

## 1. Aritmética de Máquina

As leis e os mecanismos internos que regem as operações aritméticas, quando executadas em um computador, não obedecem à mesma estrutura dos números reais. Isso acontece porque uma máquina digital tem recursos finitos e na aritmética real, muitas vezes, é preciso trabalhar com infinitos dígitos para representar um único número.

O conjunto dos números possíveis de serem representados por uma máquina é discreto e finito, o que pode causar dois tipos de erros operacionais: *overflow* ou *underflow*. Seja  $T$  o conjunto formado por estes números e considere que as seguintes afirmações são válidas:

1.  $\forall x \in T, -x \in T$ ;
2.  $t_1 = \inf x \mid x \in T \text{ e } x > 0$ ;  
 $t_2 = \sup x \mid x \in T \text{ e } x > 0$ .

Se  $y$  é um valor tal que  $|y| < t_1$ , ocorre um erro de *underflow* e se  $|y| > t_2$ , de *overflow*.

**Exemplo 1:** Considere os números  $x_1 = 0,3491 \times 10^4$  e  $x_2 = 0,2345 \times 10^0$ . Admita que as operações sejam realizadas em uma máquina com precisão de quatro dígitos significativos. Veja que:

$$(x_2 + x_1) - x_1 = (0,2345 \times 10^0 + 0,2394 \times 10^4) - 0,2394 \times 10^4 = 0,000$$

mas, por outro lado,

$$\begin{aligned}x_2 + (x_1 - x_1) &= 0,2345 \times 10^0 + (0,2394 \times 10^4 - 0,2394 \times 10^4) \\ &= 0,2345 \times 10^0.\end{aligned}$$

O exemplo mostra que as operações executadas em uma máquina digital não só estão sujeitas a erros, mas podem inclusive gerar erros.

## 2. Fontes de Erro (Si no se comete errores, generalmente no se hace nada)

Dado um problema, para se chegar a um resultado numérico é necessário percorrer uma seqüência pré-estabelecida de passos isolados. Em cada um destes passos pode existir uma parcela de erro que se acumula ao montante final do processo. Estes erros surgem basicamente em duas áreas: aqueles inerentes à formulação matemática do problema (relacionados à aproximação da situação física e a erros nos dados) e aqueles que aparecem no processo de solução numérica (erros de truncamento e de arredondamento).

Os *erros de truncamento* surgem, em geral, pela substituição de um processo infinito (somas ou integrais) ou infinitesimal (diferenciação) por outro finito. Erros também podem surgir do fato que

as operações aritméticas quase nunca podem ser efetuadas com precisão completa. Estes são denominados de *erros de arredondamento*. A maioria dos números têm representações decimais infinitas que devem ser arredondadas. Mesmo se os dados de um problema podem ser expressos exatamente por representações decimais finitas, a divisão pode introduzir números que devem ser arredondados e a multiplicação pode produzir mais dígitos do que podem ser mantidos.

Os tipos de arredondamento mais utilizados são:

- tipo corte: as casas em excesso são simplesmente abandonadas;
- para o número de máquina mais próximo: se a máquina trabalha com  $d$  algarismos significativos para a mantissa de um número, então analisa-se o algarismo de ordem  $d + 1$ . Se este for maior ou igual a 5, soma-se uma unidade ao algarismo de ordem  $d$ ; caso contrário, o algarismo de ordem  $d$  permanece inalterado.

**Exemplo 2:** A função  $\text{sen}(x)$  é dada por

$$\text{sen}(x) = \sum_{i=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}.$$

Para obter o valor de  $\text{sen}(x)$  por esta série é preciso efetuar o cálculo de várias parcelas e depois parar, ou seja, truncar a série, cometendo então um erro causado pelo abandono das parcelas que não foram somadas. Substituindo  $\text{sen}(x)$  pelo polinômio

$$P(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}$$

é possível calcular o valor numérico de  $P(x)$  e usá-lo como aproximação de  $\text{sen}(x)$ . O erro de truncamento neste caso é definido por  $E = \text{sen}(x) - P(x)$ .

**Exemplo 3:** Se  $x = 0,6666\dots$ , os dois arredondamentos não produzem o mesmo resultado. Se a máquina em que se está trabalhando opera com 4 algarismos significativos, então  $\bar{x} = 0,6666 \times 10^0$ , por corte, ou  $\bar{x} = 0,6667 \times 10^0$ , por arredondamento para o número mais próximo.

A diferença entre o valor arredondado e o valor exato pode ser medida pelo erro absoluto ou pelo relativo. O erro absoluto, indicado por  $E_A$ , é dado por:

$$E_A = |\bar{x} - x|$$

e o erro relativo, indicado por  $E_R$ , é

$$E_R = \frac{|\bar{x} - x|}{|\bar{x}|} \quad \text{ou} \quad E_R = \frac{|\bar{x} - x|}{|x|}$$

Existem alguns procedimentos inexatos que podem levar a situações de erro, como a soma de grandezas bastante desproporcionais e a subtração de grandezas muito próximas em condições de precisão limitada (precisão definida  $n$ ).

**Exemplo 4:** Sejam os números

$$\begin{aligned}x &= 3,91543782 & \text{e} & \quad \bar{x} = 3,915438 \\y &= 3,91542534 & \text{e} & \quad \bar{y} = 3,915425.\end{aligned}$$

A diferença entre eles é

$$d_{x,y} = 0,00001248 \quad \text{e} \quad d_{\bar{x},\bar{y}} = 0,000013.$$

O erro absoluto é

$$|E_A| = |d_{x,y} - d_{\bar{x},\bar{y}}| = 0,5200 \times 10^{-6}$$

e o erro relativo,

$$|E_R| = \frac{|d_{x,y} - d_{\bar{x},\bar{y}}|}{d_{x,y}} = 4,16666 \dots \times 10^{-2} \approx 0,0417 \quad \text{ou} \quad 4,17\%.$$

**Exemplo 5:** Quantas casas decimais devemos trabalhar para termos uma resposta com determinado erro percentual (1 por cento)?

Solução: Se o número (resposta) for da ordem de 1 podemos trabalhar com três casas decimais. Se o citado número for da ordem de 10 com 2 e se for da ordem de 0,1 com 4 casas decimais.

**Exemplo 6:** A solução exata da soma

$$S = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} \quad \text{é} \quad S = 1,$$

mas a solução calculada, supondo uma máquina que opere com cinco dígitos de precisão, é

$$S = 0,33333 + 0,33333 + 0,33333 = 0,99999.$$

Neste caso, verifica-se que o erro de arredondamento, se calculado por qualquer uma das duas maneiras indicadas acima, é igual a  $1 - 0,99999 = 0,00001$ .

Na prática, quando se obtém um resultado de uma expressão aritmética avaliada em uma máquina e não se conhece o seu valor exato, torna-se complicado calcular o erro relativo ou o absoluto. Por isto, trata-se com os dígitos significativos exatos (DIGSE) de um determinado número.

Dado um número  $\bar{x}$  aproximado de um valor exato  $x$ , diz-se que esta aproximação tem pelo menos  $n$  dígitos significativos exatos se

$$|E_R| \leq \frac{1}{2} 10^{-n}.$$

Uma outra maneira de calcular o DIGSE de uma aproximação é

$$DIGSE(\bar{x}, x) = - \left[ 0,3 + \log \left( \mu + \frac{|\bar{x} - x|}{|x|} \right) \right],$$

onde  $\mu$  é a unidade de arredondamento da máquina. Se o arredondamento for para o número de máquina mais próximo,  $\mu = \frac{1}{2}10^{1-m}$  ( $m$ , número de algarismos da mantissa da máquina).

Na prática, geralmente não se conhece o valor exato de  $x$  para que as comparações sejam feitas. Convencionando que  $x = \lim_{k \rightarrow \infty} x_k$ , pode-se modificar a fórmula acima para

$$DIGSE(x_k, x_{k+1}) = - \left[ 0, 3 + \log \left( \mu + \frac{|x_{k+1} - x_k|}{x_k} \right) \right],$$

que fornece o número de dígitos significativos de  $x_{k+1}$  em relação a  $x_k$ .

Dois conceitos relacionados à qualidade dos resultados obtidos computacionalmente são *precisão* e *exatidão*. A *precisão* de uma máquina digital é definida como o número de dígitos da mantissa desta máquina e *exatidão* é uma medida da perfeição do resultado. Sendo assim, a precisão é algo claro, não variável de máquina para máquina, mas a exatidão, pelo contrário, depende da precisão da máquina e do método utilizado para obtenção deste resultado.

**Exemplo 7:** Considere o número irracional  $\pi = 3,14159265\dots$

- 3,1415926 é mais preciso e mais exato do que 3,14159;
- 3,1415929 é mais preciso e menos exato do que 3,14159;

Além da existência de erros, existem outros problemas que devem ser levados em conta ao se resolver um problema numericamente. Eles podem ser analisados do ponto de vista da *instabilidade*.

### 3. Instabilidade

Para muitos problemas não importa qual o procedimento de solução utilizado, o resultado é sempre o mesmo. Entretanto, em muitos casos, diferentes modos de solução podem conduzir a diferentes resultados. Isto caracteriza a *instabilidade*, que pode ser entendida como uma sensibilidade a perturbações e pode ocorrer tanto no problema em si como no algoritmo, isto é, na maneira de resolver o problema.

Observe que o exemplo 8 e o exercício 1 referem-se à instabilidade do algoritmo, ao passo que o exemplo 9 e o exercício 2 ilustram a instabilidade do problema.

**Exemplo 8:** Considere o problema da determinação das raízes da equação

$$ax^2 + bx + c = 0.$$

Da álgebra, sabe-se que as raízes são fornecidas pela fórmula

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{e} \quad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

Se  $4ac$  for muito pequeno comparado com  $b^2$ , existe a possibilidade de acontecer a subtração de grandezas muito próximas. Para um procedimento mais correto, usa-se o algoritmo:

1. Se  $b > 0$  então  $x_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$  e  $x_2 = \frac{c}{ax_1}$ ;
2. Se  $b < 0$  então  $x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$  e  $x_2 = \frac{c}{ax_1}$ .

Este processo impede que se faça uma segunda subtração na fórmula além daquela que ocorre inevitavelmente no radical.

**Exercício 1:** Calcular as raízes da equação  $x^2 + 80x + 1 = 0$  com precisão de três dígitos significativos e arredondamento por corte. Utilizar o procedimento usual e o proposto acima. Comparar e discutir os resultados.

Com o exemplo 8, percebe-se que os algoritmos podem gerar, sob certos aspectos, erros consideráveis de modo a tornar inútil o resultado. Outro tipo de instabilidade é a que pode ser observada, por exemplo, em polinômios de grau elevado e com coeficientes altos, na derivação numérica e em sistemas mal condicionados.

**Exemplo 9:** Um exemplo clássico em Matemática Numérica é dado pelo polinômio

$$P(x) = x^{20} - 210x^{19} + \dots = (x - 1)(x - 2)\dots(x - 20) = 0,$$

de onde se conclui claramente que as raízes de  $P(x)$  são  $1, 2, \dots, 20$ . Entretanto, introduzindo uma pequena perturbação em  $P(x)$  e calculando as raízes de  $P(x) + 2^{-23} \cdot 210x^{19} = 0$  em um computador de sistema de ponto flutuante com base 2 e mantissa com 90 algarismos binários, tem-se

$$\begin{aligned} R_1 &= 1 \\ R_2 &= 2 \\ R_3 &= 3 \\ R_4 &= 4 \\ R_5 &= 4,999999928 \\ R_6 &= 6,000006944 \\ R_7 &= 6,999697234 \\ R_8 &= 8,007267603 \\ R_9 &= 8,917250249 \\ R_{10}, R_{11} &= 10,095266145 \pm 0,643500904i \\ R_{12}, R_{13} &= 11,793633881 \pm 1,652329728i \\ R_{14}, R_{15} &= 13,992358137 \pm 2,518830070i \\ R_{16}, R_{17} &= 16,730737466 \pm 2,812624894i \\ R_{18}, R_{19} &= 19,502439400 \pm 1,940330347i \\ R_{20} &= 20,846908101 \end{aligned}$$

Mediante a ligeira modificação no coeficiente de apenas um termo da equação, o resultado apresenta mudanças significativas nas raízes; algo inesperado. A razão desta mudança drástica não é nem o arredondamento nem o algoritmo, mas um problema de condicionamento (pequenas variações nos dados de entrada provocam grandes variações nos resultados).

**Exercício 2:** Calcule a solução do sistema de equações lineares:

$$\begin{cases} x + 5y = 17 \\ 1,5x + 7,501y = 25,503 \end{cases}$$

e depois, do mesmo sistema, afetado de um erro de 0,002 no elemento 25,503 do vetor independente, ou seja,

$$\begin{cases} x + 5y = 17 \\ 1,5x + 7,501y = 25,501 \end{cases}$$

Compare os resultados obtidos.

## 4. Breves Comentários sobre Processos Iterativos e Algoritmos

Processo iterativo corresponde à uma seqüência de operações  $x_1, x_2, \dots, x_n$  para encontrar a solução desejada. O cálculo de uma nova aproximação é feito em função das aproximações anteriores. Uma seqüência de operações apropriadamente ordenada conduz a algoritmos eficientes.

Diz-se que todo algoritmo pode ser otimizado; talvez isto seja verdade, mas os aspectos gerais importantes num algoritmo são:

- simplicidade: facilidade de uso e entendimento
- baixo custo
- resultados confiáveis

É importante lembrar que algoritmos muito sofisticados não possuem geralmente todas estas características. Estes, portanto, não são os mais eficazes. Em termos práticos, um algoritmo numérico de boa qualidade deve ter as seguintes características:

- inexistência de erro lógico
- inexistência de erro operacional
- quantidade finita de cálculos
- existência de um critério de exatidão
- independência da máquina
- precisão infinita (limite do erro tende a zero)
- eficiência (resposta correta no menor tempo ou custo possível)

Desta forma, pode-se dizer que a otimização de algoritmos não é uma tarefa trivial, sendo geralmente adotada em problemas iterativos de grande porte, exigindo geralmente profissionais especializados. O entendimento destes detalhes é importante principalmente quando da solução numérica de equações diferenciais.

## 5. Solução Numérica de Equações Diferenciais Ordinárias

Equações diferenciais aparecem com grande freqüência em modelos que descrevem quantitativamente fenômenos em diversas áreas, como por exemplo em mecânica dos fluidos, fluxo de calor, vibrações, reações químicas e nucleares, economia, biologia, etc. A motivação para a construção dos primeiros computadores foi ocasionada, em grande parte, pela necessidade de serem calculadas trajetórias balísticas de uma forma precisa e rápida. Hoje em dia, os computadores estão sendo muito empregados na solução de equações relacionadas com foguetes balísticos, satélites artificiais, redes elétricas, curvatura de vigas, estabilidade de aviões, vibrações e outras aplicações.

**Exemplo 10:** Considere a equação

$$\frac{dy}{dt} = 1 - e^{-t}. \quad (1)$$

Esta é uma equação diferencial porque envolve a derivada  $\frac{dy}{dt}$  de  $y = y(t)$ . Apenas a variável independente  $t$  aparece do lado direito desta equação. Portanto, uma solução desta é a antiderivada de  $1 - e^{-t}$  e as regras de integração podem ser empregadas para determinar  $y(t)$ :

$$y(t) = t + e^{-t} + c, \quad (2)$$

onde  $c$  é a constante de integração.

**Exemplo 11:** Considere a temperatura  $y(t)$  de um objeto sob processo de resfriamento. A taxa de variação desta está relacionada com a diferença de temperatura do corpo e do meio que o cerca. Este fenômeno pode ser expresso pela equação diferencial

$$\frac{dy}{dt} = -k(y - A)$$

onde  $A$  é a temperatura do meio,  $y$  a temperatura do objeto no tempo  $t$  e  $k$  uma constante positiva. O sinal negativo é necessário para garantir que  $\frac{dy}{dt}$  será negativo quando a temperatura do corpo superar a do meio.

Se a temperatura do objeto é conhecida no tempo  $t = 0$ , diz-se que esta é uma condição inicial e inclui-se esta informação na formulação do problema, que fica:

$$\frac{dy}{dt} = -k(y - A) \quad \text{com} \quad y(0) = y_0.$$

Pode-se usar a técnica de separação de variáveis para encontrar a solução

$$y = A + (y_0 - A)e^{-kt}.$$

Para cada escolha de  $y_0$  a curva solução será diferente, como mostra a figura. Pode-se observar que, à medida que o tempo passa, a temperatura do objeto se aproxima da temperatura do meio. Se  $y_0 < A$  então o objeto está sendo aquecido, e não resfriado.

Uma equação envolvendo uma relação entre uma função desconhecida e uma ou mais de suas derivadas é denominada *equação diferencial*. Assim, uma equação diferencial ordinária (que tem apenas uma variável independente) de ordem  $n$  tem a forma

$$y^{(n)} = f(x, y, y', y'' \dots, y^{(n-1)}). \quad (3)$$

Sua solução é uma função  $\phi(x)$   $n$  vezes diferenciável em um intervalo determinado e que satisfaz (3), isto é,

$$\phi^{(n)} = f(x, \phi, \phi', \phi'' \dots, \phi^{(n-1)}).$$

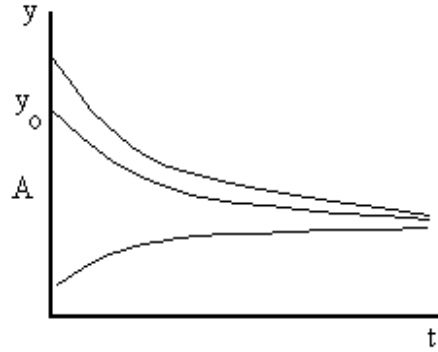


Figura 1: Temperatura de um objeto sob processo de resfriamento ou aquecimento

**Exemplo 12:** Equações diferenciais ordinárias e parciais

$$\frac{dy}{dx} = x + y \quad \text{equação diferencial ordinária}$$

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad \text{equação diferencial parcial}$$

Diz-se que a ordem de uma equação diferencial é a ordem da mais alta derivada que aparece na equação. Uma equação diferencial é linear se a função e suas derivadas aparecem linearmente na equação.

**Exemplo 13:** Equação diferencial (não) linear

$$xy' = x - y \quad \text{equação diferencial linear}$$

$$y'' + (1 - y^2)y' + y = 0 \quad \text{equação diferencial não linear}$$

Se, dada uma equação de ordem  $m$ , a função, assim como as suas derivadas até ordem  $m - 1$  são especificadas em um mesmo ponto, então tem-se um *problema de valor inicial - PVI*. Se, em problemas envolvendo equações diferenciais ordinárias de ordem  $m \geq 2$ , as  $m$  condições fornecidas para busca da solução única não são todas dadas em um mesmo ponto, então tem-se um *problema de valor de contorno - PVC*.

**Exemplo 14:** Problemas de valor inicial e de contorno

$$\begin{cases} y'(x) = y \\ y(0) = 1 \end{cases} \quad \text{é um PVI}$$

$$\begin{cases} y^{(4)}(x) + ky(x) = q \\ y(0) = y'(0) = 0 \\ y(L) = y''(L) = 0 \end{cases} \quad \text{é um PVC}$$

Embora existam várias técnicas para solucionar, de forma aproximada, algumas classes selecionadas de equações diferenciais, a grande maioria das equações encontradas na prática não podem ser solucionadas analiticamente. Não existe, por exemplo, nenhuma “expressão fechada” para a solução de  $y' = x^3 + y^2$  com  $y(0) = 0$ . Neste caso, os recursos disponíveis são os métodos numéricos, que aproximam a solução desejada.

Um procedimento numérico para calcular a solução de um dado PVI é um algoritmo para calcular os valores aproximados  $y_0, y_1, y_2, \dots, y_n, \dots$  da solução  $y = \phi(t)$  em um conjunto de pontos  $t_0 < t_1 < t_2 < \dots < t_n \dots$ , conforme a figura 2:

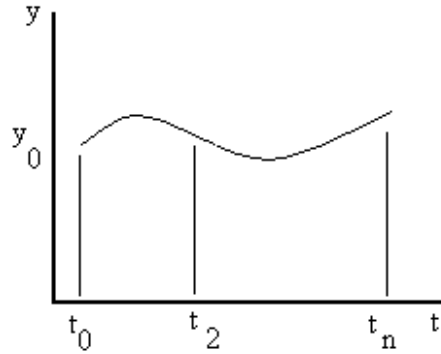


Figura 2: Aproximação da solução de um PVI

A primeira tentativa de resolução numérica de uma equação diferencial foi provavelmente feita por Euler, aproximadamente em 1768. Ele usou o que se chama hoje de *método da tangente* ou também *método de Euler*.

### 5.1 Método de Euler:

Este serve atualmente como base para o entendimento de métodos mais elaborados. Seu uso é limitado, pois o erro acumulado à medida que o processo se desenvolve é grande.

Conhecendo-se  $t_0$  e  $y_0$ , o coeficiente angular da reta tangente ao gráfico da solução em  $t = t_0$ , ou seja,  $\phi'(t_0) = f(t_0, y_0)$ , também é conhecido. Portanto, é possível construir a tangente à solução em  $t_0$  e obter um valor aproximado  $y_1$  de  $\phi(t_1)$  mediante um deslocamento sobre a reta tangente desde  $t_0$  até  $t_1$ , conforme a figura.

Assim,

$$\begin{aligned} y_1 &= y_0 + \phi'(t_0) (t_1 - t_0) \\ &= y_0 + f(t_0, y_0) (t_1 - t_0). \end{aligned}$$

Conhecendo-se  $y_1$ , pode-se calcular  $f_1 = f(t_1, y_1)$  e usar este valor como coeficiente angular da nova aproximação, percorrendo o caminho de  $t_1$  até  $t_2$ . Desta forma, obtém-se

$$\begin{aligned} y_2 &= y_1 + f_1 (t_2 - t_1) \\ &= y_1 + f(t_1, y_1) (t_2 - t_1). \end{aligned}$$

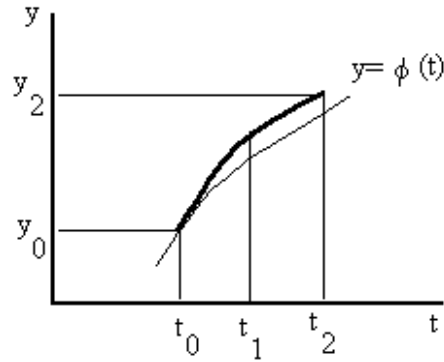


Figura 3: Aproximação da solução via tangente

Continuando desta maneira, utiliza-se o valor de  $y$  calculado em cada passo para determinar o coeficiente angular da aproximação do passo seguinte. A fórmula geral do método de Euler é

$$y_{n+1} = y_n + f(t_n, y_n) (t_{n+1} - t_n).$$

Admitindo que o espaçamento  $h$  entre os pontos  $t_0, t_1, \dots$  seja uniforme, então  $t_{n+1} = t_n + h$  e a fórmula de Euler pode ser escrita como:

$$\begin{aligned} y_{n+1} &= y_n + h f(t_n, y_n) \\ &= y_n + h f_n, \quad n = 0, 1, 2, \dots \end{aligned}$$

**Exemplo 15:** Resolver o problema

$$y' = 1 - t + 4y \tag{4}$$

$$y(0) = 1 \tag{5}$$

**Solução:** Esta equação é do tipo diferencial linear de primeira ordem, sendo fácil verificar que a solução que cumpre a condição inicial (5) é

$$y = \phi(t) = \frac{1}{4}t - \frac{3}{16} + \frac{19}{16}e^{4t}.$$

Assim, com a fórmula de Euler e um incremento  $h = 0,1$  é possível determinar um valor aproximado para a solução em  $t = 0,2$  deste PVI. Neste caso,  $f(t, y) = 1 - t + 4y$ . Para usar a aproximação de Euler calcula-se inicialmente  $f_0 = f(0, 1) = (1 - 0 + 4 * 1) = 5$ . Então:

$$\begin{aligned} y_1 &= y_0 + h f(0, 1) \\ &= 1 + (0,1) (5) \\ &= 1,5. \end{aligned}$$

Na etapa seguinte,

$$\begin{aligned}
y_2 &= y_1 + h f(t_1, y_1) \\
&= 1,5 + (0,1) f(0,1, 1,5) \\
&= 1,5 + (0,1) (6,9) \\
&= 2,19.
\end{aligned}$$

Este resultado pode ser comparado ao valor exato de  $\phi(0,2)$  que é 2,5053299. Logo, o erro é aproximadamente  $2,51 - 2,19 = 0,32$ . Um erro desta grandeza (de 12%) não é normalmente aceitável.

*Sugestão:* Experimente refazer seus cálculos considerando espaçamentos progressivamente menores (como 0,05 e 0,01) e observe o que acontece.

**Exercício 3:** Considere que o decaimento radioativo de uma substância segue a equação diferencial

$$\frac{dQ}{dt} = -0,0525 Q$$

Se 50 mg desta estiverem presentes numa amostra no dia de hoje, determine quanto existirá daqui a 2 anos. Considere  $h = 1$  e 0,5. Discuta os resultados.

**Exercício 4:** Um capital de R\$ 1000,00 é depositado em uma conta. Sabendo que sobre ele incide uma taxa de 10% de juros compostos ao ano, determine qual será o saldo na conta ao final de 5 anos.

**Erros nos procedimentos numéricos:** A adoção de um procedimento numérico como a fórmula de Euler para resolver um problema de valor inicial conduz a diversas questões. Uma delas é a da *convergência*. Será que, à medida que o incremento  $h$  tende para zero os valores da solução numérica,  $y_1, y_2, \dots$ , se aproximam dos valores correspondentes aos da solução exata? Se a resposta for afirmativa, ainda assim permanece a importante questão prática da velocidade com que a solução numérica se aproxima da solução exata. Em outras palavras, qual o tamanho necessário do incremento para que se garanta um nível de precisão dado? É importante que seja usado um incremento suficientemente pequeno para assegurar a precisão desejada, mas não tão pequeno que torne os cálculos mais lentos ou que, em alguns casos, cause perda de precisão.

Há duas fontes fundamentais de erro na resolução numérica de um problema de valor inicial: erro de truncamento e erro de arredondamento.

A diferença  $E_n$  entre a solução exata  $y = \phi(x)$  e a solução aproximada, dada por

$$E_n = \phi(t_n) - y_n$$

é o *erro de truncamento global*. Este é proveniente de duas causas: em cada passo usa-se uma fórmula aproximada para determinar  $y_{n+1}$  e os dados de entrada, em cada passo, não concordam, em geral, com a solução exata. Se se admite que os dados de entrada são corretos, o único erro, quando se avança um passo no processo, se deve ao uso de uma fórmula aproximada. Este é o *erro de truncamento local*,  $|e_n|$ .

A segunda fonte fundamental de erro é que os cálculos são feitos em aritmética com apenas um número finito de dígitos. Isto leva a um *erro de arredondamento*  $R_n$  definido por

$$R_n = y_n - Y_n$$

onde  $Y_n$  é o valor “realmente” calculado pelo procedimento numérico dado. O valor absoluto do erro total no cálculo de  $\phi(t_n)$  é dado por

$$|\phi(t_n) - Y_n| \leq |E_n| + |R_n|$$

Os erros de arredondamento têm natureza mais aleatória do que os de truncamento, pois dependem do tipo de computador utilizado, da seqüência de execução dos cálculos, do método de arredondamento, etc.

**Erro de truncamento para o método de Euler:** Quando a fórmula de Euler é obtida usando a idéia da tangente, na verdade alguns termos são desprezados em cada passo, dando origem ao erro de truncamento local, de forma que:

$$|e_n| \leq \frac{h^2}{2} \max_{t \in I} \phi''(t) \quad (6)$$

Assim, se o único erro cometido a cada passo é o de truncamento local, então depois de  $m$  passos o erro acumulado, ou seja, o erro de truncamento global, no intervalo  $I = [a, b]$  considerado é

$$|E_n| = \phi(b) - y_m$$

e

$$|E_n| \leq C h$$

onde  $C$  é uma constante. Isto mostra que, se o espaçamento  $h$  no método de Euler é dividido pela metade, pode-se esperar que o erro de truncamento global do método também seja reduzido pela metade.

**Exemplo 16:** Como exemplo da forma de usar os resultados mostrados anteriormente quando se dispõe de informações a respeito da solução de uma problema de valor inicial, resolva

$$\begin{aligned} y' &= 1 - t + 4y \\ y(0) &= 1 \end{aligned}$$

no intervalo  $[0, 1]$ .

**Solução:** Seja  $y = \phi(t)$  a solução do problema. Neste caso,

$$\phi(t) = \frac{t}{4} - \frac{3}{16} + \frac{19}{16} e^{4t}$$

e portanto,

$$\phi''(t) = 19e^{4t}.$$

De acordo com a equação (6), o erro no primeiro passo (para  $\Delta t = 0.1$ ) é dado por

$$|e_1| \leq \frac{19e^{0,4}(0,1)^2}{2},$$

pois

$$\max_{t \in [0;0,1]} \phi''(t) = e^{4 \times 0,1}.$$

Isto mostra que o erro aumenta progressivamente com  $t$ .

**Exemplo 17:** Compare o erro de truncamento global quando o método de Euler é empregado para resolver

$$\begin{aligned} y' &= \frac{t-y}{2} \\ y(0) &= 1 \end{aligned}$$

no intervalo  $[0, 3]$  com  $h = 1, \frac{1}{2}, \dots, \frac{1}{64}$ .

**Solução:** A tabela a seguir<sup>1</sup> fornece o erro de truncamento global para vários espaçamentos e mostra que o erro na aproximação de  $y(3)$  é dividido pela metade quando o espaçamento é dividido pela metade.

$h$	$m$	$y_m$	$ \phi(3) - y_m $
1	3	1,375	0,294390
1/2	6	1,533936	0,135454
1/4	12	1,604252	0,065138
1/8	24	1,637429	0,031961
1/16	48	1,653557	0,015833
1/32	96	1,661510	0,007880
1/64	192	1,665459	0,003931

**Exercício 5:** Considere o problema de valor inicial

$$\begin{aligned} y' &= \cos 5\pi t \\ y(0) &= 1 \end{aligned}$$

1. Determine a solução exata  $y = \phi(t)$  no domínio  $[0, 1]$ .
2. Determine os valores aproximados de  $\phi(t)$  em  $t = 0, 2; 0, 4; 0, 6$  pelo método de Euler com  $h = 0, 2$ . Desenhe o gráfico de segmentos retilíneos da solução aproximada e compare-a com o gráfico da solução exata.

---

<sup>1</sup> $m$  é o número de passos e  $y_m$  é a aproximação para  $\phi(3)$ .

3. Repetir o cálculo da parte (1) no domínio  $[0; 0, 4]$  tomando  $h = 0, 1$ .
4. Mostre, pelo cálculo do erro de truncamento local, que nenhum dos dois incrementos é suficientemente pequeno. Determine um valor de  $h$  que assegure que o erro de truncamento local seja menor do que  $0, 05$  em todo o intervalo  $[0, 1]$ .

## 5.2 Método de Taylor:

O método de Taylor é bem geral e fornece uma maneira de comparar os vários métodos numéricos existentes para a solução de um problema de valor inicial. Ele pode ser construído de maneira a ter um alto grau de precisão.

Se  $y(t)$  é uma função analítica e tem derivadas contínuas de todas as ordens, então ela pode ser representada por uma série de Taylor:

$$y(t+h) = y(t) + h y'(t) + \frac{h^2}{2!} y''(t) + \frac{h^3}{3!} y^{(3)}(t) + \dots$$

Para fins numéricos, deve-se fixar o número  $N$  e usar uma série finita para aproximar  $y(t+h)$ . Então, esta fórmula fica

$$y(t+h) \approx y(t) + h y'(t) + \frac{h^2}{2!} y''(t) + \dots + \frac{h^N}{N!} y^{(N)}(t) \quad (7)$$

O erro nesta aproximação é chamado de *erro de truncamento local* e é dado por

$$|e| \leq \frac{h^{N+1}}{(N+1)!} \max_{t \in I} y^{(N+1)}(t),$$

de forma que o erro de truncamento global satisfaz

$$|E| \leq C h^N,$$

onde  $C$  é uma constante.

A solução do problema de valor inicial  $y' = f(t, y)$  em  $[t_0, t_m]$  é obtida por uma adaptação de (3) em cada subintervalo  $[t_n, t_{n+1}]$ . As derivadas  $y^{(j)}(t)$  são obtidas pela diferenciação de  $f(t, y)$ . A fórmula geral para o método de Taylor de ordem  $N$  é

$$y_{n+1} = y_n + d_1 h + \frac{d_2 h^2}{2!} + \frac{d_3 h^3}{3!} + \dots + \frac{d_N h^N}{N!}, \quad (8)$$

onde  $d_j = y^{(j)}(t_k)$  para  $j = 1, 2, \dots, N$ ,  $n = 0, 1, 2, \dots, m-1$ .

*Observação:* Como as fórmulas de Taylor exigem o cálculo da função  $f$  e de suas derivadas em cada ponto, em geral é mais vantajoso escolher um outro método que tenha precisão comparável, mas que não exija o cálculo das derivadas (que em alguns casos podem ser bastante complicadas).

**Exemplo 18:** Use o método de Taylor de ordem  $N = 4$  para calcular  $y_1$  e  $y_2$  com  $h = 0, 25$ , sabendo que

$$y' = \frac{t-y}{2}$$

$$y(0) = 1$$

**Solução:** O primeiro passo é determinar as derivadas de  $y(t)$ , ou seja,

$$y'(t) = \frac{t-y}{2},$$

$$y''(t) = \frac{d}{dt} \frac{t-y}{2} = \frac{1-y'}{2} = \frac{2-t+y}{4},$$

$$y^{(3)}(t) = \frac{d}{dt} \frac{2-t+y}{4} = \frac{0-1+y'}{4} = \frac{-2+t-y}{8},$$

$$y^{(4)}(t) = \frac{d}{dt} \frac{-2+t-y}{8} = \frac{0+1-y'}{8} = \frac{2-t+y}{16}.$$

Para determinar  $y_1$ , estas derivadas devem ser calculadas no ponto  $(t_0, y_0) = (0, 1)$ , o que resulta:

$$d_1 = y'(0) = -0,5,$$

$$d_2 = y''(0) = 0,75,$$

$$d_3 = y^{(3)}(0) = -0,375,$$

$$d_4 = y^{(4)}(0) = 0,1875$$

Agora, as derivadas  $\{d_j\}$  são substituídas em (8) com  $h = 0,25$  e  $N = 4$ , de forma que

$$y_1 = 0,8974915$$

Para calcular  $y_2$ , as derivadas  $\{d_j\}$  são calculadas no ponto  $(t_1, y_1) = (0,25; 0,8974915)$  e, repetindo o processo mostrado, obtém-se

$$y_2 = 0,8364037$$

**Exercício 6:** Um paraquedista salta de um avião e até o momento que ele abre o pára-quedas, a resistência do ar é proporcional a  $v^{\frac{3}{2}}$ . Assuma que o intervalo de tempo é  $[0, 2]$  e que a equação diferencial para a velocidade vertical seja

$$v' = 32 - 0,032 v^{\frac{3}{2}}$$

$$v(0) = 0$$

Use o método de Taylor de segunda ordem com  $h = 0,5$  e encontre a solução deste problema.

**Exercício 7:** Assuma que a curva  $P(t)$  para uma determinada população obedeça a equação diferencial  $P' = aP - bP^2$ . Seja  $t$  o tempo em anos e  $h = 10$  o passo. Os valores  $a = 0,02$  e  $b = 0,00004$  produzem um modelo para a população. Considerando que no ano de 1990 a população era de 76,1 milhões, obtenha, usando o método de Taylor de segunda ordem, uma estimativa para esta população no ano de 2010.

### 5.3 Métodos de Runge-Kutta:

Os métodos de Taylor têm uma boa característica: o erro de truncamento global é  $O(h^N)$ , onde  $N$  pode ser escolhido tão grande tal que o erro seja pequeno. Entretanto, o problema destes métodos é a necessidade de cálculo de derivadas de ordem mais alta, as quais podem ser complicadas. Cada método de Runge-Kutta é derivado de um método de Taylor apropriado de tal maneira que o erro de truncamento global seja  $O(h^N)$ ; para eliminar o cálculo das derivadas, faz-se várias avaliações da função  $f$  a cada passo. Estes métodos podem ser construídos para qualquer ordem  $N$ .

O método de Runge-Kutta de segunda ordem é definido pelas equações:

$$y_{n+1} = y_n + \frac{h}{2} (k_1 + k_2) \quad (9)$$

onde

$$\begin{aligned} k_1 &= f(x_n, y_n) \\ k_2 &= f(x_n + h, y_n + h k_1) \end{aligned}$$

**Exemplo 19:** Calcular a solução aproximada da equação diferencial

$$\begin{aligned} \frac{dy}{dx} &= -x y \\ y(0) &= 1 \end{aligned}$$

no intervalo  $[0, 1]$ , utilizando o método de Runge-Kutta de segunda ordem com  $h = 0,1$ .

**Solução:** A fórmula iterativa (9) neste caso fica

$$y_{n+1} = y_n + \frac{0,1}{2} (k_1 + k_2)$$

onde

$$\begin{aligned} k_1 &= -x_n y_n \\ k_2 &= -(x_n + 0,1)(y_n + 0,1 k_1) \end{aligned}$$

Como  $x_0 = 0$  e  $y_0 = 1$ , os resultados obtidos são os seguintes:

O método de Runge-Kutta de ordem  $N = 4$  se constitui num dos métodos clássicos. Ele é em geral uma boa escolha, pois é bastante preciso, estável e fácil de programar. Sua forma é a seguinte:

$$y_{n+1} = y_n + \frac{h}{6} (k_1 + 2k_2 + 2k_3 + k_4) \quad (10)$$

onde

$$\begin{aligned}
k_1 &= f(t_n, y_n) \\
k_2 &= f\left(t_n + \frac{h}{2}, y_n + \frac{h}{2} k_1\right) \\
k_3 &= f\left(t_n + \frac{h}{2}, y_n + \frac{h}{2} k_2\right) \\
k_4 &= f(t_n + h, y_n + h k_3)
\end{aligned}$$

Ao invés de mostrar o desenvolvimento rigoroso das fórmulas anteriores, é interessante observar algumas características importantes. Considere o gráfico (Fig. 4) de  $y = y(t)$  no primeiro subintervalo  $[t_0, t_1]$ . Os valores em (10) são aproximações para as derivadas desta curva em determinados pontos. Aqui,  $k_1$  é a derivada na extremidade esquerda do intervalo,  $k_2$  e  $k_3$  são duas estimativas para a derivada no ponto central e  $k_4$  é a derivada na extremidade direita.

O ponto seguinte,  $(t_1, y_1)$ , é obtido pela integração da função derivada:

$$y(t_1) - y(t_0) = \int_{t_0}^{t_1} f(t, y(t)) dt. \quad (11)$$

Se a regra de Simpson é aplicada à equação (11) com espaçamento  $h/2$ , então a aproximação para a integral fica

$$\int_{t_0}^{t_1} f(t, y(t)) dt \approx \frac{h}{6} [f(t_0, y(t_0)) + 4f(t_{1/2}, y(t_{1/2})) + f(t_1, y(t_1))], \quad (12)$$

onde  $t_{1/2}$  é o ponto médio do intervalo (conforme figura 4).

Três valores da função  $f$  são necessários; portanto, faz-se a escolha:  $f(t_0, y(t_0)) = k_1$  e  $f(t_1, y(t_1)) \approx k_4$  e para o valor na metade do intervalo, escolhe-se uma média entre  $k_2$  e  $k_3$ , isto é,

$$f(t_{1/2}, y(t_{1/2})) \approx \frac{k_2 + k_3}{2}.$$

Estes valores são substituídos em (12), que é usada na equação (11) para determinar  $y_1$ :

Tabela 1: *Resultados via método de Runge-Kutta ordem 2*

$n$	$x_n$	$y_n$	$k_1$	$k_2$	$y_{n+1}$
0	0,0	1,0	0,0	-0,1	0,995
1	0,1	0,995	-0,0995	-0,19701	0,9801745
2	0,2	0,9801745	-0,1960349	-0,288171303	0,95596419
3	0,3	0,95596419	-0,286789257	-0,370914106	0,923079022
4	0,4	0,923079022	-0,369231609	-0,443077930	0,882463545
5	0,5	0,882463545	-0,441231772	-0,503004221	0,835251745
6	0,6	0,835251745	-0,501151047	-0,549595648	0,78271441
7	0,7	0,78271441	-0,547900087	-0,582339521	0,72620243
8	0,8	0,72620243	-0,580961944	-0,601295612	0,667089552
9	0,9	0,667089552	-0,600380597	-0,607051492	0,606717947

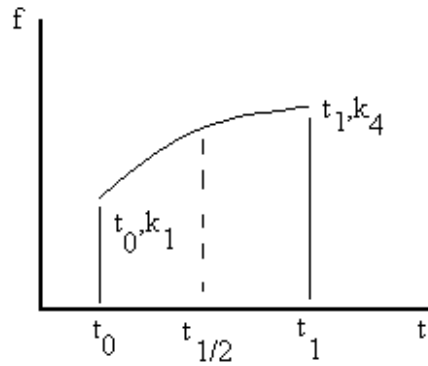


Figura 4: Aproximação integral da função  $f[t,y(t)]$

$$y_1 = y_0 + \frac{h}{6} \left[ k_1 + \frac{4(k_2 + k_3)}{2} + k_4 \right].$$

**Exemplo 20:** Calcular a solução aproximada da equação diferencial

$$\begin{aligned} \frac{dy}{dx} &= -xy \\ y(0) &= 1 \end{aligned}$$

no intervalo  $[0, 1]$ , utilizando o método de Runge-Kutta de quarta ordem com  $h = 0,1$ .

**Solução:** A fórmula iterativa (10) neste caso fica

$$y_{n+1} = y_n + \frac{0,1}{2} (k_1 + 2k_2 + 2k_3 + k_4)$$

onde

$$\begin{aligned} k_1 &= -x_n y_n \\ k_2 &= -(x_n + 0,05)(y_n + 0,05 k_1) \\ k_3 &= -(x_n + 0,1)(y_n + 0,1 k_2) \\ k_4 &= -(x_n + 0,1)(y_n + 0,1 k_3) \end{aligned}$$

Como  $x_0 = 0$  e  $y_0 = 1$ , os resultados obtidos são:

Tabela 2: Resultados via método de Runge-Kutta, ordem 4

$n$	$x_n$	$y_n$	$k_1$	$k_2$
0	0,0	1,0	0,0	-0,05
1	0,1	0,995012479	-0,099501248	-0,148505613
2	0,2	0,980198673	-0,196039735	-0,242599172
3	0,3	0,955997481	-0,286799244	-0,329580132
4	0,4	0,923116345	-0,369246538	-0,407094308
5	0,5	0,882496901	-0,44124845	-0,4732238963
6	0,6	0,83527021	-0,501162126	-0,526637868
7	0,7	0,782704542	-0,547893179	-0,566482412
8	0,8	0,726149051	-0,580919241	-0,592537626
9	0,9	0,666976845	-0,600279160	-0,605114742

$k_3$	$k_4$	$y_{n+1}$
-0,049875	-0,09950125	0,995012479
-0,14813808	-0,196039734	0,980198673
-0,242017199	-0,286799087	0,955997481
-0,328831466	-0,369245734	0,923116345
-0,406242733	-0,441246036	0,882496901
-0,472359224	-0,501156587	0,83527021
-0,525809906	-0,547882454	0,782704542
-0,565785316	-0,580900808	0,726149051
-0,592043844	-0,6002502	0,666976845
-0,604885052	-0,606488339	0,606530727

**Exercício 8:** Em uma reação química, uma molécula de  $A$  se combina com uma molécula de  $B$  para formar uma molécula do produto químico  $C = A + B$ . Sabe-se que a concentração,  $y(t)$ , no tempo  $t$ , é a solução do problema de valor inicial:

$$y' = k(a - y)(b - y)$$

$$y(0) = 0$$

onde  $k$  é uma constante positiva e  $a$  e  $b$  são as concentrações iniciais de  $A$  e de  $B$ , respectivamente. Suponha que  $k = 0,01$ ,  $a = 70$  milimoles/litro e  $b = 50$  milimoles/litro. Use os métodos de Runge-Kutta de ordem  $N = 2$  e  $4$  com  $h = 0,5$  para encontrar a solução em  $[0, 2]$ .

#### 5.4 Métodos de passo múltiplo:

Os métodos de Euler, Taylor e Runge-Kutta são chamados de métodos de passo simples, pois o valor  $y_{n+1}$  calculado em qualquer ponto depende apenas dos dados no ponto precedente,  $y_n$ . Os métodos que utilizam informação em mais de um ponto, além do último, são denominados *métodos de passo múltiplo*.

### 5.4.1 Métodos de Adams:

De acordo com o teorema fundamental do cálculo, sabe-se que

$$y(t_{n+1}) - y(t_n) = \int_{t_n}^{t_{n+1}} y'(t) dt \quad (13)$$

A idéia básica dos métodos de Adams é aproximar  $y'(t)$  por um polinômio  $p_k(t)$  de grau  $k$  e usar o polinômio para determinar a integral do lado direito da equação (13). Os coeficientes de  $p_k(t)$  são determinados usando os pontos calculados anteriormente.

As fórmulas de Adams mais precisas podem ser obtidas usando polinômios de grau mais elevado e, portanto, um número maior de pontos. Suponha que seja utilizado um polinômio de grau 3. Os coeficientes são determinados a partir dos quatro pontos  $(t_n, y_n)$ ,  $(t_{n-1}, y_{n-1})$ ,  $(t_{n-2}, y_{n-2})$  e  $(t_{n-3}, y_{n-3})$ . Substituindo  $y'(t)$  por este polinômio na equação (13), calculando a integral e simplificando o resultado, obtém-se a fórmula de Adams-Bashforth de quarta ordem:

$$y_{n+1} = y_n + \frac{h}{24} (55 f_n - 59 f_{n-1} + 37 f_{n-2} - 9 f_{n-3}), \quad (14)$$

cujo erro de truncamento global é proporcional a  $h^4$ .

Uma pequena modificação na dedução das fórmulas de Adams-Bashforth resulta em outro conjunto de fórmulas denominadas de fórmulas de Adams-Moulton de quarta ordem, para a qual o erro de truncamento global também é proporcional a  $h^4$ , sendo dada por

$$y_{n+1} = y_n + \frac{h}{24} (9 f_{n+1} + 19 f_n - 5 f_{n-1} + f_{n-2}), \quad (15)$$

que corresponde a uma fórmula implícita, já que  $f_{n+1} = f(t_{n+1}, y_{n+1})$ .

Embora os erros de truncamento locais para as fórmulas de Adams-Bashforth e de Adams-Moulton de mesma ordem sejam proporcionais à mesma potência de  $h$ , as fórmulas de Adams-Moulton são mais precisas. Surge, portanto, a questão: deve-se usar a fórmula de Adams-Bashforth, que é explícita e mais rápida, ou a de Adams-Moulton, que é implícita e mais lenta? Para responder a esta pergunta, é preciso determinar se o uso de uma fórmula mais precisa permite aumentar o incremento e, portanto, diminuir o número de passos, o suficiente para compensar as computações adicionais necessárias em cada passo.

Na verdade, o *método de predição-correção* é uma tentativa de obter simplicidade e precisão combinando as duas fórmulas. Uma vez que  $y_{n-3}$ ,  $y_{n-2}$ ,  $y_{n-1}$  e  $y_n$  são conhecidas, pode-se calcular  $f_{n-3}$ ,  $f_{n-2}$ ,  $f_{n-1}$  e  $f_n$  e, em seguida, usar a fórmula de Adams-Bashforth (preditor), para obter um primeiro valor para  $y_{n+1}$ :

$$y_{n+1}^{(0)} = y_n + \frac{h}{24} (55 f_n - 59 f_{n-1} + 37 f_{n-2} - 9 f_{n-3}). \quad (16)$$

Em seguida, calcula-se  $f_{n+1} = f(t_{n+1}, y_{n+1}^{(0)})$  e usa-se a fórmula de Adams-Moulton (corretor), para obter um valor mais preciso de  $y_{n+1}$ :

$$y_{n+1} = y_n + \frac{h}{24} (9 f_{n+1} + 19 f_n - 5 f_{n-1} + f_{n-2}). \quad (17)$$

Pode-se continuar usando a fórmula de correção (17) se a variação de  $y_{n+1}$  for muito grande. Entretanto, se for necessário usar a fórmula de correção mais do que uma ou duas vezes, isto significa que o incremento  $h$  é excessivamente grande e deve ser reduzido.

Para usar qualquer um dos métodos de passos múltiplos, é necessário primeiro calcular alguns  $y_n$  usando outro método. Por exemplo: o método de Adams-Moulton de quarta ordem requer valores para  $y_1$  e  $y_2$ , enquanto o método de Adams-Bashforth de quarta ordem também requer um valor para  $y_3$ . Uma maneira de contornar esta dificuldade é usar um método de passo simples de precisão comparável para determinar os valores iniciais necessários. Assim, no caso de métodos de passos múltiplos de quarta ordem, pode-se empregar o método de Runge-Kutta de quarta ordem para obter os valores iniciais.

**Exemplo 21:** Considere o problema de valor inicial

$$\begin{aligned}y' &= 1 - t + 4y \\ y(0) &= 1\end{aligned}$$

Determine um valor aproximado da solução  $y(t)$  com um incremento  $h = 0,1$  em  $t = 0,4$ . Use as fórmulas de quarta ordem de Adams-Bashforth, de Adams-Moulton e de predição-correção.

**Solução:** Como dados iniciais, usa-se os valores iniciais  $y_1$ ,  $y_2$  e  $y_3$  determinados com o auxílio do método de Runge-Kutta de quarta-ordem. Em seguida, calculando os valores correspondentes de  $f(t, y)$ , obtém-se

$y_0 = 1$	$f_0 = 5$
$y_1 = 1,6089333$	$f_1 = 7,3357332$
$y_2 = 2,5050062$	$f_2 = 10,820025$
$y_3 = 3,8294145$	$f_3 = 16,017658$

1. Usando a fórmula de Adams-Bashforth, equação (14), determina-se que  $y_4 = 5,7836305$ .
2. A fórmula de Adams-Moulton, equação (15), leva à equação

$$y_4 = 4,9251275 + 0,15 y_4,$$

da qual segue-se que  $y_4 = 5,7942676$ .

3. Finalmente, usando o resultado da fórmula de Adams-Bashforth (16) como preditor, pode-se usar a fórmula (17) como corretor. O valor do preditor,  $y_4 = 5,7836305$ , leva a  $f_4 = 23,734522$  e, de acordo com a equação (17), o valor corrigido de  $y_4$  é  $5,7926721$ .

Considerando que o valor da solução exata no ponto  $t = 0,4$  seja  $\phi(0,4) = 5,7942260$ , o método de Adams-Bashforth, embora seja o mais simples e rápido (já que envolve o uso de uma única fórmula explícita), é o menos preciso. O emprego da fórmula de Adams-Moulton como correção aumenta o número de cálculos necessários, mas o método continua a ser explícito (usa-se  $y_{n+1}^{(0)}$ ). Neste problema, o erro no valor corrigido de  $y_4$  ( $y_4 = 5,7926721$ ) é reduzido aproximadamente sete vezes em relação ao erro no valor do preditor ( $y_4 = 5,7836305$ ). O método de Adams-Moulton sozinho ( $y_4 = 5,7942676$ ) é

o que fornece o resultado mais preciso, com erro 40 vezes menor que o associado ao método preditor-corretor. É necessário não esquecer, porém, que o método de Adams-Moulton é implícito, o que significa que é preciso resolver uma equação em cada passo. *Neste problema em questão a equação é linear, de modo que a solução não é difícil de encontrar. Entretanto, em outros problemas esta parte do processo pode ser muito mais demorada.*

**Exercício 9:** Uma das equações básicas dos circuitos elétricos é

$$L \frac{di}{dt} + Ri = E$$

onde  $L$  é a indutância,  $R$  a resistência,  $i$  a corrente e  $E$  a força eletromotriz. Considere  $L = 3 H$ ,  $R = 15 \Omega$ ,  $E = 110 V$  e  $i = 0$  quando  $t = 0$ . Determine o valor da corrente quando  $t = 0,5 s$  tomando  $h = 0,1 s$  e usando o método de predição-correção de quarta ordem. Seria o valor de  $h$  adotado adequado?

#### 5.4.2 Método de Milne-Simpson:

Este é um outro método de predição-correção. Seu preditor é baseado na integração de  $f(t, y(t))$  no intervalo  $[t_{n-3}, t_{n+1}]$ :

$$y(t_{n+1}) = y(t_{n-3}) + \int_{t_{n-3}}^{t_{n+1}} f(t, y(t)) dt.$$

O preditor usa uma aproximação por polinômios de Lagrange para  $f(t, y(t))$  baseada nos pontos  $(t_{n-3}, f_{n-3})$ ,  $(t_{n-2}, f_{n-2})$ ,  $(t_{n-1}, f_{n-1})$  e  $(t_n, f_n)$ . Este é integrado no intervalo  $[t_{n-3}, t_{n+1}]$ , produzindo o preditor de Milne:

$$p_{n+1} = y_{n-3} + \frac{4h}{3} (2f_{n-2} - f_{n-1} + 2f_n).$$

O corretor é obtido de maneira similar. Como o valor de  $p_{n+1}$  pode ser utilizado, um segundo polinômio de Lagrange para  $f(t, y(t))$ , baseado nos pontos  $(t_{n-1}, f_{n-1})$ ,  $(t_n, f_n)$  e  $(t_{n+1}, f(t_{n+1}, p_{n+1}))$ , é construído. O polinômio é integrado sobre  $[t_{n-1}, t_{n+1}]$ , produzindo a regra de Simpson:

$$y_{n+1} = y_{n-1} + \frac{h}{3} (f_{n-1} + 4f_n + f_{n+1}).$$

## 6. Sistemas de Equações Diferenciais

Considere o seguinte problema de valor inicial:

$$\begin{aligned}\frac{dx}{dt} &= x + 2y \\ \frac{dy}{dt} &= 3x + 2y\end{aligned}\tag{18}$$

com  $x(0) = 6$  e  $y(0) = 4$ . A solução deste problema é

$$\begin{aligned}x(t) &= 4e^{4t} + 2e^{-t} \\ y(t) &= 6e^{4t} - 2e^{-t}\end{aligned}$$

Uma solução numérica para um sistema do tipo

$$\begin{aligned}\frac{dx}{dt} &= f(t, x, y) \\ \frac{dy}{dt} &= g(t, x, y)\end{aligned}\tag{19}$$

com  $x(t_0) = x_0$  e  $y(t_0) = y_0$  no intervalo  $a \leq t \leq b$  pode ser encontrada com os mesmos métodos estudados anteriormente.

O método de Euler para a solução de sistemas é fácil de formular. O intervalo  $[a, b]$  é dividido em  $M$  subintervalos de tamanho  $h = (b - a)/M$ , onde  $t_{n+1} = t_n + h$ , de forma que as fórmulas recursivas são as seguintes:

$$\begin{aligned}x_{n+1} &= x_n + h f(t_n, x_n, y_n) \\ y_{n+1} &= y_n + h g(t_n, x_n, y_n) \quad \text{para } n = 0, 1, \dots, M - 1\end{aligned}$$

Um método de ordem mais alta também pode ser usado para que se obtenha um nível razoável de precisão. Um deles poderia ser o método de Runge-Kutta de ordem 4:

$$\begin{aligned}x_{n+1} &= x_n + \frac{h}{6} (k_1 + 2k_2 + 2k_3 + k_4) \\ y_{n+1} &= y_n + \frac{h}{6} (l_1 + 2l_2 + 2l_3 + l_4)\end{aligned}$$

onde

$$\begin{aligned}k_1 &= f(t_n, x_n, y_n) & l_1 &= g(t_n, x_n, y_n) \\ k_2 &= f\left(t_n + \frac{h}{2}, x_n + \frac{h}{2}k_1, y_n + \frac{h}{2}l_1\right) & l_2 &= g\left(t_n + \frac{h}{2}, x_n + \frac{h}{2}k_1, y_n + \frac{h}{2}l_1\right) \\ k_3 &= f\left(t_n + \frac{h}{2}, x_n + \frac{h}{2}k_2, y_n + \frac{h}{2}l_2\right) & l_3 &= g\left(t_n + \frac{h}{2}, x_n + \frac{h}{2}k_2, y_n + \frac{h}{2}l_2\right) \\ k_4 &= f(t_n + h, x_n + hk_3, y_n + hl_3) & l_4 &= g(t_n + h, x_n + hk_3, y_n + hl_3)\end{aligned}$$

## 6.1 Equações diferenciais de ordem mais alta:

Este tipo de equação envolve derivadas do tipo  $x''(t)$ ,  $x'''(t)$  e assim por diante. A equação

$$m x''(t) + c x'(t) + k x(t) = f(t),$$

por exemplo, representa um sistema mecânico para uma mola com constante  $k$ . Assume-se que o atrito seja proporcional à velocidade e que a função  $f(t)$  seja uma força externa. Em geral, a posição  $x(t_0)$  e a velocidade  $x'(t_0)$  são conhecidas em um certo tempo  $t_0$ .

Esta equação pode ser escrita como

$$x''(t) = f(t, x(t), x'(t)) \quad (20)$$

com  $x(t_0) = x_0$  e  $x'(t_0) = y_0$ .

A equação diferencial de segunda ordem pode ser reformulada como um sistema de duas equações de primeira ordem através da substituição

$$x'(t) = y(t).$$

Desta forma,  $x''(t) = y'(t)$  e a equação diferencial (20) torna-se um sistema

$$\begin{aligned} \frac{dx}{dt} &= y \\ \frac{dy}{dt} &= f(t, x, y) \end{aligned} \quad (21)$$

com  $x(t_0) = x_0$  e  $y(t_0) = y_0$ .

Um procedimento numérico tal como o método de Runge-Kutta pode ser empregado para resolver (21) e gerar duas seqüências,  $\{x_n\}$  e  $\{y_n\}$ , sendo que a primeira delas é a solução do problema (20).

**Exemplo 22:** Considere o problema de valor inicial de segunda ordem (movimento harmônico amortecido):

$$\begin{aligned} x''(t) + 4x'(t) + 5x(t) &= 0 \\ x(0) &= 3 \\ x'(0) &= -5 \end{aligned}$$

1. Escreva o sistema de duas equações de primeira ordem equivalente.
2. Use o método de Runge-Kutta para resolver o problema reformulado no intervalo  $[0, 5]$  usando  $M = 50$  subintervalos de tamanho  $h = 0,1$ .
3. Compare a solução numérica com a analítica,

$$x(t) = 3e^{-2t} \cos t + e^{-2t} \sin t$$

Solução: A equação diferencial tem a forma

$$x''(t) = -4x'(t) - 5x(t).$$

Usando a substituição  $x'(t) = y(t)$ , chega-se ao problema reformulado:

$$\begin{aligned}\frac{dx}{dt} &= y \\ \frac{dy}{dt} &= -5x - 4y\end{aligned}$$

com  $x(0) = 3$  e  $y(0) = -5$ .

Algumas amostras dos cálculos estão na tabela 3:

Tabela 3: *Resultados via método de Runge-Kutta*

$n$	$t_n$	$x_n$	$x(t_n)$
0	0,0	3,00000000	3,00000000
1	0,1	2,52564583	2,52565822
2	0,2	2,10402783	2,10404686
3	0,3	1,73506269	1,73508427
4	0,4	1,41653369	1,41655509
5	0,5	1,14488509	1,14490455
10	1,0	0,33324303	0,33324661
20	2,0	-0,00620684	-0,00621162
30	3,0	-0,00701079	-0,00701204
40	4,0	-0,00091163	-0,00091170
48	4,8	-0,00004972	-0,00004969
49	4,9	-0,00002348	-0,00002345
50	5,0	-0,00000493	-0,00000490

**Exercício 10:** O modelo matemático de um certo circuito elétrico *RCL* é

$$\begin{aligned}q'' + 20q' + 125q &= 9 \operatorname{sen} t \\ q(0) &= 0 \\ q'(0) &= 0\end{aligned}$$

Use o método de Runge-Kutta para resolver esta equação diferencial no intervalo  $[0, 2]$ .

**Exercício 11:** Um exemplo de um sistema de equações diferenciais não lineares é o modelo presa-predador. Seja  $x(t)$  a população de coelhos no tempo  $t$  e  $y(t)$  a de raposas. O modelo presa-predador exige que  $x(t)$  e  $y(t)$  satisfaçam

$$\begin{aligned}x' &= Ax - Bxy \\ y' &= Cxy - Dy\end{aligned}$$

Para fins de simulação numérica, pode-se considerar os coeficientes:

$$A = 2 \quad B = 0,02 \quad C = 0,0002 \quad D = 0,8.$$

Use o método de Runge-Kutta para resolver a equação diferencial no intervalo  $[0, 5]$  se:

1.  $x(0) = 3000$  coelhos e  $y(0) = 120$  raposas;
2.  $x(0) = 5000$  coelhos e  $y(0) = 100$  raposas.

## 6.2 Problemas de valores no contorno

Muitos problemas físicos podem ser descritos através de equações diferenciais com condições de contorno dadas em mais de um ponto.

**Exemplo 23:** Considere um bastão delgado de comprimento unitário com propriedades de condução de calor variáveis (diferentes partes do bastão podem ser feitas de diferentes materiais, por exemplo). Assuma que o bastão esteja isolado e que as suas extremidades sejam mantidas a diferentes temperaturas constantes. Se o tempo de duração do experimento for longo o suficiente, então a temperatura do bastão será estacionária, ou seja, independente do tempo.

Seja  $y(x)$  a temperatura do bastão no ponto  $x$ . Então  $y$  satisfaz a equação diferencial

$$\frac{d}{dx} \left[ k(x) \frac{dy}{dx} \right] = 0$$

com condições de contorno

$$y(0) = \alpha \quad y(1) = \beta$$

onde  $k(x)$  é o coeficiente de condução do calor do bastão e  $\alpha$  e  $\beta$  são as temperaturas nas extremidades. Se, além disto, o bastão contém uma fonte de calor (como uma resistência elétrica), então a equação torna-se não homogênea:

$$\frac{d}{dx} \left[ k(x) \frac{dy}{dx} \right] = f(x)$$

onde  $f(x)$  descreve a produção de calor como uma função de  $x$ .

Se  $k$  e  $f$  são constantes, então o problema proposto pode ser resolvido analiticamente (se  $f(x) = 0$  e  $k$  é constante, a solução é uma linha reta, por exemplo). Para problemas com coeficientes variáveis, geralmente é necessário discretizar o domínio e computar uma solução aproximada.

### 6.3 Diferenças finitas:

Este método é baseado na divisão do intervalo  $[a, b]$  em  $N$  subintervalos de comprimento

$$h = \frac{b - a}{N}$$

de forma que  $x_{n+1} = x_n + h$ ,  $n = 0, 1, \dots, N - 1$  e  $y_n$  é uma aproximação para  $y(x_n)$ . Observe que os valores de  $y_0$  e  $y_N$  são conhecidos, pois eles são dados pelas condições de contorno.

Considere o caso especial quando a equação é

$$y'' - q(x)y = f(x), \quad q(x) \geq 0.$$

Agora, discretiza-se a equação diferencial fazendo a substituição de  $y''$  por uma aproximação a diferenças no ponto  $x = x_n$ :

$$y''(x_n) \approx \frac{y_{n+1} - 2y_n + y_{n-1}}{h^2},$$

o que pode ser feito em todos os pontos interiores  $x_n$ ,  $n = 1, 2, \dots, N - 1$ .

Inserindo na equação diferencial, obtém-se

$$\frac{y_{n+1} - 2y_n + y_{n-1}}{h^2} - q(x_n)y_n = f(x_n), \quad n = 1, 2, \dots, N - 1$$

ou equivalentemente,

$$y_{n+1} - (2 + h^2 q_n)y_n + y_{n-1} = h^2 f_n \quad n = 1, 2, \dots, N - 1,$$

onde  $f_n$  e  $q_n$  são notações para  $f(x_n)$  e  $q(x_n)$ .

Observa-se ainda o que acontece com a primeira e a última equações, já que elas dependem também de  $y_0$  e de  $y_N$ .

Para  $n = 1$ , tem-se

$$y_2 - (2 + h^2 q_1)y_1 + y_0 = h^2 f_1$$

que, usando a condição de contorno  $y_0 = \alpha$ , fica

$$y_2 - (2 + h^2 q_1)y_1 = h^2 f_1 - \alpha.$$

Analogamente, para  $n = N - 1$ , tem-se

$$-(2 + h^2 q_{N-1})y_{N-1} + y_{N-2} = h^2 f_{N-1} - \beta.$$

Coloca-se as incógnitas  $y_n$ ,  $n = 1, 2, \dots, N - 1$ , em um vetor

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{N-2} \\ y_{N-1} \end{bmatrix}$$

de forma que as equações podem ser escritas como um sistema linear  $AX = B$ , onde a matriz  $A$  é tridiagonal

$$A = \begin{bmatrix} -(2 + h^2 q_1) & 1 & 0 & 0 & \dots & 0 \\ 1 & -(2 + h^2 q_2) & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & -(2 + h^2 q_{N-2}) & 1 \\ 0 & 0 & 0 & 0 & 1 & -(2 + h^2 q_{N-1}) \end{bmatrix}$$

e o lado direito é dado por

$$B = \begin{bmatrix} h^2 f_1 - \alpha \\ h^2 f_2 \\ \vdots \\ h^2 f_{N-2} \\ h^2 f_{N-1} - \beta \end{bmatrix}$$

Este é o denominado *método de diferenças finitas* devido à maneira com que as derivadas são aproximadas, por diferenças.

**Exemplo 24:** O problema

$$\begin{aligned} y'' - y &= 0 \\ y(0) &= 0 \\ y(1) &= \operatorname{senh}(1) \end{aligned}$$

tem solução analítica  $y(x) = \operatorname{senh}x$ . Discretize o problema usando o método de diferenças finitas com  $h = 0,25$ .

**Solução:** As condições de contorno fornecem  $y_0 = 0$  e  $y_4 = \operatorname{senh}1$ . As incógnitas  $y_1, y_2, y_3$  satisfazem

$$y_{n+1} - (2 + h^2) y_n + y_{n-1} = 0 \quad n = 1, 2, 3,$$

ou, em forma matricial

$$A = \begin{bmatrix} -(2 + h^2) & 1 & 0 \\ 1 & -(2 + h^2) & 1 \\ 0 & 1 & -(2 + h^2) \end{bmatrix}$$

e o lado direito é dado por

$$B = \begin{bmatrix} 0 \\ 0 \\ -\sinh 1 \end{bmatrix}$$

A solução deste sistema linear é dada na tabela 4:

Tabela 4: *Resultados via diferenças finitas*

$x_n$	$y_n$	<i>Erro</i>
0,25	0,252803	$1,9 \times 10^{-4}$
0,5	0,521406	$3,1 \times 10^{-4}$
0,75	0,822598	$2,8 \times 10^{-4}$

**Exercício 12:** Considere uma viga apoiada de comprimento  $l$  e de peso  $w$  por unidade de comprimento. Tomando a origem do sistema de coordenadas no apoio esquerdo, tem-se que as forças externas são a reação  $\frac{wl}{2}$  e a carga  $-wx$  e seus momentos são  $\frac{wlx}{2}$  e  $\frac{-wx^2}{2}$ , respectivamente. Portanto,

$$EI y'' = \frac{w l x}{2} - \frac{w x^2}{2}$$

com condições de contorno

$$y(0) = 0 \quad y(l) = 0$$

Resolva numericamente este problema considerando  $l = 2$ ,  $EI = 10000$  e  $w = 5000$ . Interprete os resultados.

## 7. Solução Numérica de Equações Diferenciais Parciais

Uma equação diferencial parcial (EDP) pode ser escrita na forma geral

$$a \frac{\partial^2 \phi}{\partial x^2} + b \frac{\partial^2 \phi}{\partial x \partial y} + c \frac{\partial^2 \phi}{\partial y^2} + d \frac{\partial \phi}{\partial x} + e \frac{\partial \phi}{\partial y} + f \phi + g = 0 \quad (22)$$

onde  $a, b, c, d, e, f$  e  $g$  podem ser funções das variáveis independentes  $x$  e  $y$  e da variável dependente  $\phi$ , em uma região no plano  $\mathbf{R}^2$ , em coordenadas cartesianas.

As EDPs podem ser classificadas em *elípticas*, *parabólicas* ou *hiperbólicas*, dependendo do valor de  $b^2 - 4ac$  ser negativo, zero ou positivo, respectivamente. Como exemplos de EDPs elípticas podemos citar a *equação de Poisson*

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = g \quad (23)$$

e de *Laplace*

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0 \quad (24)$$

as quais são associadas geralmente a problemas em equilíbrio.

Uma maneira de se expressar o potencial de velocidade de um fluido incompressível, não-viscoso, em regime estável é através da equação de Laplace: a taxa com a qual tal fluido entra em uma determinada região é igual aquela com a qual ele sai.

Já na teoria eletromagnética, o teorema de Gauss nos diz que o fluxo elétrico que passa através de uma superfície fechada é igual à carga total dentro da superfície; isto pode ser expresso via equação de Poisson,

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\rho}{\epsilon} = 0 \quad (25)$$

onde  $V$  é o potencial elétrico associado a uma distribuição bi-dimensional de carga de densidade  $\rho$  e  $\epsilon$  é a constante dielétrica.

Até hoje, apenas um número limitado de equações elípticas foram resolvidas analiticamente, com sua utilidade restrita a casos onde a região de estudo considerada tem uma forma geométrica simples. O mesmo pode ser dito de equações parabólicas e hiperbólicas. Por essa razão, a solução dessas equações é feita, essencialmente, de forma numérica, com métodos específicos para cada tipo de EDP.

## 7.1 Equações parabólicas

Começaremos investigando a solução numérica de EDPs parabólicas com um dos exemplos mais simples: a equação adimensional

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial t^2} \quad (26)$$

a qual expressa a distribuição de temperatura  $U$  em uma barra isolada termicamente ao longo de seu comprimento,  $x$ ,  $t$  segundos após ter sido aquecida (ou resfriada). Em tal problema, as temperaturas nos dois extremos da barra são conhecidas ao longo do tempo, ou seja, as *condições de fronteira* são conhecidas. É também usual conhecer a distribuição de temperatura na barra em um certo instante, o qual é chamado de *tempo zero*; essa distribuição é chamada de *condição inicial*.

Vejamus então um diagrama que explica o processo de integração a ser efetuado. Suponha uma barra de comprimento  $l$ ,  $C$  uma curva e  $S$  a região englobada por essa curva, conforme a figura (5).

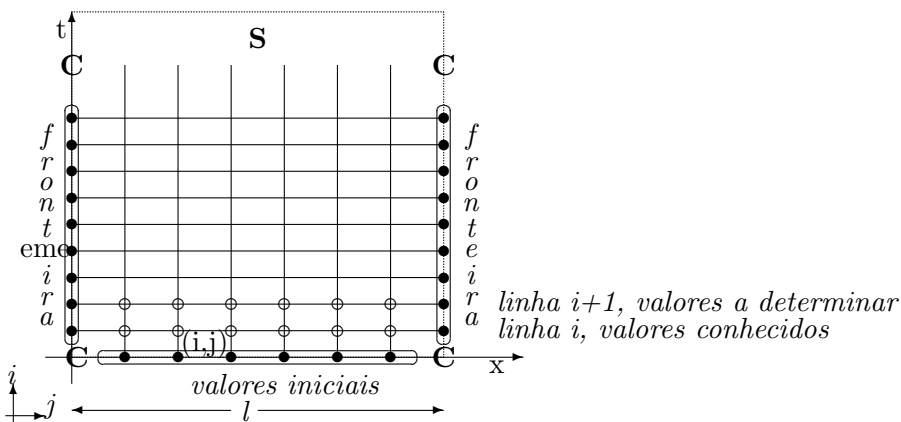


Figura 5: Malha (de pontos) para uma EDP parabólica.

Veja que  $C$  é uma curva aberta no plano  $x - t$ ; e a área  $S$  é limitada por  $0 \leq t < \infty$ ,  $0 \leq x \leq l$ . O processo de integração em  $S$  através de diferenças-finitas consiste em colocar-se uma malha com espaçamento  $k$  e  $h$  nas direções  $t$  e  $x$ , respectivamente, e aproximar as derivadas em cada um dos pontos de intersecção nessa malha.

### 7.1.1 Método explícito

Para a equação (26), aproximaremos as derivadas de primeira e de segunda ordem através de diferenças finitas, i.e.

$$\frac{\partial u}{\partial t} \approx \frac{u_{i+1,j} - u_{i,j}}{k}$$

$$\frac{\partial^2 u}{\partial t^2} \approx \frac{u_{i+1,j} - 2u_{i,j} + u_{i,j+1}}{h^2}$$

onde  $k$  e  $h$  são os espaçamentos nas direções  $t$  e  $x$ , respectivamente. Substituindo essas aproximações em (26), vem

$$\frac{u_{i+1,j} - u_{i,j}}{k} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i,j+1}}{h^2}$$

e, isolando o termo  $u_{i+1,j}$ , obtemos

$$u_{i+1,j} = ru_{i,j-1} + (1 - 2r)u_{i,j} + ru_{i,j+1}, \quad r = \frac{k}{h^2} \quad (27)$$

a qual nos dá a temperatura  $U$  em cada ponto  $j$  no  $(i + 1)$ -ésimo tempo. Note que os pontos discretos são  $x_j = jh$  e  $t_i = ik$ . Os exemplos a seguir mostram como resolver problemas com base na equação (26).

**Exemplo 25:** Considere uma barra de metal isolada termicamente, com as suas duas extremidades em contato com blocos de gelo a  $0^\circ C$ , e aquecida instantaneamente em seu ponto médio por um maçarico. Qual a temperatura da barra após um certo tempo?

**Solução:** Note que, como a barra permanece em contato com gelo durante toda a simulação, devemos esperar que, após ter sido instantaneamente aquecida, sua temperatura dever cair até  $0^\circ C$ , depois de um certo tempo.

Suponha que a distribuição inicial de temperatura na barra seja dada por

$$u = 2x, \quad 0 \leq x \leq \frac{1}{2}; \quad u = 2(1 - x), \quad \frac{1}{2} \leq x \leq 1$$

Pela formulação do problema, vemos que os extremos da barra estão a  $0^\circ C$ . Então, podemos especificar as *condições iniciais* (CI) e *de fronteira* (CF) como:

$$CI : \begin{cases} u = 2x, & 0 \leq x \leq \frac{1}{2} \\ u = 2(1 - x), & \frac{1}{2} \leq x \leq 1 \end{cases}, \quad t = 0$$

$$CF : \begin{cases} u = 0, & x = 0 \\ u = 0, & x = 1 \end{cases}, \quad t > 0$$

Dividamos então a barra em dez pedaços e integremos a equação (26) em mil passos, i.e., tomemos  $h = 1/10$  e  $k = 1/1000$ , tal que  $r = k/h^2 = 1/10$ . A equação (27) pode ser simplificada e escrita como

$$u_{i+1,j} = \frac{1}{10}(u_{i,j-1} + 8u_{i,j} + u_{i,j+1}) \quad (28)$$

e, identificando as variáveis envolvidas na malha, vemos que o cálculo dos  $u_{i+1,j}$  corresponde à seguinte “malha”, onde os números dentro de cada círculo são os fatores multiplicadores de  $u$  nos pontos da malha.

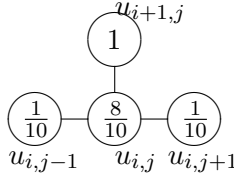


Figura 6: Malha computacional para o método explícito,  $r = 1/10$ .

Agora, aplicando a equação (28), obtemos alguns valores, mostrados na tabela (5).

Tabela 5: Alguns valores para a equação (28), com  $r = 1/10$ .

$x$	0.1	0.2	0.3	0.4	0.4	0.5	0.6
$t = 0$	0	0.2	0.4	0.6	0.8	1.0	0.8
$t = 0.001$	0	0.2	0.4	0.6	0.8	0.96	0.8
$t = 0.002$	0	0.2	0.4	0.6	0.7960	0.9280	0.7960
$t = 0.003$	0	0.2	0.4	0.5996	0.7896	0.9016	0.7896

Considerando que a solução analítica do problema, sujeita àquelas CI e CF, seja dada por

$$u(x, t) = \frac{8}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \left( \sin \frac{n \pi x}{2} \right) (\sin n \pi x) e^{-n^2 \pi^2 t}$$

podemos calcular os erros correspondentes à solução numérica para alguns pontos, conforme mostra a tabela (6).

Tabela 6: Erros na aproximação numérica da equação (28), com  $r = 1/10$ .

$x = 0.3$	$t$	numérica	analítica	erro (%)
	0.005	0.5971	0.5966	0.08
	0.01	0.5822	0.5799	0.4
	0.02	0.5373	0.5334	0.7
	0.1	0.2472	0.2444	1.1
<hr/>				
$x = 0.5$				
	0.005	0.8597	0.8404	2.3
	0.01	0.7867	0.7743	1.6
	0.02	0.6891	0.6809	1.2
	0.1	0.3056	0.3021	1.2

Nota-se que, para  $x = 0.3$ , o erro é pequeno; já para  $x = 0.5$ , o erro é maior, particularmente devido às CI, pois

$$\frac{\partial u}{\partial x} \Big|_{x=\frac{1}{2}^-} = 2 \neq -2 = \frac{\partial u}{\partial x} \Big|_{x=\frac{1}{2}^+}.$$

No entanto, à medida que o processo de integração prossegue, o erro diminui.

Richtmeyer e Morton (1967) mostraram que, para o esquema em diferenças-finitas utilizado no exemplo (25), se a função inicial e suas  $p - 1$  primeiras derivadas são contínuas, e a  $p$ -ésima derivada é descontínua de forma ordinária (i.e., varia em saltos finitos); então, para  $k$  pequeno, o erro é menor ou igual a

$$k^{\frac{p+2}{p+4}}$$

No exemplo (25),  $p = 1$ , então o erro é  $k^{\frac{3}{5}} = 0.016$ ; comparando com os valores presentes na tabela (6), vemos que essa condição é satisfeita.

**Exemplo 26:** Para o mesmo problema do exemplo (25), utilize  $h = 1/10$  e  $k = 5/1000$ .

**Solução:** Nesse caso,  $r = 1/2$ . A equação (26) reduz-se a

$$u_{i+1,j} = \frac{1}{2}(u_{i,j-1} + u_{i,j+1}) \quad (29)$$

Alguns dos valores calculados são mostrados na tabela (7); os erros em  $x = 0.3$  são mostrados na tabela (8).

Tabela 7: *Alguns valores para a equação (29), com  $r = 1/2$ .*

$x$	0.1	0.2	0.3	0.4	0.4	0.5	0.6
$t = 0$	0	0.2	0.4	0.6	0.8	1.0	0.8
$t = 0.005$	0	0.2	0.4	0.6	0.8	0.8	0.8
$t = 0.010$	0	0.2	0.4	0.6	0.7	0.8	0.7
$t = 0.015$	0	0.2	0.4	0.55	0.7	0.7	0.7

Tabela 8: *Erros na aproximação numérica da equação (29), com  $r = 1/2$ .*

$x = 0.3$	$t$	numérica	analítica	erro (%)
	0.005	0.6	0.5966	0.57
	0.01	0.6	0.5799	3.5
	0.02	0.55	0.5334	3.1
	0.1	0.2484	0.2444	1.6

Comparando com os dados na tabela (6), vemos que os erros aumentaram; isso nos leva a supor que o valor de  $k$  está relacionado ao erro existente na solução.

**Exemplo 27:** Para o exemplo (25), utilize  $h = 1/10$  e  $k = 1/100$ .

**Solução:** Nesse caso,  $r = 1$ . A equação (26) reduz-se a

$$u_{i+1,j} = u_{i,j-1} - u_{i,j} + u_{i,j+1} \quad (30)$$

A tabela (9) mostra alguns dos valores calculados os quais, obviamente, não representam o comportamento físico esperado.

Tabela 9: Alguns valores para a equação (30), com  $r = 1$ .

$x$	0.1	0.2	0.3	0.4	0.4	0.5	0.6
$t = 0$	0	0.2	0.4	0.6	0.8	1.0	0.8
$t = 0.01$	0	0.2	0.4	0.6	0.8	0.6	0.8
$t = 0.02$	0	0.2	0.4	0.6	0.4	1.0	0.4
$t = 0.03$	0	0.2	0.4	0.2	1.2	-0.2	1.2
$t = 0.04$	0	0.2	0.0	1.4	-1.2	2.6	-1.2

O método explícito expresso pela equação (30) é computacionalmente simples, porém ele apresenta uma restrição. O passo de integração temporal,  $k$ , deve ser necessariamente pequeno, uma vez que o processo de integração só é válido quando

$$k \geq \frac{h^2}{2} \quad (31)$$

Além disso, o particionamento em  $x$  deve ser suficientemente grande, levando a um  $h$  pequeno, para que as aproximações das derivadas espaciais, usando diferenças-finitas, sejam aceitáveis. No exemplo (27) a condição (31) foi violada, levando a resultados incorretos (do ponto de vista físico).

A fim de remover essa restrição no passo de integração temporal, devemos recorrer a um outro método, conforme descrito a seguir.

### 7.1.2 Método de Crank-Nicolson

Em 1947, Crank e Nicolson propuseram um método válido para quaisquer valores finitos de  $r$ . Eles consideraram que a EDP (26) é satisfeita no ponto médio  $(i + \frac{1}{2}k, jh)$ , ou seja,

$$\left(\frac{\partial u}{\partial t}\right)_{i+\frac{1}{2},j} = \left(\frac{\partial^2 u}{\partial x^2}\right)_{i+\frac{1}{2},j}$$

e, substituindo estas derivadas pelas expressões em diferenças finitas já vistas, temos

$$\frac{u_{i+1,j} - u_{i,j}}{k} = \frac{1}{2} \left( \frac{u_{i+1,j-1} - 2u_{i+1,j} + u_{i+1,j+1}}{h^2} + \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h^2} \right)$$

ou

$$-ru_{i+1,j-1} + (2 + 2r)u_{i+1,j} - ru_{i+1,j+1} = ru_{i,j-1} + (2 - 2r)u_{i,j} + ru_{i,j+1} \quad (32)$$

onde  $r = k/h^2$ . Na equação (32) temos três termos conhecidos e três a determinar, os quais referem-se ao tempo  $i + 1$ ; daí, o método de Crank-Nicolson é dito ser *implícito*.

Se cada linha da malha tiver  $n$  pontos, então (32) é um sistema de equações lineares de  $n$  equações a  $n$  variáveis, sendo o termo independente de cada sistema, onde calculamos  $u$  no tempo  $i + 1$ , composto pelos valores de  $u$  no tempo  $i$ . Note que a resolução do sistema de equações deve ser efetuada *a cada passo de integração no tempo*. O exemplo a seguir mostra como utilizar o método.

**Exemplo 28:** Resolva o exemplo (26) com o método de Crank-Nicolson, usando  $h = 1/10$ .

**Solução:** Ao selecionarmos  $k$ , devemos verificar que uma escolha cuidadosa de  $r$  permite simplificar a equação (32); p. ex., com  $r = 1$ , o termo  $u_{i,j}$  é removido. Para esse valor de  $r$ , temos  $k = 1/100$ . Então, a equação governante, nesse caso, é

$$-u_{i+1,j-1} + 4u_{i+1,j} - u_{i+1,j+1} = u_{i,j-1} + u_{i,j+1}$$



O principal problema, ao se resolver uma equação diferencial parcial, é o tratamento adequado das condições de contorno. Condições simples, como as de Dirichlet ( $u = 0$ ), são facilmente tratadas. Por exemplo, considere a aproximação central por diferenças finitas da derivada  $\partial u/\partial x$ ,

$$\frac{u_{j+1} - u_{j-1}}{2h}.$$

Se a condição é  $u = 0$  em  $x = 0$ , então para o ponto na malha em  $j = 0$ ,  $u_{j=0} = 0$ . Daí, a aproximação para a derivada no primeiro ponto,  $x = h$ , i.e.  $j = 1$ , passa a ser

$$\frac{u_1 - 0}{2h}$$

O mesmo procedimento é válido para outras aproximações em diferenças finitas.

Condições de fronteira envolvendo *derivadas* ocorrem com frequência para indicar a inexistência de fluxo numa região em estudo, ou a taxa com a qual uma dada quantidade varia. Por exemplo, quando a superfície de um objeto condutor de calor é termicamente isolado, podemos dizer que

$$\frac{\partial U}{\partial n} = 0$$

i.e., a derivada de  $U$  na direção *normal* à superfície (indicada por  $n$ ) é nula.

Um outro exemplo refere-se à transmissão de calor. A taxa com que o calor é transferido por radiação de uma superfície quente para o meio ao seu redor, a uma temperatura  $v$ , é normalmente assumida como proporcional a  $U - v$ . Como na teoria de condução de calor a premissa fundamental é de que o calor que flui através de uma superfície é igual a  $-\kappa \frac{\partial U}{\partial n}$  unidades de calor por unidade de tempo na direção normal à superfície, podemos escrever a condição de contorno como

$$-\kappa \frac{\partial U}{\partial n} = H(U - v)$$

onde  $\kappa$  é a constante de condutividade térmica do material e  $H$  a constante de transferência de calor da superfície. O sinal é tomado de forma a indicar que o fluxo de calor é na direção oposta àquela em que  $U$  cresce algebricamente. Podemos, então, simplificar a condição de fronteira para

$$\frac{\partial U}{\partial n} = -s(U - v), \quad s > 0.$$

Considere, então, uma barra fina, termicamente isolada, que irradia calor no seu extremo  $x = 0$ . A temperatura nesse extremo, em  $t = 0$ , é, desconhecida, e necessitamos de uma outra equação; essa pode ser a própria condição de fronteira, se usarmos uma aproximação frontal para a derivada em  $U$ , pois

$$-\frac{\partial U}{\partial x} = -s(U - v)$$

será aproximado por

$$\frac{u_{i,1} - u_{i,0}}{h} = s(u_{i,0} - v)$$

Um sinal negativo deve ser associado à derivada pois a normal, apontando para fora da barra, em  $x = 0$ , tem a direção  $-x$ .

Se desejarmos aproximar a derivada  $\partial U/\partial x$  por uma aproximação central, então é necessário introduzir uma célula fictícia  $u_{i,-1}$ , imaginando-se que a barra estende-se até  $-h$ , conforme a figura (9).

Podemos, então, escrever

$u_{-1}$	$u_0$	$u_1$	$u_2$	$\dots$	$u_{n-1}$	$u_n$	$u_{n+1}$	$u_{n+2}$
$-h$	$0$	$h$	$2h$		$(n-1)h$	$nh$	$(n+1)h$	$(n+2)h$

Figura 9: Malha com pontos fictícios.

$$\frac{u_{i,1} - u_{i,-1}}{2h} = s(u_{i,0} - v)$$

porém, agora, devemos eliminar  $u_{i,-1}$ , já que ele é fictício, conforme mostra o exemplo a seguir.

**Exemplo 29:** Seja uma barra fina, termicamente isolada, a uma temperatura diferente da ambiente, a qual irradia calor através das suas duas extremidades. A EDP é

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2}$$

sujeita às condições:

$$\begin{aligned}
 CI & : U = 1, \quad 0 \leq x \leq 1, \quad t = 0 \\
 CF & : \begin{cases} \frac{\partial U}{\partial x} = U, & x = 0, \quad t > 0 \\ \frac{\partial U}{\partial x} = -U, & x = 1, \quad t > 0 \end{cases}
 \end{aligned}$$

Obtenha as equações governantes de acordo com o método explícito.

**Solução:** Usando a equação (27), podemos escrever, em  $x = 0$ ,

$$u_{i+1,0} = u_{i,0} + r(u_{i,-1} - 2u_{i,0} + u_{i,1})$$

A CF em  $x = 0$ , usando uma aproximação central, é

$$\frac{u_{i,1} - u_{i,-1}}{2h} = u_{i,0}$$

de onde  $u_{i,-1} = u_{i,1} - 2hu_{i,0}$ ; substituindo na equação para  $u_{i,0}$ , temos

$$\begin{aligned}
 u_{i+1,0} & = u_{i,0} + r(u_{i,1} - 2hu_{i,0} - 2u_{i,0} + u_{i,1}) = \\
 & = u_{i,0} + 2r(u_{i,1} - (1+h)u_{i,0})
 \end{aligned}$$

De forma similar, em  $x = 1$ , temos como CF

$$\frac{u_{i,n+2} - u_{i,n}}{2h} = u_{i,n+1}$$

e, como a equação (27) é expressa no ponto  $x = 1$  como

$$u_{i+1,n+1} = u_{i,n+1} + r(u_{i,n} - 2u_{i,n+1} + u_{i,n+2})$$

podemos eliminar  $u_{i,n+2}$ , tal que

$$u_{i+1,n+1} = u_{i,n+1} + 2r(u_{i,n} - (1+h)u_{i,n+1})$$

Para resolver esse problema, então, vamos utilizar três equações: uma para o ponto  $u_{i+1,0}$ , outra para o ponto  $u_{i+1,n+1}$  e a equação (27) para os pontos  $u_{i+1,j}$ ,  $1 \leq j \leq n$ .

## 7.2 Equações diferenciais parciais elípticas

As EDPs *elípticas* são normalmente relacionadas a problemas em equilíbrio e as suas soluções correspondem à um máximo ou mínimo da integral que representa a energia do sistema. As mais conhecidas e importantes são as equações de Poisson (23) e de Laplace (24). Suas aplicações são as mais variadas: p.ex., a equação de Poisson pode representar o movimento de um fluido viscoso incompressível, a baixa velocidade; a equação de Laplace é empregada para descrever o potencial eletromagnético, dentre outras.

O intervalo de integração de uma EDP elíptica é sempre uma área  $S$  cercada por uma curva *fechada*  $C$ . As condições de contorno especificam ou o valor da função ou de sua derivada em cada ponto de  $C$ ; é comum, também, que em certas regiões de  $C$  seja especificado o valor da função e, noutras, o da sua derivada.

A solução de uma EDP elíptica através de sua discretização em diferenças finitas leva à solução de um sistema de equações lineares, tipicamente grande e esparso, para os quais os métodos iterativos são particularmente indicados.

Considere, então, a equação

$$\nabla^2 u = f(x, y) \quad (34)$$

na região  $R = [0, 1] \times [0, 1]$  (i.e., o quadrado unitário), sujeita às condições de fronteira  $u = 0$  em  $\Omega R$ .

Discretizamos, então, a região com uma malha cartesiana com espaçamento  $h = 1/m$ , idêntico nas direções  $x$  e  $y$ , conforme o diagrama da figura (10).

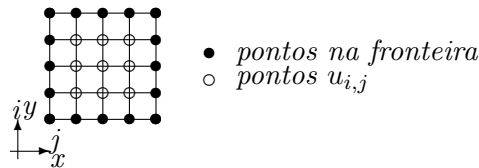


Figura 10: Malha (de pontos) para uma EDP elíptica.

Note que temos  $n = m - 1$  pontos ao longo de cada direção, totalizando  $n^2$  pontos  $u_{i,j}$  onde obteremos a aproximação para a EDP. Aproximando as derivadas parciais de segunda ordem por diferenças finitas, temos

$$\frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h^2} + \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2} = f(jh, ih), \quad 0 < i, j < m \quad (35)$$

onde  $jh = x_j$  e  $ih = y_i$ . Usando um ordenamento natural dos pontos  $u_{i,j}$  na malha, ao longo das linhas verticais, obtemos um número  $k = (i - 1)n + j$ , tal que  $u_{i,j} \equiv u_k$ , conforme mostrado na figura (11). O ordenamento natural ao longo das linhas horizontais é equivalente: nesse caso, teríamos  $k = (j - 1)n + i$ .

Avaliando a equação (35) em cada um dos pontos  $(i, j)$ , obteremos um sistema de equações lineares, conforme mostra o exemplo a seguir:

**Exemplo 30:** Calcule  $u$  que satisfaça a EDP

$$-\nabla^2 u = -2(x^2 + y^2)$$

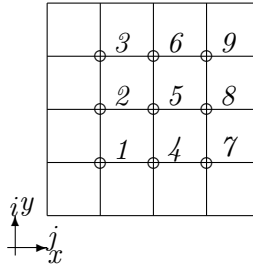


Figura 11: Ordenação natural dos pontos na malha.

na região  $R = [0, 1] \times [0, 1]$ , sujeita a  $u = 0$  nas linhas  $x = 0$  e  $y = 0$ ;  $u = y^2$  na linha  $x = 1$  e  $u = x^2$  na linha  $y = 1$ .

**Solução:** Apenas para fins de explanação, vamos utilizar uma malha quadrada com  $m = 4$ , mas cabe ressaltar que, tipicamente, devemos utilizar  $m > 100$ . Dessa forma, estaremos avaliando a EDP em  $n^2 = (m - 1)^2 = 9$  pontos, o que nos levará a um sistema de equações lineares de 9 equações a 9 incógnitas. O espaçamento é  $h = 1/4$ .

Escrevendo, então, a equação (35) para cada um dos pontos, utilizando o ordenamento natural conforme mostrado na figura (11), temos:

$$\begin{aligned}
 1 & : (0 + 2u_1 - u_4) + (0 + 2u_1 - u_2) = -2h^2(h^2 + h^2) \\
 2 & : (0 + 2u_2 - u_5) + (u_1 + 2u_2 - u_3) = -2h^2(h^2 + 4h^2) \\
 3 & : (0 + 2u_3 - u_6) + (u_2 + 2u_3 - h^2) = -2h^2(h^2 + 9h^2) \\
 4 & : (u_1 + 2u_4 - u_7) + (0 + 2u_4 - u_5) = -2h^2(4h^2 + h^2) \\
 5 & : (u_2 + 2u_5 - u_8) + (u_4 + 2u_5 - u_6) = -2h^2(4h^2 + 4h^2) \\
 6 & : (u_3 + 2u_6 - u_9) + (u_5 + 2u_6 - 4h^2) = -2h^2(4h^2 + 9h^2) \\
 7 & : (u_4 + 2u_7 - h^2) + (0 + 2u_7 - u_8) = -2h^2(9h^2 + h^2) \\
 8 & : (u_5 + 2u_8 - 4h^2) + (u_7 + 2u_8 - u_9) = -2h^2(9h^2 + 4h^2) \\
 9 & : (u_6 + 2u_9 - 9h^2) + (u_8 + 2u_9 - 9h^2) = -2h^2(9h^2 + 9h^2)
 \end{aligned}$$

onde os valores 0 correspondem àqueles pontos no contorno, pois  $u = 0$  é a condição de fronteira. Multiplicando por  $-1$  todas as equações e combinando os termos, obtemos o seguinte sistema de equações, na forma matricial:

$$\begin{bmatrix}
 4 & -1 & & & & & & & \\
 -1 & 4 & -1 & & & & & & \\
 & -1 & 4 & & & & & & \\
 -1 & & & 4 & -1 & & & & \\
 & -1 & & -1 & 4 & -1 & & & \\
 & & -1 & & -1 & 4 & & & \\
 & & & -1 & & & 4 & -1 & \\
 & & & & -1 & & -1 & 4 & -1 \\
 & & & & & -1 & & -4 & 4
 \end{bmatrix}
 \begin{bmatrix}
 u_1 \\
 u_2 \\
 u_3 \\
 u_4 \\
 u_5 \\
 u_6 \\
 u_7 \\
 u_8 \\
 u_9
 \end{bmatrix}
 =
 \begin{bmatrix}
 -0.0156 \\
 -0.0391 \\
 -0.0156 \\
 -0.0391 \\
 -0.0625 \\
 0.1484 \\
 -0.0156 \\
 0.1484 \\
 0.9144
 \end{bmatrix}
 \quad (36)$$

cuja solução é

$$u = (0.0039, 0.0156, 0.0352, 0.0156, 0.0625, 0.1406, 0.0352, 0.1406, 0.3164)$$

Note que, pela definição do problema, temos  $u = x^2y^2$ ; podemos confirmar que os valores obtidos como solução do sistema (36) satisfaz a EDP, como mostra a tabela (10).

Tabela 10: *Comparação entre os valores calculados para  $u$  pela solução direta do sistema e a solução exata.*

$k$	$x$	$y$	$u = x^2y^2$	$u_k$	erro relativo
1	0.25	0.25	0.0039	0.0039	$0.6661 \times 10^{-15}$
2	0.25	0.50	0.0156	0.0156	$0.4441 \times 10^{-15}$
3	0.25	0.75	0.0352	0.0352	0
4	0.50	0.25	0.0156	0.0156	$0.2220 \times 10^{-15}$
5	0.50	0.50	0.0625	0.0625	$0.1110 \times 10^{-15}$
6	0.50	0.75	0.1406	0.1406	0
7	0.75	0.25	0.0352	0.0352	0
8	0.75	0.50	0.1406	0.1406	$0.1974 \times 10^{-15}$
9	0.75	0.75	0.3164	0.3164	0

Como as matrizes que surgem da discretização em diferenças finitas de uma EDP são esparsas, é comum se utilizar métodos iterativos na solução do sistema de equações lineares, como mostra o exemplo a seguir.

**Exemplo 31:** Resolvendo o sistema (36) através do método dos Gradientes-Conjugados para tolerância de  $10^{-10}$ , obtém-se como solução

$$u = (0.0039, 0.0156, 0.0352, 0.0156, 0.0625, 0.1406, 0.0352, 0.1406, 0.3164)$$

A tabela (11) mostra que os erros relativos dessa solução também são aceitáveis.

Tabela 11: *Comparação entre os valores calculados para  $u$  pela solução iterativa do sistema e pelo método dos Gradientes-Conjugados com a solução exata.*

$k$	$x$	$y$	$u = x^2y^2$	$u_k$	erro relativo
1	0.25	0.25	0.0039	0.0039	$0.1332 \times 10^{-14}$
2	0.25	0.50	0.0156	0.0156	$0.0555 \times 10^{-14}$
3	0.25	0.75	0.0352	0.0352	$0.0197 \times 10^{-14}$
4	0.50	0.25	0.0156	0.0156	$0.0555 \times 10^{-14}$
5	0.50	0.50	0.0625	0.0625	$0.0222 \times 10^{-14}$
6	0.50	0.75	0.1406	0.1406	$0.0197 \times 10^{-14}$
7	0.75	0.25	0.0352	0.0352	$0.0395 \times 10^{-14}$
8	0.75	0.50	0.1406	0.1406	$0.0197 \times 10^{-14}$
9	0.75	0.75	0.3164	0.3164	$0.0175 \times 10^{-14}$

Assim como nas equações parabólicas, pode-se especificar condições de Neumann na fronteira.

## 8.1 Exercícios

**Exercício 13:** Considere a equação

$$\frac{\partial U}{\partial t} = \kappa \frac{\partial^2 U}{\partial x^2}, \quad \kappa < 1$$

Estabeleça a condição necessária para convergência de um método explícito e mostre o que acontece quando  $\kappa = 0.5$  e  $\kappa = 10^{-6}$ .

**Exercício 14:** Calcule a solução aproximada de

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \cos(y) \frac{\partial U}{\partial x} - \sin(x) \frac{\partial U}{\partial y} = 0$$

no retângulo unitário, com condições de Dirichlet na fronteira. Resolva o problema através de uma aproximação em diferenças-finitas, utilizando diferenças ascendentes para as derivadas de primeira ordem.

**Exercício 15:** Resolva o exercício (14) utilizando diferenças centrais para as derivadas de primeira ordem. Explique o que ocorre.

## Referências Bibliográficas

Conte, S.D., Boor, C. De, "Elementary Numerical Analysis", McGraw-Hill Book Company, London, 1981.

Mathews, J. H., "Numerical Methods for Computer Science, Engineering and Mathematics", Prentice-Hall, London, 1987.

Burden, R. L, Faires, JD, "Numerical Analysis", PWS Publishing, 1993.

Butcher, JC, "The Numerical Analysis of Ordinary Differential Equations - Runge-Kutta and General Linear Methods", John Wiley and Sons, 1987.

Ortega, JM., Poole, WG., "An Introduction to Numerical Methods for Differential Equations", Potman Publishing Inc, Massachusetts, 1981.