Elementos de Cálculo Numérico

Juan Pablo Pinasco (jpinasco@dm.uba.ar//gmail.com)

Departamento de Matemática e IMAS,
FCEyN, UBA - CONICET

2020

Parte I

Próxima clase

Hoy:

- Convergencia: 179-184

Próxima:

- Diferencias finitas
- PDEs

(no está en el apunte de DLR, sí en el Burden)

Parte II

Repaso

Hasta ahora, nos preocupó el error cometido al pasar de $x(t)$ a $x(t+h)$ usando Euler, Taylor o los métodos de R-K.

Sabemos que en Taylor tiramos un término que se acota $O(h^k)$ si usamos hasta la derivada $k-1$, pero no sabemos cómo afecta la solución, y cómo influye más adelante.

En cada paso se genera un nuevo error, agravado porque no evaluamos en el punto correcto, y, además, sólo aproximamos esa evaluación, así que es de esperar que el error se propague.

Esencialmente, podemos pensar que se van a acumular errores de orden $O(h^{k+1})$ al truncar el desarrollo de Taylor.

Pero como iteramos $1/h$ veces, yo esperaría -como mínimo- un error de orden

$$O(h^{k+1}) \times 1/h = O(h^k).$$

Sea $x_{j+1} = x_j + h\Phi(t_j, x_j, h)$

Def: Error de truncamiento local. Llamamos τ_j al error que se comete en el paso j al aproximar $x(jh + h)$ si reemplazamos por la solución exacta $x(t_j)$,

$$x(t_j + h) = x(t_j) + h\Phi(t_j, x(t_j), h) + h\tau_j.$$

Def: Decimos que un método es de orden k si el error de truncamiento local satisface

$$\tau = O(h^k).$$

Teorema

Tenemos los siguientes órdenes:

- *Euler:* $\tau = \frac{h}{2}x''(y) = O(h)$.
- *Taylor:* $\tau = \frac{h^k}{(k+1)!}x^{(k+1)}(y) = O(h^k)$.
- *Runge-Kutta:* $\tau = O(h^4)$ *(no se si lo hicimos, pero lo dejé de ejercicio).*

Def: Error global. Sea $x(t)$ la solución de $x'(t) = F(t, x(t))$, con $x(t_0) = x_0$. Llamamos error global al que se comete al aproximar $x(T)$ con nuestro método, es decir, $x(t_n) - x_n$.

Teorema

Sean $t_j = t_0 + jh$, con $h = (T - t_0)/n$. Para el método de un paso

$$x_{j+1} = x_j + h\Phi(t_j, x_j, h),$$

donde Φ es una función Lipschitz en x,

$$\Phi(t, x, h) - \Phi(t, y, h) \leq K|x - y|$$

para todo $x, y \in \mathbb{R}$, $t \in [t_0, T]$. Entonces

$$|x(T) - x_n| \leq \frac{\tau_{max}}{K}\left(e^{K(T-t_0)} - 1\right)$$

donde

$$\tau_{max} = \max_{1 \leq j \leq n} |\tau_j|$$

con τ_j el error de truncamiento local del método en el paso j.

6.- Demostración:

Sea $e_i = x(t_i) - x_i$ el error global hasta el paso i.

$$\begin{aligned} x_{i+1} &= x_i + h\Phi(t_i, x_i, h) \\ x(t_{i+1}) &= x(t_i) + h\Phi(t_i, x(t_i), h) + h\tau_i \end{aligned}$$

$$e_{i+1} = x(t_i) - x_i + h\Big(\Phi(t_i, x(t_i), h) - \Phi(t_i, x_i, h)\Big) + h\tau_i$$

$$e_{i+1} = e_i + h\Big(\Phi(t_i, x(t_i), h) - \Phi(t_i, x_i, h)\Big) + h\tau_i$$

Como Φ es Lipschitz,

$$h\Big(\Phi(t_i, x(t_i), h) - \Phi(t_i, x_i, h)\Big) \leq hK|x(t_i) - x_i| = hK|e_i|$$

Entonces,

$$|e_{i+1}| \leq |e_i| + hK|e_i| + h\tau_i = (1 + hK)|e_i| + h\tau_{max}$$

7.- Demostración:

Tenemos $|e_{i+1}| \leq (1+hK)|e_i| + h\tau_{max}$. Sea $A = (1+hK)$.

$$
\begin{aligned}
|e_1| &\leq h\tau_{max} \\
|e_2| &\leq A|e_1| + h\tau_{max} \leq Ah\tau_{max} + h\tau_{max} \\
&= (A+1)h\tau_{max} \\
|e_3| &\leq A|e_2| + h\tau_{max} \leq A\Big(Ah\tau_{max} + h\tau_{max}\Big) + h\tau_{max} \\
&= (A^2 + A + 1)h\tau_{max} \\
&\cdots \\
|e_n| &\leq (A^{n-1} + \cdots + A^2 + A + 1)h\tau_{max}
\end{aligned}
$$

$$
|e_n| \leq h\tau_{max} \sum_{j=1}^{n-1} A^j
$$

Sumando la serie geométrica,

$$
|e_n| \leq h\tau_{max} \frac{A^n - 1}{A - 1} = h\tau_{max} \frac{(1+hK)^n - 1}{hK} =
$$

Entonces

$$|e_n| \leq h\tau_{max}\frac{(1+Kh)^n - 1}{Kh} = \tau_{max}\frac{(1+Kh)^n - 1}{K}$$

Usamos $(1+a)^n \leq e^{na}$, con lo cual

$$|e_n| \leq \frac{\tau_{max}}{K}(e^{nKh} - 1) = \frac{\tau_{max}}{K}(e^{K(T-t_0)} - 1)$$

$$|x_n - x(T)| \leq \frac{\tau_{max}}{K}(e^{K(T-t_0)} - 1)$$

Y demostramos el teorema.

Ejercicio: (p/matemáticos) encuentre tres demostraciones (o más) de $(1+a)^n \leq e^{na}$.

9.- Local-global:

- La cuenta que hicimos funciona en muchos contextos.
- Va a reaparecer en los métodos para encontrar ceros de funciones (bisección, Newton-Raphson,...), cuando queramos resolver sistemas lineales iterando (Jacobi, Gauss-Seidel,...), etc.
- Tenemos un método iterativo, y una estimación para el error en cada paso.
- El propio método nos permite iterar la estimación del error, que en parte se puede amplificar, y se agregarán nuevos errores.
- Con suerte, la nueva recurrencia se resuelve de manera exacta o se acota.
- Con eso controlamos el error global.

La fórmula para el error global,

$$|x_n - x(T)| \leq \frac{\tau_{max}}{K}(e^{K(T-t_0)} - 1)$$

nos dice que, si $\tau_{max} = O(h)$, $O(h^2)$, etcétera, entonces cuando $h \to 0$, $e_n \to 0$.

Def: Decimos que un método es *convergente* si $x(n) \to x(T)$.

Es decir, si $\tau \to 0$ cuando $h \to$, tenemos que $x(n) \to x(T)$, probamos la convergencia del método.

- ¿Cuándo $\tau_{max} \to 0$ para $h \to 0$?
- ¿Cómo calculo la constante de Lipschitz?
- ¿Qué h tomar si quiero un error global 10^{-d}?
- ¿Euler, Taylor, R-K... cuál es mejor?

¿Cuándo vale que $\tau_{max} \to 0$ para $h \to 0$?

Teorema

Sea un método de un paso $x_{j+1} = x_j + h\Phi(t_j, x_j)$ *para la ecuación* $x' = F(t, x)$. *Entonces*

$$\lim_{h\to 0} \tau_{max} = 0 \iff \Phi(t, x, 0) = F(t, x).$$

Esta condición se llama **consistencia**. Observemos que

$$\lim_{h\to 0} \frac{x(t+h) - x(t)}{h} = x'(t) = f(t, x),$$

y en el método numérico,

$$\lim_{h\to 0} \frac{x(t+h) - x(t)}{h} = \Phi(t, x, 0)$$

La relación entre estabilidad, consistencia y convergencia es el teorema de Lax, pero queda para otro curso!

¿Cómo calculo la constante de Lipschitz?

Dada g, hallar la constante de Lipschitz puede ser complicado. Si es derivable, es el máximo de $|g'|$, porque el Teorema del Valor Medio nos dice

$$|g(x) - g(y)| = |g'(z)||x - y| \leq \max |g'||x - y|,$$

y si en algún x se alcanza el máximo, cuando $y \to x$,

$$\frac{|g(x) - g(y)|}{|x - y|} \to |g'(x)|.$$

14.- Veamos qué pasa con Euler modificado

$$x(t+h) = x(t) + hF\Big(t + \frac{h}{2}, x(t) + \frac{h}{2}F(t, x(t))\Big)$$

Tenemos que ver si $\Phi(t,x,h) = F\Big(t + \frac{h}{2}, x(t) + \frac{h}{2}F(t,x(t))\Big)$ es Lipschitz.

$$|\Phi(t,x,h) - \Phi(t,y,h)| = \Big|F\Big(t + \frac{h}{2}, x + \frac{h}{2}F(t,x))\Big) - F\Big(t + \frac{h}{2}, y + \frac{h}{2}F(t,y)\Big)\Big|$$

Como F es Lipschitz, $|F(t,a) - F(t,b)| \leq L|a-b|$. Entonces,

$$|\Phi(t,x,h) - \Phi(t,y,h)| \leq L\Big|x + \frac{h}{2}F(t,x) - y(t) - \frac{h}{2}F(t,y)\Big|$$

Por la desigualdad triangular

$$L\Big|x + \frac{h}{2}F(t,x) - y(t) - \frac{h}{2}F(t,y)\Big| \leq L|x-y| + L\Big|\frac{h}{2}F(t,x) - \frac{h}{2}F(t,y)\Big|$$

$$\leq L|x-y| + L^2\frac{h}{2}|x-y|$$

Entonces

$$|\Phi(t,x,h) - \Phi(t,y,h)| \leq (L + \frac{h}{2}L^2)|x-y|$$

¿No podía usar ese teorema, $\Phi(t,x,0) = F(t,x)$ y ahorrarme la cuenta?

Si hago $h = 0$ queda

$$\Phi(t,x,0) = F\Big(t + \frac{0}{2}, x(t) + \frac{0}{2}F(t,x(t))\Big) = F(t,x)$$

Pasa lo mismo con Euler común, Taylor, Heun, y RK 4:

$$\begin{aligned} k_1 &= = F(t,x) = F(t,x) \\ k_2 &= F(t+0/2, x+0/2k_1) = F(t,x) \\ k_3 &= F(t+0/2, x+0/2k_2) = F(t,x) \\ k_4 &= F(t+0, x+0k_3) = F(t,x) \end{aligned}$$

$$\Phi(t,x,0) = \frac{1}{6}\Big(k_1 + 2k_2 + 2k_3 + k_4\Big) = \frac{6}{6}F(t,x) = F(t,x)$$

¿Y para qué calculamos la constante de Lipschitz K de Φ?

Para responder la siguiente pregunta: ¿Qué h tomar si quiero un error global 10^{-d}?

Veamos un ejemplo para Euler. Tenemos

$$x'(t) = t\cos(t \cdot x), \qquad x(0) = 1, \quad x(\pi) = ???$$

Para Euler, $\Phi(t, x, h) = F(t, x) = t\cos(t \cdot x)$.

La constante de Lipschitz de Φ es la misma que la de F, y usando

$$\frac{d}{dx}\Big(t\cos(t \cdot x)\Big) = -t^2 \text{sen}(t \cdot x),$$

tenemos

$$|t\cos(t \cdot x) - t\cos(t \cdot y)| \leq t^2 \text{sen}(t \cdot z)|x - y| \leq \pi^2 \cdot 1 \cdot |x - y|.$$

Luego $K = L = \pi^2$.

Veamos ahora τ_{max}. Tenemos

$$x(t_j + h) = x(t_j) + hx'(t_j) + h \cdot \frac{h}{2} x''(s)$$

El error de truncado es $\frac{h}{2}x''(s)$, tenemos que acotarlo en $[0, \pi]$. Necesitamos x'',

$$x''(t) = \frac{d}{dt} F(t, x(t)) = \frac{d}{dt}\Big(t\cos(t \cdot x(t))\Big)$$

$$= \cos(t \cdot x(t)) + t\text{sen}(t \cdot x(t))\Big(x(t) + t \cdot x'(t)\Big)$$

Ahora, como $|t| \leq \pi$,

$$x''(s) \leq 1 + \pi \cdot 1 \cdot \Big(x(t) + \pi x'(t)\Big)$$

Podemos acotar $|x'(s)| = |s\cos(t \cdot x)| \leq \pi$.

Nos falta acotar $|x(t)|$.

Tenemos

$$x''(s) \leq 1 + \pi\left(x(t) + \pi^2\right)$$

Para acotar $x(s)$ sabemos que su derivada, $x'(t) = t\cos(tx)$, es menor o igual π, con lo cual tendríamos por el teorema de Valor Medio,

$$|x(t) - x(0)| = r\cos(r \cdot (x(r)) \cdot t$$

para todo $t \in [0, \pi]$, es decir

$$|x(t) - 1| \leq \pi^2$$

lo cual nos da

$$|x(t)| \leq |x(t) - 1 + 1| \leq |x(t) - 1| + 1 \leq \pi^2 + 1$$

Entonces

$$\begin{aligned} x''(s) \leq & 1 + \pi\left(\pi^2 + 1 + \pi^2\right) \\ = & 1 + \pi + 2\pi^3 \\ \leq & 68 \end{aligned}$$

Tenemos

$$K = L = \pi^2$$

$$\tau_{max} \leq \max_{t \in [0,\pi]} \frac{h}{2} x''(s) \leq 34h$$

Reemplazando,

$$\begin{aligned} |x_n - x(T)| \leq & \frac{\tau_{max}}{K}(e^{K(T-t_0)} - 1) \\ \leq & \frac{34h}{\pi^2}(e^{\pi^3} - 1) \end{aligned}$$

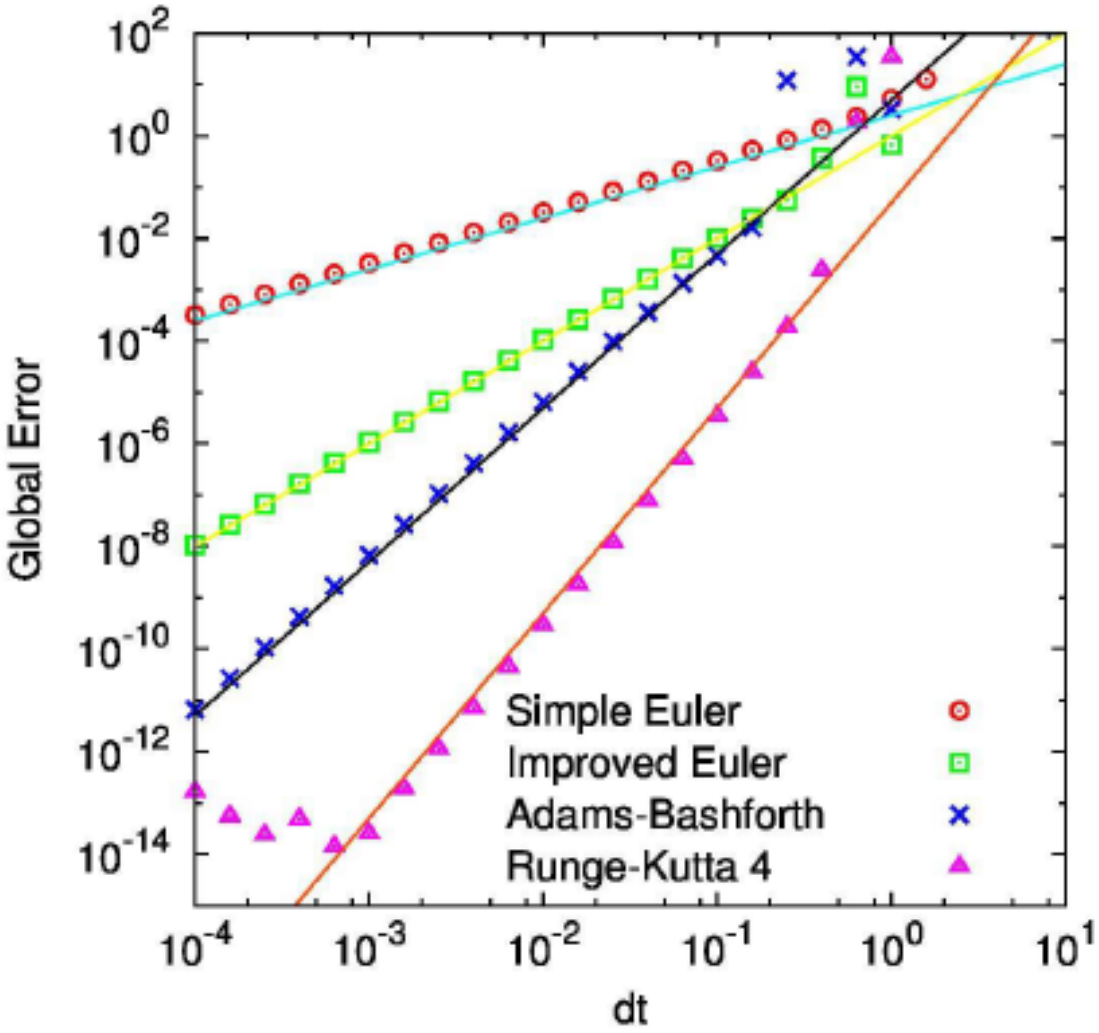
Pedimos

$$\frac{34h}{\pi^2}(e^{\pi^3} - 1) \leq 10^{-d}$$

y despejamos h.

El h que sale es imposible (culpa del e^{π^3}), aunque $d = 1$.

¿Qué método es mejor?



Queremos un error global $O(10^{-8})$

Euler: el error global es $O(h)$, así que necesitamos $h = 10^{-8}$, lo cual son $O(10^8)$ pasos del algoritmo, y evaluar $O(10^8)$ veces la F.

Heun: el error global es $O(h^2)$, así que necesitamos $h = 10^{-4}$, lo cual son $O(10^4)$ pasos del algoritmo, y evaluar $O(2 \times 10^4)$ veces la F.

RK 4: el error global es $O(h^4)$, así que necesitamos $h = 10^{-2}$, lo cual son $O(10^2)$ pasos del algoritmo, y evaluar $O(4 \times 10^2)$ veces la F.

Calcular la constante de Lipschitz K de la Φ nos permite determinar la constante en los $O(\cdot)$.