Ecuaciones Diferenciales Ordinarias y Simulación con Matlab

L. Héctor Juárez Valencia y Mª Luisa Sandoval Solís

Departamento de Matemáticas, Universidad Autónoma Metropolitana-Iztapalapa,
D. F., México

5° Coloquio del Departamento de Matemáticas
Enero de 2012. Metepec, Atlixco, Puebla

Contenido

- Conceptos básicos
- Solución numérica de ecuaciones diferenciales ordinarias
- Teoría cualitativa de ecuaciones diferenciales
- Simulación de sistemas de ecuaciones diferenciales ordinarias

Definición

Un problema de valor inicial es una ecuación diferencial lineal de primer orden sujeta a una condición inicial, y se escribe como

$$\frac{dx(t)}{dt} = f(t,x),\ t \in [a,b]\,,\ x(a) = x_0. \tag{1}$$

Definición

Se dice que una función $f(t,x)$ satisface la condición de Lipschitz en la variable x en un conjunto $\mathbf{D} \subset \mathbb{R}^2$, si existe una constante $L > 0$ con la propiedad siguiente:

$$|f(t,x_1) - f(t,x_2)| \leq L\,|x_1 - x_2|\,,$$

siempre que (t,x_1), $(t,x_2) \in \mathbf{D}$. A L se le llama constante de Lipschitz para f.

TEOREMA

Supongamos que $\mathbf{D} = \{(t,x) |\, a \leq t \leq b, \ -\infty < x < \infty$ *y que* $f(t,x)$ *es continua en* $\mathbf{D}$. *Si* f *satisface una condición de Lipschitz en* $\mathbf{D}$ *en la variable* x, *entonces el problema de valor inicial (1) tiene una solución única* $x(t)$, *para* $a \leq t \leq b$.

TEOREMA

Si $\mathbf{D}$ *es convexo y* $\left|\frac{\partial f}{\partial x}(t,x)\right| \leq \mathbf{L}\, \forall (t,x) \in \mathbf{D}$, *entonces* f *es Lipschitz en* $\mathbf{D}$.

Definición

El problema de valor inicial (1) está bien planteado si:

1. *Tiene solución única;*
2. *Para cualquier* $\varepsilon > 0$, *existe una constante positiva* $k(\varepsilon)$ *tal que* $|\varepsilon_0| < \varepsilon$ *y* $\delta(t)$ *continua con* $|\delta(t)| < \varepsilon$ *en* $[a, b]$, *el problema*

$$\frac{dz}{dt} = f(t, z) + \delta(t), \qquad a \leq t \leq b, \;\; z(a) = x_0 + \varepsilon_0 \tag{2}$$

   tiene solución única $z(t)$, *además* $|z(t) - x(t)| < k(\varepsilon)\varepsilon, \forall t \in [a, b]$.

Teorema

Sea $\mathbf{D} = \{(t,x) | a \leq t \leq b, \ -\infty < x < \infty\}$. *Si* f *es continua y es Lipschitz en* $\mathbf{D}$*, entonces el problema de valor inicial (1) estará bien planteado.*

La solución aproximada del problema (1) consistirá de aproximaciones individuales a los valores de x en un conjunto discreto de tiempos, tales como

$$a = t_0 < t_1 < t_2 < \cdots < t_{N-1} < t_N = b.$$

Notación: $x(t_i)$ denota la solución real en $t = t_i$ y x_i la solución aproximada.

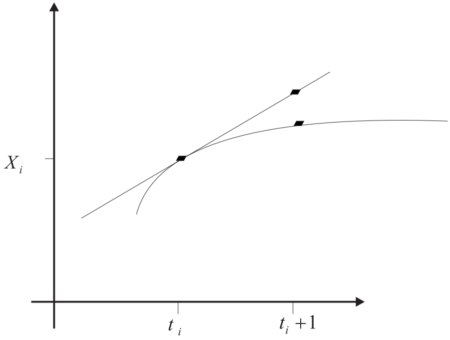
Definición

La forma general para un método de un paso explícito es

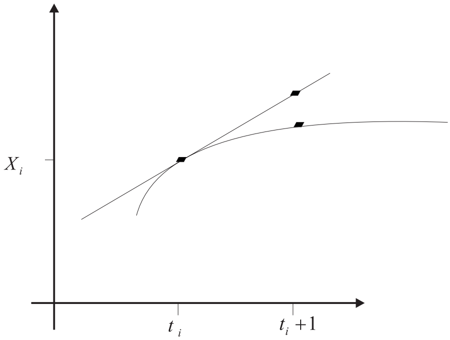
$$\frac{x_{i+1} - x_i}{h_i} = \phi(t_i, x_i; h_i)$$

donde $h_i = t_{i+1} - t_i$.

$$x_{i+1} = x_i + hf(t_i, x_i) \qquad \forall\, i = 0, 1, ..., N-1.$$



$$\overline{x}_{i+1} = x_i + hf(t_i, x_i)$$
$$x_{i+1} = x_i + \frac{h}{2}\left[f(t_i, x_i) + f(t_{i+1}, \overline{x}_{i+1}))\right] \qquad \forall i = 0, 1, \ldots, N-1.$$



Si $x(t)$ es solución exacta del problema (1) e infinitamente diferenciable, entonces

$$\begin{aligned} x(t) &= x(t_0) + x'(t_0)(t-t_0) + \frac{x''(t_0)(t-t_0)^2}{2} + \cdots + \frac{x^{(k)}(t_0)(t-t_0)^k}{k!} + \cdots \\ &= x_0 + f(t_0,x_0)(t-t_0) + \frac{f'(t_0,t_0)(t-t_0)^2}{2} + \cdots \\ &\qquad + \frac{f^{(k-1)}(t_0,x_0)(t-t_0)^k}{k!} + \cdots \end{aligned}$$

donde $t_0 \in [a,b]$ y $f'(t,x) = f_t + f_x x' = f_t + f_x f$; similarmente se calculan las derivadas de orden más alto, las cuales cada vez serán más complicadas.

Si $a = t_0$, $h = \frac{b-a}{N}$ y $t_i = t_0 + ih$ se tiene

$$\begin{aligned} x(t_1) = x(t_0) + f(t_0,x_0)h + \frac{f'(t_0,x_0)h^2}{2} + \cdots + \frac{f^{(k-1)}(x_0,t_0)h^k}{k!} \\ + \frac{h^{k+1}}{(k+1)!} f^{(k)}(\xi, x(\xi)) \end{aligned}$$

donde ξ está entre t_0 y t_1.

$$x(t_1) = x(t_0) + h\mathbf{T}_k(t_0, x_0) + \frac{h^{k+1}}{(k+1)!} f^{(k)}\left(\xi, x(\xi)\right)$$

donde

$$\mathbf{T}_k(t, x) = f(t, x) + \frac{1}{2!} f'(t, x) + \cdots \frac{h^{k-1}}{k!} f^{(k-1)}(t, x) \tag{3}$$

es el **Polinomio de Taylor de orden k.**

Error local de truncamiento de $O\left(h^k\right)$

Método de Taylor de orden k

Algoritmo

Escoger $h = \frac{b-a}{N}$ *y calcular* $t_i = a + ih, \forall i = 0, 1, ..., N$

La aproximación a la solución $x(t)$ *se construye generando los puntos de la sucesión* $\{x_{i+1}\}$ *recursivamente mediante*

$$x_{i+1} = x_i + h\mathbf{T}_k\left(t_i, x_i\right) \qquad \forall i = 0, 1, \ldots, N-1$$

donde $\mathbf{T}_k(t, x)$ *está definido en* (3)

Desventaja: cálcular las derivadas $f'(t,x)$, $f''(t,x), \ldots, f^{(k-1)}(t,x)$

El método de Euler modificado se puede reescribir como

$$x_{i+1} = x_i + h\left[\frac{1}{2}f(t_i, x_i) + \frac{1}{2}f(t_{i+h}, x_i + hf(t_i, x_i))\right]$$

Generalizando:

$$x_{i+1} = x_i + \left[w_1 hf(t_i, x_i) + w_2 hf(t_i + ah, x_i + bhf(t_i, x_i))\right]. \qquad (4)$$

Es decir,

$$x_{i+1} = x_i + [w_1 hk_1 + w_2 hk_2]$$

donde

$$\begin{aligned} k_1 &= f(t_i, x_i) \\ k_2 &= f(t_i + ah, x_i + bk_1) \end{aligned}$$

con w_1, w_2, a y b constantes a determinar.

Igualando la expresión (4) para x_{i+1} con el algoritmo de Taylor del mayor orden posible, por ejemplo de $O(h^3)$

$$x_{i+1} = x_i + hf + \frac{h^2}{2}[f_t + f_x f] + \frac{h^3}{4}\left[f_{tt} + 2f_{tx}f + f_x^2 f + f_t f_y\right] + O(h^4) \tag{5}$$

Esto es, primero se realiza el desarollo de Taylor para la función f

$$\begin{aligned} f(t_i + ah, x_i + bk_1) = f(t_i, x_i) &+ ahf_t(t_i, x_i) + bk_1 f_x(t_i, x_i) \\ &+ \frac{a^2h^2}{2} f_{tt}(t_i, x_i) + abk_1 hf_{tx}(t_i, x_i) \\ &+ \frac{b^2k_1^2}{2} f_{xx}(t_i, x_i) + O(h^3); \end{aligned}$$

y se sustituye la relación anterior en la ecuación (4)

$$\begin{aligned} x_{i+1} = x_i + w_1 hf(t_i, x_i) + w_2 h[f(t_i, x_i) + ahf_t(t_i, x_i) + bk_1 f_x(t_i, x_i) \\ + \frac{a^2h^2}{2} f_{tt}(t_i, x_i) + abk_1 hf_{tx}(t_i, x_i) + \frac{b^2k_1^2}{2} f_{xx}(t_i, x_i) + \cdots]. \end{aligned}$$

Ahora, se comparan los coeficientes de la igualdad anterior con la ecuación (5), y se obtiene el siguiente sistema

$$
\begin{aligned}
w_1 h + w_2 h &= h \\
a w_2 h^2 &= \frac{h^2}{2} \\
b k_1 h w_2 &= \frac{h^2}{2} f
\end{aligned}
$$

Si $w_1 = w_2 = \frac{1}{2}$ se recupera Euler modificado. De igual forma, si $w_1 = \frac{1}{4}$ y $w_2 = \frac{3}{4}$ se tiene el *método de Heun*

$$
x_{i+1} = x_i + \frac{h}{4}\left[f(t_i, x_i) + 3f\left(t_i + \frac{2}{3}h, x_i + \frac{2}{3}hf(t_i, x_i) \right) \right]
$$

Ambas técnicas, Euler modificado y Heun, son métodos de Runge-Kutta de segundo orden, es decir, el error local es proporcional a h^2.

$$x_{i+1} = x_i + \frac{h}{6}\left(k_1 + 2k_2 + 2k_3 + k_4\right) \quad \forall i = 0, 1, 2, \ldots, N-1$$

donde

$$\begin{aligned}
k_1 &= f\left(t_i, x_i\right)\\
k_2 &= f\left(t_i + \frac{h}{2}, x_i + \frac{1}{2}k_1\right)\\
k_3 &= f\left(t_i + \frac{h}{2}, x_i + \frac{1}{2}k_2\right)\\
k_4 &= f\left(t_{i+1}, x_i + k_3\right)
\end{aligned}$$

Error local de truncamiento: $O(h^4)$.

Método de un paso.

Desventaja: varias evaluaciones de la función por paso. Por ejemplo en RK4 se evalua a f cuatro veces.

Idea: generar dos métodos encajados

$$
\begin{aligned}
y(t+h) &\approx y(t) + h\,\phi(t,y;h), \qquad &\text{(orden } p)\\
y^*(t+h) &\approx y(t) + h\,\phi^*(t,y;h), \qquad &\text{(orden } p+1).
\end{aligned}
$$

Dado que

$$
\begin{aligned}
\phi(t,y;h) - \frac{y(t+h)-y(t)}{h} &= T(t,y;h) = \tau(t,y)h^p + \mathcal{O}(h^{p+1}),\\
\phi^*(t,y;h) - \frac{y(t+h)-y(t)}{h} &= T^*(t,y;h) = \mathcal{O}(h^{p+1}),
\end{aligned}
$$

al restar y dividir por h^p, se obtiene

$$
\frac{1}{h^p}[\phi(t,y;h) - \phi^*(t,y;h)] = \tau(t,y) + \mathcal{O}(h), \tag{6}
$$

de tal forma que

$$
r(t,y;h) = \frac{1}{h^p}[\phi(t,y;h) - \phi^*(t,y;h)] \tag{7}
$$

es una aproximación de orden $\mathcal{O}(h)$ de la función principal del error $\tau(t,y)$ para el método de orden p con $\phi(t,y;h)$.

Consideremos

$$
\begin{aligned}
y_{i+1} &= y_i + h\,\phi(t_i, y_i; h), \qquad \text{con} \quad T(t, y; h) \sim \mathcal{O}(h^2),\\
y^*_{i+1} &= y_i + h\,\phi^*(t_i, y_i; h), \quad \text{con} \quad T^*(t, y; h) \sim \mathcal{O}(h^3),
\end{aligned}
$$

donde

$$
\begin{aligned}
\phi(t, y; h) &= \alpha_1\, k_1(t, y; h) + \alpha_2\, k_2(t, y; h) + \alpha_3\, k_3(t, y; h),\\
\phi^*(t, y; h) &= \alpha_1^*\, k_1(t, y; h) + \alpha_2^*\, k_2(t, y; h) + \alpha_3^*\, k_3(t, y; h) + \alpha_4^*\, k_4(t, y; h).
\end{aligned}
$$

Nótese que en la construcción de ϕ^* se necesita un valor adicional k_4. Así,

$$
\begin{aligned}
k_1 &= f(t, y),\\
k_2 &= f(t + \mu_2 h, y + h\lambda_{21} k_1),\\
k_3 &= f(t + \mu_3 h, y + h[\lambda_{31} k_1 + \lambda_{32} k_2]),\\
k_4 &= f(t + \mu_4 h, y + h[\lambda_{41} k_1 + \lambda_{42} k_2 + \lambda_{43} k_3]),
\end{aligned}
$$

Las constantes μ_r, λ_{rj}, con $r = 2$, 3, 4, y $1 \leq j \leq r - 1$, así como α_r, α_r^*, $r = 1$, 2, 3, 4, se calculan de tal forma que los errores de truncamiento sean $\mathcal{O}(h^2)$ para $\phi(t, y; h)$ y $\mathcal{O}(h^3)$ para $\phi^*(t, y; h)$.

Además $\mu_r = \sum_{j=1}^{r-1} \lambda_{rj}$, $r = 2,\ 3,\ 4$

Resolver,

$$\alpha_1 + \alpha_2 + \alpha_3 = 1, \qquad \alpha_1^* + \alpha_2^* + \alpha_3^* + \alpha_4^* = 1,$$

$$\mu_2\alpha_2 + \mu_3\alpha_3 = \frac{1}{2}, \qquad \mu_2\alpha_2^* + \mu_3\alpha_3^* + \mu_4\alpha_4^* = \frac{1}{2},$$

$$\mu_2^2\alpha_2^* + \mu_3^2\alpha_3^* + \mu_4^2\alpha_4^* = \frac{1}{3}, \qquad \mu_2\lambda_{32}\alpha_3^* + (\mu_2\lambda_{42} + \mu_3\lambda_{43})\alpha_4^* = \frac{1}{6}.$$

Restricciones adicionales: k_4 del i–ésimo paso debe usarse como k_1 en el $(i+1)$–ésimo paso. Es decir,

$$f(t+h, y+h\Phi) \equiv f(t+\mu_4 h, y + h[\lambda_{41}k_1 + \lambda_{42}k_2 + \lambda_{43}k_3]),$$

entonces

$$\mu_4 = 1, \quad \lambda_{41} = \alpha_1, \quad \lambda_{42} = \alpha_2, \quad \lambda_{43} = \alpha_3.$$

Se pueden imponer restricciones adicionales para minimizar los coeficientes de la función principal del error.

Método de Runge–Kutta–Fehlberg de tercer orden

r	μ_r	λ_{r1}	λ_{r2}	λ_{r3}	α_r	α_r^*
1	0	0	$--$	$--$	$214/891$	$533/2106$
2	$1/4$	$1/4$	$--$	$--$	$1/33$	0
3	$27/40$	$-189/800$	$729/800$	$--$	$650/891$	$800/1053$
4	1	$214/891$	$1/33$	$650/891$	$--$	$-1/78$

Método *RKF-23*:

$$
\begin{aligned}
k_1 &= f(t_i, y_i),\\
k_2 &= f(t_i + \frac{h}{4}, y_i + \frac{h}{4}k_1),\\
k_3 &= f(t_i + \frac{27}{40}h, y_i - \frac{189}{800}hk_1 + \frac{729}{800}hk_2),\\
k_4 &= f(t_i + h, y_i + \frac{214}{819}hk_1 + \frac{1}{33}hk_2 + \frac{650}{891}hk_3),\\
y_{i+1} &= y_i + \frac{h}{891}(214k_1 + 27k_2 + 650k_3),\\
y_{i+1}^* &= y_i + \frac{h}{2106}(533k_1 + 1600k_3 - 27k_4).
\end{aligned}
$$

Observe que solo se requieren tres evaluaciones de f por paso:

$$k_4(t_i, y_i; h) = f(t_i + h, y_i + h\Phi(t_i, y_i; h)) = f(t_{i+1}, y_{i+1}) = k_1(t_{i+1}, y_{i+1}; h)$$

Si

$$y_{i+1} - y^*_{i+1} = h[\Phi(t_i, y_i; h)) - \Phi^*(t_i, y_i; h))] = h^3 r(t_i, y_i; h),$$

entonces

$$\|y_{i+1} - y^*_{i+1}\| = h^3 \|r(t_i, y_i; h)\| \leq \epsilon. \tag{8}$$

Para que el nuevo paso del tiempo h_{new} sea un error menor a ϵ, se debe satisfacer

$$\|r(t_{i+1}, y_{i+1}; h)\| \, h^3_{new} \leq \epsilon \,. \tag{9}$$

Pero, salvo errores de orden h,

$$r(t_{i+1}, y_{i+1}; h) \approx r(t_i, y_i; h) = \frac{\|y_{i+1} - y^*_{i+1}\|}{h^3}$$

por (8). Sustituyendo esta expresión en (9), obtenemos

$$\frac{\|y_{i+1} - y^*_{i+1}\|}{h^3} h^3_{new} \leq \epsilon$$

Por lo tanto, para garantizar que el error global sea menor a ϵ en cada paso debemos escoger

$$h_{new} \approx h \left(\frac{\epsilon}{\|y_{i+1} - y^*_{i+1}\|} \right)^{1/3}$$

Para el caso general,

$$h_{new} \approx h \left(\frac{\epsilon}{\|y_{i+1} - y^*_{i+1}\|} \right)^{1/(p+1)} \tag{10}$$

Se recomiendan hacer el siguiente ajuste

$$h_{new} \approx q\, h$$

con

$$q = \alpha \left(\frac{\epsilon h}{\|y_{i+1} - y^*_{i+1}\|} \right)^{1/p} \quad \text{y} \quad \alpha \approx 0.9$$

$$
\begin{aligned}
k_1 &= f(t_i,\, y_i)\\
k_2 &= f\left(t_i + \frac{h}{4}\,, y_i + \frac{h}{4} k_1\right)\\
k_3 &= f\left(t_i + \frac{3h}{8},\, y_i + \frac{3}{32} h k_1 + \frac{9}{32} h k_2\right)\\
k_4 &= f\left(t_i + \frac{12h}{13},\, y_i + h\left[\frac{1932}{2197} k_1 - \frac{7200}{2197} k_2 + \frac{7296}{2197} k_3\right]\right)\\
k_5 &= f\left(t_i + h,\, y_i + h\left[\frac{439}{216} k_1 - 8 k_2 + \frac{3680}{513} k_3 - \frac{845}{4104} k_4\right]\right)\\
k_6 &= f\left(t_i + \frac{h}{2},\, y_i + h\left[-\frac{8}{27} k_1 + 2 k_2 - \frac{3544}{2565} k_3 + \frac{1859}{4104} k_4 - \frac{11}{40} k_5\right]\right)\\
y_{i+1} &= y_i + h\left(\frac{25}{216} k_1 + \frac{1408}{2565} k_3 + \frac{2197}{4104} k_4 - \frac{1}{5} k_5\right)\\
y_{i+1}^{*} &= y_i + h\left(\frac{16}{135} k_1 + \frac{6656}{12825} k_3 + \frac{28561}{56430} k_4 - \frac{9}{50} k_5 + \frac{2}{55} k_6\right)
\end{aligned}
$$

El paso de tiempo nuevo en cada iteración se calcula por medio de

$$h_{new} = 0.84\, h \left(\frac{\epsilon h}{\|y_{i+1} - y^*_{i+1}\|} \right)^{1/4}$$

donde ϵ es la tolerancia deseada para el error global.

En ocasiones, en lugar de calcular y^*_{i+1} se calcula directamente la diferencia $e_{i+1} = \|y_{i+1} - y^*_{i+1}\|$. Por ejemplo, en el caso del método *RKF-45* (Runge–Kutta–Fehlberg de cuarto orden), se tiene

$$e_{i+1} = h \left\| \frac{1}{360} k_1 - \frac{128}{4275} k_3 - \frac{2197}{75240} k_4 + \frac{1}{50} k_5 + \frac{2}{55} k_6 \right\|$$

El problema de valor inicial de orden n

$$
\begin{gathered}
x^{(n)}(t) = f(t, x, x', x'', \cdots, x^{(n-1)}) \\
x(a) = \alpha_1, x'(a) = \alpha_2, x''(a) = \alpha_3, \cdots, x^{(n-1)}(a) = \alpha_n
\end{gathered}
\tag{11}
$$

Se puede transformar a un sistema de n ecuaciones diferenciales de primer orden realizando el siguiente cambio de variables:

$$
\begin{aligned}
x_1(t) &= x(t) \\
x_2(t) &= x'(t) \\
x_3(t) &= x''(t) \\
&\vdots \\
x_n(t) &= x^{(n-1)}(t)
\end{aligned}
$$

Luego,

$$
\begin{aligned}
x_1'(t) &= x'(t) = x_2(t)\\
x_2'(t) &= x''(t) = x_3(t)\\
x_3'(t) &= x'''(t) = x_4(t)\\
&\vdots\\
x_n'(t) &= x^{(n)}(t) = f(t, x_1, x_2, x_3, \cdots, x_n),
\end{aligned}
$$

junto con las condiciones iniciales asociadas

$$x_1(a) = \alpha_1, x_2(a) = \alpha_2, x_3(a) = \alpha_3, \ldots, x_n(a) = \alpha_n.$$

En forma de vector, el problema (11) se escribe como

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x})$$

donde

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} \quad \text{y} \quad \mathbf{f}(t, \mathbf{x}) = \begin{pmatrix} x_2(t) \\ x_3(t) \\ \vdots \\ f(t, x_1, \ldots, x_n). \end{pmatrix}$$