Métodos de Runge–Kutta, estabilidade

Apresentação 3 — Aula Teórica 3

Física Computacional

Departamento de Física
Universidade de Aveiro

18 ou 19 de fevereiro de 2019

Métodos de Runge–Kutta

Vimos que o erro local em cada passo do método de Euler é proporcional a h^2 e que o erro global é proporcional a h^1. Diz-se que

A ordem (de convergência) do método de Euler é 1.

É evidente que é conveniente usar métodos numéricos de maior ordem. Tem que ser encontrado um equilíbrio entre a complexidade e a ordem do método a usar para cada problema. O método tem também que ser estável para esse problema, mas vamos deixar essa discussão para mais tarde.

Uma alternativa evidente para obter um método de ordem 2, por exemplo, seria usar um método de série de Taylor, em que usaríamos mais um termo na expansão,

$$y_{k+1} = y_k + h f_1(t_k, y_k) + \frac{h^2}{2!} f_2(t_k, y_k),$$

onde

$$f_1(t, y) = \frac{\mathrm{d}y}{\mathrm{d}t} \qquad f_2(t, y) = \frac{\mathrm{d}^2 y}{\mathrm{d}t^2}$$

Métodos de Runge–Kutta

Não há nada de fundamentalmente errado com este método, mas as expressões das derivadas de maior ordem tornam-se frequentemente muito complexas e os cálculos podem ficar muito pesados.

Para introduzirmos os métodos de Runge–Kutta, vamos começar por relembrar o método de Crank–Nicolson:

$$y_{k+1} = y_k + \Big[f(t_k, y_k) + f(t_{k+1}, y_{k+1}) \Big] \frac{h}{2}$$

Foi mencionado na aula anterior que este é um método de ordem 2. Ora, esta ordem é alcançada sem o cálculo de derivadas de ordem superior a 1. Em termos de complexidade, o problema deste método é que ele é implícito: não podemos calcular explicitamente $f(t_{k+1}, y_{k+1})$, temos que resolver uma equação que, no caso geral, pode ser não linear.

Métodos de Runge–Kutta

Vamos tentar encontrar um método que use a primeira derivada calculada explicitamente em t_k e em t_{k+1}. Vamos chamar r_1 e r_2 a esses dois valores:

$$
\begin{aligned}
r_1 &= f(t_k, y_k) \\
r_2 &= f(t_{k+1}, y_k + r_1 h)
\end{aligned}
$$

O que fizemos foi usar o valor de r_1 para obter um estimativa intermédia de y em t_{k+1}, $y_k + r_1 h$, que depois é usada para calcular o valor de r_2. A nossa estimativa para o valor de y_{k+1} é então dada por

$$
y_{k+1} = y_k + \left[\frac{1}{2} r_1 + \frac{1}{2} r_2\right] h
$$

Parece demasiado fácil para funcionar e de facto não temos garantia nenhuma que o método (que admitimos ser convergente) seja de ordem 2, como nós queríamos.

Métodos de Runge–Kutta

Veremos mais tarde se a ordem é a desejada. Vamos introduzir de uma forma mais genérica os métodos de Runge–Kutta explícitos.

Em cada passo, a função $f(t, y)$ é calculada s vezes, obtendo-se sucessivamente $r_1, r_2, \ldots, r_s$.

No nosso exemplo, $s = 2$.

Os valores de $r_1, r_2, \ldots r_s$ são calculados para valores da variável independente t entre t_k e t_{k+1}, inclusive, dados por $t_k + c_1 h, t_k + c_2 h, \ldots, t_k + c_s h$.

No nosso exemplo, $c_1 = 0, c_2 = 1$.

Note que $c_1 = 0$ é inevitável para um método explícito, mas em relação aos outros c, só sabemos para já que não serão superiores a 1.

Métodos de Runge–Kutta

O valor de r_i, para cada $0 < i \leq s$, é calculado a partir dos $i - 1$ valores de r calculados previamente:

$$r_i = f\left(t_k + c_i h, y_k + \left[a_{i1} r_1 + \cdots + a_{i(i-1)} r_{i-1}\right] h\right)$$

A notação pode vonfundir um pouco. Para tentar simplificar, vejamos o exemplo para $i = 3$:

$$r_3 = f\left(t_k + c_3 h, y_k + \left[a_{31} r_1 + a_{32} r_2\right] h\right)$$

No nosso exemplo, só existe o elemento

$$a_{21} = 1$$

Métodos de Runge–Kutta

Finalmente, quando todos os s valores de r já foram calculados, a estimativa para y_{k+1} obtém-se a partir de uma média ponderada dos valores de r,

$$y_{k+1} = y_k + (b_1 r_1 + b_2 r_2 + \cdots + b_s r_s)\, h$$

No nosso exemplo,

$$b_1 = \frac{1}{2}$$
$$b_2 = \frac{1}{2}$$

É fácil de perceber que a soma dos pesos tem que dar 1: $b_1 + b_2 + \cdots + b_s = 1$.

Métodos de Runge–Kutta

Cada método de Runge–Kutta pode ser representado sinteticamente por um quadro de Butcher:

$$
\begin{array}{c|c}
\boldsymbol{c} & A \\
\hline
 & \boldsymbol{b}^{\mathrm{T}}
\end{array}
$$

Um método explícito com $s = 3$, por exemplo, terá um quadro genérico

$$
\begin{array}{c|ccc}
0 & & & \\
c_2 & a_{2_1} & & \\
c_3 & a_{3_1} & a_{3_2} & \\
\hline
 & b_1 & b_2 & b_3
\end{array}
$$

Métodos de Runge–Kutta

Para ordens p iguais ou inferiores a 4 é possível construir métodos de Runge–Kutta em que o número m de valores de r a determinar é igual à ordem do método, mas isso exige que os valores dos vetores $\boldsymbol{c}$ e $\boldsymbol{b}$ e dos elementos da matriz A obedeçam a certas relações cujo estudo não se enquadra nesta unidade curricular. Para construir um método de ordem 5, por exemplo, já é necessário calcular e usar pelo menos 6 valores diferentes de r.

Quando estiver a programar métodos de Runge–Kutta tem que ser muito cuidadoso. Pequenos enganos podem significar que está a escrever um método de ordem inferior à desejada e isso pode não ser observável nos gráficos dos resultados. Como exemplo, pode acontecer escrever r_2 num sítio onde deveria estar r_3 e o método passa a estar completamente errado, no sentido em que passa a ter uma ordem inferior à pretendida.

Métodos de Runge–Kutta

Antes de continuar, convém dizer que o método proposto no início da aula satisfaz as condições necessárias e é de facto um método de segunda ordem. O seu quadro de Butcher, como vimos, é

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

O quadro de Butcher do método de Euler é simplesmente

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

Métodos de Runge–Kutta

Antes de prosseguir, é uma boa ideia partir de um quadro de Butcher de um método de Runge–Kutta de terceira ordem e tentar compreendê-lo em termos gráficos.

$$\begin{array}{c|ccc}
0 & & & \\
\frac{1}{2} & \frac{1}{2} & & \\
1 & -1 & 2 & \\
\hline
 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6}
\end{array}$$

A figura que vair ser feita no quadro durante a aula não será acrescentada a este documento.

Um método de Runge–Kutta de segunda ordem

No Trabalho 3 vamos usar o método de Runge–Kutta de segunda ordem com o seguinte quadro de Butcher:

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

Descrito de outra forma,

$$
\begin{aligned}
r_1 &= f(t_k, y_k) \\
r_2 &= f(t_k + \frac{h}{2}, y_k + r_1 \frac{h}{2}) \\
y_{k+1} &= y_k + r_2 h
\end{aligned}
$$

Este método de Runge–Kutta de segunda ordem também é conhecido por método de Euler–Richardson.

RK de 2ª ordem aplicado a uma eq. de 2ª ordem

Para uma equação de 2ª ordem

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = f\left(t, y, \frac{\mathrm{d}y}{\mathrm{d}t}\right)$$

obtemos duas equações de 1ª ordem

$$\frac{\mathrm{d}y}{\mathrm{d}t} = v, \qquad \frac{\mathrm{d}v}{\mathrm{d}t} = f(t, y, v)$$

$$\begin{cases} r_{1v} = f(t_k, y_k, v_k) \\ r_{1y} = v_k \\ r_{2v} = f\left(t_k + \frac{h}{2}, y_k + r_{1y}\frac{h}{2}, v_k + r_{1v}\frac{h}{2}\right) \\ r_{2y} = v_k + r_{1v}\frac{h}{2} \end{cases}$$

RK de 2ª ordem aplicado a uma eq. de 2ª ordem

Note que tem que calcular primeiro r_{1v} e r_{1y} antes de calcular r_{2v} e r_{2y} (e assim sucessivamente, se o método for de ordem p superior a 2).

Por fim,

$$\begin{cases} v_{k+1} = v_k + r_{2v}h \\ y_{k+1} = y_k + r_{2y}h \end{cases}$$

Um método de Runge–Kutta de quarta ordem

No Trabalho 3 vamos usar o seguinte método de Runge–Kutta de quarta ordem:

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

$$
\begin{aligned}
r_1 &= f\,(t_k, y_k) \\
r_2 &= f\left(t_k + \frac{1}{2}h, y_k + r_1\frac{h}{2}\right) \\
r_3 &= f\left(t_k + \frac{1}{2}h, y_k + r_2\frac{h}{2}\right) \\
r_4 &= f\,(t_k + h, y_k + r_3 h) \\
y_{k+1} &= y_k + \frac{1}{6}\Big(r_1 + 2r_2 + 2r_3 + r_4\Big)h
\end{aligned}
$$

Este método terá que ser aplicado a um sistema de duas ODE de primeira ordem, de forma análoga à dos dois slides anteriores.

Métodos de Runge–Kutta de passo adaptativo

Os métodos de Runge–Kutta utilizados na prática são algo diferentes daqueles que apresentámos aqui, pois a diferença h entre valores consecutivos da variável independente não é mantida constante. Em zonas do domínio em que a função y tem um comportamento ”errático”, justifica-se o uso de um h bastante pequeno, mas quando a função tem uma variação previsível, pode-se aumentar bastante o passo.

Estes métodos exigem alguns cálculos extra em cada passo para comparar a exatidão com a exatidão desejada, mas a possibilidade que nos dão de reduzir o número de passos permite tornar os algoritmos mais rápidos por fatores que podem ser até de uma ou duas ordens de grandeza.

A função ode45 do MATLAB

Uma maneira de verificar se o passo deve ser aumentado ou diminuído é estimar $y(t+h)$ de duas maneiras: usando um passo $h/2$ e um passo h. A diferença entre as duas estimativas dá-nos uma indicação do erro cometido e permite-nos decidir se devemos aumentar, diminuir, ou manter o passo.

A função **ode45** do MATLAB, que vai usar nos trabalhos práticos, usa um método chamado de Dormand–Prince, com uma abordagem diferente. Em cada passo, é calculada uma estimativa através de um método de Runge–Kutta de quarta ordem e outra estimativa através de um método de quinta ordem. O erro do método de Runge–Kutta de quarta ordem é avaliado a partir da diferença entre as duas estimativas. Por outras palavras, o método é de facto de quarta ordem, mas corre-se em simultâneo um método de quinta ordem que nos permite ir alterando o passo de maneira a obter a exatidão desejada.

Estabilidade

Nas aulas anteriores, foi dito que, desde que satisfeitas algumas condições, como o método de Euler tem um erro local da ordem de h^2, o seu erro global é da ordem de h.

O argumento que foi usado para o método de Euler pode ser estendido para outros métodos: se o erro local é de ordem h^{p+1}, o erro global é da ordem de h^p.

No entanto, isto só acontece se, para o problema em estudo, e dado o passo h escolhido, o método for **estável**.

Nos slides seguintes vai ser feita uma discussão sucinta da estabilidade dos métodos até agora estudados para problemas de valor inicial.

Consistência e convergência

Vamos começar por introduzir algumas definições.

Consistência

Um método diz-se consistente com uma equação diferencial se o erro local vai para zero quando h tende para zero.

É óbvio que os métodos que estudámos são consistentes.

Convergência

Um método diz-se convergente se, quando aplicado a qualquer ODE com função f Lipschitz contínua, a diferença entre a solução numérica e a solução exata em cada ponto tende para zero quando h tende para zero.

Pode-se mostrar que todos os métodos que estudámos até agora são convergentes quer para problemas lineares, quer para problemas não lineares.

Estabilidade

O facto de um método ser convergente não nos garante que o erro global se comporte da forma que desejamos quando usamos um h maior que zero. Temos que verificar se o erro se mantém *bem comportado* para um dado valor de h. Para que isso aconteça, o método tem que ser **estável**[1].

A discussão que se segue é sempre baseada no problema de teste

$$\frac{dy}{dt} = \lambda y$$

cuja solução é

$$y(t) = y_0 e^{\lambda(t-t_0)}$$

que não cresce com t se $\lambda \leq 0$. Esta análise não se aplica diretamente a sistemas não lineares, mas pode ajudar a escolher valores de h para esses casos.

[1]Também se diz absolutamente estável para distinguir de outras formas de estabilidade.

Estabilidade

Iremos verificar mais tarde que os valores de λ podem ser complexos. Assim, consideremos um caso geral em que

$$\lambda = \lambda_r + i\lambda_i$$

e $\lambda_r \leq 0$ para que a solução da equação não cresça com t.

O método de Euler aplicado à equação diferencial com $f(y)$ é da forma

$$y_k = y_{k-1} + f(y_{k-1})h = y_{k-1} + \lambda y_{k-1}h = y_{k-1}(1+\lambda h) = y_0(1+\lambda h)^k$$

A solução numérica y_k não cresce se

$$|1+\lambda h| \leq 1 \quad \rightarrow \quad |1+\lambda h|^2 \leq 1 \quad \rightarrow \quad (1+\lambda_r h)^2 + \lambda_i^2 h^2 \leq 1.$$

Estabilidade do método de Euler

Analisemos 2 casos particulares:

- Se λ é puramente real, $\lambda \equiv \lambda_r \leq 0$. Para

$$h \leq \frac{2}{|\lambda_r|}$$

  o método de Euler é estável. Neste caso o método de Euler é condicionalmente estável.

- Se λ é puramente imaginário, $\lambda \equiv \lambda_i$, o método de Euler é instável.

Estabilidade do método de Euler implícito

Aplicando o método de Euler implícito à equação $dy/dt = \lambda y$, obtemos

$$y_k = y_{k-1} + \lambda y_k h \quad \rightarrow \quad y_k = (1-\lambda h)^{-1} y_{k-1} = y_0 (1-\lambda h)^{-k}$$

A solução numérica não cresce se

$$\frac{1}{|1-\lambda h|} \leq 1 \quad \rightarrow \quad \frac{1}{(1-\lambda_r h)^2 + \lambda_i^2 h^2} \leq 1$$

que é verdade sempre que $\lambda_r \leq 0$

O método é **incondicionalmente estável** para $\lambda_r \leq 0$.

Estabilidade do método de Crank-Nicolson

Aplicando o método de Crank-Nicolson à equação $dy/dt = \lambda y$, obtemos

$$y_k = \frac{1 + \frac{\lambda h}{2}}{1 - \frac{\lambda h}{2}} y_{k-1} \quad \rightarrow \quad y_k = y_0 \left(\frac{1 + \frac{\lambda h}{2}}{1 - \frac{\lambda h}{2}} \right)^k$$

A solução não cresce se

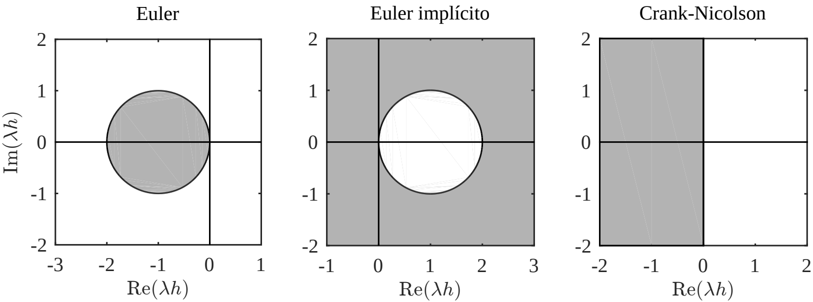
$$\left| \frac{1 + \frac{\lambda h}{2}}{1 - \frac{\lambda h}{2}} \right| \leq 1$$

ou seja

$$\frac{\left(1 + \frac{\lambda_r h}{2}\right)^2 + \frac{\lambda_i^2 h^2}{4}}{\left(1 - \frac{\lambda_r h}{2}\right)^2 + \frac{\lambda_i^2 h^2}{4}} \leq 1$$

Sempre que $\lambda_r \leq 0$, o numerador é menor ou igual ao denominador e a condição acima é verdadeira. O método é **incondicionalmente estável** para $\lambda_r \leq 0$.

Regiões de estabilidade (a sombreado)



Estabilidade do RK2

Para estudar a estabilidade, voltemos a considerar a equação diferencial linearizada:

$$\frac{dy}{dt} = \lambda y.$$

Aplicando o método de RK de 2ª ordem com $r_1 = \lambda y_{k-1}$, $r_2 = \lambda \left(y_{k-1} + \frac{h}{2} r_1 \right)$, vem

$$y_k = y_{k-1} + r_2 h = y_{k-1} \left(1 + \lambda h + \frac{\lambda^2 h^2}{2} \right)$$

$$y_k = y_0 \left(1 + \lambda h + \frac{\lambda^2 h^2}{2} \right)^k$$

A solução numérica y_k não cresce se

$$\left| 1 + \lambda h + \frac{\lambda^2 h^2}{2} \right|^2 \leq 1$$

Estabilidade do RK2

A solução numérica y_k não cresce se

$$\left|1 + \lambda h + \frac{\lambda^2 h^2}{2}\right|^2 \leq 1$$

Analisemos de novo os 2 casos particulares:

- Se λ é puramente real, $\lambda \equiv \lambda_r \leq 0$, para

$$h \leq \frac{2}{|\lambda_r|}$$

  o método é estável. Neste caso o método RK de 2ª ordem é condicionalmente estável.
- Se λ é puramente imaginário, $\lambda \equiv \lambda_i$, o método RK de 2ª ordem é instável.

Estabilidade do RK2

Estes resultados particulares são idênticos aos obtidos para o método de Euler. No entanto:

- A região de estabilidade no plano $(h\lambda_r, h\lambda_i)$ é maior no caso do RK de 2ª ordem.
- No caso dos movimentos oscilatórios (λ puramente imaginários), o aumento da amplitude é muito menor que num método de Euler.

Estabilidade do RK4

Voltemos a considerar a equação diferencial linear:

$$\frac{dy}{dt} = \lambda y.$$

Aplicando o método de RK de 4ª ordem descrito acima

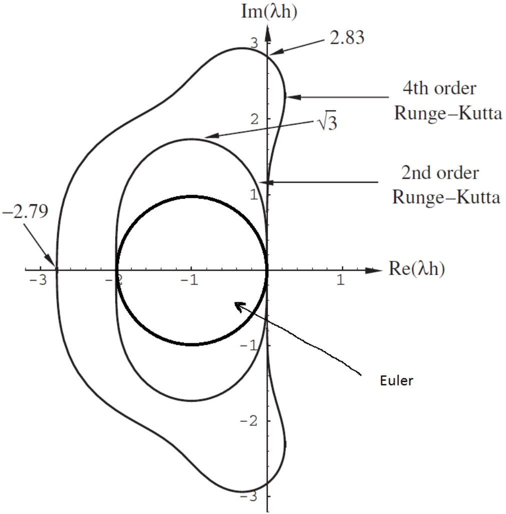
$$y_k = y_0 \left(1 + \lambda h + \frac{\lambda^2 h^2}{2} + \frac{\lambda^3 h^3}{6} + \frac{\lambda^4 h^4}{24} \right)^k$$

A solução numérica y_k não cresce se

$$\left| 1 + \lambda h + \frac{\lambda^2 h^2}{2} + \frac{\lambda^3 h^3}{6} + \frac{\lambda^4 h^4}{24} \right|^2 \leq 1$$

A figura seguinte mostra as regiões de estabilidade no plano $(h\lambda_r, h\lambda_i)$ de ambos os métodos de Runge–Kutta agora estudados. Na verdade, todos os métodos explícitos de Kunge–Kutta de uma dada ordem ordem $p \leq 4$, com $s = p$, têm as mesmas regiões de estabilidade.

Regiões de estabilidade de Euler e Runge–Kutta



Estabilidade para sistemas de ODE

Até agora, estudámos a estabilidade para uma equação diferencial escalar linear,

$$\frac{dy}{dt} = \lambda y.$$

Para um sistema linear de equações deste tipo,

$$\frac{dy}{dt} = Ay.$$

onde A é uma matriz quadrada, a solução pode ser escrita como

$$y(t) = y_0 \, \mathrm{e}^{A(t-t_0)}.$$

Uma condição necessária para a estabilidade é que o produto λh se encontre na região de estabilidade para todos os valores próprios λ de A.

Exemplo: oscilador harmónico

A equação para uma mola de constante K (oscilador harmónico)

$$F = -Ky \quad \rightarrow \quad m\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = -Ky$$

ou seja

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = -\frac{K}{m}y = -\omega^2 y, \qquad \omega^2 = \frac{K}{m}$$

que escrita como um sistema de duas equações diferenciais de 1ª ordem fica

$$\frac{\mathrm{d}y}{\mathrm{d}t} = v$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\omega^2 y$$

Exemplo: oscilador harmónico

O sistema linear pode ser escrito na forma matricial

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} y \\ v \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}\begin{pmatrix} y \\ v \end{pmatrix}$$

para o qual os λ são os valores próprios da matriz

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}$$

que são

$$\lambda = \pm i\omega$$

Como os valores próprios são imaginários puros, o método de Euler é, neste caso, instável.

O mesmo acontece para muitos outros movimentos oscilatórios.