Métodos Numéricos

Rui F. Vigelis

`rfvigelis@gmail.com`

Universidade Federal do Ceará – UFC

Versão:
2023-05-21 05:01:31

Objetivos:

- Capacitar o aluno a identificar e enfrentar os problemas de Engenharia que possam ser resolvidos com técnicas de Métodos Numéricos.

Frequência:

- $\geq 75\%$, que equivale a um máximo de 16 horas em faltas.

Avaliação:

- 3 avaliações progressivas distribuídas durante o semestre.

Critério de aprovação:

- Se $7 \leq \text{MAPs}$, o aluno é aprovado por média.
- Se $4 \leq \text{MAPs} < 7$, o aluno faz a prova de avaliação final.
- Se $4 \leq \text{NAF}$ e $5 \leq \text{MAF} = \dfrac{\text{MAPs} + \text{NAF}}{2}$, o aluno é aprovado.
- Caso contrário, o aluno é reprovado.

Bibliografia:

- Franco, Neide Bertoldi. **Cálculo Numérico**, 1a. ed. São Paulo: Pearson Prentice Hall, 2006.
- Chapra, Steven C. & Canale, Raymond P. **Métodos Numéricos para Engenharia**, 5a. ed. Porto Alegre: AMGH, 2011.

Conteúdo:

- Análise de arredondamento (AP1)
- Raízes reais de funções reais (AP1)
- Solução de sistemas lineares (AP2)
- Interpolação polinomial (AP3)
- Integração numérica (AP3)
- Solução numérica de equações diferenciais ordinárias (AP3)

Definição (Sistema de Ponto Flutuante)

Um número real $x \neq 0$ é chamado de ponto flutuante (normalizado) se pode ser expresso como

$$x = \pm 0{,}d_1 d_2 \cdots d_t \times \beta^e,$$

em que

- β é a **base**,
- t é o número de dígitos na **mantissa**, com $d_1 \neq 0$ e $0 \leq d_j \leq \beta - 1$, para $j = 1, \ldots, t$, e
- e é o **exponente**, com $-m \leq e \leq M$.

Usamos a notação $F(\beta, t, m, M)$ para o conjunto de todos os pontos flutuantes, fixados β, t, m e M, e adicionando algumas exceções como o zero.

- Muitos programas de computação numérica adotam o padrão IEEE (754 2008) para precisão dupla com 64 bits: 1 para o sinal, 11 para o expoente, 52 para a mantissa.
- No padrão IEEE para precisão dupla, pode-se representar números positivos entre $2{,}23 \times 10^{-308}$ e $1{,}79 \times 10^{308}$, aproximadamente.
- O padrão IEEE possui uma representação especial para o zero, $\pm\infty$ (obtido após a divisão por zero), e NaN (Not a Number, que se obtém em certas operações como $0/0$).

- Seja x um número real dentro dos limites de representação do sistema em ponto flutuante.
- Se x não pertence ao conjunto $F(\beta, t, m, M)$, ele é representado pelo seu arredondamento em ponto flutuante, que consiste em encontrar $\overline{x} \in F(\beta, t, m, M)$ tal que $|x - \overline{x}|$ seja o menor possível.
- Seja $\mathrm{fl}(\cdot)$ a função que associa um número real x ao seu arredondamento em ponto flutuante.
- O valor $|x - \overline{x}|$ é chamado **erro absoluto** de arredondamento, e $|x - \overline{x}|/|x|$ é chamado de **erro relativo** de arredondamento.

- Utilizaremos a seguinte regra para arredondamento em ponto flutuante.
- Se $x = 0$, então $\overline{x} = 0$.
- Se $x \neq 0$, escolhemos f_x e g_x tais que

$$|x| = f_x \times \beta^e + g_x \times \beta^{e-t}, \quad \text{em que} \quad \beta^{-1} \leq f_x < 1 \text{ e } 0 \leq g_x < 1.$$

- O valor absoluto do número arredondado é então dado por

$$|\overline{x}| = \begin{cases} f_x \times \beta^e, & \text{se } g_x < \frac{1}{2}, \\ f_x \times \beta^e + \beta^{e-t}, & \text{se } g_x \geq \frac{1}{2}, \end{cases}$$

e com isso $\overline{x} = (\operatorname{sinal} x)|\overline{x}|$.

Exemplo

Represente no sistema $F(10, 3, 5, 5)$ os números:

$$x_1 = 1234{,}56, \qquad x_2 = -0{,}00054962, \qquad x_3 = 0{,}9995,$$
$$x_4 = 123456{,}7, \qquad x_5 = 0{,}0000001.$$

R.: $\overline{x}_1 = 0{,}123 \times 10^4$, $\overline{x}_2 = -0{,}550 \times 10^{-3}$, $\overline{x}_3 = 0{,}100 \times 10^1$, $|x_4| = 0{,}1234567 \times 10^6$, $|x_5| = 0{,}1 \times 10^{-6}$.

Definição (Épsilon da Máquina)

O **épsilon da máquina**, denotado por ε_{mach}, é a metade da distância entre 1 e o menor número em ponto flutuante estritamente maior que 1.

- O épsilon da máquina de um sistema $F(\beta, t, m, M)$ é

$$\varepsilon_{mach} = \frac{1}{2}\beta^{1-t}.$$

- No padrão IEEE para precisão dupla, tem-se

$$\varepsilon_{mach} = 2^{-52} \approx 2{,}2 \times 10^{-16}.$$

- O épsilon da máquina ε_{mach} fornece um limitante superior para o erro relativo do arredondamento em ponto flutuante.

Proposição

Seja x qualquer número real dentro dos limites de representação do sistema. Então existe ε com $|\varepsilon| \leq \varepsilon_{mach}$ tal que

$$\mathrm{fl}(x) = x(1 + \varepsilon).$$

- Se $g_x < \frac{1}{2}$, então

$$\begin{aligned}|x - \overline{x}| &= |f_x \times \beta^e + g_x \times \beta^{e-t} - f_x \times \beta^e| \\ &= g_x \times \beta^{e-t} < \frac{1}{2}\beta^{e-t}.\end{aligned}$$

- Se $g_x \geq \frac{1}{2}$, então

$$\begin{aligned}|x - \overline{x}| &= |f_x \times \beta^e + g_x \times \beta^{e-t} - f_x \times \beta^e - \beta^{e-t}| \\ &= |g_x - 1| \times \beta^{e-t} \leq \frac{1}{2}\beta^{e-t}.\end{aligned}$$

- Em ambos os casos,

$$\frac{|x - \overline{x}|}{|x|} \leq \frac{\frac{1}{2}\beta^{e-t}}{\beta^{-1} \times \beta^e} = \frac{1}{2}\beta^{1-t},$$

visto que $|x| \geq \beta^{-1} \times \beta^e$.

- As operações aritméticas básicas $+$, $-$, $\times$ e $/$ com números reais, quando realizadas no computador com sistema $F(\beta, t, m, M)$ serão denotadas por $\oplus$, $\ominus$, $\otimes$ e $\oslash$.
- As operações aritméticas de ponto flutuante são definidas de modo a satisfazer o axioma:

Axioma das Operações de Ponto Flutuante

Seja $*$ uma operação aritmética básica, e $\circledast$ a respectiva operação em ponto flutuante. A operação $\circledast$ satisfaz

$$x \circledast y = \mathrm{fl}(x * y),$$

para quaisquer $x, y \in F(\beta, t, m, M)$.

- As operações de ponto flutuante não são nem associativas e nem distributivas!

Exemplo

Considere o sistema $F(10,3,5,5)$. Sejam $x = \mathrm{fl}(11{,}4)$, $y = \mathrm{fl}(3{,}18)$ e $z = \mathrm{fl}(5{,}05)$. Efetue as operações:

(a) $(x \oplus y) \oplus z$ e $x \oplus (y \oplus z)$;

(b) $(y \otimes x) \oslash z$ e $(y \oslash z) \otimes x$;

(c) $y \otimes (z \oplus x)$ e $(y \otimes z) \oplus (y \otimes x)$.

R.: (a) $0{,}197 \times 10^2$ e $0{,}196 \times 10^2$; (b) $0{,}719 \times 10^1$ e $0{,}718 \times 10^1$; (c) $0{,}525 \times 10^1$ e $0{,}524 \times 10^1$.

- Em vista do axioma das operações de ponto flutuante, tem-se:

Proposição

Para quaisquer $x, y \in F(\beta, t, m, M)$, existe ε com $|\varepsilon| \leq \varepsilon_{mach}$ tal que

$$x \circledast y = (x * y)(1 + \varepsilon),$$

em que $*$ denota uma operação aritmética básica, e $\circledast$ a respectiva operação em ponto flutuante.

- A seguir veremos dois efeitos numéricos que contribuem para que o resultado obtido não tenha crédito: o cancelamento subtrativo, e a propagação de erro.
- Sejam x e y dois números reais dentro dos limites de representação do sistema.
- Na soma desses números em ponto flutuante, tem-se:

$$\begin{aligned} \mathrm{fl}(x) \oplus \mathrm{fl}(y) &= [\mathrm{fl}(x) + \mathrm{fl}(y)](1 + \varepsilon_1) \\ &= [x(1+\varepsilon_2) + y(1+\varepsilon_3)](1+\varepsilon_1) \\ &= x + x(\varepsilon_1 + \varepsilon_2 + \varepsilon_1\varepsilon_2) + y + y(\varepsilon_1 + \varepsilon_3 + \varepsilon_1\varepsilon_3) \end{aligned}$$

com $|\varepsilon_1|, |\varepsilon_2|, |\varepsilon_3| \leq \varepsilon_{mach}$.

- Deste modo, para o erro absoluto (EA), podemos escrever

$$\mathrm{EA} = |\mathrm{fl}(x) \oplus \mathrm{fl}(y) - (x+y)| \leq (|x| + |y|)(2\varepsilon_{mach} + \varepsilon_{mach}^2).$$

- Já para o erro relativo (ER), temos

$$\mathrm{ER} = \frac{|\mathrm{fl}(x) \oplus \mathrm{fl}(y) - (x+y)|}{|x+y|} \leq \frac{|x| + |y|}{|x+y|}(2\varepsilon_{mach} + \varepsilon_{mach}^2).$$

- Como consequência, o erro relativo pode ser grande se $|x+y|$ for pequeno, ou seja, se $x \approx -y$.
- Neste caso, tem-se o chamado **cancelamento subtrativo**.

Exemplo

No sistema $F(10, 10, 10, 10)$, calcule

$$x = \sqrt{9876} - \sqrt{9875}.$$

- Além do cancelamento, pode ocorrer a **propagação de erro**, que é observada em uma sequência de operações aritméticas.

Exemplo

Considerando o sistema $F(10, 3, 5, 5)$, calcule o polinômio

$$P(x) = x^3 - 6x^2 + 4x - 0.1,$$

no ponto 5,24, e compare o resultado com o valor exato $P(5{,}24) = -0{,}007776$.

- Um algoritmo $\mathcal{A}$ é uma sequência de operações aritméticas que gera uma saída y dada uma entrada x.
- Seja $\mathcal{P}$ o problema ao qual o algoritmo $\mathcal{A}$ se propõe a resolver.
- O algoritmo $\mathcal{A}$ é chamado de **preciso** (em inglês, *accurate*) se o seu erro relativo é pequeno, ou seja, se

$$\frac{|\mathcal{P}(x) - \mathcal{A}(x)|}{|\mathcal{P}(x)|}$$

  é da ordem de ε_{mach}.

- Dependendo do contexto, um algoritmo preciso pode ser difícil de ser implementado, pois sempre há a presença de erros de arredondamento em ponto flutuante.
- Podemos usar como requisito outra noção que nos permite analisar a influência de erros de arredondamento em um algoritmo.

Algoritmo Regressivamente Estável

Um algoritmo $\mathcal{A}$, usado para resolver um problema $\mathcal{P}$, é chamado de **regressivamente estável** (em inglês, *backward stable*) se

$$\mathcal{P}(\widetilde{x}) = \mathcal{A}(x),$$

para algum $\widetilde{x}$ com

$$\frac{|\widetilde{x} - x|}{|x|}$$

da ordem de ε_{mach}.

Exemplo

Mostre que a operação $\text{fl}(x) \oplus \text{fl}(y)$, que é usado para resolver a soma $x + y$, é regressivamente estável num sistema de ponto flutuante com épsilon da máquina ε_{mach}.

Definição

Dada uma função $f\colon [a, b] \to \mathbb{R}$, um número $\alpha \in [a, b]$ tal que

$$f(\alpha) = 0$$

é chamado de raiz (ou zero) de f. Dizemos também que α é uma solução da equação $f(x) = 0$.

- Nem sempre é possível encontrar uma expressão analítica para a raiz de uma função.
- Neste caso, recorremos a métodos numéricos para aproximar a raiz da função.

- Para garantir a existência de uma raiz de f em $[a, b]$, podemos fazer uso do seguinte resultado:

Teorema de Bolzano

Seja $f : [a, b] \to \mathbb{R}$ uma função contínua. Se $f(a)f(b) < 0$, i.e., $f(a)$ e $f(b)$ tem sinais opostos, então existe pelo menos um $\alpha \in (a, b)$ tal que $f(\alpha) = 0$.

- O Teorema de Bolzano é uma consequência do Teorema do Valor Intermediário, visto em Cálculo I.
- Além de garantir a existência da raiz, o Teorema de Bolzano é a base para o método da bissecção.

- Seja f uma função real contínua definida no intervalo $[a, b]$ tal que $f(a)f(b) < 0$.
- O método da bissecção consiste no seguinte procedimento:
  - Calcule o ponto médio do intervalo:
    $$m = \frac{a+b}{2}.$$
  - Substitua a ou b por m de modo que o novo intervalo contém uma raiz:
    $$\text{se } f(m)f(b) < 0, \text{ então } a \leftarrow m, \text{ senão } b \leftarrow m.$$
  - Repita os passos anteriores até $(b - a) \leq 2\delta$.
- O ponto médio entre a e b consiste na estimativa da raiz de f.

- Na n-ésima iteração, o intervalo resultante terá comprimento $\dfrac{b-a}{2^n}$, que converge para zero com $n \to \infty$.
- A condição de parada é satisfeita se

$$n \geq \log_2\left(\frac{b-a}{\delta}\right) - 1$$

- Neste caso, o erro absoluto da aproximação satisfaz $|m - \alpha| \leq \delta$.

- Inicialização:
  - Escolha a_0 e b_0 de modo que $f(a_0)f(b_0) < 0$.
- Para $n \geq 0$:
  - Faça $x_n = \dfrac{a_n + b_n}{2}$.
  - Pare se $f(x_n) = 0$ ou $\dfrac{b_n - a_n}{2} \leq \delta$.
  - Se $f(a_n)f(x_n) < 0$, escolha $[a_{n+1}, b_{n+1}] = [a_n, x_n]$, caso contrário, $[a_{n+1}, b_{n+1}] = [x_n, b_n]$.

Exemplo

Use o método da bissecção para encontrar uma estimativa para a raiz da função

$$f(x) = e^x - 2x - 1,$$

no intervalo $[1, 2]$, considerando uma tolerância $(b_n - a_n)/2 < \delta = 5 \times 10^{-2}$.

Exemplo

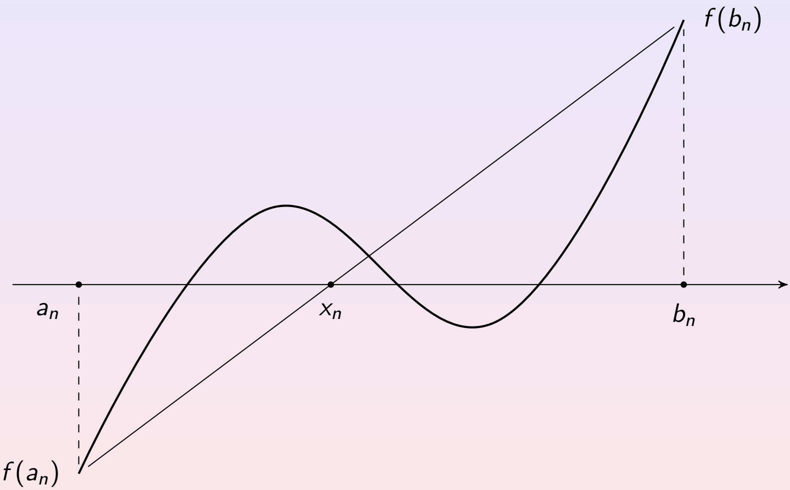
Aplique o método da bissecção para encontrar a raiz da função

$$f(x) = \text{sen}(2x) - \cos(3x),$$

no intervalo $[0, 1]$, com tolerância $(b_n - a_n)/2 < \delta = 5 \times 10^{-2}$.

- O método da bisseção, que considera apenas o sinal de f nos extremos dos intervalos, pode ser modificado para consideramos os valores de f.
- No método da posição falsa, em vez de escolhermos o ponto médio do intervalo, adotamos a intersecção do eixo x com a reta que passa pelos pontos $(a, f(a))$ e $(b, f(b))$.
- Essa intersecção é dada por

$$m = \frac{af(b) - bf(a)}{f(b) - f(a)}.$$


$f(b_n)$
a_n
x_n
b_n
$f(a_n)$

- Inicialização:
  - Escolha a_0 e b_0 de modo que $f(a_0)f(b_0) < 0$.
- Para $n \geq 0$:
  - Faça $x_n = b_n - f(b_n)\dfrac{b_n - a_n}{f(b_n) - f(a_n)}$.
  - Pare se $|f(x_n)| < \varepsilon$.
  - Se $f(a_n)f(x_n) < 0$, escolha $[a_{n+1}, b_{n+1}] = [a_n, x_n]$, caso contrário, $[a_{n+1}, b_{n+1}] = [x_n, b_n]$.

Exemplo

Use o método da posição falsa para encontrar uma estimativa para a raiz da função

$$f(x) = e^x - 2x - 1,$$

no intervalo $[1, 2]$, considerando uma tolerância $|f(x_n)| < \varepsilon = 5 \times 10^{-2}$.

Exemplo

Usando o método da posição falsa, encontre a raiz da função

$$f(x) = \text{sen}(2x) - \ln(x),$$

no intervalo $[1, 3]$, com tolerância $|f(x_n)| < \varepsilon = 10^{-4}$.

- Seja f é uma função contínua em $[a, b]$.
- No lugar da equação $f(x) = 0$, consideramos o problema na forma

$$x = g(x),$$

  em que g é tal que $f(\alpha) = 0$ se e somente se $\alpha = g(\alpha)$.
- Uma função g como acima é chamada de função de iteração, e um número α satisfazendo $\alpha = g(\alpha)$ é chamado ponto fixo de g.
- Dada uma aproximação inicial x_0, o método da iteração de ponto fixo define as aproximações sucessivas

$$x_{n+1} = g(x_n), \qquad \text{para } n \geq 0.$$

- Espera-se que $x_n \to \alpha$ com $n \to \infty$.

Exemplo

Considere a função $f(x) = e^x - 2x - 1$. Mostre que $g_1(x) = (e^x - 1)/2$ e $g_2(x) = \ln(2x + 1)$ são funções de iteração para $f(x)$.

Teorema (∗)

Seja $g\colon [a,b] \to \mathbb{R}$ uma função de iteração contínua, com ponto fixo $\alpha \in (a,b)$. Assuma que g é derivável em (a,b), e a derivada $g'(x)$ satisfaz $|g'(x)| \leq M < 1$, para todo $x \in (a,b)$. Seja $x_0 \in [a,b]$ qualquer. Então a iteração de ponto fixo

$$x_{n+1} = g(x_n), \qquad n \geq 0,$$

está bem definida ($x_n \in [a,b]$ para todo $n \geq 1$), e converge para α.

- Pelo Teorema do Valor Médio, existe η entre x_{n-1} e α tal que

$$g(x_{n-1}) - g(\alpha) = g'(\eta)(x_{n-1} - \alpha).$$

- Deste modo, podemos escrever

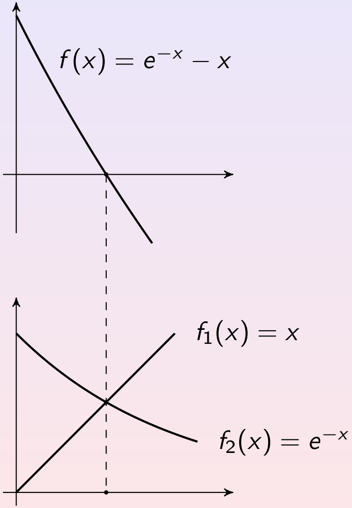
$$|x_n - \alpha| = |g(x_{n-1}) - g(\alpha)| = |g'(\eta)||x_{n-1} - \alpha| \leq M|x_{n-1} - \alpha|.$$

- Com isso, obtemos

$$|x_n - \alpha| \leq M^n|x_0 - \alpha|, \qquad \text{para todo } n \geq 0.$$

- Como $M < 1$, concluímos que $x_n \in [a, b]$ para todo $n \geq 1$, e

$$\lim_{n\to\infty} |x_n - \alpha| \leq \lim_{n\to\infty} M^n|x_0 - \alpha| = 0.$$


$f(x) = e^{-x} - x$
$f_1(x) = x$
$f_2(x) = e^{-x}$

- Inicialização:
  - Escolha x_0.
- Para $n \geq 0$:
  - Faça $x_{n+1} = g(x_n)$.
  - Pare se $|f(x_{n+1})| < \varepsilon$.

Exemplo

Considere a função $f(x) = e^x - 2x - 1$. Usando a aproximação inicial $x_0 = 1$, determine as aproximações $\{x_n\}$ considerando as funções de iteração:

(a) $g_1(x) = (e^x - 1)/2$;

(b) $g_2(x) = \ln(2x + 1)$.

Exemplo

Aplique o método da iteração de ponto fixo para encontrar a raiz da função $f(x) = x^3 - x^2 - 3$ no intervalo $[1, 3]$, com função de iteração $g(x) = (x^2 + 3)^{1/3}$, ponto inicial $x_0 = 2{,}5$, e tolerância $|f(x_{n+1})| < \varepsilon = 10^{-1}$. Verifique as hipóteses que garantem a convergência do método.

Definição

Seja $\{x_n\}$ uma sequência convergindo para α, e seja $e_n = x_n - \alpha$ seu erro. Considere a condição

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^p} = \mu,$$

para números reais $p \geq 1$ e $\mu \geq 0$.

- Para $p = 1$,
    - se $\mu = 0$, a convergência é **superlinear**,
    - se $0 < \mu < 1$, a convergência é **linear** com **taxa de convergência** μ,
    - se $\mu = 1$, a convergência é **sublinear**.
- Para $p > 1$ e $\mu > 0$, a convergência é dita ser de **ordem** p. Em particular, se $p = 2$, a convergência é **quadrática**, e se $p = 3$, a convergência é **cúbica**.

Definição

A **ordem de convergência** e a **taxa de convergência** de um método numérico são definidas como a ordem de convergência e a taxa de convergência da sequência $\{x_n\}$ resultante da aplicação do método numérico.

Teorema

Suponha que as condições do Teorema $(*)$ são satisfeitas. Assuma também que g' é contínua em $[a, b]$, e $x_n \neq \alpha$ para todo $n \geq 0$. Seja $e_n = x_n - \alpha$. Se $g'(\alpha) \neq 0$, então

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|} = |g'(\alpha)|.$$

- Pelo Teorema do Valor Médio, existe ξ_n entre x_n e α tal que

$$g(x_n) - g(\alpha) = g'(\xi_n)(x_n - \alpha), \qquad \text{para } n \geq 0.$$

- Portanto, $\lim_{n\to\infty} \xi_n = \alpha$, e, por g' ser contínua,

$$\lim_{n\to\infty} g'(\xi_n) = g'(\alpha).$$

- Deste modo, temos

$$\begin{aligned}
\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|} &= \lim_{n\to\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|} = \lim_{n\to\infty} \frac{|g(x_n) - g(\alpha)|}{|x_n - \alpha|} \\
&= \lim_{n\to\infty} \frac{|g'(\xi_n)(x_n - \alpha)|}{|x_n - \alpha|} \\
&= \lim_{n\to\infty} |g'(\xi_n)| = |g'(\alpha)|.
\end{aligned}$$

Teorema

Suponha que as condições do Teorema $(*)$ são satisfeitas. Assuma também que g'' é contínua em $[a, b]$, e $x_n \neq \alpha$ para todo $n \geq 0$. Seja $e_n = x_n - \alpha$. Se $g'(\alpha) = 0$ e $g''(\alpha) \neq 0$, então

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^2} = \frac{1}{2}|g''(\alpha)|.$$

- Pela expansão de Taylor, existe ξ_n entre x_n e α tal que

$$g(x_n) = g(\alpha) + g'(\alpha)(x_n - \alpha) + \frac{1}{2}g''(\xi_n)(x_n - \alpha)^2,$$

que implica

$$g(x_n) - g(\alpha) = \frac{1}{2}g''(\xi_n)(x_n - \alpha)^2.$$

- Deste modo, temos

$$\begin{aligned}
\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^2} &= \lim_{n\to\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^2} = \lim_{n\to\infty} \frac{|g(x_n) - g(\alpha)|}{|x_n - \alpha|^2} \\
&= \lim_{n\to\infty} \frac{|\frac{1}{2}g''(\xi_n)(x_n - \alpha)^2|}{|x_n - \alpha|^2} \\
&= \lim_{n\to\infty} \frac{1}{2}|g''(\xi_n)| = \frac{1}{2}|g''(\alpha)|.
\end{aligned}$$

- Considere a função de iteração

$$g(x) = x + A(x)f(x), \qquad \text{com } f'(x) \neq 0,$$

onde a função $A(x)$ é escolhida de modo que $A(\alpha) \neq 0$.
- Se escolhermos $A(x)$ tal que $g'(\alpha) = 0$, teremos que $|g'(x)| < 1$ para todo x numa vizinhança de α, o que garante a convergência do método.
- Derivando $g(x)$, obtemos

$$g'(x) = 1 + A'(x)f(x) + A(x)f'(x).$$

- Escolhendo

$$A(x) = -\frac{1}{f'(x)},$$

segue que $g'(\alpha) = 0$.

- Assim, a escolha

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

define o processo iterativo

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

chamado de **método de Newton**.

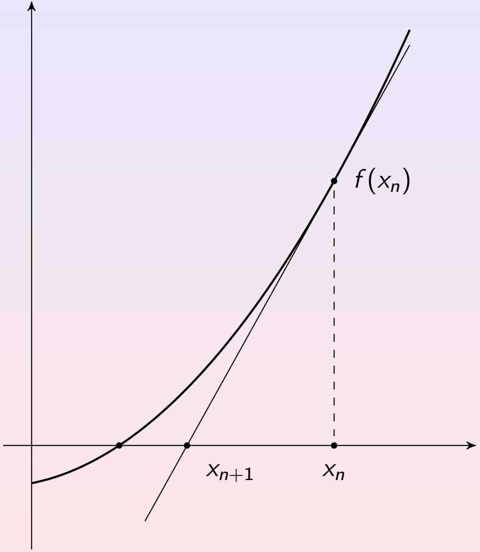
- O método de Newton sempre converge se $|x_0 - \alpha|$ for suficientemente pequeno.

Teorema

Assuma que f'' é contínua em $[a, b]$, com $f'(\alpha) \neq 0$ e $f''(\alpha) \neq 0$. Seja $\{x_n\}$ a sequência obtida no método de Newton. Suponha que $x_n \neq \alpha$ para todo $n \geq 0$. Seja $e_n = x_n - \alpha$. Então

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^2} = \frac{1}{2}\left|\frac{f''(\alpha)}{f'(\alpha)}\right|.$$

- Esse resultado é uma consequência de $g'(\alpha) = 0$ e $g''(\alpha) = \dfrac{f''(\alpha)}{f'(\alpha)}$.


$f(x_n)$
x_{n+1}
x_n

- Inicialização:
  - Escolha x_0.
- Para $n \geq 0$:
  - Faça $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$.
  - Pare se $|f(x_{n+1})| < \varepsilon$.

Exemplo

Use o método de Newton para encontrar uma estimativa para a raiz positiva da função

$$f(x) = e^x - 2x - 1,$$

com aproximação inicial $x_0 = 1$ e tolerância $\varepsilon = 10^{-5}$.

Exemplo

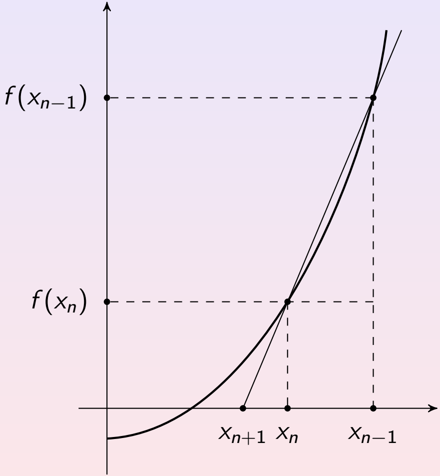
Use o método de Newton para encontrar a raiz da função $f(x) = \operatorname{sen}(x) - e^{-x}$ no intervalo $[0, 1]$, com ponto inicial $x_0 = 0{,}0$ e tolerância $|f(x_{n+1})| < \varepsilon = 10^{-5}$.

- Uma modificação do método de Newton consiste em substituir a derivada $f'(x_n)$ pelo quociente

$$f'(x_n) \simeq \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

- O método de Newton, com essa modificação, é conhecido como **método das secantes**.
- Com a substituição, obtemos a iteração

$$x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}.$$


$f(x_{n-1})$
$f(x_n)$
x_{n+1}
x_n
x_{n-1}

- Inicialização:
  - Escolha x_0 e x_1.
- Para $n \geq 1$:
  - Faça $x_{n+1} = x_n - \dfrac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}$.
  - Pare se $|f(x_{n+1})| < \varepsilon$.

Teorema

Assuma que f' é contínua num intervalo I contendo a raiz α de f satisfazendo $f'(\alpha) \neq 0$. Seja $\{x_n\}$ a sequência obtida no método das secantes, com x_0 e x_1 suficientemente próximos a α. Suponha que $x_n \neq \alpha$ para todo $n \geq 0$. Seja $e_n = x_n - \alpha$. Então

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^p} = c,$$

em que $c > 0$ e $p = (1 + \sqrt{5})/2 \simeq 1{,}618$.

Exemplo

Use o método das secantes para encontrar uma estimativa para a raiz positiva da função

$$f(x) = e^x - 2x - 1,$$

com aproximações iniciais $x_0 = 1$ e $x_1 = 2$ e tolerância $\varepsilon = 10^{-3}$.

Exemplo

Aplique o método das secantes para encontrar a raiz da função $f(x) = \cos(x) - e^{-x}$ no intervalo $[1, 2]$, com pontos iniciais $x_0 = 1{,}0$ e $x_1 = 1{,}2$, e tolerância $|f(x_{n+1})| < \varepsilon = 10^{-3}$.

Um sistema de n equações lineares é escrito como:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

em que a_{ij} são os coeficientes, b_i são os termos independentes, e x_i são as incógnitas.

Uma forma conveniente de representar um sistema de n equações lineares é em sua forma matricial:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

ou, simplesmente,

$$Ax = b$$

em que A é chamada de matriz dos coeficientes, b é o vetor do termo independente, e x é o vetor solução.

Métodos numéricos para solução de sistemas de equações lineares podem ser classificados como:

- Métodos Exatos: Executados com um número finito de operações, fornecem a solução exata. Porém, devido a erros de arredondamento, problemas com propagação de erros podem acontecer.
- Métodos Iterativos: A solução é obtida com uma dada precisão, como resultado de um processo iterativo convergente.

Definição

Seja $C = (c_{ij})$ uma matriz quadrada. A matriz C é chamada de triangular superior (triangular inferior) se $c_{ij} = 0$ para $i > j$ ($i < j$).

Definição

Um sistema de equações lineares é chamado de triangular superior (triangular inferior) se a matriz dos coeficientes for triangular superior (triangular inferior).

Assumindo $a_{ii} \neq 0$, o sistema triangular superior

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix}$$

pode ser resolvido usando a seguinte fórmula, chamada de substituição reversa:

$$\begin{cases} x_n = \dfrac{b_n}{a_{nn}} \\ x_i = \dfrac{b_i - \sum_{j=i+1}^{n} a_{ij} x_j}{a_{ii}}, \qquad i = n-1, \ldots, 1 \end{cases}$$

Já o sistema triangular inferior, com $a_{ii} \neq 0$,

$$\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix}$$

pode ser resolvido usando a seguinte fórmula, chamada de substituição direta:

$$\begin{cases} x_1 = \dfrac{b_1}{a_{11}} \\ x_i = \dfrac{b_i - \sum_{j=1}^{i-1} a_{ij} x_j}{a_{ii}}, \qquad i = 2, \ldots, n \end{cases}$$

Exemplo

Resolva o sistema triangular superior

$$\begin{pmatrix} 2 & 3 & 4 \\ 0 & -1 & -2 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ -3 \end{pmatrix}.$$

R.: $(-2, 0, 1)^T$.

Exemplo

Resolva o sistema triangular superior

$$\begin{pmatrix} 4 & -1 & 3 & 4 \\ 0 & 2 & -3 & 5 \\ 0 & 0 & 7 & -3 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 12 \\ 7 \\ -13 \\ 4 \end{pmatrix}.$$

R.: $(1, -3, -1, 2)^T$.

Exemplo

Resolva o sistema triangular inferior

$$\begin{pmatrix} -7 & 0 & 0 \\ 3 & 1 & 0 \\ 9 & -5 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -21 \\ 7 \\ 39 \end{pmatrix}.$$

R.: $(3, -2, 1)^T$.

Exemplo

Resolva o sistema triangular inferior

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 4 & 2 & 0 & 0 \\ -1 & 3 & 3 & 0 \\ 2 & -1 & -2 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ 16 \\ -5 \end{pmatrix}.$$

R.: $(-1, 2, 3, -1)^T$.

Considere a matriz aumentada

$$\tilde{A}^{(1)} = (A^{(1)}, b^{(1)}) = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ a_{31}^{(1)} & a_{32}^{(1)} & a_{33}^{(1)} & \cdots & a_{3n}^{(1)} & b_3^{(1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & a_{n3}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{pmatrix},$$

em que $a_{ij}^{(1)} = a_{ij}$ e $b_i^{(1)} = b_i$.
Em cada estágio do método da eliminação de Gauss, os elementos da k-ésima coluna abaixo da diagonal são zerados.

Na primeiro estágio, obtemos

$$\tilde{A}^{(2)} = (A^{(2)}, b^{(2)}) = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} & \cdots & a_{1n}^{(2)} & b_1^{(2)} \\ & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ & a_{32}^{(2)} & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} & b_3^{(2)} \\ & \vdots & \vdots & \ddots & \vdots & \vdots \\ & a_{n2}^{(2)} & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{pmatrix},$$

Já no segundo,

$$\tilde{A}^{(3)} = (A^{(3)}, b^{(3)}) = \begin{pmatrix} a_{11}^{(3)} & a_{12}^{(3)} & a_{13}^{(3)} & \cdots & a_{1n}^{(3)} & b_1^{(3)} \\ & a_{22}^{(3)} & a_{23}^{(3)} & \cdots & a_{2n}^{(3)} & b_2^{(3)} \\ & & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & b_3^{(3)} \\ & & \vdots & \ddots & \vdots & \vdots \\ & & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} & b_n^{(3)} \end{pmatrix}$$

Continuamos até chegarmos à matriz triangular superior

$$\tilde{A}^{(n)} = (A^{(n)}, b^{(n)}) = \begin{pmatrix} a_{11}^{(n)} & a_{12}^{(n)} & a_{13}^{(n)} & \cdots & a_{1n}^{(n)} & b_1^{(n)} \\ & a_{22}^{(n)} & a_{23}^{(n)} & \cdots & a_{2n}^{(n)} & b_2^{(n)} \\ & & a_{33}^{(n)} & \cdots & a_{3n}^{(n)} & b_3^{(n)} \\ & & & \ddots & \vdots & \vdots \\ & & & & a_{nn}^{(n)} & b_n^{(n)} \end{pmatrix}$$

Podemos então encontrar a solução de $Ax = b$ como solução do sistema triangular superior $A^{(n)}x = b^{(n)}$.

- Para zerar os elementos da k-ésima coluna abaixo da diagonal, substituímos a i-ésima linha pela diferença entre a i-ésima linha e a k-ésima linha multiplicada por $m_{ik} = \dfrac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$.
- Equivalentemente, para $k = 1, \ldots, n-1$,

$$\begin{cases} m_{ik} = \dfrac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \\ a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik} a_{kj}^{(k)} \\ b_i^{(k+1)} = b_i^{(k)} - m_{ik} b_k^{(k)} \end{cases}$$

para $i = k+1, \ldots, n$ e $j = k, \ldots, n$.
- O termo $a_{kk}^{(k)}$ é chamado de pivô.

Resultado

Para garantirmos que $a_{kk}^{(k)} \neq 0$, para $k = 1, \ldots, n$, devemos assumir que $\det(A_k) \neq 0$, em que $A_k = (a_{ij})_{k \times k}$ é o menor principal de A de ordem k.

Para calcularmos o número de operações no método da eliminação de Gauss, escrevemos

$$\begin{aligned}\#\ \text{operações} &= \sum_{k=1}^{n-1} (\#\ \text{operações no estágio } k) \\ &= \sum_{k=1}^{n-1} \sum_{i=k+1}^{n} (\#\ \text{operações na linha } i)\end{aligned}$$

- Com isso, temos

$$\#\ \text{divisões} = \sum_{k=1}^{n-1} \sum_{i=k+1}^{n} 1 = \frac{(n-1)n}{2}$$

e

$$\begin{aligned}\#\ \text{multiplicações} &= \sum_{k=1}^{n-1} \sum_{i=k+1}^{n} \{[n-(k-1)]+1\} \\ &= \frac{2n^3+3n^2-5n}{6}\end{aligned}$$

- O número de subtrações é igual ao número de multiplicações.
- Portanto, o método da eliminação de Gauss possui uma complexidade aritmética da ordem de n^3.

Exemplo

Use o método da eliminação de Gauss para resolver o sistema

$$\begin{pmatrix} 1 & 3 & -2 \\ -1 & 0 & 0 \\ 2 & -3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}.$$

R.: $(1, 1, 1)^T$.

Exemplo

Use o método da eliminação de Gauss para resolver o sistema

$$\begin{pmatrix} 1 & -3 & 2 \\ 2 & -2 & 3 \\ 3 & -13 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 12 \\ 15 \\ 47 \end{pmatrix}.$$

R.: $(4, -2, 1)^T$.

Exemplo

Usando o método da eliminação de Gauss, resolva o sistema

$$\begin{pmatrix} 1 & -1 & 2 & -1 \\ -3 & 5 & -3 & 2 \\ 1 & -5 & -6 & 2 \\ 3 & -11 & -8 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -8 \\ 21 \\ 3 \\ -8 \end{pmatrix}.$$

R.: $(-1, 2, -2, 1)^T$.

Exemplo

Usando o método da eliminação de Gauss, resolva o sistema

$$\begin{pmatrix} 3 & -2 & 1 & -1 \\ 6 & -6 & 1 & 1 \\ -9 & -2 & -8 & 17 \\ -6 & -2 & -4 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -15 \\ -30 \\ 48 \\ 25 \end{pmatrix}.$$

R.: $(-3, 2, -1, 1)^T$.

Exemplo

Resolva o sistema

$$\begin{cases} 0{,}0003x_1 + 3{,}0000x_2 = 2{,}0001 \\ 1{,}0000x_1 + 1{,}0000x_2 = 1{,}0000 \end{cases}$$

usando o método da eliminação de Gauss, com quatro e cinco algarismos significativos.

R.: $(1{,}0000; 0{,}6666)^T$, $(0{,}40000; 0{,}66666)^T$, $(1/3, 2/3)^T$.

Exemplo

Resolva o sistema

$$\begin{cases} 1{,}0000x_1 + 1{,}0000x_2 = 1{,}0000 \\ 0{,}0003x_1 + 3{,}0000x_2 = 2{,}0001 \end{cases}$$

usando o método da eliminação de Gauss, com quatro e cinco algarismos significativos.

R.: $(0{,}3334; 0{,}6666)^T$, $(0{,}33334; 0{,}66666)^T$, $(1/3, 2/3)^T$.

Pivotamento parcial

No método da eliminação de Gauss com pivotamento parcial, antes de iniciar o k-ésimo estágio, as linhas da matriz $A^{(k)}$ são permutadas de modo que $|a_{kk}^{(k)}| \geq |a_{ik}^{(k)}|$, para $i = k, \ldots, n$. O pivô é escolhido como sendo um dos elementos de maior valor absoluto dentre $a_{kk}^{(k)}, a_{k+1,k}^{(k)}, \ldots, a_{nk}^{(k)}$.

Exemplo

Resolva o sistema $Ax = b$, em que

$$A = \begin{pmatrix} 1 & -2 & -1 \\ 2 & 2 & 1 \\ 3 & 1 & -8 \end{pmatrix}, \qquad \text{e} \qquad b = \begin{pmatrix} -1 \\ 10 \\ -6 \end{pmatrix},$$

usando o método da eliminação de Gauss com pivotamento parcial.

R.: $x = (3, 1, 2)^T$.

Exemplo

Resolva o sistema abaixo, com precisão de duas casas decimais, usando o método da eliminação de Gauss com pivotamento parcial:

$$\begin{pmatrix} -2 & 1 & 1 \\ 2 & 3 & 3 \\ 3 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}.$$

(*Obs.*: O arredondamento deve ser aplicado a cada operação aritmética.)

R.: $(-0{,}74; 1{,}24; -1{,}74)^T$.

- Seja A uma matriz quadrada. A fatoração LU da matriz A se refere ao produto

$$A = LU,$$

em que L é uma matriz triangular inferior, com os termos da diagonal iguais a 1, e U é uma matriz triangular superior.

- Uma aplicação do método da eliminação de Gauss fornece a decomposição LU da matriz A.

- Organizamos os multiplicadores $m_{ik} = a_{ik}^{(k)}/a_{kk}^{(k)}$ na matriz triangular inferior
$$L = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_{21} & 1 & 0 & \cdots & 0 \\ m_{31} & m_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & m_{n3} & \cdots & 1 \end{pmatrix}.$$
- Seja $U = A^{(n)}$ a matriz triangular superior obtida no final do método da eliminação de Gauss.
- Essa matrizes satisfazem
$$A = LU.$$

- Uma vez encontrada a fatoração LU da matriz A, o sistema $Ax = b$ é resolvido da seguinte forma:
    - Primeiro, usando substituição direta, resolve-se $Ly = b$.
    - Depois, usando substituição reversa, resolve-se $Ux = y$.
- O procedimento acima é útil no caso em que o sistema linear $Ax = b$ precisa ser resolvido, com a mesma matriz A, e diferentes vetores b, pois não é necessário fatorar a matriz A novamente.

Exemplo

Encontre a fatoração LU da matriz

$$A = \begin{pmatrix} -1 & 2 & 3 \\ -2 & 7 & 4 \\ 3 & 9 & -15 \end{pmatrix},$$

e posteriormente resolva o sistema $Ax = b$ para $b = (5, 18, 21)^T$ e $b = (5, 9, -12)^T$.

R.: $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & 5 & 1 \end{pmatrix} \begin{pmatrix} -1 & 2 & 3 \\ 0 & 3 & -2 \\ 0 & 0 & 4 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}.$

Exemplo

Encontre a fatoração LU da matriz

$$A = \begin{pmatrix} 2 & 3 & -2 & -1 \\ 4 & 4 & -1 & -6 \\ 4 & 14 & -19 & 16 \\ -4 & 2 & -17 & 22 \end{pmatrix},$$

e posteriormente resolva o sistema $Ax = b$ para $b = (-3, -11, 21, 43)^T$ e $b = (-8, -20, 8, 54)^T$.

R.: $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & -4 & 1 & 0 \\ -2 & -4 & 3 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 & -2 & -1 \\ 0 & -2 & 3 & -4 \\ 0 & 0 & -3 & 2 \\ 0 & 0 & 0 & -2 \end{pmatrix}, \begin{pmatrix} 3 \\ -3 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ -3 \\ -2 \\ 1 \end{pmatrix}.$

Isolando a variável x_i na i-ésima linha do sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

podemos escrever

$$\begin{cases} x_1 = (b_1 - a_{12}x_2 - \cdots - a_{1n}x_n)/a_{11} \\ x_2 = (b_2 - a_{21}x_1 - \cdots - a_{2n}x_n)/a_{22} \\ \qquad\vdots \\ x_n = (b_n - a_{n1}x_1 - \cdots - a_{n,n-1}x_{n-1})/a_{nn} \end{cases}$$

No método de Jacobi, a sequência de aproximações $x^{(k)}$ para a solução do sistema $Ax = b$ é dada pela fórmula recursiva

$$\begin{cases} x_1^{(k+1)} = (b_1 - a_{12}x_2^{(k)} - \cdots - a_{1n}x_n^{(k)})/a_{11} \\ x_2^{(k+1)} = (b_2 - a_{21}x_1^{(k)} - \cdots - a_{2n}x_n^{(k)})/a_{22} \\ \qquad\vdots \\ x_n^{(k+1)} = (b_n - a_{n1}x_1^{(k)} - \cdots - a_{n,n-1}x_{n-1}^{(k)})/a_{nn} \end{cases}$$

em que assumimos $a_{ii} \neq 0$.

Podemos reescrever a fórmula recursiva para o método de Jacobi na forma matricial

$$x^{(k+1)} = D^{-1}(b - Mx^{(k)}), \qquad k \geq 0,$$

em que D e M são matrizes satisfazendo $A = D + M$, com

$$D = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \qquad \text{e} \qquad M = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{pmatrix}.$$

- Para a parada do método de Jacobi podemos usar o critério

$$\frac{\|x^{(k+1)} - x^{(k)}\|_\infty}{\|x^{(k+1)}\|_\infty} < \varepsilon,$$

dada uma tolerância $\varepsilon > 0$.

- A norma $\|\cdot\|_\infty$ corresponde ao maior dos valores absolutos dos componentes do vetor, i.e., se $x = (x_1, \ldots, x_n)^T$ então

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Exemplo

Use o método de Jacobi, com aproximação inicial $x^{(0)} = (0,0)^T$, e $\varepsilon = 5 \times 10^{-2}$ para o critério de parada, para encontrar a solução do sistema linear

$$\begin{cases} 3x_1 - 2x_2 = 1 \\ -x_1 + 4x_2 = 1 \end{cases}$$

R.:

k	$x^{(k)}$	$\frac{\|x^{(k)} - x^{(k-1)}\|_\infty}{\|x^{(k)}\|_\infty}$
0	$(0{,}00000; 0{,}00000)^T$	
1	$(0{,}33333; 0{,}25000)^T$	1,00000
2	$(0{,}50000; 0{,}33333)^T$	0,33333
3	$(0{,}55555; 0{,}37500)^T$	0,09999
4	$(0{,}58333; 0{,}38888)^T$	0,04762

Exemplo

Considere o sistema linear

$$\begin{pmatrix} 4 & -2 & 1 \\ 1 & -3 & -1 \\ -1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix}.$$

Dada a aproximação inicial $x^{(0)} = (0, 0, 0)^T$, encontre as aproximações sucessivas $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ e $x^{(4)}$, usando o método de Jacobi.

R.: Solução: $(0{,}04; -1{,}16; -0{,}48)^T$. Aproximações:

k	$x^{(k)}$
0	$(0{,}00000; 0{,}00000; 0{,}00000)^T$
1	$(0{,}50000; -1{,}33333; -0{,}50000)^T$
2	$(-0{,}04166; -1{,}00000; -0{,}25000)^T$
3	$(0{,}06250; -1{,}26388; -0{,}52083)^T$
4	$(-0{,}00173; -1{,}13889; -0{,}46875)^T$

No método de Gauss–Seidel, a sequência de aproximações $x^{(k)}$ para a solução do sistema $Ax = b$ é dada pela fórmula recursiva (assumimos $a_{ii} \neq 0$)

$$\begin{cases} x_1^{(k+1)} = (b_1 - a_{12}x_2^{(k)} - \cdots - a_{1n}x_n^{(k)})/a_{11} \\ x_2^{(k+1)} = (b_2 - a_{21}x_1^{(k+1)} - \cdots - a_{2n}x_n^{(k)})/a_{22} \\ \qquad \vdots \\ x_n^{(k+1)} = (b_n - a_{n1}x_1^{(k+1)} - \cdots - a_{n,n-1}x_{n-1}^{(k+1)})/a_{nn} \end{cases}$$

ou, equivalentemente,

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}\right),$$

para $i = 1, \ldots, n$.

Vamos reescrever a fórmula recursiva na forma matricial. Para isso, passamos todos o termos do estágio $k+1$ para o lado esquerdo da igualdade. Assim, obtemos

$$\left(\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)}\right) + a_{ii}x_i^{(k+1)} = b_i - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)},$$

e então

$$\sum_{j=1}^{i} a_{ij}x_j^{(k+1)} = b_i - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}.$$

Sejam L e U tais que $A = L + U$, com

$$L = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \qquad \text{e} \qquad U = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Com isso, obtemos a seguinte fórmula na forma matricial:

$$Lx^{(k+1)} = b - Ux^{(k)}, \tag{$*$}$$

que resulta em

$$x^{(k+1)} = L^{-1}(b - Ux^{(k)}).$$

- O sistema $(*)$ pode ser resolvido usando substituição direta.
- Usamos como critério de parada no método de Gauss–Seidel o mesmo critério usado no método de Jacobi.

Exemplo

Use o método de Gauss–Seidel, com aproximação inicial $x^{(0)} = (1,1)^T$, e $\varepsilon = 5 \times 10^{-2}$ para o critério de parada, para encontrar a solução do sistema linear

$$\begin{cases} 5x_1 + 2x_2 = 1 \\ -x_1 + 4x_2 = 2 \end{cases}$$

R.:

k	$x^{(k)}$	$\frac{\Vert x^{(k)} - x^{(k-1)} \Vert_\infty}{\Vert x^{(k)} \Vert_\infty}$
0	$(1{,}00000; 1{,}00000)^T$	
1	$(-0{,}20000; 0{,}45000)^T$	2,66670
2	$(0{,}02000; 0{,}50500)^T$	0,43564
3	$(-0{,}00200; 0{,}49950)^T$	0,04404

Exemplo

Considere o sistema linear

$$\begin{pmatrix} 4 & -2 & 1 \\ 1 & -3 & -1 \\ -1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix}.$$

Dada a aproximação inicial $x^{(0)} = (1, 1, 1)^T$, use o método de Gauss–Seidel, para encontrar as aproximações sucessivas $x^{(1)}$, $x^{(2)}$ e $x^{(3)}$.

R.: Solução: $(0{,}04; -1{,}16; -0{,}48)^T$. Aproximações:

k	$x^{(k)}$
0	$(1{,}00000; 1{,}00000; 1{,}00000)^T$
1	$(0{,}75000; -1{,}41666; -0{,}12500)^T$
2	$(-0{,}17708; -1{,}35069; -0{,}58854)^T$
3	$(-0{,}02821; -1{,}14655; -0{,}51410)^T$

Os métodos de Jacobi e Gauss–Seidel podem ser expressos na forma matricial como

$$x^{(k+1)} = Cx^{(k)} + g, \qquad \text{para } k \geq 0, \tag{$*$}$$

com x satisfazendo a igualdade $x = Cx + g$ se, e somente se, $Ax = b$.

- Com respeito ao método de Jacobi, temos

$$x^{(k+1)} = D^{-1}(b - Mx^{(k)}) = -D^{-1}Mx^{(k)} + D^{-1}b,$$

e a fórmula $(*)$ é satisfeita para $C = -D^{-1}M$ e $g = D^{-1}b$.

- Para esses valores de C e g, vemos que

$$Ax = (D + M)x = b$$

é equivalente a

$$x = D^{-1}(b - Mx) = Cx + g.$$

- Já no método de Gauss–Seidel, o desenvolvimento

$$x^{(k+1)} = L^{-1}(b - Ux^{(k)}) = -L^{-1}Ux^{(k)} + L^{-1}b$$

implica que $(*)$ é satisfeita para $C = -L^{-1}U$ e $g = L^{-1}b$.

- Com esses valores de C e g, temos que

$$Ax = (L + U)x = b$$

é equivalente a

$$x = L^{-1}(b - Ux) = Cx + g.$$

Subtraindo as equações $x^{(k+1)} = Cx^{(k)} + g$ e $x = Cx + g$, obtemos

$$x^{(k+1)} - x = C(x^{(k)} - x)$$

Usando a desigualdade triangular, chegamos a

$$\begin{aligned} |x_i^{(k+1)} - x_i| &= \left| \sum_{j=1}^{n} c_{ij}(x_j^{(k)} - x_j) \right| \leq \sum_{j=1}^{n} |c_{ij}| \cdot |x_j^{(k)} - x_j| \\ &\leq \left(\sum_{j=1}^{n} |c_{ij}| \right) \cdot \max_{1 \leq j \leq n} |x_j^{(k)} - x_j| \end{aligned}$$

Deste modo,

$$\max_{1 \leq i \leq n} |x_i^{(k+1)} - x_i| \leq \left[\max_{1 \leq i \leq n} \left(\sum_{j=1}^{n} |c_{ij}| \right) \right] \cdot \max_{1 \leq j \leq n} |x_j^{(k)} - x_j|$$

Definindo

$$\|C\| = \max_{1\le i\le n}\left(\sum_{j=1}^{n}|c_{ij}|\right)$$

podemos escrever

$$\|x^{(k+1)} - x\|_\infty \le \|C\| \cdot \|x^{(k)} - x\|_\infty$$

Consequentemente,

$$\|x^{(k)} - x\|_\infty \le \|C\|^k \cdot \|x^{(0)} - x\|_\infty.$$

Portanto, se $\|C\| < 1$ então $\|x^{(k)} - x\|_\infty \to 0$, que implica $x_i^{(k)} \to x_i$ para $i = 1, \dots, n$.

Teorema (Critério de convergência)

Se $\|C\| < 1$ *então a sequência*

$$x^{(k+1)} = Cx^{(k)} + g, \qquad k \geq 0,$$

converge para $x = Cx + g$.

No método de Jacobi, a matriz C é dada por

$$C = -D^{-1}M = -\begin{pmatrix} 0 & \frac{a_{12}}{a_{11}} & \cdots & \frac{a_{1n}}{a_{11}} \\ \frac{a_{21}}{a_{22}} & 0 & \cdots & \frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n1}}{a_{nn}} & \frac{a_{n2}}{a_{nn}} & \cdots & 0 \end{pmatrix}.$$

A desigualdade

$$\|C\| = \max_{1\leq i\leq n}\left(\sum_{j\neq i}\left|\frac{a_{ij}}{a_{ii}}\right|\right) < 1$$

é satisfeita se e somente se

$$\sum_{j\neq i}|a_{ij}| < |a_{ii}|, \qquad \text{para } i = 1, \ldots, n. \tag{$*$}$$

Uma matriz A satisfazendo $(*)$ é chamada de diagonalmente estritamente dominante.

É possível verificar que, se A é diagonalmente estritamente dominante, então o método de Gauss–Seidel também converge.

Teorema (Critério das linhas)

Se a matriz A é diagonalmente estritamente dominante, i.e.,

$$\sum_{j \neq i} |a_{ij}| < |a_{ii}|, \qquad i = 1, \ldots, n,$$

então os métodos de Jacobi e Gauss–Seidel geram uma sequência que converge para a solução do sistema linear $Ax = b$, qualquer que seja a aproximação inicial $x^{(0)}$.

Exemplo

Verifique se o critério das linhas é válido para o sistema linear

$$\begin{cases} 8x_1 + 2x_2 - 2x_3 - 3x_4 = 1 \\ 2x_1 + 9x_2 + 3x_3 + 4x_4 = 5 \\ 3x_1 - 2x_2 + 7x_3 - 3x_4 = 3 \\ x_1 + 2x_2 + 3x_3 + 8x_4 = 2 \end{cases}$$

R.: Não é satisfeito.

Identidade de Newton

Considere o polinômio

$$P(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n,$$

cujas raízes são $x_1, x_2, \ldots, x_n$. Se denotarmos

$$s_k = \sum_{i=1}^{n} x_i^k, \qquad 1 \leq k \leq n,$$

então vale a relação

$$\sum_{i=0}^{k-1} a_i s_{k-i} + k a_k = 0, \qquad 1 \leq k \leq n.$$

Exemplo

Sejam $s_1 = 6$, $s_2 = 14$ e $s_3 = 36$ as somas das potências das raízes de um polinômio $P(x)$. Assumindo que o coeficiente líder é igual a 1 ($a_0 = 1$), determine $P(x)$.

R.: $P(x) = x^3 - 6x^2 + 11x - 6$.

- O polinômio característico de uma matriz A, de dimensão $n \times n$, é definida como
$$P(\lambda) = \det(\lambda I - A).$$
- Adotaremos a seguinte notação para o polinômio característico:
$$P(\lambda) = \lambda^n - p_1\lambda^{n-1} - p_2\lambda^{n-2} - \cdots - p_{n-1}\lambda - p_n.$$
- Sejam $\lambda_1, \ldots, \lambda_n$ os autovalores da matriz A. Isto é, $\lambda_1, \ldots, \lambda_n$ são as raízes do polinômio característico $P(\lambda)$.

- Pela identidade de Newton, se

$$s_k = \sum_{i=1}^{n} \lambda_i^k, \qquad 1 \leq k \leq n,$$

então

$$kp_k = s_k - p_1 s_{k-1} - \cdots - p_{k-1} s_1, \qquad 1 \leq k \leq n.$$

- Portanto, conhecendo $s_1, \ldots, s_n$, pode-se determinar os coeficientes $p_1, \ldots, p_n$ de $P(\lambda)$.
- Os números $s_1, \ldots, s_n$ são encontrados como

$$s_k = \mathrm{tr}(A^k), \qquad 1 \leq k \leq n,$$

em que $\mathrm{tr}(\cdot)$ denota o traço da matriz.
- O procedimento acima para determinar os coeficientes $p_1, \ldots, p_n$ é chamado de Método de Leverrier.

Exemplo

Encontre o polinômio característico e os autovalores da matriz

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix},$$

usando o método de Leverrier.

R.: $P(\lambda) = -\lambda^3 + 2\lambda^2 - 2\lambda + 2 = (\lambda - 1)(\lambda + \sqrt{2})(\lambda - \sqrt{2})$.

- Defina uma sequência de matrizes $B_1, \ldots, B_n$ do seguinte modo:

$$B_1 = A, \qquad q_1 = \mathrm{tr}(B_1), \qquad D_1 = B_1 - q_1 I,$$

e

$$B_k = AD_{k-1}, \qquad q_k = \frac{1}{k}\,\mathrm{tr}(B_k), \qquad D_k = B_k - q_k I,$$

para $2 \leq k \leq n$.

- Os termos q_k obtidos acima constituem os coeficientes do polinômio característico, isto é,

$$q_k = p_k, \qquad 1 \leq k \leq n.$$

- Tal procedimento é chamado de Método de Leverrier–Faddeev.

Exemplo

Encontre o polinômio característico e os autovalores da matriz

$$A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix},$$

usando o método de Leverrier–Faddeev.

R.: $P(\lambda) = \lambda^3 - 12\lambda - 16 = (\lambda - 4)(\lambda + 2)^2$.

Método das Potências

Seja A uma matriz real de dimensão $n \times n$, com autovalores $\lambda_1, \ldots, \lambda_n$ e respectivos autovetores $v^{(1)}, \ldots, v^{(n)}$. Suponha que os autovetores são linearmente independentes, e que $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$. Assuma que os autovetores são normalizados de forma que $||v^{(i)}||_\infty = v^{(i)}_{q_i} = 1$ para algum índice q_i. Seja $x^{(0)}$ um vetor cuja representação $x^{(0)} = \sum_{i=1}^n c_i v^{(i)}$ satisfaz $c_1 \neq 0$. Considere as sequências de vetores $x^{(k)}$ e $y^{(k)}$ definidas por

$$y^{(k)} = Ax^{(k-1)} \quad e \quad x^{(k)} = \frac{y^{(k)}}{y^{(k)}_{p_k}}, \qquad \text{para } k \geq 1,$$

em que p_k é o menor inteiro tal que $|y^{(k)}_{p_k}| = ||y^{(k)}||_\infty$. Então $x^{(k)}$ converge para $v^{(1)}$, e $\lambda^{(k)} = y^{(k)}_{p_{k-1}}$ converge para λ_1.

- Em vista de

$$x^{(k)} = \frac{Ax^{(k-1)}}{y^{(k)}_{p_k}}, x^{(k-1)} = \frac{Ax^{(k-2)}}{y^{(k-1)}_{p_{k-1}}}, \dots, x^{(1)} = \frac{Ax^{(0)}}{y^{(1)}_{p_1}},$$

obtemos

$$x^{(k)} = \frac{A^k x^{(0)}}{y^{(k)}_{p_k} \cdots y^{(1)}_{p_1}}$$

- Como $||x^{(k)}||_\infty = 1 = x^{(k)}_{p_k}$, inferimos que $y^{(k)}_{p_k} \cdots y^{(1)}_{p_1} = (A^k x^{(0)})_{p_k}$.
- Com isso,

$$x^{(k)} = \frac{A^k x^{(0)}}{(A^k x^{(0)})_{p_k}}$$

- Em vista de $A^k v^{(i)} = \lambda_i^k v^{(i)}$, segue que

$$A^k x^{(0)} = \sum_{i=1}^{n} c_i \lambda_i^k v^{(i)} = \lambda_1^k \left(c_1 v^{(1)} + \sum_{i=2}^{n} c_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v^{(i)} \right)$$

- Como $|\lambda_i / \lambda_1| < 1$ para $2 \leq i \leq n$, obtemos que $p_k = q_1$, e assim $v_{p_k}^{(1)} = v_{q_1}^{(1)} = 1$, para todo k suficientemente grande.
- Logo,

$$x^{(k)} = \frac{c_1 v^{(1)} + \sum_{i=2}^{n} c_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v^{(i)}}{c_1 v_{p_k}^{(1)} + \sum_{i=2}^{n} c_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_{p_k}^{(i)}} \to v^{(1)}$$

- Em vista das manipulações:

$$\begin{aligned} y^{(k)}_{p_{k-1}} &= (Ax^{(k-1)})_{p_{k-1}} \\ &= \left(A\frac{A^{k-1}x^{(0)}}{(A^{k-1}x^{(0)})_{p_{k-1}}}\right)_{p_{k-1}} \\ &= \frac{(A^k x^{(0)})_{p_{k-1}}}{(A^{k-1}x^{(0)})_{p_{k-1}}}, \end{aligned}$$

podemos escrever

$$y^{(k)}_{p_{k-1}} = \lambda_1 \frac{c_1 v^{(1)}_{p_{k-1}} + \sum_{i=2}^n c_i \left(\frac{\lambda_i}{\lambda_1}\right)^k v^{(i)}_{p_{k-1}}}{c_1 v^{(1)}_{p_{k-1}} + \sum_{i=2}^n c_i \left(\frac{\lambda_i}{\lambda_1}\right)^{k-1} v^{(i)}_{p_{k-1}}} \to \lambda_1$$

- A seguir expomos uma ilustração do método das potências
- A matriz
$$A = \begin{pmatrix} -2 & -3 \\ 6 & 7 \end{pmatrix}$$
tem autovalores $\lambda_1 = 4$ e $\lambda_2 = 1$ com respectivos autovetores $v^{(1)} = (-0{,}5; 1)^T$ e $v^{(2)} = (1, -1)^T$.
- Começando com $x^{(0)} = (1, 1)^T$, e multiplicando sucessivamente por A, obtemos

$$Ax^{(0)} = \begin{pmatrix} -5 \\ 13 \end{pmatrix}, \quad A^2x^{(0)} = \begin{pmatrix} -29 \\ 61 \end{pmatrix}, \quad A^3x^{(0)} = \begin{pmatrix} -125 \\ 253 \end{pmatrix},$$
$$A^4x^{(0)} = \begin{pmatrix} -509 \\ 1021 \end{pmatrix}, \quad A^5x^{(0)} = \begin{pmatrix} -2045 \\ 4093 \end{pmatrix}, \quad A^6x^{(0)} = \begin{pmatrix} -8189 \\ 16381 \end{pmatrix}$$

- Como consequência, aproximações para o autovetor dominante $\lambda_1 = 4$ são

$$\lambda^{(1)} = \frac{13}{1} = 13{,}00000, \qquad \lambda^{(2)} = \frac{61}{13} = 4{,}6923,$$
$$\lambda^{(3)} = \frac{253}{61} = 4{,}14754, \qquad \lambda^{(4)} = \frac{1021}{253} = 4{,}03557,$$
$$\lambda^{(5)} = \frac{4093}{1021} = 4{,}00881, \qquad \lambda^{(6)} = \frac{16381}{4093} = 4{,}00200$$

- Uma aproximação para o autovetor relativo ao autovalor dominante $\lambda_1 = 4$ é

$$x^{(6)} = \frac{A^6 x^{(0)}}{(A^6 x^{(0)})_{p_6}} = \frac{1}{16381}\begin{pmatrix} -8189 \\ 16381 \end{pmatrix} = \begin{pmatrix} -0{,}49908 \\ 1 \end{pmatrix}$$

Exemplo

Considere a matriz

$$A = \begin{pmatrix} -15 & 23 & -9 \\ -22 & 33 & -12 \\ -28 & 41 & -14 \end{pmatrix}.$$

Dado o vetor inicial $x^{(0)} = (1, 0, 0)^T$, aplicando o método das potências, encontre as aproximações $\lambda^{(k)}$ e $x^{(k)}$, até a iteração $k = 5$, para o autovalor dominante λ_1 e o respectivo autovetor $v^{(1)}$ da matriz A.

R.: $\lambda_1 = 4$, $v^{(1)} = (1/3, 2/3, 1)^T$. Aproximações:

k	$\lambda^{(k)}$	$x^{(k)}$
1	$-15{,}00000$	$(0{,}53571; 0{,}78571; 1{,}00000)^T$
2	$3{,}21428$	$(0{,}32222; 0{,}66666; 1{,}00000)^T$
3	$4{,}31111$	$(0{,}34793; 0{,}67525; 1{,}00000)^T$
4	$3{,}94329$	$(0{,}33267; 0{,}66666; 1{,}00000)^T$
5	$4{,}01830$	$(0{,}33425; 0{,}66720; 1{,}00000)^T$

- O método QR é usado para encontrar os autovalores de uma matriz A.
- Um número $\lambda \in \mathbb{R}$ é dito ser um autovalor da matriz A se existir um vetor v tal que
$$Av = \lambda v.$$
- O vetor v que satisfaz a equação acima é chamado de autovetor.
- Numa matriz triangular (superior ou inferior), seus autovalores correspondem aos elementos de sua diagonal.

- Na fatoração QR a matriz A é escrita como o produto

$$A = QR,$$

em que Q é uma matriz ortonormal, e R é uma matriz triangular superior.

- Uma matriz Q é chamada de ortonormal se

$$Q^T Q = QQ^T = I.$$

- Se A é não singular, então a fatoração QR é única se é requerido que os elementos da diagonal de R são positivos.

- Para encontramos a fatoração QR de uma matriz $A = (a_{ij})_{n \times n}$, pode-se usar o processo de Gram–Schmidt.
- Para isso, expressamos a matriz A como

$$A = (a_1, \ldots, a_n),$$

em que $a_j = (a_{1j}, \ldots, a_{nj})^T$ são os vetores coluna de A.
- No processo de Gram–Schmidt, um conjunto ortonormal de vetores coluna são gerados a partir dos vetores coluna de A.
- Dado um vetor coluna a, denotamos $\|a\| = (a^T a)^{1/2}$.

- Os vetores q_i são obtidos como

$$\begin{cases} a'_i = a_i - \sum_{k=1}^{i-1} (q_k^T a_i) q_k, \\ q_i = \dfrac{a'_i}{\|a'_i\|}. \end{cases}$$

- A matriz Q é então composta pelos vetores coluna q_i:

$$Q = (q_1, \ldots, q_n).$$

- Já as entradas da matriz $R = (r_{ij})$ são dadas como

$$r_{ij} = \begin{cases} q_i^T a_j, & i \leq j, \\ 0, & i > j. \end{cases}$$

- Os valores de r_{ij} são encontrados durante o cálculo dos vetores coluna q_i.

Exemplo

Usando o processo de Gram–Schmidt, encontre a fatoração QR da matriz

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}.$$

R.: $A = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \sqrt{5} & 2\sqrt{5} \\ 0 & \sqrt{5} \end{pmatrix}.$

Exemplo

Usando o processo de Gram–Schmidt, encontre a fatoração QR da matriz

$$A = \begin{pmatrix} 2 & 10 & 3 \\ 1 & -4 & -3 \\ 2 & 1 & -6 \end{pmatrix}.$$

R.: $A = \begin{pmatrix} 2/3 & 2/3 & 1/3 \\ 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \end{pmatrix} \begin{pmatrix} 3 & 6 & -3 \\ 0 & 9 & 6 \\ 0 & 0 & 3 \end{pmatrix}.$

Exemplo

Usando o processo de Gram–Schmidt, encontre a fatoração QR da matriz

$$A = \begin{pmatrix} -2 & 2 & 1 & 2 \\ 2 & 4 & 0 & 3 \\ -2 & 2 & 3 & 2 \\ 2 & 4 & 2 & 7 \end{pmatrix}.$$

R.: $A = \begin{pmatrix} -1/2 & 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 4 & 2 & -1 & 3 \\ 0 & 6 & 3 & 7 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$

- O método QR usa transformações de similaridade para transformar a matriz A na forma triangular.
- Uma transformação de similaridade é definida como $A' = M^{-1}AM$.
- As matrizes A e A' são ditas ser similares.
- Os autovalores de matrizes similares são idênticos, porém com autovetores diferentes.

- No método QR o produto $A = QR$, obtido na fatoração QR, é invertido resultando na matriz

$$A' = RQ.$$

- A matrizes A e A' são similares, e portanto têm os mesmos autovalores.
- Este resultado segue das igualdades abaixo:

$$Q^{-1}AQ = Q^{-1}QRQ = RQ = A'.$$

- No método QR, dada a matriz $A^{(k)}$ no k-ésimo estágio, encontramos sua decomposição QR como
$$A^{(k)} = Q^{(k)} R^{(k)},$$
e então calculamos
$$A^{(k+1)} = R^{(k)} Q^{(k)}.$$
- A inicialização é dada por $A^{(0)} = A$.
- Os elementos da diagonal de $A^{(k)}$ convergem para os autovalores de A, a medida que $A^{(k)}$ se aproxima da forma triangular superior.

Teorema

Seja A uma matriz não singular tendo autovalores

$$\lambda_1, \ldots, \lambda_n, \quad \textit{com} \quad |\lambda_1| > \cdots > |\lambda_n| > 0.$$

Escolha a matriz X tal que

$$X^{-1}AX = \Lambda := \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Suponha que X^{-1} admite uma decomposição LU. Seja $\{A^{(k)}\}$, com $A^{(0)} = A$, a sequência gerada no método QR. Denote por $a_{ij}^{(k)}$ a entrada de posição ij de $A^{(k)}$. Então

$$a_{ij}^{(k)} \to 0, \qquad \textit{para } i > j,$$

e

$$a_{ii}^{(k)} \to \lambda_i, \qquad \textit{para } i = 1, \ldots, n,$$

Exemplo

Aplique o método QR, com 3 iterações, à matriz

$$A = \begin{pmatrix} 1 & 5 \\ 3 & 2 \end{pmatrix}.$$

R.:

Exemplo

Aplique o método QR, com 3 iterações, à matriz

$$A = \begin{pmatrix} 25 & 32 & 40 \\ 32 & 1 & 8 \\ 40 & 8 & 37 \end{pmatrix}.$$

R.: $\Lambda = \operatorname{diag}(81, -27, 9)$.

- No problema de interpolação, dado um conjunto de pares ordenados, $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$, com $x_0, x_1, \ldots, x_n$ distintos, que organizamos na tabela

x	x_0	x_1	$\cdots$	x_n
y	y_0	y_1	$\cdots$	y_n

  procuramos uma função φ que interpola os pontos tabelados, ou seja,
$$\varphi(x_k) = y_k, \qquad k = 0, 1, \ldots, n.$$
- Na interpolação polinomial, a função φ é um polinômio de grau $\leq n$.
- Desta forma, desejamos encontrar um polinômio p_n de grau $\leq n$ tal que
$$p_n(x_k) = y_k, \qquad k = 0, 1, \ldots, n.$$

- Denotando
$$p_n(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n,$$
queremos então encontrar $\alpha_0, \alpha_1, \ldots, \alpha_n$ tais que
$$\alpha_0 + \alpha_1 x_k + \cdots + \alpha_n x_k^n = y_k, \qquad k = 0, \ldots, n,$$
que corresponde a um sistema linear com $n+1$ equações e $n+1$ incógnitas.
- Na forma matricial, escrevemos
$$V\alpha = y,$$
em que
$$V = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix}, \quad \alpha = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \quad \text{e} \quad y = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

- A matriz V é chamada de matriz de Vandermonde.
- O sistema $V\alpha = y$ admite uma única solução, ou seja, $\det(V) \neq 0$, se os pontos $x_0, x_1, \ldots, x_n$ forem distintos.

Teorema (Existência e Unicidade)

Considere o conjunto $\{(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)\}$, *com* $x_i \neq x_j$ *para* $i \neq j$. *Então existe um único polinômio* p_n *de grau* $\leq n$ *tal que* $p_n(x_k) = y_k$, *para* $k = 0, 1, \ldots, n$.

Exemplo

Resolva um sistema linear para encontrar o polinômio de grau ≤ 2 que interpola os pontos dados na tabela

x	-1	2	3
y	-4	5	12

R.: $p_2(x) = -3 + 2x + x^2$.

- A matriz de Vandermonde apresenta problemas de mal condicionamento.
- Portanto, o polinômio obtido da solução do sistema $V\alpha = y$ pode conter erros de arredondamento.
- Veremos formas de encontrar o polinômio p_n que não utilizam a matriz de Vandermonde.

- Considere os polinômios

$$L_i(x) = \frac{(x-x_0)(x-x_1)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_0)(x_i-x_1)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)}$$
$$= \prod_{k\neq i} \frac{(x-x_k)}{(x_i-x_k)},$$

  para $i = 0, 1, \ldots, n$.
- Note que L_i é um polinômio de grau n com raízes $x_0, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$.
- Além disso, esses polinômios satisfazem

$$L_i(x_k) = \begin{cases} 1, & i = k, \\ 0, & i \neq k. \end{cases}$$

- No método de Lagrange, o polinômio interpolador é dado como

$$p_n(x) = y_0 L_0(x) + y_1 L_1(x) + \cdots + y_n L_n(x)$$
$$= \sum_{k=0}^{n} y_k L_k(x).$$

Exemplo

Use o método de Lagrange para encontrar o polinômio de grau ≤ 2 que interpola os pontos dados na tabela

x	-1	2	3
y	-4	5	12

R.: $p_2(x) = -3 + 2x + x^2$.

Exemplo

Usando o Método de Lagrange, encontre o valor do polinômio em $x = 1$, passando pelos pontos dados na tabela abaixo:

x	-2	-1	3	4
y	12	6	42	126

R.: $p_3(1) = -12$.

Exemplo

Usando o Método de Lagrange, encontre o valor do polinômio em $x = 1$, passando pelos pontos dados na tabela abaixo:

x	-3	-2	-1	0	2
y	-41	-8	1	2	4

R.: $p_4(1) = 3$.

- O cálculo dos polinômios $L_0, L_1, \ldots, L_n$ é computacionalmente custoso.
- Como alternativa, temos o método de Newton em que o polinômio interpolador é dado por

$$p_n(x) = \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)(x - x_1) + \cdots$$
$$\cdots + \alpha_n(x - x_0)(x - x_1)\cdots(x - x_{n-1}),$$

em que $\alpha_0, \alpha_1, \ldots, \alpha_n$ são obtidos da solução do sistema linear

$$\begin{pmatrix} 1 & & & & \\ 1 & (x_1 - x_0) & & & \\ 1 & (x_2 - x_0) & (x_2 - x_0)(x_2 - x_1) & & \\ 1 & \vdots & \vdots & \ddots & \\ 1 & (x_n - x_0) & (x_n - x_0)(x_n - x_1) & \cdots & (x_n - x_0)\cdots(x_n - x_{n-1}) \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

- Os coeficientes $\alpha_0, \alpha_1, \ldots, \alpha_n$ podem ser obtidos usando substituição direta.

Alternativamente, podemos calcular os coeficientes $\alpha_0, \alpha_1, \ldots, \alpha_n$ usando o operador diferenças divididas:

- Ordem 0: $f[x_k] = f(x_k) = y_k$;
- Ordem 1: $f[x_{k-1}, x_k] = \dfrac{f[x_k] - f[x_{k-1}]}{x_k - x_{k-1}}$;
- Ordem 2: $f[x_{k-2}, x_{k-1}, x_k] = \dfrac{f[x_{k-1}, x_k] - f[x_{k-2}, x_{k-1}]}{x_k - x_{k-2}}$;
- Ordem l: $f[x_{k-l}, \ldots x_k] = \dfrac{f[x_{k-l+1}, \ldots, x_k] - f[x_{k-l}, \ldots, x_{k-1}]}{x_k - x_{k-l}}$.

Os cálculos podem ser organizados na seguinte tabela:

x	Ordem 0	Ordem 1	Ordem 2	Ordem 3	$\cdots$	Ordem n
x_0	$y_0 = f[x_0]$					
		$f[x_0, x_1]$				
x_1	$y_1 = f[x_1]$		$f[x_0, x_1, x_2]$			
		$f[x_1, x_2]$		$f[x_0, x_1, x_2, x_3]$		
x_2	$y_2 = f[x_2]$		$f[x_1, x_2, x_3]$			
		$f[x_2, x_3]$		$f[x_1, x_2, x_3, x_4]$	$\ddots$	
x_3	$y_3 = f[x_3]$		$f[x_2, x_3, x_4]$			$f[x_0, x_1, \ldots, x_n]$
		$f[x_3, x_4]$	$\vdots$	$\vdots$		
x_4	$y_4 = f[x_4]$	$\vdots$		$f[x_{n-3}, x_{n-2}, x_{n-1}, x_n]$	$\iddots$	
$\vdots$	$\vdots$	$f[x_{n-1}, x_n]$	$\iddots$			
x_n	$y_n = f[x_n]$	$\iddots$				

- Em termos do operador diferenças divididas, os coeficientes $\alpha_0, \alpha_1, \ldots, \alpha_n$ são dados por

$$\alpha_k = f[x_0, x_1, \ldots, x_k], \qquad k = 0, \ldots, n.$$

- Portanto,

$$\begin{aligned} p_n(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \cdots \\ \cdots + f[x_0, x_1, \ldots, x_n](x - x_0)(x - x_1) \cdots (x - x_{n-1}). \end{aligned}$$

- Na prática, o valor do polinômio $p_n(x)$ para determinado x é obtido usando parênteses encaixados.
- O polinômio

$$p_n(x) = \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)(x - x_1) + \cdots \\ \cdots + \alpha_n(x - x_0)(x - x_1)\cdots(x - x_{n-1})$$

pode ser reescrito como

$$p_n(x) = \alpha_0 + (x - x_0)(\alpha_1 + (x - x_1)(\alpha_2 + \\ + (x - x_2)(\alpha_3 + \cdots + (x - x_{n-1})\alpha_n)\cdots).$$

- Deste modo, o polinômio no método de Newton resulta em

$$p_n(x) = f[x_0] + (x - x_0)(f[x_0, x_1] + (x - x_1)(f[x_0, x_1, x_2] + \\ + (x - x_2)(f[x_0, x_1, x_2, x_3] + \cdots + (x - x_{n-1})f[x_0, x_1, \ldots, x_n])\cdots).$$

Exemplo

Use o método de Newton para encontrar o polinômio de grau ≤ 2 que interpola os pontos dados na tabela

x	-1	2	3
y	-4	5	12

R.: $p_2(x) = -3 + 2x + x^2$.

Exemplo

Usando o Método de Newton, encontre o valor do polinômio em $x = 3$, passando pelos pontos dados na tabela abaixo:

x	-3	-2	-1	0
y	-60	-9	2	-3

R.: $p_3(3) = 126$.

Exemplo

Usando o Método de Newton, encontre o valor do polinômio em $x = 1$, passando pelos pontos dados na tabela abaixo:

x	-3	-2	-1	0	2
y	-9	-6	1	2	3

R.: $p_4(1) = -1{,}6$.

- Vamos assumir que os pontos $(x_0, y_0), \ldots, (x_n, y_n)$ satisfazem

$$y_k = f(x_k), \qquad k = 0, \ldots, n,$$

em que f é a função a ser aproximada por um polinômio p_n de grau $\leq n$.

- O erro $\mathcal{E}_n$ em $x \in [x_0, x_n]$, resultante da interpolação polinomial por p_n, é definido como

$$\mathcal{E}_n(x) = |f(x) - p_n(x)|.$$

Teorema

Considere $n+1$ pontos $x_0 < x_1 < \cdots < x_n$, com $n \geq 0$. Seja f uma função com derivada de ordem $n+1$ contínua no intervalo $[x_0, x_n]$. Se p_n é o polinômio que interpola f nos pontos $x_0, \ldots, x_n$, então

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{k=0}^{n} (x - x_k), \qquad x \in [x_0, x_n],$$

em que $x_0 \leq \xi \leq x_n$.

Consequentemente, o erro da interpolação polinomial satisfaz

$$\mathcal{E}_n(x) \leq \frac{M_{n+1}}{(n+1)!} \left| \prod_{k=0}^{n} (x - x_k) \right|, \qquad x \in [x_0, x_n],$$

em que

$$M_{n+1} = \max_{x \in [x_0, x_n]} \left| f^{(n+1)}(x) \right|.$$

- Vamos mostrar que

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\pi(x), \qquad x \in [x_0, x_n],$$

em que

$$\pi(x) = (x - x_0)(x - x_1)\cdots(x - x_n).$$

- Fixado $x \neq x_k$ em $[x_0, x_n]$, defina a função

$$\mathcal{A}(t) = f(t) - p_n(t) - \frac{f(x) - p_n(x)}{\pi(x)}\pi(t), \qquad t \in [x_0, x_n].$$

- Note que

$$\mathcal{A}(x) = 0 \qquad \text{e} \qquad \mathcal{A}(x_0) = 0, \mathcal{A}(x_1) = 0, \ldots, \mathcal{A}(x_n) = 0,$$

já que $f(x_k) = p_n(x_k)$, para $k = 0, \ldots, n$.

- Consequentemente, $\mathcal{A}$ possui $n + 2$ raízes em $[x_0, x_n]$.

- Pelo Teorema de Rolle, $\mathcal{A}'$ possui $n+1$ raízes em (x_0, x_n).
- Novamente, pelo Teorema de Rolle, $\mathcal{A}''$ possui n raízes em (x_0, x_n).
- Aplicando repetidas vezes o Teorema de Rolle, concluímos que $\mathcal{A}^{(n+1)}$ possui uma raiz $\xi \in (x_0, x_n)$.
- Logo, temos

$$\mathcal{A}^{(n+1)}(\xi) = f^{(n+1)}(\xi) - p_n^{(n+1)}(\xi) - \frac{f(x) - p_n(x)}{\pi(x)}(n+1)! = 0.$$

- Como $p_n^{(n+1)}(t) = 0$, chegamos então a

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\pi(x), \qquad x \in [x_0, x_n].$$

Corolário

Considere $n+1$ *pontos igualmente espaçados:*

$$x_k = x_0 + kh, \qquad k = 0, 1, \ldots, n.$$

Seja f *uma função com derivada de ordem* $n+1$ *contínua no intervalo* $[x_0, x_n]$*. Se* p_n *é o polinômio que interpola* f *nos pontos* $x_0, \ldots, x_n$*, então o erro da interpolação polinomial satisfaz*

$$\mathcal{E}_n(x) \leq \frac{M_{n+1} h^{n+1}}{4(n+1)}, \qquad x \in [x_0, x_n],$$

em que

$$M_{n+1} = \max_{x \in [x_0, x_n]} |f^{(n+1)}(x)|.$$

- Fixado $x \in [x_0, x_n]$, seja $0 \leq l \leq n$ tal que $x \in (x_{l-1}, x_l)$.
- Usando as desigualdades

$$
\begin{aligned}
(x - x_k) &\leq h(l - k), && \text{para } 0 \leq k \leq l-2,\\
(x - x_{l-1})(x_l - x) &\leq \frac{h^2}{4}, &&\\
(x_k - x) &\leq h(k - l + 1), && \text{para } l+1 \leq k \leq n,
\end{aligned}
$$

e

$$
(n - l + 1) \leq \binom{n}{l} \Rightarrow l! \cdot (n - l + 1)! \leq n!,
$$

podem escrever

$$
\begin{aligned}
|\pi(x)| &= \prod_{k=0}^{n} |x - x_k|\\
&= \prod_{k=0}^{l-2} (x - x_k)(x - x_{l-1})(x_l - x) \prod_{k=l+1}^{n} (x_k - x)
\end{aligned}
$$

$$
\begin{aligned}
&\leq \left[\prod_{k=0}^{l-2} h(l-k)\right] \frac{h^2}{4} \left[\prod_{k=l+1}^{n} h(k-l+1)\right] \\
&= \frac{h^{n+1}}{4} \left[\prod_{k=0}^{l-2} (l-k)\right] \left[\prod_{k=l+1}^{n} (k-l+1)\right] \\
&= \frac{h^{n+1}}{4} \cdot l! \cdot (n-l+1)! \\
&\leq \frac{h^{n+1}}{4} n!
\end{aligned}
$$

- Com isso, chegamos a

$$
\begin{aligned}
\mathcal{E}_n(x) = |f(x) - p_n(x)| &= \left|\pi(x) \frac{f^{(n+1)}(\xi)}{(n+1)!}\right| \\
&\leq \frac{h^{n+1}}{4} n! \frac{M_{n+1}}{(n+1)!} = \frac{M_{n+1} h^{n+1}}{4(n+1)}.
\end{aligned}
$$

Exemplo

Obtenha uma aproximação para $\ln(2{,}3)$ e uma estimativa para o seu erro usando interpolação polinomial. Para isso, considere a tabela

x	1	2	3	4
$\ln(x)$	0	0,6931	1,0986	1,3863

R.: $p_3(2{,}3) = 0{,}8373$, $\mathcal{E}_3(2{,}3) \leq 0{,}375$.

- Seja $f\colon [a,b] \to \mathbb{R}$ uma função contínua e, geralmente, suave.
- Numa quadratura numérica de $n+1$ pontos, a integral é aproximada da seguinte forma:

$$I(f) = \int_a^b f(x)dx \simeq \sum_{k=0}^{n} w_k f(x_k),$$

  em que
  - $x_0, x_1, \ldots, x_n \in [a,b]$ são os nós de integração, e
  - $w_0, w_1, \ldots, w_n$ são os pesos.
- A quadratura numérica é usada quando a função f é conhecida em apenas alguns pontos, ou quando não é possível calcular $I(f)$ analiticamente.

- Nas fórmulas de Newton–Cotes, os pesos w_k são obtidos usando interpolação polinomial ou interpolação polinomial por partes em pontos $x_0 < x_1 < \cdots < x_n$ igualmente espaçados no intervalo $[a, b]$.
- Uma fórmula de Newton–Cotes é chamada de fechada se $x_0 = a$ e $x_n = b$.
- Por outro lado, na fórmula aberta de Newton–Cotes, temos $x_0, x_1, \ldots, x_n \in (a, b)$.

- Seja p_n o polinômio de grau n que interpola f nos pontos

$$x_k = a + \frac{b-a}{n}k, \qquad k = 0, 1, \ldots, n.$$

- Pela forma de Lagrange, temos

$$p_n(x) = \sum_{k=1}^{n} f(x_k)L_k(x),$$

em que

$$L_k(x) = \prod_{i \neq k} \frac{(x - x_i)}{(x_k - x_i)}.$$

- Deste modo,

$$\begin{aligned} I(f) = \int_a^b f(x)dx &\simeq \int_a^b p_n(x)dx \\ &= \int_a^b \left(\sum_{k=0}^{n} f(x_k) L_k(x) \right) dx \\ &= \sum_{k=0}^{n} \left(\int_a^b L_k(x)dx \right) f(x_k) \\ &= \sum_{k=0}^{n} w_k f(x_k), \end{aligned}$$

em que os pesos $w_0, w_1, \ldots, w_n$ são dados por

$$w_k = \int_a^b L_k(x)dx, \qquad k = 0, 1, \ldots, n.$$

- Na Regra dos Trapézios, consideramos $n = 1$ com $x_0 = a$ e $x_1 = b$.
- Sendo as bases de Lagrange dadas por

$$L_0(x) = \frac{x - x_1}{x_0 - x_1} = \frac{x_1 - x}{h}, \quad \text{e} \quad L_1(x) = \frac{x - x_0}{x_1 - x_0} = \frac{x - x_0}{h},$$

em que $h = x_1 - x_0$, temos

$$w_0 = \int_a^b L_0(x)dx = \int_{x_0}^{x_1} \frac{x_1 - x}{h} dx = \frac{h}{2},$$
$$w_1 = \int_a^b L_1(x)dx = \int_{x_0}^{x_1} \frac{x - x_0}{h} dx = \frac{h}{2}.$$

- Com isso, a Regra dos Trapézios resulta em

$$I(f) \simeq T_1(f) = \frac{h}{2}[f(x_0) + f(x_1)].$$

- Para melhorar a aproximação da integral pela Regra dos Trapézios, podemos fazer uma subdivisão do intervalo $[a, b]$ com pontos

$$x_k = a + \frac{b-a}{n} k, \qquad k = 0, 1, \ldots, n,$$

e aplicar a Regra dos Trapézios em cada subintervalo $[x_{k-1}, x_k]$, para $k = 1, \ldots, n$.

- Nesse caso, a função f é aproximada por um polinômio linear por partes Π_1 que interpola f em $x_0, x_1, \ldots, x_n$.

- Com isso, temos a aproximação

$$I(f) = \int_a^b f(x)dx \simeq \int_a^b \Pi_1(x)dx = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} \Pi_1(x)dx = T_n(f).$$

- Considerando a aplicação da Regra dos Trapézios em cada subintervalo $[x_{k-1}, x_k]$, obtemos

$$I(f) \simeq T_n(f) = \sum_{k=1}^{n} \frac{h}{2}[f(x_{k-1}) + f(x_k)].$$

- A Regra dos Trapézios Composta é dada então como

$$I(f) \simeq T_n(f) = \frac{h}{2}[f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)].$$

- Seja
$$R_T = I(f) - T_1(f) = \int_{x_0}^{x_1} f(x)dx - \int_{x_0}^{x_1} p_1(x)dx.$$
- Considere a função contínua
$$\mathcal{A}(x) = f(x) - p_1(x) - \frac{R_T}{W}\pi(x),$$
em que $\pi(x) = (x - x_0)(x - x_1)$,e
$$W = \int_{x_0}^{x_1} \pi(x)dx = -\frac{h^3}{6} \neq 0.$$
- Como
$$\int_{x_0}^{x_1} \mathcal{A}(x)dx = 0,$$
podemos inferir que existe $u \in (x_0, x_1)$ tal que $\mathcal{A}(u) = 0$.

- Portanto, a função $\mathcal{A}(x)$ tem três zeros distintos em $[x_0, x_1]$.
- Aplicando repetidas vezes o Teorema de Rolle, concluímos que $\mathcal{A}''(x)$ possui uma raiz $\xi \in (x_0, x_1)$.
- Logo, temos

$$0 = \mathcal{A}''(\xi) = f''(\xi) - \frac{R_T}{W} 2!,$$

de onde chegamos então a

$$R_T = \frac{f''(\xi)}{2!} W = \frac{f''(\xi)}{2!}\left(-\frac{h^3}{6}\right) = -\frac{h^3}{12} f''(\xi).$$

- No caso da Regra dos Trapézios Composta, considerando a expressão para o resto, obtemos

$$\begin{aligned} I(f) &= \sum_{k=1}^{n} \left(\frac{h}{2}[f(x_{k-1}) + f(x_k)] - \frac{h^3}{12} f''(\xi_k) \right) \\ &= T_n(f) \underbrace{- \frac{h^3}{12} \sum_{k=1}^{n} f''(\xi_k)}_{=R_T} . \end{aligned}$$

- Supondo que f'' é contínua em $[a, b]$, uma aplicação Teorema do Valor Intermediário fornece a existência de $\xi \in (a, b)$ tal que

$$R_T = -\frac{h^3}{12} \sum_{k=1}^{n} f''(\xi_k) = -\frac{h^3}{12} n f''(\xi) = -n \frac{h^3}{12} f''(\xi).$$

- Portanto, o erro na Regra dos Trapézios Composta, supondo f'' contínua em $[a, b]$, satisfaz

$$|R_T| = |I(f) - T_n(f)| \leq \frac{(b-a)^3}{12n^2} M_2,$$

em que

$$M_2 = \max_{a \leq \xi \leq b} |f''(\xi)|.$$

Exemplo

Usando a Regra dos Trapézios Composta, calcule o valor da integral

$$\int_0^1 e^x\,dx,$$

com erro $|R_T| < 10^{-2}$.

R.: $n = 5$, $T_5(f) = 1{,}724005$.

Exemplo

Usando a Regra dos Trapézios Composta, calcule o valor da integral

$$\int_0^1 \ln(\cos x) dx,$$

com erro $|R_T| < 10^{-2}$.

R.: $n = 6$, $T_6(f) = -0{,}191132$.

Exemplo

Usando a Regra dos Trapézios Composta, calcule o valor da integral

$$\int_0^1 (x-2)e^x \, dx,$$

com erro $|R_T| < 7 \times 10^{-3}$.

R.: $n = 6$, $T_6(f) = -2{,}434253$.

- A Regra 1/3 de Simpson é obtida considerando $n = 2$, com

$$x_0 = a, \quad x_1 = (a+b)/2 \quad \text{e} \quad x_2 = b.$$

- Neste caso, as integrais dos polinômios de Lagrange resultam em

$$w_0 = \int_a^b L_0(x)dx = \int_{x_0}^{x_2} \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}dx = \frac{h}{3},$$
$$w_1 = \int_a^b L_1(x)dx = \int_{x_0}^{x_2} \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}dx = \frac{4h}{3},$$
$$w_2 = \int_a^b L_2(x)dx = \int_{x_0}^{x_2} \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}dx = \frac{h}{3},$$

que podem ser calculadas considerando $z = x - x_0$.

- Portanto, a Regra 1/3 de Simpson é dada por

$$I(f) \simeq S_2(f) = \frac{h}{3}[f(x_0) + 4f(x_1) + f(x_2)].$$

- Para melhorar a aproximação da integral pela Regra 1/3 de Simpson, podemos fazer uma subdivisão do intervalo $[a, b]$ com pontos

$$x_k = a + \frac{b-a}{n}k, \qquad k = 0, 1, \dots, n,$$

e aplicar a Regra 1/3 de Simpson em cada subintervalo $[x_{2k-2}, x_{2k}]$, para $k = 1, \dots, n/2$.
- Nesse caso, a função f é aproximada por um polinômio quadrático por partes Π_2 que interpola f em $x_0, x_1, \dots, x_n$.

- Com isso, temos a aproximação

$$I(f) = \int_a^b f(x)dx \simeq \int_a^b \Pi_2(x)dx = \sum_{k=1}^{n/2} \int_{x_{2k-2}}^{x_{2k}} \Pi_2(x)dx = S_n(f).$$

- Considerando a aplicação da Regra 1/3 de Simpson em cada subintervalo $[x_{2k-2}, x_{2k}]$, obtemos

$$I(f) \simeq S_n(f) = \sum_{k=1}^{n/2} \frac{h}{3}[f(x_{2k-2}) + 4f(x_{2k-1}) + f(x_{2k})].$$

- Deste modo, a Regra 1/3 de Simpson Composta é dada como

$$S_n(f) = \frac{h}{3}[f(x_0)+4f(x_1)+2f(x_2)+4f(x_3)+\cdots+2f(x_{n-2})+4f(x_{n-1})+f(x_n)]$$

- Assim como na Regra dos Trapézios, pode-se mostrar que o resto na Regra 1/3 de Simpson, supondo $f^{(4)}$ contínua em $[a, b]$, satisfaz

$$R_S = I(f) - S_2(f) = -\frac{h^5}{90} f^{(4)}(\xi), \qquad \xi \in (a, b).$$

- Para isso, denotamos

$$\pi(x) = (x - x_0)(x - x_1)(x - x_2),$$
$$\widetilde{\pi}(x) = (x - x_0)(x - x_1)^2(x - x_2),$$

e assim temos

$$W = \int_{x_0}^{x_2} \pi(x) dx = 0,$$
$$\widetilde{W} = \int_{x_0}^{x_2} \widetilde{\pi}(x) dx = -\frac{4}{15} h^5 \neq 0.$$

- Considere a função contínua

$$\mathcal{A}(x) = f(x) - p_2(x) - k\pi(x) - \frac{R_S}{\widetilde{W}}\widetilde{\pi}(x),$$

em que a constante $k \in \mathbb{R}$ é escolhida de modo que

$$\int_{x_0}^{x_1} \mathcal{A}(x)dx = 0,$$

visto que $\int_{x_0}^{x_1} \pi(x)dx \neq 0$.

- Como $\int_{x_0}^{x_2} \mathcal{A}(x)dx = 0$, concluímos que

$$\int_{x_1}^{x_2} \mathcal{A}(x)dx = 0.$$

- Com isso, existem $u \in (x_0, x_1)$ e $v \in (x_1, x_2)$ tais que $\mathcal{A}(u) = \mathcal{A}(v) = 0$.

- Portanto,a função $\mathcal{A}(x)$ tem cinco zeros distintos em $[x_0, x_2]$.
- Aplicando repetidas vezes o Teorema de Rolle, concluímos que $\mathcal{A}^{(4)}(x)$ possui uma raiz $\xi \in (x_0, x_2)$.
- Logo, temos

$$0 = \mathcal{A}^{(4)}(\xi) = f^{(4)}(\xi) - \frac{R_S}{\widetilde{W}} 4!,$$

de onde chegamos então a

$$R_S = \frac{f^{(4)}(\xi)}{4!} \widetilde{W} = \frac{f^{(4)}(\xi)}{4!} \left(-\frac{4}{15} h^5 \right) = -\frac{h^5}{90} f^{(4)}(\xi).$$

- No caso da Regra 1/3 de Simpson Composta, considerando a expressão para o resto, obtemos

$$I(f) = \sum_{k=1}^{n/2}\left(\frac{h}{3}[f(x_{2k-2}) + 4f(x_{2k-1}) + f(x_{2k})] - \frac{h^5}{90}f^{(4)}(\xi_k)\right)$$
$$= S_n(f)\underbrace{-\frac{h^5}{90}\sum_{k=1}^{n/2} f^{(4)}(\xi_k)}_{=R_S}.$$

- Supondo que $f^{(4)}$ é contínua em $[a, b]$, o Teorema do Valor Intermediário fornece a existência de $\xi \in (a, b)$ tal que

$$R_S = -\frac{h^5}{90}\sum_{k=1}^{n/2} f^{(4)}(\xi_k) = -\frac{h^5}{90}\left(\frac{n}{2}f^{(4)}(\xi)\right) = -n\frac{h^5}{180}f^{(4)}(\xi).$$

- Portanto, o resto na Regra 1/3 de Simpson Composta, supondo $f^{(4)}$ contínua em $[a, b]$, satisfaz

$$|R_S| = |I(f) - S_n(f)| \leq \frac{(b-a)^5}{180n^4} M_4,$$

em que

$$M_4 = \max_{a \leq \xi \leq b} |f^{(4)}(\xi)|.$$

Exemplo

Calcule o valor da integral

$$\int_0^1 e^x \, dx,$$

aplicando a Regra 1/3 de Simpson Composta, com erro $|R_S| < 10^{-4}$.

Exemplo

Calcule o valor da integral

$$\int_1^3 \ln(x+1)dx,$$

aplicando a Regra 1/3 de Simpson Composta, com erro $|R_S| < 10^{-4}$.

R.: $n = 6$, $S_6(f) = 2{,}158868$.

Exemplo

Calcule o valor da integral

$$\int_0^1 (x \operatorname{sen} x + 4\cos x)dx,$$

aplicando a Regra 1/3 de Simpson Composta, com erro $|R_S| < 10^{-5}$.

R.: $n = 6$, $S_6(f) = 3{,}6670538$.

- Na quadratura Gaussiana dada por

$$I(f) = \int_a^b f(x)dx \simeq G_n(f) = \sum_{k=0}^{n} w_k f(x_k),$$

os nós de integração $x_0, x_1, \ldots, x_n$ não são pontos igualmente espaçados em $[a, b]$.
- Tanto os nós de integração como os pesos são escolhidos de modo que $R_{n+1} = 0$ se f for um polinômio de grau $\leq 2n + 1$.

- Para $n = 1$ e $[a, b] = [-1, 1]$, devemos ter

$$G_1(f) = \int_{-1}^{1} f(t)dt = w_0 f(t_0) + w_1 f(t_1),$$

sempre que f for um polinômio de grau ≤ 3.

- Deste modo,

$$\int_{-1}^{1} 1dx = w_0 f(t_0) + w_1 f(t_1) \Rightarrow w_0 + w_1 = 2,$$

$$\int_{-1}^{1} xdx = w_0 f(t_0) + w_1 f(t_1) \Rightarrow w_0 t_0 + w_1 t_1 = 0,$$

$$\int_{-1}^{1} x^2 dx = w_0 f(t_0) + w_1 f(t_1) \Rightarrow w_0 t_0^2 + w_1 t_1^2 = \frac{2}{3},$$

$$\int_{-1}^{1} x^3 dx = w_0 f(t_0) + w_1 f(t_1) \Rightarrow w_0 t_0^3 + w_1 t_1^3 = 0.$$

- Como consequência, obtemos o seguinte sistema não linear:

$$\begin{cases} w_0 + w_1 = 2 \\ w_0 t_0 + w_1 t_1 = 0 \\ w_0 t_0^2 + w_1 t_1^2 = \dfrac{2}{3} \\ w_0 t_0^3 + w_1 t_1^3 = 0 \end{cases}$$

cuja solução é

$$t_0 = -\frac{1}{\sqrt{3}}, \qquad t_1 = +\frac{1}{\sqrt{3}}, \quad \text{e} \quad w_0 = w_1 = 1.$$

- No caso de um intervalo $[a, b]$ qualquer, a mudança de variável

$$x = \frac{1}{2}(a + b + t(b - a))$$

resulta na quadradura Gaussiana dada por

$$G_1(f) = \frac{b-a}{2}\left[f\left(\frac{a+b}{2} - \frac{b-a}{2\sqrt{3}}\right) + f\left(\frac{a+b}{2} + \frac{b-a}{2\sqrt{3}}\right)\right].$$

Exemplo

Estime $I(f) = \int_0^1 e^x dx$ usando a quadratura Gaussiana $G_1(f)$ e determine o erro resultante.

Exemplo

Estime $I(f) = \int_0^1 (x-2)e^x dx$ usando a quadratura Gaussiana $G_1(f)$ e determine o erro resultante.

Exemplo

Estime $I(f) = \int_1^3 \ln(x+1) dx$ usando a quadratura Gaussiana $G_1(f)$ e determine o erro resultante.

- Um problema de valor inicial (PVI) é uma equação diferencial dada na forma
$$\begin{cases} \dfrac{dy}{dx} = f(x,y), \\ y(x_0) = y_0, \end{cases}$$
em que f é uma função dada, com condição inicial $y(x_0) = y_0$, para x_0 e y_0 fixados.
- Assumiremos que o PVI possui uma única solução $y(x)$.

- Assuma que pontos $x_0, x_1, \ldots$ são igualmente espaçados, ou seja,

$$x_{k+1} = x_k + h, \qquad k = 0, 1, \ldots$$

para dado passo h.

- O método de Euler é dado pela fórmula

$$y_{k+1} = y_k + hf(x_k, y_k), \qquad k = 0, 1, \ldots$$

- Se $y''(x)$ for contínua, então, pela série de Taylor centrada em x_k temos

$$y(x_k + h) = y(x_k) + y'(x_k)h + \frac{1}{2}y''(\xi)h^2, \qquad \xi \in (x_k, x_k + h),$$

que reescrevemos como

$$y(x_{k+1}) = y(x_k) + hf(x_k, y(x_k)) + \frac{1}{2}y''(\xi)h^2,$$

onde usamos $x_{k+1} = x_k + h$ e $y'(x_k) = f(x_k, y(x_k))$.

- Assumindo $y_k = y(x_k)$, o erro do método de Euler em x_{k+1} é então

$$e(x_{k+1}) = |y(x_{k+1}) - y_{k+1}| = \frac{1}{2}|y''(\xi)|h^2 \leq \frac{h^2}{2}M_2,$$

em que

$$M_2 = \max_{x_k \leq \xi \leq x_{k+1}} |y''(\xi)|.$$

Definição

Um método numérico para um PVI é dito ser de ordem p se existe uma constante C tal que o erro $e(x_{k+1}) = |y(x_{k+1}) - y_{k+1}|$, onde se assume $y_k = y(x_k)$, satisfaz

$$e(x_{k+1}) < Ch^{p+1},$$

em que C é uma constante.

Exemplo

Usando o Método de Euler, estabeleça aproximações para o PVI

$$\begin{cases} \dfrac{dy}{dx} = \dfrac{1}{x+1}, \\ y(0) = 1, \end{cases}$$

no intervalo $[0; 1{,}2]$, com $h = 0{,}4$.

Exemplo

Usando o Método de Euler, estabeleça aproximações para o PVI

$$\begin{cases} \dfrac{dy}{dx} = \dfrac{1}{1+xy}, \\ y(0) = 0, \end{cases}$$

no intervalo $[0, 1]$, com $h = 0{,}25$.

Exemplo

Usando o Método de Euler, estabeleça aproximações para o PVI

$$\begin{cases} \dfrac{dy}{dx} = x^2 - 2y, \\ y(0) = 1, \end{cases}$$

no intervalo $[0, 1]$, com $h = 0{,}2$.

- Os chamados métodos de Runge–Kutta de s estágios apresentam a forma
$$y_{k+1} = y_k + h\phi(x_k, y_k), \qquad k = 0, 1, \dots,$$
em que $\phi = b_1 k_1 + \cdots + b_s k_s$, e
$$\begin{aligned} k_1 &= f(x, y), \\ k_2 &= f(x + c_2 h, y + h a_{21} k_1), \\ k_3 &= f(x + c_3 h, y + h(a_{31} k_1 + a_{32} k_2)), \\ &\ \vdots \\ k_s &= f(x + c_s h, y + h(a_{s1} k_1 + \cdots + a_{s,s-1} k_{s-1})), \end{aligned}$$
com as constantes c_i, a_{ij}, b_j sendo definidas para cada método particular.
- O método de Euler, em que $\phi = f$, é um método de Runge–Kutta de estágio $s = 1$ e ordem $p = 1$.

- As constantes do método de Runge–Kutta podem ser organizados como na tabela

$$
\begin{array}{c|ccccc}
0 & & & & & \\
c_2 & a_{21} & & & & \\
c_3 & a_{31} & a_{32} & & & \\
\vdots & \vdots & \vdots & \ddots & & \\
c_s & a_{s1} & a_{s2} & \cdots & a_{s,s-1} & \\
\hline
 & b_1 & b_2 & \cdots & b_{s-1} & b_s
\end{array}
$$

- Pode-se mostrar que o método de Runge–Kutta de 2 estágios tem ordem $p = 2$ se

$$b_1 + b_2 = 1, \qquad b_2 c_2 = \frac{1}{2}, \qquad b_2 a_{21} = \frac{1}{2}.$$

- Dois métodos que satisfazem a relação acima são o método de Euler modificado, cujas constantes são dadas na tabela

$$\begin{array}{c|cc} 0 & & \\ 1/2 & 1/2 & \\ \hline & 0 & 1 \end{array}$$

e o método de Euler melhorado, com constantes dadas abaixo

$$\begin{array}{c|cc} 0 & & \\ 1 & 1 & \\ \hline & 1/2 & 1/2 \end{array}$$

- O método de Euler modificado apresenta a forma

$$y_{k+1} = y_k + hf\left(x_k + \frac{h}{2}, y_k + \frac{h}{2} f(x_k, y_k)\right).$$

- Já o método de Euler melhorado tem a forma

$$y_{k+1} = y_k + \frac{h}{2}[f(x_k, y_k) + f(x_k + h, y_k + hf(x_k, y_k))].$$

Exemplo

Compare a solução do PVI

$$\begin{cases} \dfrac{dy}{dx} = -2xy^2, \\ y(0) = 0{,}5, \end{cases}$$

no intervalo $[0, 1]$, com 10 subintervalos, usando os métodos de Euler, Euler modificado e Euler melhorado, sabendo que a solução é

$$y(x) = \frac{1}{x^2 + 2}.$$

x_i	$\lvert y_i^{\mathrm{E}} - y(x_i)\rvert$	$\lvert y_i^{\mathrm{Emod}} - y(x_i)\rvert$	$\lvert y_i^{\mathrm{Emel}} - y(x_i)\rvert$
0,0	0	0	0
0,1	$2{,}49 \times 10^{-3}$	$1{,}24 \times 10^{-5}$	$1{,}24 \times 10^{-5}$
0,2	$4{,}80 \times 10^{-3}$	$4{,}76 \times 10^{-5}$	$2{,}32 \times 10^{-5}$
0,3	$6{,}73 \times 10^{-3}$	$9{,}83 \times 10^{-5}$	$2{,}97 \times 10^{-5}$
0,4	$8{,}11 \times 10^{-3}$	$1{,}55 \times 10^{-4}$	$2{,}89 \times 10^{-5}$
0,5	$8{,}88 \times 10^{-3}$	$2{,}06 \times 10^{-4}$	$1{,}91 \times 10^{-5}$
0,6	$9{,}04 \times 10^{-3}$	$2{,}45 \times 10^{-4}$	$2{,}60 \times 10^{-8}$
0,7	$8{,}69 \times 10^{-3}$	$2{,}67 \times 10^{-4}$	$2{,}70 \times 10^{-5}$
0,8	$7{,}94 \times 10^{-3}$	$2{,}71 \times 10^{-4}$	$5{,}96 \times 10^{-5}$
0,9	$6{,}93 \times 10^{-3}$	$2{,}59 \times 10^{-4}$	$9{,}48 \times 10^{-5}$
1,0	$5{,}77 \times 10^{-3}$	$2{,}35 \times 10^{-4}$	$1{,}30 \times 10^{-4}$

- Um método bastante comum de ordem 4 é o método de Runge–Kutta com coeficientes dados na tabela

$$\begin{array}{c|cccc} 0 & & & & \\ 1/2 & 1/2 & & & \\ 1/2 & 0 & 1/2 & & \\ 1 & 0 & 0 & 1 & \\ \hline & 1/6 & 1/3 & 1/3 & 1/6 \end{array}$$

- Para esses coeficientes o método de Runge–Kutta apresenta a forma

$$y_{k+1} = y_k + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

em que

$$k_1 = f(x_k, y_k), \qquad k_3 = f\left(x_k + \frac{h}{2}, y_k + \frac{h}{2}k_2\right),$$

$$k_2 = f\left(x_k + \frac{h}{2}, y_k + \frac{h}{2}k_1\right), \qquad k_4 = f(x_k + h, y_k + hk_3).$$

Exemplo

Considere o PVI

$$\begin{cases} \dfrac{dy}{dx} = y, \\ y(0) = 1. \end{cases}$$

Use o método de Euler melhorado e o método de Runge–Kutta de ordem 4 para estimar $y(0{,}04)$ com $h = 0{,}04$. Compare com a solução $y(0{,}04) = e^{0{,}04}$.