais não lineares. Para isto, essencialmente, generalizamos o conceito de derivada.

O método de Newton é uma poderosa ferramenta na obtenção de soluções de equações dos mais variados tipos. A principal vantagem deste método e sua versátil aplicabilidade é ter convergência quadrática. Além disso, o método de Newton também pode usado para mostrar outros teoremas importantes na matemática, como por exemplo, o Teoremas de existência e unicidade de soluções para certas equações diferenciais e o Teorema da função inversa e implícita, entre outros.

Palavras-chaves: Diferencial de Fréchet. Teorema do Ponto fixo. Método de Newton.

Sumário

Introdução	1
1 Preliminares	3
1.1 Tópicos de Análise Funcional	3
1.2 Soluções de Equações Lineares	8
1.3 Teorema do ponto fixo de Banach	10
1.4 Aplicações do Teorema de ponto fixo de Banach	13
1.5 Equações diferenciais em espaços de Banach	18
2 Cálculo em espaços de Banach	20
2.1 Cálculo Diferencial	20
2.2 Cálculo Integral	31
3 Método de Newton generalizado	34
3.1 Caso unidimensional	34
3.2 Caso multidimensional	38
3.3 Método de Newton generalizado	42
4 Aplicações do método de Newton generalizado	46
4.1 Sistemas não lineares	46
4.2 Equações integrais não lineares	47
4.3 Equações diferenciais ordinárias de 1ª ordem	48
4.4 Equações diferenciais de 2ª ordem	50
5 Conclusão	52
Referências Bibliográficas	53

Introdução

Este trabalho consiste em análise matemática teórica do método de Newton em espaços de Banach, conhecido como *Método de Newton Generalizado*. A principal idéia do método é, dada a função $f : X \rightarrow X$, encontrar a solução da equação de operadores

$$f(x) = 0 \tag{1}$$

com X um espaço de Banach.

Nossa abordagem será transformar a equação (1) numa equação de ponto fixo

$$x = Tx, \tag{2}$$

com $T : X \rightarrow X$ dado por $T(x) = x - [Df(x)]^{-1}f(x)$.

O método descrito acima foi proposto inicialmente por Isaac Newton em 1669 para encontrar raízes de funções polinomiais. Pouco tempo depois, em 1690 J. Raphson estendeu o método para funções reais quaisquer. Por isso é muito comum, na literatura, o método ser chamado método de Newton-Raphson. A consolidação do método está ligada a famosos matemáticos como J. Fourier, L. A. Cauchy entre outros. Em 1818, Fourier provou que o método convergia quadraticamente desde que o ponto inicial fosse tomado em uma vizinhança da solução procurada, enquanto Cauchy (1829-1847) mostrou que o método se estende naturalmente para funções para funções de várias variáveis e usou o método para provar a existência de raízes de algumas equações. Em 1916, os matemáticos Fine e Bennet deram mais algumas contribuições para o método. Fine provou a convergência para o caso n-dimensional sem a hipótese de existência de solução. Bennet estendeu o resultado para o caso de dimensão infinita. Mais recentemente, em 1948, L. V. Kantorovich provou a existência de solução e a convergência do método para operadores $T : \mathcal{B}_1 \rightarrow \mathcal{B}_2$, onde \mathcal{B}_1 e \mathcal{B}_2 são espaços de Banach e T é um operador diferenciável qualquer.

O método de Newton é uma poderosa ferramenta na obtenção de soluções de equações dos mais variados tipos. A grande importância desse método reside no fato de que sob algumas hipóteses é garantida a convergência, a uma taxa relativamente alta, para uma solução. O método de Newton também é usado para mostrar outros teoremas importantes na matemática.

Por exemplo: teoremas de existência e unicidade de soluções para certas equações diferenciais, o teorema da função inversa e implícita entre outros.

Esta dissertação está organizada da seguinte forma.

No capítulo 1, apresentamos alguns conceitos e resultados que serão utilizados neste trabalho.

No capítulo 2, tratamos do conceito de diferenciabilidade em espaços de dimensão finita e espaços de Banach, e os principais resultados sobre diferenciação.

No capítulo 3, abordamos o método de Newton e suas principais propriedades utilizando o teorema do ponto fixo de Banach.

No capítulo 4, aplicamos o método de Newton em equações integrais do tipo: Fredholm e Volterra, lineares e não lineares e equações diferenciais.

Capítulo 1

Preliminares

Neste capítulo, descrevemos os principais resultados usados neste trabalho.

1.1 Tópicos de Análise Funcional

Nesta seção, enunciamos e demonstramos alguns resultados de Análise Funcional que serão usados no trabalho. Para mais detalhes consulte [2], [3] e [5].

Definição 1.1 *Seja X um espaço vetorial sobre o corpo \mathbb{K} , onde $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . Um produto interno em X é uma aplicação*

$$\begin{aligned} ((\cdot, \cdot)) : X \times X &\rightarrow \mathbb{K} \\ (x, y) &\rightarrow ((x, y)) \end{aligned}$$

que satisfaz as seguintes propriedades para quaisquer $x, y \in X$ e $\alpha \in \mathbb{R}$:

- i) $((x, y)) \geq 0$;
- ii) $((x, x)) = 0 \Leftrightarrow x = 0$
- iii) $((\alpha x, y)) = \alpha((x, y))$
- iv) $((x + y, z)) = ((x, z)) + ((y, z))$
- v) $((x, y)) = \overline{((y, x))}$

Definição 1.2 *Seja X um espaço vetorial sobre o corpo \mathbb{K} onde $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . Uma norma em X é uma função real*

$$\begin{aligned} \|\cdot\| : X &\rightarrow \mathbb{R} \\ x &\rightarrow \|x\| \end{aligned}$$

que satisfaz as seguintes propriedades para quaisquer $x, y \in X$ e $\alpha \in \mathbb{R}$:

$$i) \|x\| \geq 0;$$

$$ii) \|x\| = 0 \Leftrightarrow x = 0$$

$$iii) \|\alpha x\| = |\alpha| \|x\|$$

$$iv) \|x + y\| \leq \|x\| + \|y\|.$$

Definição 1.3 Um espaço vetorial X munido com uma norma $\|\cdot\|$ é chamado de Espaço Normado, o qual denotamos por $(X, \|\cdot\|)$.

A convergência em Espaços Normados é dada pela seguinte definição.

Definição 1.4 Seja X um espaço normado. Dizemos que a sequência (x_n) de elementos de X converge para $x \in X$ se

$$\lim_{n \rightarrow \infty} \|x_n - x\| = 0 \Leftrightarrow \forall \epsilon > 0, \exists n_0 \in \mathbb{N}; n \geq n_0 \Rightarrow \|x_n - x\| < \epsilon.$$

Às vezes, usaremos a notação $x_n \rightarrow x$.

Seja X um espaço normado. Uma sequência (x_n) de elementos de X é chamada sequência de Cauchy se

$$\lim_{n \rightarrow \infty} \|x_m - x_n\| = 0 \Leftrightarrow \forall \epsilon > 0, \exists n_0 \in \mathbb{N}; m, n \geq n_0 \Rightarrow \|x_m - x_n\| < \epsilon.$$

Podemos observar que toda sequência convergente é de Cauchy, mas nem toda sequência de Cauchy é convergente. Veja[6].

Definição 1.5 O espaço normado X é um espaço de Banach se, e somente se, toda sequência de Cauchy de elementos de X converge para um elemento de X .

Definição 1.6 Sejam X e Y espaços vetoriais sobre o corpo \mathbb{K} . Dizemos que $L : X \rightarrow Y$ é um operador linear se

$$L(\alpha u + v) = \alpha L(u) + L(v), \quad \forall u, v \in X, \alpha \in \mathbb{K}$$

Quando $Y = \mathbb{R}$, o operador linear L chama-se funcional linear.

Definição 1.7 Sejam X e Y espaços normados e $T : X \rightarrow Y$ um operador linear. Dizemos que T é um operador limitado se existe uma constante $C > 0$ tal que, para todo $u \in X$,

$$\|Tu\| \leq C\|u\|.$$

Definição 1.8 *Sejam X e Y espaços normados e $T : X \rightarrow Y$ um operador linear. Dizemos que T é um operador contínuo se dado $\epsilon > 0$, $\exists \delta > 0$ tal que, para $u, v \in X$,*

$$\|u - v\| < \delta \Rightarrow \|Tu - Tv\| < \epsilon.$$

Se $\delta = \delta(\epsilon)$ dizemos que T é uniformemente contínuo.

Teorema 1.1 *Sejam X e Y espaços normados. Se X tem dimensão finita, então todo operador linear $T : X \rightarrow Y$ é limitado.*

A prova deste teorema pode ser encontrada em ([5], pg 96, teorema 2.7-8).

Teorema 1.2 *Sejam X e Y espaços normados e $T : X \rightarrow Y$ linear. Então T é limitado se, e somente se, T é contínuo.*

A demonstração deste resultado pode ser encontrada em ([5], pg 97, teorema 2.7-9)

Definição 1.9 *Dizemos que um operador $T : X \rightarrow X$ é uma contração com constante de contraticidade $\alpha \in [0, 1)$ se*

$$\|Tu - Tv\|_X \leq \alpha \|u - v\|_X, \forall u, v \in X.$$

O operador T é chamado não-expansivo se

$$\|Tu - Tv\|_X \leq \|u - v\|_X, \forall u, v \in X.$$

e Lipschitz contínuo se existe uma constante $L \geq 0$ tal que

$$\|Tu - Tv\|_X \leq L \|u - v\|_X, \forall u, v \in X.$$

Assim, temos as seguintes implicações

$$\begin{aligned} \text{Contractividade} &\Rightarrow \text{não-expansividade} \\ &\Rightarrow \text{Lipschitz continuidade} \\ &\Rightarrow \text{continuidade.} \end{aligned}$$

Considere o espaço vetorial $\mathcal{L}(X, Y)$, espaço dos operadores lineares contínuos com as operações usuais. Então temos o seguinte resultado.

Teorema 1.3 *Sejam X e Y espaços normados. Então, $\mathcal{L}(X, Y)$ é um espaço normado, com*

a norma dada por

$$\|T\| = \sup_{x \in X - \{0\}} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\| \quad (1.1)$$

para todo $T \in \mathcal{L}(X, Y)$.

Temos agora a definição de Espaço de Hilbert.

Definição 1.10 Dizemos que um espaço com produto interno é um espaço de Hilbert se é um espaço normado completo.

Observe que todo Espaço de Hilbert é um Espaço de Banach porém a recíproca não é verdadeira, pois existem Espaços de Banach cuja norma não provém de um produto interno.

Definição 1.11 Sejam H_1 e H_2 espaços de Hilbert com produto interno $((,))$ e $T : H_1 \rightarrow H_2$ um operador linear limitado. Chamamos $T^* : H_2 \rightarrow H_1$ de operador adjunto de T se vale a seguinte propriedade:

$$((Tu, v)) = ((u, T^*v)), \quad \forall u \in H_1, v \in H_2. \quad (1.2)$$

Teorema 1.4 (Teorema da Representação de Riesz) Seja H um espaço de Hilbert com produto interno $((,))$ e norma $\|\cdot\|$. Dado $\phi \in H'$, existe um único $f \in H$ tal que

$$\langle \phi, v \rangle_{H', H} = ((f, v)), \quad \text{para todo } v \in H.$$

Além disso,

$$\|f\|_H = \|\phi\|_{H'}.$$

Prova: Consideremos a aplicação

$$\begin{aligned} T : H &\rightarrow H' \\ f &\mapsto Tf, \end{aligned} \quad (1.3)$$

definida por

$$\langle Tf, v \rangle_{H', H} = ((f, v)) \quad \text{para todo } v \in H.$$

$Tf : H \rightarrow \mathbb{R}$ é claramente linear e contínua, pois

$$|\langle Tf, v \rangle_{H', H}| = |((f, v))| \leq \|f\|_H \|v\|_H,$$

o que implica que $Tf \in H'$.

Assim, $T : H \rightarrow H'$ está bem definida e é linear, pois dados $f, g, v \in H$ e $\alpha, \beta \in \mathbb{R}$ temos:

$$\begin{aligned} \langle T(\alpha f + \beta g), v \rangle &= ((\alpha f + \beta g), v) \\ &= \alpha((f), v) + \beta((g), v) \\ &= \alpha \langle Tf, v \rangle + \beta \langle Tg, v \rangle \\ &= \langle \alpha Tf + \beta Tg, v \rangle, \end{aligned}$$

o que implica que $T(\alpha f + \beta g) = \alpha Tf + \beta Tg$ provando a linearidade de T . A seguir provaremos que

$$\|Tf\|_{H'} = \|f\|_H, \forall f \in H \quad (1.4)$$

De fato, dados $f, v \in H$ de (1.3) vem que

$$|\langle Tf, v \rangle| \leq \|f\|_H \|v\|_H \Rightarrow \|Tf\|_{H'} \leq \|f\|_H. \quad (1.5)$$

Por outro lado, notemos que se $f \neq 0$ (não identicamente nula), então

$$\begin{aligned} \|f\|_H^2 &= ((f), f) = \langle Tf, f \rangle = \langle Tf, \frac{f}{\|f\|} \rangle \cdot \|f\|_H \\ &\leq \|f\|_H \cdot \sup_{v \in H, \|v\|=1} |\langle Tf, v \rangle| = \|f\|_H \|Tf\|_{H'}, \end{aligned}$$

ou seja,

$$\|f\|_H \leq \|Tf\|_{H'}. \quad (1.6)$$

Observe que se $f \equiv 0$ a desigualdade (1.6) segue trivialmente. Combinando (1.5) e (1.6) obtemos o desejado em (1.4).

Assim, a aplicação $T : H \rightarrow H'$ é uma aplicação linear isométrica, portanto injetora.

Resta-nos provar que

$$TH = H', \quad (1.7)$$

isto é, T é sobrejetora. Com efeito afirmamos que

$$TH \text{ é um espaço fechado de } H' \quad (1.8)$$

pois se $(T_\nu) \subset TH$ é tal que $T_\nu \rightarrow w$ em H' , então pelo fato de

$$\|v_\nu - v_\mu\| = \|Tv_\nu - Tv_\mu\|_{H'} \rightarrow 0 \text{ quando } \nu, \mu \rightarrow \infty,$$

segue que a sequência $(v_\nu)_{\nu \in \mathbb{N}}$ é de Cauchy em H e portanto é convergente, digamos, existe $v \in H$ tal que $v_\nu \rightarrow v$ em H . Pela continuidade da aplicação $T : H \rightarrow H'$ resulta que $Tv_\nu \rightarrow Tv$ em H' e, portanto face a unicidade do limite em H' , concluímos que $w = Tv \in TH$,

o que prova (1.8). Logo se mostrarmos que

$$TH \text{ é denso em } H' \quad (1.9)$$

então, por (1.8) e (1.9) resulta que $TH = \overline{TH} = H'$, ou seja, $TH = H'$, ficando provado (1.7). Logo, basta mostrarmos (1.9). Seja então $\xi \in H''$ tal que

$$\langle \xi, Tf \rangle_{H'', H'} = 0, \forall f \in H.$$

Queremos provar que $\xi \equiv 0$ em H'' . Com efeito, como H é reflexivo (posto que é Hilbert) segue que $H'' \equiv H$. Assim, $\xi \in H'' \equiv H$ o que implica que

$$\langle Tf, \xi \rangle_{H', H} = ((f, \xi)) = 0, \quad \forall f \in H.$$

Em particular se $f = \xi$ obtemos

$$((\xi, \xi)) = \|\xi\|^2 = 0$$

o que implica que $\xi \equiv 0$, o que prova o desejado.

Na demonstração anterior usamos uma consequência do seguinte teorema (para maiores detalhes consulte [3]).

Teorema 1.5 (Teorema de Hahn-Banach-2^o Forma Geométrica) *Sejam E um espaço vetorial normado, $A, B \subset E$ subconjuntos convexos de E , disjuntos e não-vazios. Se A for fechado e B for compacto, então existe um hiperplano fechado que separa A e B no sentido estrito.*

Corolário 1.1 *Sejam E um espaço vetorial normado e F um subespaço vetorial de E . Se para toda forma $f \in E'$ tal que $\langle f, x \rangle = 0, \forall x \in F$ se tem $f \equiv 0$ (isto é, $\langle f, x \rangle = 0, \forall x \in E$) então F é denso em E (ou seja, $\bar{F} = E$).*

1.2 Soluções de Equações Lineares

Um dos problemas computacionais centrais da análise funcional linear é a solução de equações lineares. Este problema surge também no estudo de equações de operadores não-lineares, uma vez que alguns métodos de ataque de equações não-lineares são baseados na resolução de uma equação linear aproximada ou uma sequência dessas equações.

Em um espaço de Banach X o problema de resolver uma equação linear pode ser dada da seguinte forma: Seja $L : X \rightarrow X$ um operador linear limitado e $x \in X$ tais que

$$Lx = y, \quad (1.10)$$

Para algum $y \in X$. Se existir algum elemento x , será chamado solução da equação linear (1.10).

É conveniente introduzir a noção do produto (ou composição) de operadores (não necessariamente lineares) de X em X , então o produto PQ é o operador definido por

$$PQ(x) = P(Q(x))$$

para todo $x \in X$.

A equação escalar linear $ax = y$ é resolvida para $a \neq 0$ multiplicando ambos os lados por $\frac{1}{a}$ e obtendo $x = \frac{1}{a}y$.

Essa idéia pode ser estendida para operadores lineares limitados.

Definição 1.12 Se $L : X \rightarrow X$ um operador linear limitado, e existe um operador linear limitado L^{-1} tal que

$$L^{-1}L = LL^{-1} = I,$$

com I é no operador identidade. Então L^{-1} é chamado de operador inverso de L .

O inverso de um operador linear limitado, é portanto, uma generalização da noção de inverso de um escalar diferente de 0.

Operadores lineares não-nulos, no entanto, não precisam ter inversos: por exemplo, a matriz

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

é limitada, linear e não-nula em \mathbb{R}^2 , mas A não possui inversa.

Se L^{-1} existir, então a equação (1.10) tem uma única solução

$$x = L^{-1}y \tag{1.11}$$

para cada $y \in X$, e

$$\|x\| \leq \|L^{-1}\| \|y\|. \tag{1.12}$$

Se a equação (1.10) tem uma única solução x para cada $y \in X$ dado, isso define um operador linear L_1 pela correspondência $x = L_1y$. Se L_1 é limitado, então L^{-1} existe e $L^{-1} = L_1$.

O teorema fundamental que fornece condições necessárias e suficientes para existência de um operador linear limitado inverso de um operador linear limitado $L : X \rightarrow X$ é o seguinte, e baseia-se na observação de que se $|1 - a| < 1$, então

$$\frac{1}{a} = 1 + (1 - a) + (1 - a)^2 + \dots = \sum_{n=0}^{\infty} (1 - a)^n. \tag{1.13}$$

A demonstração do seguinte teorema encontra-se em ([10], pg 50, teorema 10.1).

Teorema 1.6 *Seja L é um operador linear limitado em X . L é um operador invertível, com inverso L^{-1} , se somente se existe um operador linear limitado $K : X \rightarrow X$ tal que K^{-1} existe e*

$$\|I - KL\| < 1. \quad (1.14)$$

se L^{-1} existe, então

$$L^{-1} = \sum_{n=0}^{\infty} (I - KL)^n K \quad (1.15)$$

e

$$\|L^{-1}\| \leq \frac{\|K\|}{1 - \|I - KL\|}. \quad (1.16)$$

1.3 Teorema do ponto fixo de Banach

Nesta seção, consideramos equações de operadores da forma

$$u = T(u), \quad u \in X \quad (1.17)$$

com X é um espaço de Banach e $T : X \rightarrow X$.

As soluções da equação (1.17) são chamadas de pontos fixos do operador T . O resultado mais importante da teoria de soluções de tais equações para o operador T que são uma contração é o conhecido Teorema do ponto fixo de Banach.

1.3.1 Teorema do Ponto fixo de Banach

Teorema 1.7 (Teorema do Ponto Fixo de Banach) *Sejam X um espaço de Banach e $K \subset X$ um subconjunto não-vazio, fechado. Se $T : K \rightarrow K$ é uma contração com constante de contraticidade $\alpha \in [0, 1)$. Então, as seguintes afirmações são verdadeiras.*

(1) **Existência e Unicidade:** *Existe um único $u \in K$ tal que*

$$T(u) = u.$$

(2) **Convergência e Estimativas do erro:** *Para algum $u_0 \in K$, a sequência $u_n \subset K$ definida por*

$$u_{n+1} = T(u_n), \quad n = 0, 1, 2, \dots \quad (1.18)$$

converge para u :

$$\|u_n - u\|_X \rightarrow 0 \text{ para } n \rightarrow \infty.$$

Para o erro os seguintes limites são válidos:

$$\|u_n - u\|_X \leq \frac{\alpha^n}{1 - \alpha} \|u_0 - u_1\|_X \quad (1.19)$$

$$\|u_n - u\|_X \leq \frac{\alpha}{1 - \alpha} \|u_{n-1} - u_n\|_X \quad (1.20)$$

$$\|u_n - u\|_X \leq \alpha \|u_{n-1} - u\|_X. \quad (1.21)$$

Prova:

Sendo $T : K \rightarrow K$, a sequência $\{u_n\}$ está bem definida. Inicialmente mostraremos que $\{u_n\}$ é uma sequência de Cauchy. Usando a contratividade da aplicação T temos

$$\|u_n - u_{n+m}\|_X \leq \alpha \|u_{n-1} - u_{n+m-1}\|_X \leq \dots \leq \alpha^n \|u_0 - u_m\|_X. \quad (1.22)$$

Agora, temos que

$$\|u_0 - u_m\|_X = \|u_0 - u_1 + u_1 - u_2 + u_2 - \dots + u_{m-1} - u_m\|_X \leq \|u_0 - u_1\|_X + \|u_1 - u_2\|_X + \dots + \|u_{m-1} - u_m\|_X. \quad (1.23)$$

Usando novamente a contratividade do operador T , agora em (1.23), concluímos que

$$\|u_0 - u_m\|_X \leq \|u_0 - u_1\|_X + \alpha^2 \|u_0 - u_1\|_X + \alpha^3 \|u_0 - u_1\|_X + \dots + \alpha^m \|u_0 - u_1\|_X \quad (1.24)$$

$$\leq (1 + \alpha + \alpha^2 + \alpha^3 + \dots + \alpha^{m-1}) \|u_0 - u_1\|_X \quad (1.25)$$

Como $\sum_{j=0}^{m-1} \alpha^j = \frac{1}{1 - \alpha}$, pois $0 \leq \alpha < 1$, temos:

$$\|u_0 - u_m\|_X \leq \frac{\|u_0 - u_1\|_X}{1 - \alpha}$$

Logo de (1.22) e da desigualdade anterior temos

$$\|u_n - u_{n+m}\|_X \leq \frac{\alpha^n}{1 - \alpha} \|u_0 - u_1\|_X. \quad (1.26)$$

Passando limite em (1.26) com $n \rightarrow \infty$, segue que

$$\lim_{n \rightarrow \infty} \|u_n - u_{n+m}\|_X \leq \lim_{n \rightarrow \infty} \frac{\alpha^n}{1 - \alpha} \|u_0 - u_1\|_X = 0, \quad (1.27)$$

pois $\alpha \in [0, 1)$. E portanto, $\{u_n\}$ é uma sequência de Cauchy. Como K é um subespaço fechado de um espaço de Banach X , existe $u \in K$ tal que $\lim_{n \rightarrow \infty} u_n = u$.

Resta mostrarmos agora, que u é um ponto fixo de T , assim passando limite em $u_{n+1} = T(u_n)$ quando $n \rightarrow \infty$, e observando que sendo T um operador contrativo, ele também é

contínuo segue que $T(u_n) \rightarrow T(u)$, e da unicidade de limite concluímos que $T(u) = u$.

Suponha que $u_1, u_2 \in K$ são pontos fixos de T . Então $u_1 = T(u_1)$ e $u_2 = T(u_2)$, e

$$u_1 - u_2 = T(u_1) - T(u_2).$$

Portanto,

$$\begin{aligned} \|u_1 - u_2\|_X &= \|T(u_1) - T(u_2)\|_X \leq \alpha \|u_1 - u_2\|_X \\ &\Rightarrow (1 - \alpha) \|u_1 - u_2\|_X \leq 0. \end{aligned}$$

Como $\alpha \in [0, 1)$ segue que

$$\|u_1 - u_2\|_X \leq 0 \Rightarrow u_1 = u_2.$$

Portanto, um ponto fixo de um operador contrativo é único.

Provaremos agora, as estimativas do erro. Passando limite quando $m \rightarrow \infty$ em (1.26) obtemos a estimativa (1.19). De

$$\|u_n - u\|_X = \|T(u_{n-1}) - T(u)\|_X \leq \alpha \|u_{n-1} - u\|_X$$

obtemos a estimativa (1.21). Esta estimativa junto com

$$\|u_{n-1} - u\|_X \leq \|u_{n-1} - u_n\|_X + \|u_n - u\|_X$$

implicam na estimativa (1.20).

O Teorema 1.7 tem a seguinte generalização:

Teorema 1.8 *Seja X um espaço de Banach e $T : X \rightarrow X$ um operador contínuo em X . Se T^n é uma contração para $n > 1$ então T tem um único ponto fixo.*

Prova: *Sejam $f_1 \in X$ qualquer e $f = \lim_{n \rightarrow \infty} T^n f_1$. Então pela continuidade de T tem-se*

$$Tf = \lim_{n \rightarrow \infty} TT^n f_1.$$

Mas, T^n é uma contração e logo

$$\|T^n T f_1 - T^n f_1\| \leq \alpha \|T^{(n-1)n} T f_1 - T^{(n-1)n} f_1\| \leq \dots \leq \alpha^n \|T f_1 - f_1\|$$

com $\alpha < 1$.

Portanto,

$$\|Tf - f\| = \lim_{n \rightarrow \infty} \|TT^n f_1 - T^n f_1\| = 0,$$

ou seja, $Tf = f$.

Para provarmos a unicidade, observe que se T tem mais de um ponto fixo então T^n também terá, o que é uma contradição pelo Teorema 1.7, pois T^n é uma contração.

Assim, para resolvermos o problema

$$f(x) = 0 \tag{1.28}$$

para algum operador $f : X \rightarrow X$ podemos transformá-lo num problema de ponto fixo equivalente a (1.17) definindo

$$T(u) = u - F(f(u))$$

com $F : X \rightarrow X$ um operador satisfazendo

$$F(w) = 0 \Leftrightarrow w = 0.$$

Como os problemas são equivalentes temos que a solução do problema de ponto fixo (1.17) é solução do problema (1.28) e vice-versa. Além disso, o processo iterativo (1.18) é um método de aproximação para resolução da equação (1.28).

1.4 Aplicações do Teorema de ponto fixo de Banach

Nesta Seção apresentaremos algumas aplicações importantes do Teorema de ponto fixo de Banach (para mais detalhes consulte [1], [8] e [10]).

1.4.1 Equação não-linear em espaço de Hilbert

Como uma aplicação do teorema de ponto fixo de Banach, consideramos a existência e unicidade de solução de uma equação não-linear em um espaço de Hilbert.

Teorema 1.9 *Seja H um espaço de Hilbert. Assuma que $T : H \rightarrow H$ é fortemente monótona e Lipschitz contínuo, isto é, existem duas constantes $c_1, c_2 > 0$ tais que, para $v_1, v_2 \in H$*

$$((T(v_1) - T(v_2), v_1 - v_2)) \geq c_1 \|v_1 - v_2\|^2, \tag{1.29}$$

$$\|T(v_1) - T(v_2)\| \leq c_2 \|v_2 - v_1\|. \tag{1.30}$$

Então, para qualquer $b \in H$, existe um único $u \in H$ tal que

$$T(u) = b. \tag{1.31}$$

Além disso, a solução u depende Lipschitz continuamente de b : se $T(u_1) = b_1$ e $T(u_2) = b_2$, então

$$\|u_1 - u_2\| \leq \frac{1}{c_1} \|b_1 - b_2\|. \quad (1.32)$$

Prova: A equação $T(u) = b$ é equivalente a

$$u = u - \theta[T(u) - b]$$

para qualquer $\theta \neq 0$. Definamos um operador $T_\theta : H \rightarrow H$ pela fórmula

$$T_\theta(v) = v - \theta[T(v) - b].$$

Vamos mostrar que para $\theta > 0$ suficientemente pequeno, o operador T_θ é uma contração. Escreva

$$T_\theta(v_1) - T_\theta(v_2) = (v_1 - v_2) - \theta[T(v_1) - T(v_2)].$$

Então,

$$\|T_\theta(v_1) - T_\theta(v_2)\|^2 = \|v_1 - v_2\|^2 - 2\theta((T(v_1) - T(v_2), v_1 - v_2)) + \theta^2 \|T(v_1) - T(v_2)\|^2.$$

Usando (1.29) e (1.30) obtemos

$$\|T_\theta(v_1) - T_\theta(v_2)\|^2 \leq (1 - 2c_2\theta + c_1^2\theta^2) \|v_1 - v_2\|^2.$$

Para $\theta \in (0, 2c_2/c_1^2)$

$$1 - 2c_2\theta + c_1^2\theta^2 < 1$$

e T_θ é uma contração. Então, pelo Teorema do ponto fixo de Banach, T_θ tem um único ponto fixo $u \in H$.

Portanto, a equação (1.31) tem uma única solução.

Agora provaremos a dependência Lipschitz contínua da solução.

Sejam $T(u_1) = b_1$ e $T(u_2) = b_2$, obtemos

$$T(u_1) - T(u_2) = b_1 - b_2.$$

Então,

$$((T(u_1) - T(u_2))) = ((b_1 - b_2, u_1 - u_2)).$$

Aplicando (1.29) e a desigualdade de Cauchy-Schwarz,

$$c_1 \|u_1 - u_2\|^2 \leq \|b_1 - b_2\| \|u_1 - u_2\|,$$

o que implica (1.32).

No que segue, aplicaremos o Teorema de ponto fixo de Banach para resolução de integrais do tipo Fredholm e Volterra.

1.4.2 Equação integral de Fredholm

Nesta Seção provaremos existência e unicidade da equação de Fredholm linear de segunda espécie dada por

$$\varphi(x) = f(x) + \lambda \int_a^b k(x, y)\varphi(y)dy \quad (1.33)$$

com as funções $k : [a, b] \times [a, b] \rightarrow \mathbb{R}$ e $f : [a, b] \rightarrow \mathbb{R}$ contínuas.

Para tratarmos as questões de existência e unicidade vamos escrever (1.33) como uma equação de operadores da seguinte forma:

$$\varphi(x) = f(x) + \lambda K[\varphi](x)$$

com K é definido como $K[\varphi](x) = \int_a^b k(x, y)\varphi(y)dy$.

Agora, se considerarmos o operador $T : C([a, b]) \rightarrow C([a, b])$ dado por

$$T[\varphi](x) = f + \lambda K[\varphi](x)$$

temos que o problema de determinar existência e unicidade da equação (1.33) é equivalente ao problema de determinar um único ponto fixo de T , ou seja, obter φ tal que

$$T\varphi = \varphi \Leftrightarrow \varphi = f + \lambda K\varphi.$$

Vamos aplicar o Teorema do ponto fixo de Banach. Como o espaço $C([a, b])$ é um espaço de Banach com a norma do supremo $\|\cdot\|_\infty$ devemos prova que o operador T é uma contração. Para isto, escrevemos

$$\|T\varphi_1 - T\varphi_2\|_\infty = |\lambda| \|K[\varphi_1] - K[\varphi_2]\|_\infty.$$

Como $k(x, y)$ contínua tem-se $|k(x, y)| \leq M$ em $[a, b] \times [a, b]$ e K é um operador linear tem-se

$$\begin{aligned} |K[\varphi_1](x) - K[\varphi_2](x)| &= |K[\varphi_1 - \varphi_2](x)| \leq \int_a^b |k(x, y)| |\varphi_1(y) - \varphi_2(y)| dy \\ &\leq M \int_a^b |\varphi_1(y) - \varphi_2(y)| dy \leq M(b-a) \|\varphi_1 - \varphi_2\|_\infty. \end{aligned}$$

Logo,

$$\|K[\varphi_1] - K[\varphi_2]\|_\infty \leq M(b-a) \|\varphi_1 - \varphi_2\|_\infty$$

e obtemos

$$\|T\varphi_1 - T\varphi_2\|_\infty \leq |\lambda|M(b-a)\|\varphi_1 - \varphi_2\|_\infty.$$

Portanto, para

$$|\lambda| < \frac{1}{M(b-a)} \quad (1.34)$$

temos que T é um operador de contração e a equação (1.33) tem uma única solução em $C([a, b])$. Além disso, as sucessivas aproximações $\varphi_1, \varphi_2, \dots, \varphi_n, \dots$ desta solução são dadas por

$$\varphi_{n+1}(x) = f(x) + \lambda \int_a^b k(x, y)\varphi_n(y)dy$$

com φ_1 qualquer função contínua.

Ressaltamos que o método das sucessivas aproximações só poderá ser aplicado na equação (1.33) para $|\lambda|$ satisfazendo a condição (1.34).

1.4.3 Equação de Fredholm não linear

Nesta seção provaremos existência e unicidade da equação de Fredholm não linear de segunda espécie dada por

$$\varphi(x) - \lambda \int_a^b k(x, y, \varphi(y))dy = 0 \quad (1.35)$$

com a função núcleo $k : [a, b] \times [a, b] \rightarrow \mathbb{R}$ uma função *Lipschitziana na terceira variável*, ou seja, existe $\beta > 0$ tal que k satisfaz a condição

$$|k(x, y, s_1) - k(x, y, s_2)| < \beta|s_1 - s_2|.$$

Analogamente, podemos escrever a equação integral (1.35) na forma de uma equação de operadores do seguinte modo:

$$T[\varphi](x) = \lambda K[\varphi]$$

com K dado por $K[\varphi](x) = \int_a^b k(x, y, \varphi(y))dy$.

Novamente devemos provar que o operador T é uma contração. Como k é lipschitz na terceira variável tem-se

$$\begin{aligned} |K[\varphi_1](x) - K[\varphi_2](x)| &\leq \int_a^b |k(x, y, \varphi_1(y)) - k(x, y, \varphi_2(y))|dy \\ &\leq \beta \int_a^b |\varphi_1(y) - \varphi_2(y)|dy \leq \beta(b-a)\|\varphi_1 - \varphi_2\|_\infty. \end{aligned}$$

Logo,

$$\|T\varphi_1 - T\varphi_2\|_\infty \leq |\lambda|\beta(b-a)\|\varphi_1 - \varphi_2\|_\infty.$$

Portanto, para $|\lambda| < \frac{1}{\beta(b-a)}$ temos que T uma contração e a equação (1.35) tem uma única solução em $C([a, b])$. Além disso, as sucessivas aproximações $\varphi_1, \varphi_2, \dots, \varphi_n, \dots$ desta solução são dadas por

$$\varphi_{n+1}(x) = f(x) + \lambda \int_a^b k(x, y, \varphi_n(y)) dy$$

com φ_1 qualquer função contínua.

Ressaltamos que o método das sucessivas aproximações só poderá ser aplicado na equação (1.35) para $|\lambda|$ suficientemente pequeno.

1.4.4 Equação integral de Volterra

Nesta seção provaremos existência e unicidade da equação de Volterra linear dada por

$$\varphi(x) = f(x) + \lambda \int_a^x k(x, y)\varphi(y) dy \quad (1.36)$$

com as funções $k : [a, b] \times [a, b] \rightarrow \mathbb{R}$ e $f : [a, b] \rightarrow \mathbb{R}$ contínuas.

Vamos aplicar o Teorema da contração de Banach para mostrar que a equação (1.36) tem uma única solução para *qualquer* λ e não suficientemente pequeno como no caso da equação de Fredholm. De fato, considere o operador $T : C([a, b]) \rightarrow C([a, b])$ dado por

$$T[\varphi](x) = f + \lambda K[\varphi](x)$$

com $K[\varphi](x) = \int_a^x k(x, y)\varphi(y) dy$. Então,

$$\begin{aligned} |K[\varphi_1](x) - K[\varphi_2](x)| &= |K[\varphi_1 - \varphi_2](x)| \leq \int_a^x |k(x, y)| |\varphi_1(y) - \varphi_2(y)| dy \\ &\leq M \int_a^x |\varphi_1(y) - \varphi_2(y)| dy \leq M(x-a) \|\varphi_1 - \varphi_2\|_\infty. \end{aligned}$$

Assim,

$$\begin{aligned} |K[K[\varphi_1]](x) - K[K[\varphi_2]](x)| &\leq \int_a^x |k(x, y)| |K[\varphi_1](y) - K[\varphi_2](y)| dy \\ &\leq M \int_a^x M(y-a) \|\varphi_1 - \varphi_2\|_\infty dy \\ &\leq M^2 \|\varphi_1 - \varphi_2\|_\infty \int_a^x (y-a) dy \\ &\leq M^2 \frac{(x-a)^2}{2} \|\varphi_1 - \varphi_2\|_\infty. \end{aligned}$$

Procedendo sucessivamente até n obtemos

$$\begin{aligned} |K^n[\varphi_1](x) - K^n[\varphi_2](x)| &\leq M^n \frac{(x-a)^n}{n!} \|\varphi_1 - \varphi_2\|_\infty \\ &\leq M^n \frac{(b-a)^n}{n!} \|\varphi_1 - \varphi_2\|_\infty, \end{aligned}$$

o que implica

$$\|T^n \varphi_1 - T^n \varphi_2\|_\infty \leq \lambda^n M^n \frac{(b-a)^n}{n!} \|\varphi_1 - \varphi_2\|_\infty.$$

Portanto, para *qualquer* λ e n suficientemente grande tem-se

$$\lambda^n M^n \frac{(b-a)^n}{n!} < 1,$$

e, portanto T é uma contração para n suficientemente grande, então pelo Teorema 1.8 T tem um único ponto fixo. Portanto, a equação (1.36) tem uma única solução.

1.5 Equações diferenciais em espaços de Banach

Seja X um espaço de Banach e considere o problema de valor inicial

$$\begin{cases} u'(t) = f(t, u(t)) & \text{em } [t_0, b] \\ u(t_0) = u_0 \end{cases} \quad (1.37)$$

com $u_0 \in X$ e $f : [t_0, b] \times X \rightarrow X$ é contínua e Lipschitz na segunda variável:

$$|f(t, s_1) - f(t, s_2)| \leq \beta |s_1 - s_2|.$$

O problema da equação diferencial (1.37) é equivalente a equação integral

$$u(t) = u_0 + \int_{t_0}^t f(s, u(s)) ds. \quad (1.38)$$

Note que (1.38) é da forma $u = T(u)$ e (1.38) é uma equação integral de Volterra não linear.

Deste modo, temos o seguinte o processo iterativo:

$$u_n(t) = u_0 + \int_{t_0}^t f(s, u_{n-1}(s)) ds. \quad (1.39)$$

Para mostramos a existência e unicidade de (1.37) usamos a forma equivalente (1.38) e aplicamos o Teorema de ponto fixo de Banach. Para isto, considere $X = C([t_0, b])$ com a

norma do máximo $\|\cdot\|_\infty$, então

$$\begin{aligned} \|Tu_1 - Tu_2\|_\infty &\leq \max_{t \in [t_0, b]} \int_{t_0}^t |f(s, u_1(s)) - f(s, u_2(s))| ds \\ &\leq \beta b \|u_1 - u_2\|_\infty \end{aligned}$$

Se $\beta b < 1$ temos que T é uma contração e (1.39) tem uma única solução. Como (1.38) e (1.39) são equivalentes temos que o problema de Cauchy (1.38) tem uma única solução para $\beta b < 1$.

Em resumo,

Teorema 1.10 (Teorema de Picard-Lindelof em espaço de Banach) *Suponha que $f : [t_0, b] \rightarrow X$ é contínua e Lipchitz com relação na segunda variável:*

$$|f(t, s_1) - f(t, s_2)| \leq \beta |s_1 - s_2|$$

com β é uma constante independente de t .

Se $\beta b < 1$, então o problema de valor inicial (1.37) tem uma única solução continuamente diferenciável $u(\cdot)$ em $[t_0, b]$.

Além disso, qualquer valor inicial u_1 tal que $\|u_0 - u_1\|_\infty < b$, o método iterativo (1.39) converge para a solução $u(t)$, ou seja,

$$\max_{\|t-t_0\| \leq b} \|u_n(t) - u(t)\| \rightarrow 0 \text{ quando } n \rightarrow \infty.$$

Capítulo 2

Cálculo em espaços de Banach

Neste capítulo, faremos uma breve introdução do Cálculo em espaço de Banach. Os conceitos e resultados tratados neste capítulo serão importantes para generalizarmos o Método de Newton, pois precisaremos generalizar o conceito de derivada para espaços de Banach. Além disso, para provarmos a convergência do método aplicaremos uma versão do Teorema do valor médio em espaços de Banach.

Para melhor entendimento da generalização do conceito de derivada, primeiro, trataremos o caso em \mathbb{R}^n . No caso do Cálculo integral, só apresentaremos os resultados para funções do tipo $f : [a, b] \rightarrow X$.

Não é nosso objetivo fazer um estudo detalhado do Cálculo em espaço de Banach, para maiores detalhes consulte [1].

2.1 Cálculo Diferencial

Definição 2.1 *Seja $f : I \rightarrow \mathbb{R}$, I um intervalo aberto. Dizemos que f é diferenciável em $x_0 \in I$ se, e somente se existe um operador linear $L : \mathbb{R} \rightarrow \mathbb{R}$ dada por $L(h) = ah$ com $a = f'(x_0)$ tal que para todo $\epsilon > 0$ existe $\delta > 0$ satisfazendo*

$$|h| < \delta \Rightarrow |f(x_0 + h) - f(x_0) - L(h)| \leq \epsilon |h|, \quad (2.1)$$

equivalentemente,

$$\lim_{|h| \rightarrow 0} \frac{|f(x_0 + h) - f(x_0) - L(h)|}{|h|} = 0. \quad (2.2)$$

Observação 2.1 *Podemos mostrar que L é única.*

Ressaltamos que a definição de derivada dada em (2.1) (ou (2.2)) é a mais adequada para obtermos a generalização do conceito, pois não usa a divisão por um elemento do espaço, mas pela norma do elemento.

Observe que no caso unidimensional, a aplicação linear L é o diferencial da função f em x_0 . Por esta razão, usaremos a notação $L = Df(x_0)$.

Vamos generalizar o conceito de derivada para uma função $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ usando (2.1) (ou (2.2)). Esta definição de diferenciabilidade foi dada pelos matemáticos Fréchet e Stolz.

Definição 2.2 *Sejam $U \subset \mathbb{R}^n$ um aberto e $x_0 \in U$. Dizemos que uma função $f : U \rightarrow \mathbb{R}^m$ é fortemente diferenciável em x_0 , se existe um operador linear $Df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tal que dado $\epsilon > 0$ existe $\delta > 0$ satisfazendo*

$$\|h\| < \delta \Rightarrow \|f(x_0 + h) - f(x_0) - Df(x_0)(h)\| < \epsilon \|h\|, \quad \forall h \in \mathbb{R}^n, \quad (2.3)$$

equivalentemente,

$$\lim_{\|h\| \rightarrow 0} \frac{\|f(x_0 + h) - f(x_0) - L(h)\|}{\|h\|} = 0. \quad (2.4)$$

O operador $Df(x_0)$ é chamado derivada de f .

Observação 2.2 *O operador $Df(x_0)$ também é chamado de Derivada de Fréchet ou diferencial de f em x_0 e denotado por $f'(x_0)$ ou $df(x_0)$, respectivamente.*

Teorema 2.1 *Seja $U \subset \mathbb{R}^m$ um aberto. Se $f : U \rightarrow \mathbb{R}^n$ é diferenciável em $x_0 \in U$ então a aplicação $Df(x_0)$ é única.*

Exemplo 2.1 *Seja $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma transformação linear, então $DL(x_0) = L, \forall x_0 \in \mathbb{R}^n$.*

De fato, dado $\epsilon > 0$ tem-se

$$\|L(x_0 + h) - L(x_0) - L(h)\| = \|L(x_0) + L(h) - L(x_0) - L(h)\| = \|0\| \leq \epsilon \|h\|.$$

Então, para qualquer $\delta > 0$ tem-se

$$\|h\| < \delta \Rightarrow \|L(x_0 + h) - L(x_0) - L(h)\| \leq \epsilon \|h\|.$$

Logo, pela unicidade, $DL(x_0) = L$.

Teorema 2.2 *Se $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ é diferenciável em um ponto x_0 interior de U então f é contínua em x_0 .*

Prova: Como f diferenciável em x_0 temos que dado $\epsilon > 0$ existe $\delta_1 > 0$ tal que

$$\|x - x_0\| < \delta_1 \Rightarrow \|f(x) - f(x_0) - Df(x_0)(x - x_0)\| < \epsilon \|x - x_0\|.$$

Tomando $\epsilon = 1$ tem-se

$$\|x - x_0\| < \delta_1(1) \Rightarrow \|f(x) - f(x_0) - Df(x_0)(x - x_0)\| < \|x - x_0\| \quad (2.5)$$

Por outro lado, como todo operador linear definido num espaço de dimensão finita é limitado, então existe uma constante M_1 tal que

$$\|Df(x_0)(x - x_0)\| \leq M_1 \|x - x_0\| \quad (2.6)$$

Mas, pela desigualdade triangular tem-se

$$\begin{aligned} \|f(x) - f(x_0)\| &= \|f(x) - f(x_0) - Df(x_0)(x - x_0) + Df(x_0)(x - x_0)\| \\ &\leq \|f(x) - f(x_0) - Df(x_0)(x - x_0)\| + \|Df(x_0)(x - x_0)\|. \end{aligned}$$

Agora, aplicando (2.5) e (2.6) obtemos

$$\|x - x_0\| \leq \delta_1 \Rightarrow \|f(x) - f(x_0)\| \leq M \|x - x_0\|$$

com $M = 1 + M_1$. Isto mostra que f satisfaz a condição de Lipchitz.

Agora, escolhendo $\delta = \min\left(\delta_1, \frac{\epsilon}{M}\right)$ concluímos que

$$\|x - x_0\| \leq \delta \Rightarrow \|f(x) - f(x_0)\| \leq \epsilon$$

E a continuidade de f em x_0 está provada. ■

Na seguinte definição generalizaremos o conceito de derivada de direcional.

Definição 2.3 *Sejam $U \subset \mathbb{R}^n$ um aberto, $x_0 \in U$ ponto interior, $u \in \mathbb{R}^n$ e $f : U \rightarrow \mathbb{R}^m$. Dizemos que f tem derivada direcional em x_0 na direção do vetor u se dado $\epsilon > 0$, existe $\delta > 0$ tal que*

$$0 < |t| < \delta \Rightarrow \left\| \frac{f(x_0 + tu) - f(x_0) - D_u f(x_0)}{t} \right\| < \epsilon, \quad \forall t \in \mathbb{R}, \quad (2.7)$$

ou equivalentemente, se existe o seguinte limite:

$$D_u f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + tu) - f(x_0)}{t}.$$

Observação 2.3 *Usamos também as seguintes notações para derivada direcional: $\frac{\partial f}{\partial u}(x_0)$, $f'(x_0; u)$, $f_u(x_0)$.*

Quando $u = e_j$ para $j = 1, 2, \dots, n$, as derivadas direcionais de f em x_0 na direção dos

vetores canônicos são chamadas *derivadas parciais* de f em x_0 . Neste caso, as notações mais usadas são $D_j f(x_0)$ ou $\frac{\partial f}{\partial x_j}(x_0)$.

O seguinte resultado afirma que todo operador linear definido em espaços normados de dimensão finita pode ser representado por uma matriz. Sua demonstração pode ser encontrada em [6].

Proposição 2.1 *Sejam X e Y espaços de dimensão finita com $\dim X = n, \dim Y = m$ e $m, n \geq 1$. Sejam $\{e_1, e_2, \dots, e_n\}$ e $\{b_1, b_2, \dots, b_m\}$ bases de X e Y , respectivamente. Então o operador $T : X \rightarrow Y$ é linear se, e somente se, existe uma matriz $A_{m \times n}$ tal que para $x = \sum_{k=1}^n \alpha_k e_k$ temos que*

$$T(x) = \sum_{i=1}^m \beta_i b_i$$

com $\beta_i = \sum_{j=1}^n a_{ij} \alpha_j$, $A = (a_{ij})$, $a_{ij} \in \mathbb{K}$, $1 \leq i \leq n$, $1 \leq j \leq m$. Além disso, T é contínuo.

Usaremos o Teorema 2.1 para provar o seguinte resultado:

Teorema 2.3 *Se $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ é diferenciável em um ponto x_0 interior de U então as derivadas parciais $D_i f_j(x_0)$ existem para $i = 1, \dots, n$ e $j = 1, \dots, m$ e $Df(x_0)$ é dada por*

$$Df(x_0)(h) = \begin{bmatrix} D_1 f_1(x_0) & \dots & D_n f_1(x_0) \\ \vdots & \ddots & \vdots \\ D_1 f_m(x_0) & \dots & D_n f_m(x_0) \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix}$$

com $h = (h_1, \dots, h_n)$.

Prova: Como $Df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é um operador linear pelo Teorema 2.7 temos que $Df(x_0)$ pode ser representada por

$$Df(x_0)(h) = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \quad (2.8)$$

Como f diferenciável em x_0 temos

$$\lim_{\|h\| \rightarrow 0} \frac{\|f(x_0 + h) - f(x_0) - Df(x_0)(h)\|}{\|h\|} = 0.$$

Tomando $h = te_i$ e escrevendo (2.8) em termos das componentes, obtemos

$$\lim_{|t| \rightarrow 0} \frac{|f_j(x_0 + te_i) - f_j(x_0) - a_{ij}t|}{|t|} = 0,$$

com a_{ij} o (i, j) elemento da matriz $Df(x_0)$, dado pela i -ésima componente do vetor $Df(x_0).e_j = Df(x_0)$ aplicado no j -ésimo vetor canônico e_j .

Portanto,

$$\lim_{|t| \rightarrow 0} \frac{f_j(x_0 + te_i) - f_j(x_0)}{t} - a_{ij} = 0.$$

Logo, $D_j f_i(x_0)$ existe e pela unicidade da derivada parcial, concluímos que $D_j f_i(x_0) = a_{ij}$.

■

Observação 2.4 Para $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ o Teorema 2.3 afirma que

$$Df(x_0)(h) = \nabla f(x_0).h.$$

Matriz jacobiana

Sejam $U \subset \mathbb{R}^n$ um aberto e $f : U \rightarrow \mathbb{R}^m$. Se as derivadas parciais das funções componentes f_j de f existem em x_0 podemos formar a seguinte matriz

$$(D_i f_j(x_0)) = \begin{bmatrix} D_1 f_1(x_0) & \dots & D_n f_1(x_0) \\ \vdots & \ddots & \vdots \\ D_1 f_m(x_0) & \dots & D_n f_m(x_0) \end{bmatrix}_{m \times n}$$

chamada *matriz jacobiana* de f no ponto x_0 . Se f for diferenciável em x_0 então pelo Teorema 2.3, a matriz diferencial coincide com a matriz jacobiana, ou seja, $Df(x_0) = (D_i f_j(x_0))$. Se f é diferenciável em x_0 o determinante

$$Jf(x_0) = \begin{vmatrix} D_1 f_1(x_0) & \dots & D_n f_1(x_0) \\ \vdots & \ddots & \vdots \\ D_1 f_m(x_0) & \dots & D_n f_m(x_0) \end{vmatrix}$$

é chamado *Jacobiano* de f em x_0 .

Sejam $U \subset \mathbb{R}^n$ um aberto e $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ espaço vetorial das aplicações lineares de \mathbb{R}^n em \mathbb{R}^m . Se $f : U \rightarrow \mathbb{R}^m$ é uma função diferenciável em cada ponto em U então podemos considerar a função diferencial Df dada por

$$\begin{aligned} Df : U &\rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \\ x_0 &\rightarrow f'(x_0) \end{aligned}$$

com

$$\begin{aligned} Df(x_0) : \mathbb{R}^n &\rightarrow \mathbb{R}^m \\ h &\rightarrow Df(x_0)(h) \end{aligned}$$

Exemplo 2.2 *Sejam $f(x, y) = (x^2 + y^2, xy)$ e $x_0 = (a, b)$ então*

$$Df(x_0) = \begin{bmatrix} 2a & 2b \\ b & a \end{bmatrix}$$

Logo, para $h = (h, k)$ tem-se $Df(x_0)(h) = (2ah + 2bk, bh + ak)$.

O seguinte Teorema caracteriza o diferencial de uma função diferenciável em termos da derivada direcional:

Teorema 2.4 *Sejam $U \subset \mathbb{R}^n$ um aberto e $f : U \rightarrow \mathbb{R}^m$. Se f é diferenciável em x_0 , então a derivada direcional de f em x_0 na direção do vetor u existe e*

$$Df(x_0)(u) = D_u f(x_0)$$

Observação 2.5 *A existência de todas as derivadas direcionais em x_0 não implica a diferenciabilidade de f em x_0 . De fato, podemos mostrar que a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por*

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

não é diferenciável em $(0, 0)$, mas têm todas as derivadas direcionais em $(0, 0)$.

2.1.1 Funções de classe C^r

Teorema 2.5 *Sejam $U \subset \mathbb{R}^n$ um aberto e $f : U \rightarrow \mathbb{R}^m$. Se as derivadas parciais das funções componentes $D_i f_j(x_0)$ de f existem e são contínuas em U , dizemos que a função f é continuamente diferenciável ou de classe $C^1(U)$.*

Se as funções componentes $D_i f_j(x_0)$ são diferenciáveis em U , ou seja, podemos calcular suas derivadas parciais $D_k(D_i f_j(x_0))$, então definimos as *derivadas parciais de segunda ordem de f* . Analogamente, podemos definir as derivadas parciais de ordem r , $r \in \mathbb{N}$ de f .

Dizemos que uma função f é de classe $C^r(U)$ se, e somente se, cada função $D_i f_j(x_0)$ é de classe $C^{r-1}(U)$ em U .

Teorema 2.6 *Sejam $U \subset \mathbb{R}^n$ um aberto e conexo e $f : U \rightarrow \mathbb{R}^n$ de classe $C^1(U)$. Se $Df(x)$ é lipschitz contínua em $x \in U$ com constante $L > 0$. Então, para quaisquer $x + h \in U$ tem-se*

$$\|f(x+h) - f(x) - Df(x)(h)\| \leq \frac{L}{2} \|h\|^2.$$

2.1.2 Diferenciabilidade em espaços de Banach

Nesta seção apresentaremos o conceito de diferenciabilidade em espaços mais gerais, em particular espaços de Banach.

Definição 2.4 *Sejam X e Y espaços de Banach, $U \subset X$ aberto. Dizemos que a função $f : U \rightarrow Y$ é Fréchet diferenciável num ponto $x_0 \in U$ se existe um operador linear limitado $Df(x_0) : X \rightarrow Y$ tal que $\epsilon > 0$ existe $\delta > 0$ satisfazendo*

$$\|h\| < \delta \Rightarrow \|f(x_0 + h) - f(x_0) - Df(x_0)(h)\| \leq \epsilon \|h\|, \quad \forall h \in U.$$

equivalentemente,

$$\lim_{\|h\| \rightarrow 0} \frac{\|f(x_0 + h) - f(x_0) - Df(x_0)(h)\|}{\|h\|} = 0.$$

O operador $Df(x_0)$ é chamado derivada de Fréchet ou derivada forte de f em x_0 .

Se f é Fréchet diferenciável em todos os pontos de U dizemos que f é diferenciável em U . Nesse caso, a aplicação

$$x \in U \mapsto Df(x) \in \mathcal{L}(X, Y)$$

é chamada a *diferencial* de f em U e representada por Df .

Observação 2.6 *Observe que para o caso de dimensão finita, na Definição 2.1, está implícito que diferencial $Df(x_0)$ é uma aplicação linear contínua, pois tal continuidade é garantida pelo Teorema 1.1.*

O seguinte teorema é uma definição alternativa de derivada de Fréchet.

Proposição 2.2 *A função f é Fréchet diferenciável em x_0 se e somente se, existe $E \in \mathcal{L}(X, Y)$ tal que*

$$r(h) = f(x_0 + h) - f(x_0) - E(h) \tag{2.9}$$

com $\lim_{\|h\| \rightarrow 0} \frac{\|r(h)\|}{\|h\|} = 0$.

Prova: Suponha que f é Fréchet diferenciável em x_0 então, existe o operador linear contínuo $Df(x_0)$ tal que

$$\lim_{\|h\| \rightarrow 0} \frac{\|f(x_0 + h) - f(x_0) - Df(x_0)(h)\|}{\|h\|} = 0$$

Seja $r : X \rightarrow Y$ definida por

$$r(h) = f(x_0 + h) - f(x_0) - Df(x_0)(h).$$

Então,

$$\frac{\|r(h)\|}{\|h\|} = \frac{\|f(x_0 + h) - f(x_0) - Df(x_0)(h)\|}{\|h\|},$$

o que implica

$$\lim_{\|h\| \rightarrow 0} \frac{\|r(h)\|}{\|h\|} = 0. \quad (2.10)$$

Portanto, $E = Df(x_0)$.

Reciprocamente, suponha que existe $E \in \mathcal{L}(X, Y)$ tal que a equação (2.9) satisfeita e

$$\lim_{\|h\| \rightarrow 0} \frac{\|r(h)\|}{\|h\|} = \lim_{h \rightarrow 0} \frac{\|f(x_0 + h) - f(x_0) - E(h)\|}{\|h\|} = 0.$$

Então, f é diferenciável em x_0 , e logo, existe $Df(x_0)$ e $E = Df(x_0)$. ■

Exemplo 2.3 *Seja X um espaço de Banach e $f : [a, b] \rightarrow X$. Se f é diferenciável em t_0 então, pela Proposição 2.2, tem-se*

$$f(t) - f(t_0) = f'(t_0)(t - t_0) + E(t)$$

com $\lim_{t \rightarrow t_0} \frac{E(t)}{t - t_0} = 0$.

O teorema a seguir nos assegura o importante resultado que diferenciabilidade implica em continuidade. A prova é similar a do Teorema 2.2 e será omitida.

Teorema 2.7 *Seja $f : U \rightarrow Y$ com U um aberto de X . Se f é diferenciável em $x_0 \in U$ então f é contínua em x_0 .*

Exemplo 2.4 *Se $L : X \rightarrow Y$ é um operador linear então, $\forall x_0 \in X$, $DL(x_0) = L$.*

A prova é análoga ao caso n -dimensional e será omitida. Note que não estamos dizendo que DL e L são idênticas, mas que DL tem o mesmo valor que L para todos os pontos x_0 .

Exemplo 2.5 *Seja H um espaço de Hilbert com produto interno $((\cdot, \cdot))$. Considere a função $f : H \rightarrow \mathbb{R}$ dada por*

$$f(x) = \|x\|^2 = ((x, x)).$$

Então, f é Fréchet diferenciável em $x_0 \in H$ e $Df(x_0)(h) = 2((x_0, h))$.

De fato, note que

$$\begin{aligned} |f(x_0 + h) - f(x_0) - Df(x_0)(h)| &= | \|x_0 + h\|^2 - \|x_0\|^2 - 2((x_0, h)) | \\ &= | ((x_0 + h, x_0 + h)) - \|x_0\|^2 - 2((x_0, h)) | \\ &= \|h\|^2. \end{aligned}$$

Escolha $\delta = \sqrt{\epsilon}$.

Note que f não é diferenciável em $x_0 = 0$. Suponha, por contradição, que f é diferenciável em $x_0 = 0$. Então,

$$\frac{E(h)}{\|h\|} + \frac{r(h)}{\|h\|} = \|h\|$$

para $E \in \mathcal{L}(H, \mathbb{R})$. Substituindo h por $-h$ temos que

$$-\frac{E(h)}{\|h\|} + \frac{r(h)}{\|h\|} = \|h\|,$$

o que implica $\|h\| = \frac{r(h)}{\|h\|}$, o que é um absurdo!

Exemplo 2.6 Seja H um espaço de Hilbert com produto interno $((\cdot, \cdot))$. Se $f : H \rightarrow \mathbb{R}$ é Fréchet diferenciável em x_0 . Então,

$$Df(x_0)(h) = ((\nabla f(x_0), h)). \quad (2.11)$$

De fato, seja $Df(x_0) : H \rightarrow \mathbb{R}$ a diferencial de Fréchet de f em x_0 . Como $Df(x_0) \in H'$, pelo Teorema da Representação de Riesz (veja Teorema (1.4)), existe um elemento $\nabla f(x_0) \in H$ tal que $Df(x_0)(h) = ((\nabla f(x_0), h))$, $\forall h \in H$.

O elemento $\nabla f(x_0)$ é chamado gradiente de f em x_0 . Note que $\nabla f : H \rightarrow H$.

Definiremos agora a generalização do conceito de derivada direcional para espaços Banach, chamada de *Derivada de Gâteaux*.

Definição 2.5 Seja X e Y espaços de Banach. Dizemos que a função $f : X \rightarrow Y$ é Gâteaux diferenciável em $x_0 \in X$ na direção de h se, para $x_0, h \in X$, existe $\delta f(x_0; h) \in Y$ tal que dado $\epsilon > 0$ existe $\delta > 0$ satisfazendo

$$|t| < \delta \Rightarrow \left\| \frac{f(x_0 + th) - f(x_0)}{t} - \delta f(x_0, h) \right\| < \epsilon,$$

equivalentemente,

$$\delta(x_0, h) = \lim_{t \rightarrow 0} \left\| \frac{f(x_0 + th) - f(x_0)}{t} \right\| = 0.$$

Se $\delta(x_0; h)$ existe para todo $h \in X$ e a aplicação $df(x_0) : h \mapsto \delta(x_0; h)$ é linear e contínua, então $df(x_0)$ é chamada *derivada de Gâteaux* de f em x_0 .

Observação 2.7 A terminologia sobre a derivada de Gâteaux não é uniforme. Alguns autores não assume a linearidade de $df(x_0)$.

No caso de dimensão finita, mostramos que a diferencial de uma função diferenciável é caracterizado em termos da derivada direcional (veja Teorema 2.4). O seguinte Teorema é a generalização deste resultado para espaços de Banach:

Teorema 2.8 *Se $f : X \rightarrow Y$ é Fréchet diferenciável em $x_0 \in X$ então f é Gâteaux diferenciável em x_0 e $Df(x_0)(u) = \delta f(x_0; u), \forall u \in X$.*

Prova: Como f é Fréchet diferenciável em x_0 , então existe um operador linear limitado $Df(x_0) : X \rightarrow Y$ tal que dado $\epsilon > 0$, existe $\delta > 0$ satisfazendo

$$\|h\| < \delta \Rightarrow \|f(x_0 + h) - f(x_0) - Df(x_0)(h)\| < \epsilon\|h\|, \forall h \in X.$$

Como vale para todo $u \in X$, podemos tomar $h = tu$, com $t \in \mathbb{R}$ e $u \in X, t \neq 0$. Assim,

$$\|tu\| < \delta \Rightarrow \|f(x_0 + tu) - f(x_0) - Df(x_0)(tu)\| < \epsilon\|tu\|, \forall h \in X.$$

Portanto, para $|t| < \frac{\delta}{\|u\|}$ e $\|u\| \neq 0$ obtemos

$$\|f(x_0 + tu) - f(x_0) - tDf(x_0)(u)\| < \epsilon|t|\|u\|$$

pela linearidade de $Df(x_0)$, o que implica

$$|t| < \frac{\delta}{\|u\|} \Rightarrow \left\| \frac{f(x_0 + tu) - f(x_0)}{t} - Df(x_0)(u) \right\| < \epsilon\|u\|$$

Portanto, tomando $\delta_1 = \frac{\delta}{\|u\|}$ temos que f é Gâteaux diferenciável em x_0 e $\delta f(x_0, u) = Df(x_0)$.

■

Exemplo 2.7 *Sejam $X = C^1([0, 1])$ e $Y = C([0, 1])$ com a norma do máximo $\|\cdot\|$. Considere*

$$f(x) = x^2 \frac{dx}{dt}.$$

Assim,

$$\begin{aligned} f(x_0 + th) - f(x_0) &= (x_0 + th)^2 \frac{d}{dt}(x_0 + th) - x_0^2 \frac{dx_0}{dt} \\ &= t \left(x_0^2 \frac{dh}{dt} + 2x_0 \frac{dx_0}{dt} \right) + O(h^2). \end{aligned}$$

Então, a derivada de Gâteaux é dada por

$$\delta(x_0; h)[\cdot] = x_0^2 \frac{d}{dt}[\cdot] + 2x_0 \frac{dx_0}{dt}[\cdot].$$

O próximo Teorema é a generalização do Teorema 2.5 em termos da derivada de Gâteaux, sua prova pode ser encontrada em [4, p. 123]:

Teorema 2.9 *Assuma que $df(x)$ existe numa vizinhança de $x_0 \in X$. Se $x \mapsto df(x)$ é contínua em x_0 (como função de $X \rightarrow \mathcal{L}(X, Y)$) então $Df(x_0)$ existe.*

2.1.3 Regras de diferenciação

Agora, enunciaremos algumas regras de diferenciação. Se não especificarmos o tipo de derivada, então o resultado é válido para a diferencial de Fréchet.

Proposição 2.3 (Regra da Soma) *Sejam X e Y espaços normados. Se $f, g : U \subset X \rightarrow Y$ são diferenciáveis em x_0 , então para quaisquer escalares α e β , $\alpha f + \beta g$ é diferenciável em x_0 e*

$$D(\alpha f + \beta g)(x_0) = \alpha Df(x_0) + \beta Dg(x_0)$$

Proposição 2.4 (Regra da Cadeia) *Sejam X, Y e Z espaços normados. Seja $f : U \subset Y \rightarrow Z$ e $g : W \subset X \rightarrow Y$ com $g(W) \subset U$. Assuma que u_0 é um ponto interior de W , $g(u_0)$ em ponto interior de U . Se f é Fréchet diferenciável em x_0 e g é Gâteaux diferenciável em u_0 com $g(u_0) = x_0$. Então, $f \circ g$ é Gâteaux diferenciável em u_0 e*

$$\delta(f \circ g)(u_0; h) = Df(x_0)(\delta g(u_0; h)).$$

Vejamos um exemplo.

Exemplo 2.8 *Seja $X = Y = C([a, b])$ com a norma do máximo. Assuma que $g \in C([a, b])$, $k \in C([a, b] \times [a, b] \times \mathbb{R})$. Então, podemos definir o operador $G : V \rightarrow W$ pela fórmula*

$$G(u)(t) = g(t) + \int_a^b k(t, s, u(s)) ds$$

Este operador é chamado operador integral de Urysohn.

Seja $u_0 \in C([a, b])$ tal que

$$\frac{\partial k}{\partial u}(t, s, u_0(s)) \in C([a, b] \times [a, b]).$$

Então, G é Fréchet diferenciável em u_0 e

$$(G'(u_0)h)(t) = \int_a^b \frac{\partial k}{\partial u}(t, s, u_0(s))h(s) ds, \quad h \in V.$$

2.1.4 Derivadas Parciais

No que segue, vamos generalizar do Teorema 2.5 em termos das derivadas parciais de f , sua prova pode ser encontrada em

Definição 2.6 *Sejam X_1, X_2 e Y espaços de normados e $f : X_1 \times X_2 \rightarrow Y$. Para $a_2 \in X_2$ fixo, se $f_1 : x_1 \mapsto f(x_1, a_2)$ tem derivada de Gâteaux (ou Fréchet) em $a_1 \in X_1$ então $df_1(a_1)$*

(ou $Df_1(a_1)$) é chamada a derivada parcial Gâteaux (ou Fréchet) de f em (a_1, a_2) com relação a primeira variável e representada por $d_1f(a_1, a_2)$ (ou $D_1f(a_1, a_2)$).

De modo análogo, podemos definir a derivada parcial $d_2f(a_1, a_2)$ (ou $D_2f(a_1, a_2)$) com relação a segunda variável.

Se $df(a_1, a_2)$ existe então existem $d_1f(a_1, a_2)$ e $d_2f(a_1, a_2)$ e

$$df(a_1, a_2)(h_1, h_2) = d_1f(a_1, a_2)(h_1) + d_2f(a_1, a_2)(h_2). \quad (2.12)$$

Para termos a reversa desta afirmação precisamos de mais hipótese, como mostra o seguinte Teorema:

Teorema 2.10 ([4], p. 124) *Suponha que $d_1f(a_1, a_2)$ no ponto (a_1, a_2) e $d_2f(a_1, a_2)$ existem numa vizinhança do ponto (a_1, a_2) e a aplicação $d_2f : X_1 \times X_2 \rightarrow \mathcal{L}(X_2, Y)$ é contínua em (a_1, a_2) . Então, $df(a_1, a_2)$ existe e vale (2.12).*

A prova desta proposição e a demonstração que a existência de derivadas parciais em um ponto não implica a existência da derivada de Fréchet ou Gâteaux.

Corolário 2.1 ([1], p. 225) *Sejam $A \subset X_1 \times X_2$ um aberto e $f : X_1 \times X_2 \rightarrow Y$. Então, $f \in C^1(A)$ se, e somente se, $f_1, f_2 \in C^1(A)$.*

2.2 Cálculo Integral

Sejam X um espaço de Banach e $f : [a, b] \rightarrow X$. Considere uma partição qualquer dada por

$$\mathcal{P} = \{a = t_0 \leq t_1 \leq \dots \leq t_{n-1} \leq t_n = b\}$$

e os números $c_i \in [t_{i+1}, t_i]$. Considere $\Delta t_i = t_{i+1} - t_i$, $\|\Delta\| = \max\{\Delta t_0, \Delta t_1, \dots, \Delta t_{n-1}\}$ e a soma de Riemann

$$S_n(f) = \sum_{i=0}^{n-1} f(c_i) \Delta t_i$$

que é um elemento do espaço X .

Definição 2.7 *Se o limite*

$$\lim_{\|\Delta\| \rightarrow 0} S_n(f) = I$$

existe em X para toda partição \mathcal{P} e quaisquer $c_i \in [t_{i+1}, t_i]$, $0 \leq i \leq n-1$, dizemos que f é integrável em $[a, b]$ e usamos a notação

$$I = \int_a^b f(t) dt.$$

Teorema 2.11 ([12], p. 87) *Se $f(t)$ é contínua ou contínua por partes em $[a, b]$ então $f(t)$ é integrável em $[a, b]$.*

2.2.1 Regras de integração

Teorema 2.12 ([12], p. 88)

- (a) $\int_a^b \alpha f(t) dt = \alpha \int_a^b f(t) dt$ com $\alpha \in \mathbb{R}$;
- (b) $\int_a^b f(t) + g(t) dt = \int_a^b f(t) dt + \int_a^b g(t) dt$;
- (c) $\int_a^c f(t) dt + \int_c^b f(t) dt = \int_a^b f(t) dt$ com $c \in (a, b)$;
- (d) $\left\| \int_a^b f(t) dt \right\| \leq (b - a) \max_{t \in [a, b]} \|f(t)\|$;
- (e) $\left\| \int_a^b f(t) dt \right\| \leq \int_a^b \|f(t)\| dt$.

Teorema 2.13 ([12], p. 91) *Sejam X um espaço de Banach e $f : [a, b] \rightarrow X$. Se $f(t)$ é contínua por partes em $[a, b]$ e contínua em $t_0 \in [a, b]$ então a função*

$$F(t) = \int_a^t f(\tau) d\tau$$

é diferenciável em t_0 e $DF(t_0) = f(t_0)$.

2.2.2 Teorema fundamental do Cálculo

Sejam X um espaço de Banach e $f : [a, b] \rightarrow X$ contínua em $[a, b]$. Dizemos que a função contínua $G : [a, b] \rightarrow X$ é uma **primitiva** de $f(t)$ se $DG(t) = f(t), \forall t \in [a, b]$.

Podemos mostrar que duas primitivas de $f(t)$ diferem por uma constante e que qualquer primitiva de $f(t)$ é da forma

$$G(t) = \int_a^t f(\tau) d\tau + x_0$$

com x_0 um elemento fixo de X

Teorema 2.14 ([12], p. 92) *Sejam X um espaço de Banach e $f : [a, b] \rightarrow X$. Se f é de classe $C^1([a, b])$. Então,*

$$\int_a^b \delta Df(t) dt = f(b) - f(a)$$

para todo $t \in [a, b]$.

Teorema 2.15 *Seja X um espaço normado, Y um espaço de Banach e $f : X \rightarrow Y$. Suponha que f tem derivada de Gâteaux em todos os pontos do segmento que une os pontos a e b em X na direção deste segmento, isto é, $\delta f(a + t(b - a); b - a)$ existe $\forall t \in [0, 1]$. Se a aplicação $t \mapsto \delta f(a + t(b - a); b - a)$ é contínua em $[0, 1]$ então*

$$f(b) - f(a) = (b - a) \int_0^1 \delta f(a + t(b - a); b - a) dt.$$

Capítulo 3

Método de Newton generalizado

O teorema do ponto fixo de Banach contém a maioria das propriedades desejáveis de um método numérico. Sob as condições desejadas, a sequência de aproximações está bem definida, e converge para solução do problema original. Além disso, sabemos que a taxa de convergência é linear (veja expressão (1.21)), e também temos uma estimativa de erro a priori (veja (1.19)). Esta estimativa pode ser usada para determinar o número de iterações necessárias para termos uma boa solução aproximada.

Neste capítulo, descrevemos o método de Newton analisando suas principais propriedades. A nossa abordagem será aplicação direta do Teorema do ponto fixo de Banach (veja Teorema 1.7).

3.1 Caso unidimensional

Dada a função real $f : \mathbb{R} \rightarrow \mathbb{R}$ estamos interessados em encontrar as raízes reais da equação

$$f(x) = 0, \quad x \in \mathbb{R}. \quad (3.1)$$

Para aplicarmos o Teorema do ponto fixo de Banach (Teorema 1.7) vamos reformular o problema (3.1) na forma

$$x = T(x), \quad x \in \mathbb{R} \quad (3.2)$$

e gerar uma sequência de aproximações da solução do seguinte modo:

$$x_{k+1} = T(x_k). \quad (3.3)$$

Existem muitos modos de escolhermos a função $T(x)$, por exemplo $T(x) = x - c_0 f(x)$ para alguma constante $c_0 \neq 0$. Uma outra escolha poderia ser

$$T(x) = x - \frac{f(x)}{f'(x)}, \quad f'(x) \neq 0. \quad (3.4)$$

Métodos que usam o teorema do ponto fixo são chamados *Métodos iterativos*, em particular para a escolha (3.4) temos o conhecido *Método de Newton*.

A forma geral das funções $T(x)$ para $f : [a, b] \rightarrow \mathbb{R}$ contínua e ξ é uma solução da equação $f(x) = 0$ é dada por

$$T(x) = x + A(x)f(x) \text{ com } A(\xi) \neq 0.$$

Graficamente, uma raiz da equação (3.2) é a abscissa do ponto de intersecção da reta $y = x$ e da curva $y = T(x)$, como mostram as seguintes figuras:

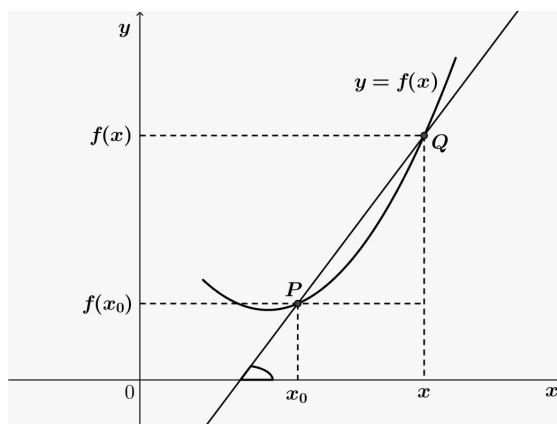


Figura 3.1:

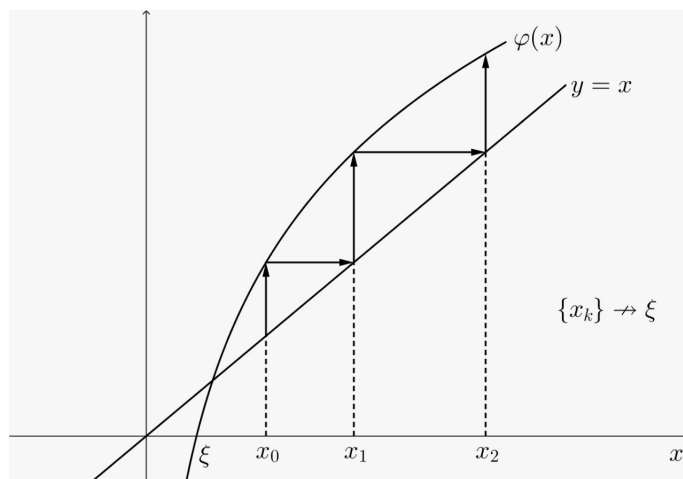


Figura 3.2:

Naturalmente, a convergência dos métodos iterativos depende das propriedades da função $T(x)$. Para aplicarmos o Teorema da Contração de Banach (Teorema 1.7) devemos mostrar que T é uma contração. Além disso, outras condições adicionais, como no caso do método de Newton, garantem maior rapidez de convergência.

No seguinte resultado apresentaremos as condições mínimas necessárias que garantem a convergência do processo iterativo (3.3).

Teorema 3.1 *Suponha que $T(x)$ pertence ao espaço $C^1([a, b])$ tal que $\sup_{x \in [a, b]} |T'(x)| < 1$. Então, o problema (3.2) tem uma única solução ξ e a sequência (x_k) gerada pelo processo iterativo (3.3) converge para ξ .*

Prova: *Vamos mostrar que T é uma contração e aplicar o Teorema 1.7.*

De fato, como $T \in C^1([a, b])$ segue do Teorema do Valor Médio que, para $x_1, x_2 \in [a, b]$, existe $c \in (x_1, x_2)$ tal que

$$|T(x_1) - T(x_2)| = |T'(c)| |x_1 - x_2|.$$

Portanto,

$$|T(x_1) - T(x_2)| \leq \sup_{x \in [x_1, x_2]} |T'(x)| |x_1 - x_2| \leq |x_1 - x_2|,$$

pois $\sup_{x \in [a, b]} |T'(x)| < 1$ por hipótese. Logo, T é uma contração e pelo Teorema da contração o problema (3.2) tem uma única solução ξ e o processo iterativo (3.3) converge para ξ . ■

Observe que, com as hipóteses do Teorema 3.1, temos que a ordem de convergência do processo iterativo (3.3) é linear (veja (1.21)). Para a escolha de $T(x)$ dada em (3.4) temos que a ordem de convergência será quadrática.

De fato, convergência dos métodos iterativos será mais rápida quanto menor for o valor de $|T'(\xi)|$. Observe que

$$T'(x) = \frac{f(x) f''(x)}{(f'(x))^2}$$

e, logo, se $f'(\xi) \neq 0$ tem-se $T'(\xi) = 0$, pois $f(\xi) = 0$ (para mais detalhes consulte [11]).

Em resumo,

Teorema 3.2 *Seja $f \in C^2([a, b])$ com $f'(x) \neq 0$ em $[a, b]$. Suponha que existe uma constante $0 \leq \lambda < 1$ tal que*

$$\left| \frac{f(x) f''(x)}{(f'(x))^2} \right| \leq \lambda, \quad \forall x \in [a, b]. \quad (3.5)$$

Então, a equação $f(x) = 0$ tem uma única solução $x = \xi$ em $[a, b]$ e a sequência $(x_k) \subset [a, b]$ gera pelo processo iterativo

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k \geq 0$$

converge para ξ com $x_0 \in [a, b]$ arbitrário.

Além disso, a ordem de convergência do processo iterativo (3.2) é quadrática, isto é,

$$\lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k^2} = C$$

com C uma constante.

Prova: Como $f \in C^2([a, b])$ com $f'(x) \neq 0$ em $[a, b]$ temos que $T(x) = x - \frac{f(x)}{f'(x)}$ e $T'(x) = \frac{f(x)f''(x)}{(f'(x))^2}$ são contínuas em $[a, b]$. Além disso,

$$|T'(x)| = \left| \frac{f(x)f''(x)}{(f'(x))^2} \right| < \lambda < 1$$

por hipótese. Logo,

$$\sup_{x \in [a, b]} |T'(x)| < 1$$

e pelo Teorema 3.1 a equação $f(x) = 0$ tem uma única solução $\xi \in [a, b]$ e o processo iterativo (3.2) converge para ξ .

Além disso,

$$x_{k+1} - \xi = x_k - \xi - \frac{f(x_k)}{f'(x_k)} \Rightarrow e_{k+1} = e_k - \frac{f(x_k)}{f'(x_k)}.$$

Aplicando a série de Taylor e efetuando alguns cálculos elementares obtemos

$$\frac{e_{k+1}}{e_k^2} = \frac{1}{2} \frac{f''(\alpha_k)}{f'(x_k)}$$

para algum $\alpha_k \in [x, x_k]$. Assim,

$$\lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k^2} = \frac{1}{2} \lim_{k \rightarrow \infty} \frac{f''(\alpha_k)}{f'(x_k)} = \frac{1}{2} \frac{f''\left(\lim_{k \rightarrow \infty} \alpha_k\right)}{f'\left(\lim_{k \rightarrow \infty} x_k\right)} = \frac{1}{2} \frac{f''(\xi)}{f'(\xi)} = \frac{1}{2} T''(\xi) = C.$$

■

3.1.1 Interpretação geométrica do método de Newton

Nesta seção faremos a interpretação geométrica do método de Newton. Para isto, considere o processo iterativo

$$x_{k+1} = T(x_k) = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Observe que pode ser escrito na forma

$$x_{k+1} - x_k = -\frac{f(x_k)}{f'(x_k)} \Leftrightarrow f'(x_k)(x_{k+1} - x_k) = -f(x_k) \Leftrightarrow f(x_k) + f'(x_k)(x_{k+1} - x_k) = 0.$$

Fazendo $x_{k+1} = x$ e $L_k(x) = f(x_k) + f'(x_k)(x - x_k)$ temos que $L_k(x)$ é a aproximação linear (reta tangente) da função $f(x)$ numa vizinhança de x_k . Portanto, o ponto x_{k+1} tal que $L_k(x_{k+1}) = 0$ é iteração do método de Newton, como mostra a seguinte figura:

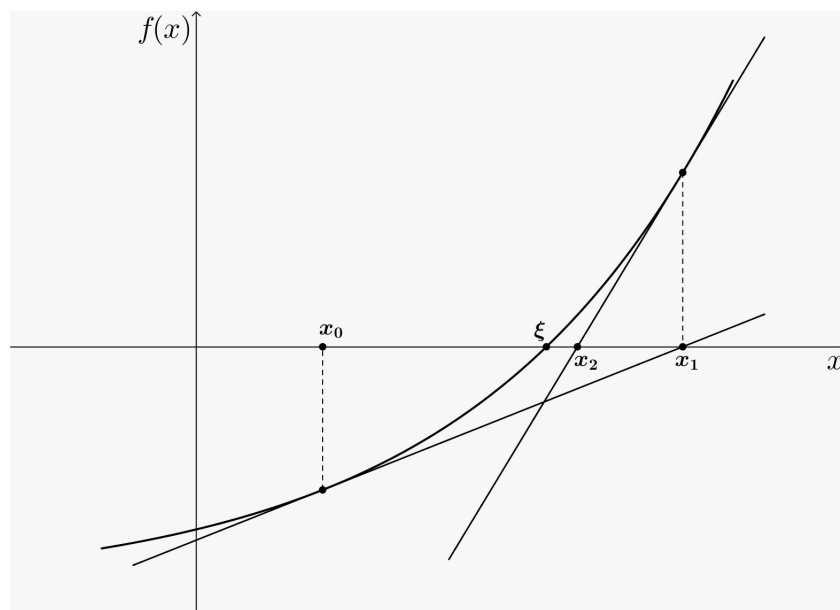


Figura 3.3:

3.2 Caso multidimensional

Nesta seção trataremos o problema de encontrar a raiz da equação

$$f(x) = 0 \tag{3.6}$$

para $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Em termos das funções coordenadas $f(\vec{x}) = (f_1(\vec{x}), f_2(\vec{x}), \dots, f_n(\vec{x}))$ a equação (3.6) torna-se o seguinte sistema de equações:

$$\begin{cases} f_1(x_1, \dots, x_n) = 0 \\ f_2(x_1, \dots, x_n) = 0 \\ \vdots = \vdots \\ f_n(x_1, \dots, x_n) = 0 \end{cases} \quad (3.7)$$

No caso unidimensional, geometricamente o processo iterativo do método de Newton significa aproximar $f(x)$ pela reta tangente $L_k(x)$ numa vizinhança de x_k , ou seja, $f(x) \simeq L_k(x)$. Usaremos esta ideia para generalizar o método para o caso multidimensional.

Sejam f é de classe $C^1(U)$ e ξ a solução da equação $f(x) = 0$ então para x numa vizinhança de ξ tem-se

$$0 = f(\xi) = f(x) + Df(x)(\xi - x).$$

Isto sugere o seguinte processo iterativo:

$$Df(x_k) x_{k+1} = Df(x_k) x_k - f(x_k)$$

ou na forma alternativa

$$x_{k+1} = x_k - [Df(x_k)]^{-1} f(x_k) \quad (3.8)$$

com

$$Df(x_k) = \left(\begin{array}{cccc} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{array} \right)_{x=x_k}$$

Do ponto de vista computacional, é conveniente escrever (3.8) como

$$x_{k+1} = x_k + \delta x_k \quad (3.9)$$

sendo δx_k a solução do sistema linear

$$Df(x_k) \delta x_k = -f(x_k). \quad (3.10)$$

Portanto, o método de Newton é um processo iterativo que a cada iteração resolve um sistema linear. Além disso, observe que a cada iteração devemos garantir de a matriz jacobiana $Df(x_k)$ é inversível.

3.2.1 Convergência local

Vamos provar um resultado de convergência local para o método de Newton. Isto significa que o resultado vale para a aproximação inicial x_0 suficientemente próxima de ξ . Usaremos o seguinte resultado, cuja prova pode ser encontrada em [1].

Lema 3.1 *Sejam $U \subset \mathbb{R}^n$ um aberto e $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de classe $C^1(U)$. Suponha que:*

(i) *$Df(x)$ é lipschitz contínua em U , isto é, existe uma constante $L > 0$ tal que*

$$\|Df(x) - Df(y)\| \leq L \|x - y\|, \quad \forall x, y \in \mathbb{R}^n.$$

(ii) *Para $z \in U$ fixo, existe $[Df(z)]^{-1}$.*

(ii) *Existe uma constante $\beta > 0$ tal que $\|[Df(z)]^{-1}\| \leq \beta$.*

Então, $\forall x \in B(z, r) = \{y \in \mathbb{R}^n; \|y - x\| < r\}$ com $0 < r < \frac{c}{L\beta}$ e $0 < c < 1$ uma constante fixa, $Df(x)$ é não singular e satisfaz

$$\|[Df(x)]^{-1}\| \leq \frac{\beta}{1 - c}$$

Teorema 3.3 *Sejam $U \subset \mathbb{R}^n$ um aberto e convexo e $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de classe $C^1(U)$. Seja $\xi \in \mathbb{R}^n$ tal que $f(\xi) = 0$. Considere $B(\xi, r) = \{y \in \mathbb{R}^n; \|y - \xi\| < r\} \subset U$.*

Suponha que:

(i) *existe $[Df(\xi)]^{-1}$*

(ii) *existe uma constante $\beta > 0$ tal que $\|[Df(x)]^{-1}\| \leq \beta$ em $B(\xi, r)$.*

(ii) *$Df(x)$ é lipschitz contínua em $B(\xi, r)$, isto é, existe uma constante $L > 0$ tal que*

$$\|Df(x) - Df(y)\| \leq L \|x - y\|, \quad \forall x, y \in B(\xi, r).$$

Então, existe $\epsilon > 0$ tal que $\forall x \in B(\xi, \epsilon)$ a sequência (x_k) gerada por (3.8) está bem definida, converge para ξ e satisfaz

$$\frac{\|e_{k+1}\|}{\|e_k\|^2} \leq \frac{L\beta}{2(1 - c)} \quad k = 0, 1, \dots \quad (3.11)$$

para alguma constante fixa $0 < c \leq \frac{2}{3}$.

Prova: Provaremos por indução. Para $0 < c \leq \frac{2}{3}$ fixa considere

$$\epsilon = \min \left(r, \frac{c}{L\beta} \right).$$

Aplicando o Lema 3.1 com $z = \xi$ e $x = x_0$ temos que $Df(x_0)$ é não singular e satisfaz

$$\|[Df(x_0)]^{-1}\| \leq \frac{\beta}{1-c}. \quad (3.12)$$

Portanto, a aproximação x_1 gerada por (3.8) está bem definida e satisfaz

$$x_1 - \xi = x_0 - \xi - [Df(x_0)]^{-1}f(x_0).$$

Assim,

$$\begin{aligned} x_1 - \xi &= x_0 - \xi - [Df(x_0)]^{-1}[f(x_0) - f(\xi)] \\ &= [Df(x_0)]^{-1}[f(x_0) - f(\xi) - Df(x_0)(\xi - x_0)] \end{aligned}$$

e, logo,

$$\|x_1 - \xi\| \leq \|[Df(x_0)]^{-1}\| \|f(x_0) - f(\xi) - Df(x_0)(\xi - x_0)\|.$$

Usando (3.12) e aplicando o Teorema 2.6, obtemos

$$\|x_1 - \xi\| \leq \frac{L\beta}{2(1-c)} \|x_0 - \xi\|^2$$

o que implica que (3.11) é válida para $k = 0$.

Como $\|x_0 - \xi\| \leq \frac{c}{L\beta}$ temos que

$$\|x_1 - \xi\| \leq \frac{c}{2(1-c)} \|x_0 - \xi\| \leq \|x_0 - \xi\| \leq \epsilon.$$

Suponha, por hipótese de indução, que (3.11) vale para k , então

$$\begin{aligned} x_{k+1} - \xi &= x_k - \xi - [Df(x_k)]^{-1}f(x_k) \\ &= x_k - \xi - [Df(x_k)]^{-1}[f(x_k) - f(\xi)] \\ &= [Df(x_k)]^{-1}[f(x_k) - f(\xi) - Df(x_k)(\xi - x_k)] \end{aligned}$$

Aplicando o Lema 3.1 com $z = \xi$ e $x = x_k$ temos que $Df(x_k)$ é não singular e satisfaz

$$\|[Df(x_k)]^{-1}\| \leq \frac{\beta}{1-c}. \quad (3.13)$$

Usando (3.13) e aplicando o Teorema 2.6, obtemos

$$\|x_{k+1} - \xi\| \leq \frac{L\beta}{2(1-c)} \|x_k - \xi\|^2$$

o que implica que (3.11) é válida para $k + 1$.

Como $\|x_k - \xi\| \leq \frac{c}{L\beta}$ temos que

$$\|x_{k+1} - \xi\| \leq \frac{c}{2(1-c)} \|x_k - \xi\| \leq \|x_k - \xi\| \leq \epsilon.$$

■

Note que, pelo Teorema 3.3 temos que a convergência do método de Newton é quadrática.

3.3 Método de Newton generalizado

Nesta seção queremos encontrar a solução da equação de operadores

$$f(x) = 0 \tag{3.14}$$

para $f : X \rightarrow Y$ com X e Y espaço de Banach.

Analogamente, como no caso multidimensional, sejam f Fréchet diferenciável em X e ξ a solução da equação $f(x) = 0$ então para x numa vizinhança de ξ tem-se

$$0 = f(\xi) = f(x) + Df(x)(\xi - x).$$

Isto sugere o seguinte processo iterativo:

$$Df(x_k)(x_{k+1}) = Df(x_k)(x_k) - f(x_k)$$

ou na forma alternativa

$$x_{k+1} = x_k - [Df(x_k)]^{-1}(f(x_k)). \tag{3.15}$$

Novamente, do ponto de vista computacional, é conveniente escrever (3.15) como

$$x_{k+1} = x_k + \delta x_k \tag{3.16}$$

sendo δx_k a solução do sistema linear

$$Df(x_k)(\delta x_k) = -f(x_k). \tag{3.17}$$

Como $Df(x_k)$ é um operador linear temos que encontrar δx_k implica em resolver um problema linear. Portanto, o método de Newton generalizado é um processo iterativo que a cada iteração envolve a resolução de um problema linear.

No que segue vamos a convergência e da ordem de convergência do método de Newton dado pelo processo (3.15):

Teorema 3.4 *Sejam X e Y espaço de Banach e ξ uma solução da equação $f(x) = 0$ e $B(\xi, r) = \{u \in X; \|u - \xi\| \leq r\}$ uma vizinhança de ξ . Suponha que:*

(i) $f : X \rightarrow Y$ é Fréchet diferenciável em X .

(ii) existe $[Df(\xi)]^{-1}$ e é um operador linear contínuo na vizinhança $B(\xi, r)$.

(ii) $Df(x)$ é localmente Lipschitz contínua na vizinhança $B(\xi, r)$, isto é, existe uma constante $L > 0$ tal que

$$\|Df(x) - Df(y)\| \leq L\|x - y\| \quad \forall x, y \in B(\xi, r).$$

Então, existe uma constante $\epsilon > 0$ tal que $\|x_0 - \xi\| < \epsilon$, a sequência (x_k) gerada pelo processo iterativo (3.15) está bem definida e converge para ξ .

Além disso, para alguma constante c tal que $c\epsilon < 1$ temos as seguintes estimativas do erro:

$$\|x_{k+1} - \xi\| \leq c\|x_k - \xi\|^2; \quad (3.18)$$

$$\|x_k - \xi\| \leq (c\epsilon^{2^k})/c. \quad (3.19)$$

Prova: Como $[Df(\xi)]^{-1}$ existe e é um operador contínuo em $B(\xi, r)$ temos

$$c_0 = \sup_{x \in B(\xi, r)} \|Df(x)^{-1}\| < \infty.$$

Considere o seguinte operador $T : B(\xi, r) \rightarrow B(\xi, r)$:

$$T(x) = x - [Df(x)]^{-1}(f(x)).$$

Note que $T(\xi) = \xi$. Assim, para $x \in B(\xi, r)$ temos

$$T(x) - T(\xi) = x - \xi - [Df(x)]^{-1}(f(x)) = [Df(x)]^{-1}[f(\xi) - f(x) - Df(x)(x - \xi)]. \quad (3.20)$$

Aplicando o Teorema 2.15 obtemos

$$f(\xi) - f(x) = (\xi - x) \int_0^1 Df(x + t(\xi - x)) dt. \quad (3.21)$$

Combinando (3.20) e (3.21) tem-se

$$T(x) - T(\xi) = [Df(x)]^{-1} \left[(\xi - x) \int_0^1 Df(x + t(\xi - x)) dt - Df(x)(x - \xi) \right].$$

Portanto,

$$\|T(x) - T(\xi)\| \leq \| [Df(x)]^{-1} \| \|\xi - x\| \int_0^1 \|Df(x + t(\xi - x)) - Df(x)\| dt.$$

Usando que $Df(x)$ é localmente Lipschitz obtemos

$$\|T(x) - T(\xi)\| \leq \| [Df(x)]^{-1} \| \|x - \xi\| \int_0^1 Lt \|x - \xi\| dt.$$

Logo,

$$\|T(x) - T(\xi)\| \leq \frac{c_0 L}{2} \|x - \xi\|^2. \quad (3.22)$$

Escolhendo $\epsilon < \frac{2}{c_0 L}$ tal que $\bar{B}(\xi, \epsilon) \subset B(\xi, r)$. Deste modo,

$$\alpha \equiv \frac{c_0 L \epsilon}{2} < 1.$$

Então (3.22) torna-se

$$\|T(x) - \xi\| = \|T(x) - T(\xi)\| \leq \alpha \|x - \xi\|, \quad x \in \bar{B}(\xi, \epsilon). \quad (3.23)$$

Assumindo que o chute inicial $x_0 \in \bar{B}(\xi, \epsilon)$. Então, (3.23) implica

$$\|x_1 - \xi\| = \|T(x_0) - \xi\| \leq \alpha \|x_0 - \xi\| \leq \alpha \epsilon < \epsilon.$$

Portanto, $x_1 \in \bar{B}(\xi, \epsilon)$. Repetindo sucessivamente este argumento obteremos $x_k \in \bar{B}(\xi, \epsilon)$ para todo $k \geq 0$.

Para mostrarmos a convergência de (x_k) considere

$$\begin{aligned} \|x_{n+1} - \xi\| &= \|T(x_n) - \xi\| \leq \alpha \|x_n - \xi\|, \quad k \geq 0 \\ &\leq \alpha^2 \|x_{k-1} - \xi\| \\ &\dots \\ \|x_k - \xi\| &\leq \alpha^k \|x_0 - \xi\| \end{aligned} \quad (3.24)$$

Como $\alpha < 1$ temos $\lim_{k \rightarrow \infty} \alpha^k = 0$, o que implica

$$\lim_{k \rightarrow \infty} x_k = \xi.$$

Retornando a (3.22), faça $c = \frac{c_0 L}{2}$ e note que $c \epsilon = \alpha < 1$. Então, temos a estimativa

$$\|x_{k+1} - \xi\| = \|T(x_k) - \xi\| \leq c \|x_k - \xi\|^2,$$

o que prova (3.18).

Agora, multiplicando ambos os membros por c , obtemos:

$$c\|x_{k+1} - \xi\| \leq (c\|x_k - \xi\|)^2.$$

Aplicando o argumento recursivo, como em (3.24), obtemos

$$c\|x_k - \xi\| \leq (c\|x_0 - \xi\|)^{2^k},$$

o que prova (3.19). ■

O Teorema 3.4 mostra claramente que o método de Newton é localmente convergente com convergência quadrática.

A principal desvantagem dos resultados dados no Teorema 3.4 é que precisamos conhecer a raiz ξ . O seguinte resultado, conhecido com *Teorema de Kantorovich* elimina esta dificuldade. Sua prova pode ser encontrada em ([1]).

Teorema 3.5 (Kantorovich) *Suponha que:*

(i) *Seja $f : X \rightarrow Y$ é Fréchet diferenciável em $U \subset X$ um aberto convexo de X com derivada $Df(x)$ é Lipschitz contínua, ou seja, existe uma constante $L > 0$ tal que*

$$\|Df(x) - Df(y)\| \leq L\|x - y\|, \quad \forall x, y \in X.$$

(ii) *Para algum $x_0 \in U$, $[Df(x_0)]^{-1}$ existe e é um operador linear contínuo de Y em U e tal que $h = a b L \leq 1/2$ para algum $a \geq \|[Df(x_0)]^{-1}\|$ e $b \geq \|[Df(x_0)]^{-1}f(x_0)\|$. Sejam*

$$t^* = \frac{1 - (1 - 2h)^{1/2}}{a L}, \quad t^{**} = \frac{1 + (1 - 2h)^{1/2}}{a L}.$$

(iii) *Escolha x_0 tal que $\overline{B}(x_1, r) \subset U$ com $r = t^* - b$.*

*Então, a equação (3.14) tem solução $\xi \in \overline{B}(x_1, r)$ e a solução ξ é única em $\overline{B}(x_0, t^{**}) \cap U$. Além disso, a sequência (x_k) converge para ξ e temos a seguinte estimativa do erro:*

$$\|x_k - \xi\| \leq \frac{[1 - (1 - 2h)^{1/2}]^{2^k}}{2^k a L}, \quad k = 0, 1, 2, \dots$$

Capítulo 4

Aplicações do método de Newton generalizado

Neste capítulo aplicaremos o Método de Newton generalizado em problemas clássicos da Análise Funcional.

4.1 Sistemas não lineares

Considere o sistema não linear

$$f \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \dots \\ f_n(x_1, x_2, \dots, x_n) \end{pmatrix} = 0$$

Sabemos que $Df(x)$ é uma matriz $n \times n$ com elementos $D_i f_j(x)$ e $D^2 f(x)$ é um arranjo $n \times n \times n$ com elementos $\frac{\partial^2 f_i}{\partial x_j \partial x_k}$. Se usamos a norma do máximo temos

$$\|Df(x)\| = \max_i \sum_{j=1}^n \left| \frac{\partial f_i}{\partial x_j} \right|.$$

e

$$\|D^2 f(x)\| \leq \max_i \sum_{j=1}^n \left| \frac{\partial^2 f_i}{\partial x_j \partial x_k} \right|.$$

Para aplicarmos o Teorema de Kantorovich devemos ser capazes que limitar as normas acima.

Exemplo 4.1 *Encontre a solução do sistema*

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 4y^2 - 3/2 \\ 4x^2 + y - 3/2 \end{pmatrix} = 0$$

Solução: *Note que*

$$Df(x) = \begin{pmatrix} 1 & 8y \\ 8x & 1 \end{pmatrix}.$$

Escolhendo

$$x_0 = \begin{pmatrix} 0.4 \\ 0.4 \end{pmatrix}$$

temos

$$Df(x_0) = \begin{pmatrix} 1 & 3.2 \\ 3.2 & 1 \end{pmatrix}, \quad f(x_0) = \begin{pmatrix} 2.54 \\ -0.46 \end{pmatrix}$$

Agora, resolvendo o sistema linear

$$\begin{pmatrix} 1 & 3.2 \\ 3.2 & 1 \end{pmatrix} \begin{pmatrix} \delta x_0 \\ \delta y_0 \end{pmatrix} = \begin{pmatrix} -2.54 \\ 0.46 \end{pmatrix}$$

obtemos

$$\begin{pmatrix} \delta x_0 \\ \delta y_0 \end{pmatrix} = \begin{pmatrix} 0.1095238 \\ 0.1095238 \end{pmatrix}$$

Continuando o processo obtemos

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0.5095238 \\ 0.5095238 \end{pmatrix}; \quad \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.5000714 \\ 0.5000714 \end{pmatrix}$$

4.2 Equações integrais não lineares

Considere a equação integral não linear

$$u(t) = \int_0^1 k(t, s, u(s)) ds \tag{4.1}$$

sobre o espaço $U = C([0, 1])$. Assuma que $k \in C([0, 1] \times [0, 1] \times \mathbb{R})$ e de classe C^1 com relação ao terceiro argumento.

Considere um operador $F : U \rightarrow U$ dado por

$$F(u)(t) = u(t) - \int_0^1 k(t, s, u(s)) ds, \quad t \in [0, 1],$$

Note que a equação integral pode ser escrita na forma

$$F(u) = 0.$$

O método de Newton para o problema é dado por

$$u_{n+1} = u_n + \delta x_n$$

com δx_n solução do problema linear

$$DF(u_n)\delta x_n = -F(u_n).$$

Vamos calcular $DF(u)$:

$$DF(u)(v)(t) = \lim_{h \rightarrow 0} \frac{F(u + hv)(t) - F(u)(t)}{h} \quad (4.2)$$

$$= \lim_{h \rightarrow 0} \frac{hv(t)}{h} - \int_0^1 [k(t, s, u(s) + hv(s)) - k(t, s, u(s))] ds \quad (4.3)$$

$$= v(t) + \int_0^1 \frac{\partial k(t, s, u(s))}{\partial u} v(s) ds. \quad (4.4)$$

Portanto, δx_n é solução da equação integral linear:

$$\delta_n(t) - \int_0^1 \frac{\partial k(t, s, u_n(s))}{\partial u} \delta_n(s) ds = -u_n(t) + \int_0^1 k(t, s, u_n(s)) u_n(s) ds. \quad (4.5)$$

4.3 Equações diferenciais ordinárias de 1ª ordem

Considere o problema de Cauchy

$$\begin{aligned} \frac{dy}{dt} &= f(t, y) \\ y(0) &= c. \end{aligned} \quad (4.6)$$

Considere o operador

$$P(y) = \frac{dy}{dt} - f(t, y), \quad (4.7)$$

para $P : C^1([0, \tau]) \rightarrow C([0, \tau])$.

Suponha que $y \in C^1([0, \tau])$ e f é contínua em $[0, \tau]$ e diferenciável com respeito a y . Então, para $y_0 = y_0(t)$ temos que $DP(y_0) : C^1([0, \tau]) \rightarrow C([0, \tau])$ é dada por

$$DP(y_0) = \frac{d}{dt} - f'_2(t, y_0(t))I, \quad (4.8)$$

com

$$f'_2(t, y_0(t)) = \left. \frac{\partial f}{\partial y}(t, y) \right|_{y=y_0(t)}, \quad 0 \leq t \leq \tau. \quad (4.9)$$

Portanto, δy_m é a solução dos problemas lineares:

$$\begin{aligned} DP(y_m)\delta y_m &= -P(y_m), \quad m = 0, 1, \dots \\ y_0 &= c. \end{aligned} \quad (4.10)$$

Definindo

$$a_m(t) = -f'_2(t, y_m(t)) \quad (4.11)$$

$$v_m(t) = -P(y_m) = -\frac{dy_m(t)}{dt} - f(t, y_m(t)) \quad (4.12)$$

(4.10) torna-se

$$\begin{aligned} \frac{d\delta_m(t)}{dt} + a_m(t)\delta_m(t) &= v_m(t) \\ \delta_m(0) &= 0. \end{aligned} \quad (4.13)$$

Sabemos que a solução do problema de valor inicial (4.13) (veja ([9])) é obtida do seguinte modo:

Seja

$$A_m(t) = \int_0^t a_m(s)ds, \quad 0 \leq t \leq \tau. \quad (4.14)$$

Multiplique os dois lados de (4.13) por $\exp^{A_m(t)}$ para obter

$$\frac{d}{dt}(\exp^{A_m(t)} \delta y_m(t)) = \exp^{A_m(t)} v_m(t), \quad 0 \leq t \leq \tau. \quad (4.15)$$

Integrando ambos os lados de (4.15) temos

$$\delta y_m(t) = \int_0^t \exp^{A_m(s)-A_m(t)} v_m(s)ds, \quad 0 \leq t \leq \tau. \quad (4.16)$$

O processo iterativo do método de Newton das aproximações $y_m(t)$ da solução de (4.6) é dado por

$$y_{m+1}(t) = y_m(t) + \delta y_m(t).$$

Logo,

$$y_{m+1}(t) = y_m(t) + \int_0^t \exp^{A_m(s)-A_m(t)} v_m(s)ds, \quad 0 \leq t \leq \tau, \quad m = 0, 1, 2, \dots, \quad (4.17)$$

onde v_m, A_m são dadas por (4.11) e (4.14) respectivamente.

Exemplo 4.2 Considere o problema de valor inicial

$$\begin{cases} \frac{dy}{dt} - (1 + y^2) = 0 \\ y(0) = 0 \end{cases} \quad (4.18)$$

Observe que

$$a_m(t) = 2y_m(t), \quad v_m(t) = \frac{dy_m}{dt} - (1 + y_m^2).$$

Logo, para $y_0(t) = t$ temos

$$v_0(t) = t^2, \quad A_0(t) = t^2.$$

Assim, por (4.17) tem-se

$$y_1(t) = t + \int_0^t e^{s^2-t^2} s^2 ds. \quad (4.19)$$

A integral (4.19) não tem uma primitiva dada em termos de funções elementares. Porém, podemos aproximar o integrando por uma série uniformemente convergente e obter

$$y_1(t) = t + \frac{t^3}{3} + \frac{2t^5}{15} + \frac{4t^7}{105} + \dots \quad (4.20)$$

O resultado (4.20) poderia ser substituído em (4.17) para obtermos $y_2(t)$, porém os cálculos são mais complexos.

Para $\tau < \pi/2$, o problema (4.18) tem uma única solução dada por

$$y^*(t) = \operatorname{tg}(t), \quad 0 \leq t \leq \tau. \quad (4.21)$$

Comparando a série (4.20) com a série de Maclaurin da solução $\operatorname{tg}(t)$ temos

$$y^*(t) - y_1(t) = \frac{t^7}{61} + \frac{38t^9}{2835} + \dots \quad 0 \leq t \leq \tau < \pi/2$$

Portanto, $y_1(t)$ será uma boa aproximação para $y^*(t)$ para t suficientemente pequeno.

4.4 Equações diferenciais de 2ª ordem

Considere o problema de 2ª ordem

$$\begin{cases} u''(t) = g(t, u(t)), \quad t \in (0, 1) \\ u(0) = u(1) = 0 \end{cases}$$

Suponha que a função $f : [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$ contínua e de classe C^1 com relação ao segundo argumento.

Seja

$$U = C^2([0, 1]) = \{v \in C^2[0, 1]; v(0) = v(1) = 0\}$$

com a norma $\|\cdot\|_{C^2[0,1]}$. Defina

$$f(u)(t) = u''(t) - g(t, u(t)), t \in [0, 1].$$

Pode-se demonstrar que

$$Df(u)(y)(t) = y''(t) - \frac{\partial g(t, u(t))}{\partial u} y(t).$$

Assim, a cada iteração devemos resolvemos o seguinte problema linear:

$$\begin{cases} u''_{n+1}(t) - \frac{\partial f}{\partial u}(t, u_n(t))u_{n+1}(t) = f(t, u_n(t)) - \frac{\partial f}{\partial u}(t, u_n(t))u_n(t), t \in (0, 1) \\ u_{n+1}(0) = u_{n+1}(1) = 0 \end{cases}$$

com $u_n = \delta u_n$.

Capítulo 5

Conclusão

Neste trabalho desenvolvemos o Cálculo Diferencial em Espaços de Banach, afim de generalizar para o mesmo, a parte teórica do Método de Newton, concluindo assim que a diferença que existe para o caso de dimensão finita para o de dimensão infinita é dado pelo conceito de diferenciabilidade, estrutura e elementos dos espaços.

Referências Bibliográficas

- [1] ATKINSON, K.; HAN, W. **Theoretical numerical analysis, volume 39 of Texts in Applied Mathematics**: Springer, Dordrecht, 2009.
- [2] BREZIS, H. **Functional analysis, Sobolev spaces and partial differential equations**: Springer Science & Business Media, 2010.
- [3] CAVALCANTI, Ma. M.; CAVALCANTI, V. N. D.; KOMORNIK, V. **Introdução análise funcional**. Eduem, Maringá, 2011.
- [4] DRÁBEK P.; MILOTA, J., **Methods of Nonlinear Analysis: Applications to Differential equations**. Birkhauser, Boston, 2007.
- [5] KREYSZIG, E. **Introductory functional analysis with applications**: New York: wiley, 1989. v. 1.
- [6] LIMA, E. L. **Curso de análise**: Projeto Euclides, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 2012. vol. 2.
- [7] LOURÊDO, A. T.; OLIVEIRA, A. M.; LIMA, O. A. **Cálculo Avançado**: Campina Grande: Eduepb, 2012.
- [8] MOISEWITSCH, B. L. **Integral equations**: Courier Corporation, 2011.
- [9] MORRIS, M.; O.E.BROWN, L. B. **Differential Equations**: Englewood Cliffs, New Jersey, 1952.
- [10] RALL, L. B. **Computational solution of nonlinear operator equations**: New York: Wiley, 1969.
- [11] RUGGIERO, M. A. G.; LOPES, V. L. R. **Cálculo numérico: aspectos teóricos e computacionais**: Makron Books do Brasil, 1997.
- [12] SHILOV, G. E.. **Elementary Functional Analysis**: Dover Publications, INC, New York: Wiley, 1974.