

Métodos Numéricos de las Ecuaciones Diferenciales

Rocío Piñera

9 de diciembre de 2021

Índice general

1. Ecuaciones Diferenciales. Introducción	5
1.1. Recordatorio: Problema de Cauchy	5
1.2. Discretización y Primeros Conceptos	7
1.2.1. Convergencia	8
1.2.2. Estimación del orden de Convergencia	9
2. Método de Euler Explícito	11
2.1. Consistencia	12
2.2. Estabilidad	14
2.3. Convergencia	16
3. Método de Euler implícito	19
3.1. Consistencia	21
3.2. Estabilidad	22
3.3. Convergencia	22
4. Método de Crank-Nicolson	23
4.1. Consistencia, Estabilidad y Convergencia	24
5. Métodos de Runge-Kutta	25
5.1. Métodos de Runge-Kutta Explícitos	26
5.1.1. Tablero de Butcher	27
5.1.2. Consistencia	29
5.1.3. Estabilidad	29
5.1.4. Convergencia	31
5.2. Métodos de Runge-Kutta Implícitos	33
6. Métodos Multipaso	35

Capítulo 1

Ecuaciones Diferenciales. Introducción

1.1. Recordatorio: Problema de Cauchy

Siendo I un intervalo de la recta real \mathbb{R} , se busca $u: I \rightarrow \mathbb{R}^m$, $u(t) \in C^1(I)$ tal que

$$(P.C.) \quad \begin{cases} u'(t) = f(t, u(t)), & t \in I \\ u(t_0) = u_0 \end{cases}$$

De manera explícita, en el caso de un sistema con $m > 1$, tenemos las ecuaciones

$$\begin{cases} u_1'(t) = f_1(t, u_1(t), u_2(t), \dots, u_m(t)) \\ u_2'(t) = f_2(t, u_1(t), u_2(t), \dots, u_m(t)) \\ \vdots \\ u_m'(t) = f_m(t, u_1(t), u_2(t), \dots, u_m(t)) \end{cases}$$

- Si f_j sólo depende sólo de (t, u_j) , el sistema se llama acoplado, en caso contrario, desacoplado.
- Suponemos que $f(t, u): I \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ es una función continua en ambas variables
- I siempre es un intervalo real de alguna de las formas

$$[t_0, t_0 + T] \quad [t_0, t_0 + T) \quad [t_0, +\infty)$$

- Cuando f no depende de la variable temporal, la ecuación es autónoma.

Método del factor integrante

Nos va a ser muy útil recordar esta técnica para resolver ecuaciones lineales, básicamente es la única que vamos a usar. Si queremos encontrar la solución de

$$u'(t) = \lambda u(t) + g(t)$$

ponemos

$$u'(t) - \lambda u(t) = g(t) \Rightarrow \text{multiplicando por } e^{-\lambda t} \quad (e^{-\lambda t} u(t))' = e^{-\lambda t} g(t)$$

Integrando llegamos a la solución fácilmente.

Campos de Velocidades

Una ecuación diferencial de primer orden nos indica en cada punto la pendiente de la curva solución que pasa por ese punto. La ecuación de la recta tangente en el punto (t_*, u_*) viene dada por

$$r(s) = u_* + (s - t_*) f(t_*, u_*)$$

Si en cada punto dibujamos el vector de componente

$$(1, f(t_*, u_*))$$

generamos el campo de velocidades que representa la edo.

Luego si consideramos un segmento pequeño en sentido positivo $t_1 > t_*$ y $u_1 > u_*$ entonces

$$\frac{u_1 - u_*}{t_1 - t_*} \approx f(t_*, u_*) \approx \tan(\alpha)$$

donde α es el ángulo formado por la recta tangente a la solución en (t_*, u_*) .

$$\text{Si } f(t_*, u_*) > 0 \Rightarrow \alpha \in \left(0, \frac{\pi}{2}\right) \quad \text{Si } f(t_*, u_*) < 0 \Rightarrow \alpha \in \left(-\frac{\pi}{2}, 0\right)$$

Orden de derivación mayor que 1

Las ecuaciones del tipo

$$u^{(q)} = \varphi(t, u, u', \dots, u^{(q-1)}) \quad (q > 1)$$

se pueden reescribir como sistemas de primer orden usando variables auxiliares, tantas como el orden de derivación. Ponemos $y_j = u^{(j-1)}$ para $j = 1, 2, 3, \dots, q$, y obtenemos las ecuaciones

$$\begin{cases} y_1' = u' \\ y_2' = u'' \\ \vdots \\ y_{q-1}' = u^{(q-1)} \\ y_q' = \varphi(t, y_1, \dots, y_{q-1}) \end{cases}$$

Existencia y Unicidad

Definición 1. Una función $f(t, u)$ cumple la condición de Lipschitz sobre su segunda variable cuando existe una constante $L_f > 0$ tal que

$$|f(t, x) - f(t, y)| \leq L_f |x - y|, \quad \forall x, y \in \mathbb{R}, \quad t \in [0, T]$$

Si se cumple la condición de Lipschitz sobre la segunda variable de $f(t, u)$ entonces la función es derivable con respecto a u con derivada acotada en casi todo punto. Mientras que en aquellos puntos donde no es derivable las derivadas laterales tienen saltos de discontinuidad finitos. Esto ocurre porque

$$\frac{|f(t, x) - f(t, y)|}{|x - y|} \leq L_f \quad \forall x, y \in \mathbb{R}, \quad t \in [0, T]$$

Ejemplo: Ser Lipschitz sobre la segunda variable es más que continua pero menos que derivable. Por ejemplo, una función en forma de dientes de sierra es Lipschitz pero no derivable en los picos o valles de la sierra.

□

Teorema 1. Dado el problema de Cauchy

$$(P.C.) \quad \begin{cases} u'(t) = f(t, u(t)), & t \in I \\ u(t_0) = u_0 \end{cases}$$

Si la función f cumple la condición de Lipschitz sobre su segunda variable entonces podemos garantizar la unicidad de la solución.

Ahora, bajo distintas hipótesis, sabemos si hay, además de unicidad, existencia:

1. $f((t, y)): \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ una función continua en ambas variables.
2. $f((t, y))$ cumple la condición de Lipschitz sobre la segunda variable
3. Controlamos el crecimiento de f con respecto la segunda variable

$$|f(t, y)| \leq C_f (1 + |y|), \quad \forall y \in \mathbb{R}, \quad t \in [0, T]$$

(1) + (2) implican la existencia y unicidad LOCAL en un subintervalo de $[0, T]$.

(1) + (2) + (3) implican la existencia y unicidad GLOBAL en $[0, T]$.

Estudio Cualitativo de las Soluciones

Dado el problema autónomo $z'(t) = f(z(t))$, si

$$f(c_*) = 0 \Rightarrow z(t) \equiv c_* \text{ solución constante}$$

que se denomina un **punto de equilibrio** del sistema.

Como $f(c_*) = 0$, su desarrollo de Taylor es

$$f(z) = (z - c_*) f'(c_*) + \dots$$

Entonces para z cercano a c_* , que supondremos, para simplificar $c_* = 0$,

$$f(z) \approx az, \quad \text{donde } a = f'(0)$$

Luego la dinámica cerca del punto de equilibrio se puede estudiar viendo el problema modelo:

$$z'(t) = az(t) \text{ con solución en un entorno del } 0 \quad z(t) = z_0 e^{at}$$

Si $a > 0$, todas las soluciones se alejan del punto de equilibrio a gran velocidad, por lo que denotamos al punto estacionario **inestable**. Por otro lado, si $a < 0$, todas las soluciones convergen a gran velocidad hacia cero y se tiene que el estado estacionario es **estable**.

1.2. Discretización y Primeros Conceptos

Se fija $T > 0$ y $t_0 \in \mathbb{R}$, trabajaremos en el intervalo $[t_0, t_0 + T]$. Denotamos una partición cualquiera como

$$\pi_h := \{t_n = t_0 + nh, \quad n = 0, 1, \dots, N\}_{h = \frac{T}{N}}$$

Y consideremos la familia de particiones $\{\pi_h\}_{h>0}$. Dado el problema de Cauchy,

$$\begin{cases} u'(t) = f(t, u(t)) & t \in [t_0, t_0 + T] \\ u(0) = u_0 \end{cases}$$

a través de un método numérico generaremos de forma iterativa valores u_n^h sobre cada π_h que aproximen a $u(t_n^h)$. Siempre trabajaremos con existencia y unicidad global de la solución, continua y derivable. Distinguimos:

$$\text{Métodos de un paso: } u_{n+1}^h(t_n) \quad \text{frente a} \quad \text{Métodos multipaso: } u_{n+1}^h(t_n, t_{n-1}, \dots, t_{n-k+1})$$

1.2.1. Convergencia

Definición 2. Diremos que una función $E(h)$ **converge a cero con orden p con respecto a h** si converge a cero de forma proporcional a como lo hace h^p , e.d.

$$E(h) = O(h^p)$$

o lo que es lo mismo, existe una constante $K > 0$ tal que

$$\lim_{h \rightarrow 0^+} \frac{E(h)}{h^p} = K \neq 0$$

Si $u(t)$ es la solución cualquiera de un P.C. y la aproximamos por los valores $\{u_n^h\}_n$ sobre la partición π_h a través de un método numérico, entendemos que el error que cometemos al aproximar la solución por los valores es una expresión de la forma

$$e_n^h = u(t_n^h) - u_n^h := \text{error global}$$

Y entendemos también que el método será convergente si cuando $h \rightarrow 0^+$, la estimación cumple

$$\max_{0 \leq n \leq N} |e_n^h| \rightarrow 0$$

concretamente, si $\max_{0 \leq n \leq N} |e_n^h| = O(h^p)$ entonces el método tiene **orden p de convergencia con respecto a h** . Pero, el anterior límite no es sencillo de estudiar por dos razones: en primer lugar, no solemos conocer la expresión del error global, sólo se tendrán los valores computacionales para casos concretos de h . Por otro lado, sabemos que

$$h \rightarrow 0^+ \Leftrightarrow N \rightarrow +\infty$$

y al mover h , las aproximaciones calculadas con distintos h van cambiando su localización. Si nos fijamos en un punto $t_* = t_0 + nh$ fijo que supondremos que está en todas las particiones π_h a partir de un cierto h . Uniendo ambas ideas, el límite anterior es:

$$\text{Simultáneo: } h \rightarrow 0^+ \text{ y } N \rightarrow +\infty \quad \text{Estacionario: } t_* = t_0 + nh \text{ permanece fijo}$$

Por lo que para estudiar la convergencia de un método numérico la entenderemos la convergencia como la *calidad* la aproximación calculada. Depende de dos nociones fundamentales que sí podemos estudiar:

1. **Consistencia:** El *tamaño* de los errores locales cometidos en cada paso, únicamente dependen del propio método numérico. Se obtienen suponiendo que

$$u(t_n^h) = u_n^h := \text{Hipótesis de Localización}$$

2. **Estabilidad:** La propagación de los errores que se cometen en cada paso de cálculo. Depende del propio método y de la longitud de h .

Teorema de Lax. Un método numérico es convergente si y sólo si es estable y consistente.

Normalmente trabajaremos con h fijo, por lo que podemos denotar $u_n \equiv u_n^h$

1.2.2. Estimación del orden de Convergencia

Antes hemos dicho que normalmente no conocemos la expresión $E(h)$, por ello tiene sentido estimar el orden de un método aun así. Podemos hacerlo de varias formas: Uso de Tablas, Recta Pendiente y una tercera forma cuando la solución es desconocida.

Uso de Tablas

Si el método es convergente de orden $p > 0$, el error es de la forma $E(h) \approx Ch^p$

$$\Rightarrow E\left(\frac{h}{2}\right) \approx C \frac{h^p}{2^p} \approx \frac{E(h)}{2^p} \Rightarrow \frac{E(h)}{E\left(\frac{h}{2}\right)} \approx 2^p \Rightarrow p \approx \log_2 \left(\frac{E(h)}{E\left(\frac{h}{2}\right)} \right) = \frac{\log \left(\frac{E(h)}{E\left(\frac{h}{2}\right)} \right)}{\log 2}$$

La idea ahora es combinar distintos valores de h para estimar p . Por ejemplo, supongamos que con el computador hemos obtenido la siguiente tabla de valores para las funciones $f(h)$ y $g(h)$.

h	$f(h)$	$g(h)$
0,0625	0,0118053	0,0803533
0,03125	0,0058242	0,0212917
0,015625	0,0028929	0,0059369
0,0078125	0,0014417	0,0015428

Podemos entonces ver que al dividir por 2 el valor de h el valor de $f(h)$ decae por un factor de 2 mientras que $g(h)$ decae por un factor $4 = 2^2$. Con esta información se puede deducir

$$f(h) = O(h) \quad \text{y} \quad g(h) = O(h^2)$$

Observación 1: Cuando los valores de los errores estén cerca del cero del computador ($\approx 10^{-16}$), este método deja de ser válido por los errores de redondeo o truncatura. Pero, si el valor de h no es lo suficientemente pequeño estos patrones tampoco se van a encontrar. Por lo tanto, siempre hay un $h_{\min} < h_{\max}$ tal que si $h \in (h_{\min}, h_{\max})$,

$$E(h) \approx O(h^p)$$

Recta Pendiente

Como antes, si el método es convergente de orden $p > 0$, el error es de la forma $E(h) \approx Ch^p$. Podemos entonces tomar distintos valores $\{h_j, E(h_j)\}_j$ y dibujar su gráfica para intentar deducir que se comporta como una curva $y = Cx^p$. PERO puede ser muy complicado en casos en los que el error pase de 10^{10} a 10^{-14} , por lo que es mejor usar logaritmos.

$$\text{Si } E(h) \approx Ch^p \Rightarrow \log(E(h)) \approx p \log(h) + \log C$$

\Rightarrow si dibujamos la recta sobre la que caen $\{\log(h_j), \log(E(h_j))\}_j \Rightarrow p$ coincide con su pendiente.

Observación 2: Teniendo en cuenta que

$$\log 10 \approx 2,3 \Rightarrow \log(10^{\pm M}) \approx \pm M \cdot 2,3$$

Podemos deducir la magnitud de los errores de los valores en el eje OY. Por ejemplo, como $\log(10^{-15}) \approx -35$, cuando en el eje OY alcanzamos el valor -35 , el error es del orden 10^{-15} .

Solución Desconocida

Como antes, si el método es convergente de orden $p > 0$, el error es de la forma $E(h) \approx Ch^p$. Veamos como aproximar p cuando la solución al problema de Cauchy $u(t)$ es desconocida. Al aplicarlo en π_h obtendríamos la aproximación u_h , que deberá ser de la forma

$$u^h = u + Ch^p + O(h^{p+1})$$

Repitiendo el mismo procedimiento para $\frac{h}{2}$ y $\frac{h}{4}$, tenemos:

$$\left\{ \begin{array}{l} u^h = u + Ch^p + O(h^{p+1}) \\ u^{\frac{h}{2}} = u + C\left(\frac{h}{2}\right)^p + O\left(\left(\frac{h}{2}\right)^{p+1}\right) \\ u^{\frac{h}{4}} = u + C\left(\frac{h}{4}\right)^p + O\left(\left(\frac{h}{4}\right)^{p+1}\right) \end{array} \right. \Rightarrow \frac{u^h - u^{\frac{h}{2}}}{u^{\frac{h}{2}} - u^{\frac{h}{4}}} = \dots = 2^p + O(h)$$

Entonces, como las particiones $\pi_{\frac{h}{2}}, \pi_{\frac{h}{4}}$ son más finas que π_h , tenemos que

$$\left\| \frac{u^h - u^{\frac{h}{2}}}{u^{\frac{h}{2}} - u^{\frac{h}{4}}} \right\| = 2^p + O(h) \text{ en } \pi_h \text{ y donde } \| \cdot \| \text{ denota la norma del máximo}$$

Capítulo 2

Método de Euler Explícito

Veamos una idea general primero. Buscamos la curva $u(t)$ que es solución del problema de Cauchy

$$\begin{cases} u'(t) = f(t, u(t)), & t \in [t_0, t_0 + T] \\ u(t_0) = u_0 \end{cases}$$

En primer lugar, la recta tangente a la curva solución u en el punto (t_0, u_0) es

$$z_0(s) = u_0 + (s - t_0) f(t_0, u_0)$$

Si la distancia $t_1 - t_0$ no es muy grande, el valor $u(t_1)$ estará cerca de $z_0(t_1) \equiv y_1$ e.d.

$$u(t_1) \approx u_0 + (t_1 - t_0) f(t_0, u_0)$$

Y ya tenemos dos puntos que son próximos a puntos de la curva

$$(t_0, u_0), \quad (t_1, u_1)$$

Repitiendo el proceso con la recta pendiente a la curva solución en el punto (t_1, u_1)

$$z_1(s) = u_1 + (s - t_1) f(t_1, u_1)$$

obtenemos $z_1(t_2) \equiv y_2$, la aproximación al punto $u(t_2)$. Ahora hemos generado tres puntos que sirven para construir una poligonal que va ajustándose a la solución buscada y que es solución al problema de Cauchy dado.

$$(t_0, u_0), \quad (t_1, u_1), \quad (t_2, u_2)$$

De forma general:

Esquema Euler Explícito

Dados (t_0, u_0) y la partición π_h obtener

$$\begin{aligned} t_{n+1} &= t_n + h \\ u_{n+1} &= u_n + hf(t_n, u_n), \quad n \in \{0, 1, \dots, N-1\}, \quad \text{donde } N = \frac{T}{h} \end{aligned}$$

Pseudocódigo Euler Explícito

Dados $t_0, T, N, y_0, f(t, u)$

Calculamos $h = \frac{T}{N}$

Para $n = 0$ hasta $N - 1$

 calcular $t_{n+1} = t_n + h$

 calcular $u_{n+1} = u_n + hf(t_n, u_n)$

Fin

Dibujar puntos (t_n, u_n)

Observación 3: El Desarrollo de Taylor nos proporciona una deducción alternativa del método. Supongamos que, como antes, tenemos el problema de Cauchy

$$\begin{cases} u'(t) = f(t, u(t)), & t \in [t_0, t_0 + T] \\ u(t_0) = u_0 \end{cases}$$

Y dada la solución $u(t)$, hacemos su desarrollo de Taylor en torno al punto t_n perteneciente a la partición π_h .

$$\begin{aligned} u(t_{n+1}) &= u(t_n) + hu'(t_n) + \frac{1}{2}h^2u''(\xi_n), \quad \text{donde } t_n < \xi_n < t_{n+1} \\ &= u(t_n) + hf(t_n, u(t_n)) + \frac{1}{2}h^2u''(\xi_n) \end{aligned}$$

Si el valor de h es lo suficientemente pequeño podemos despreciar el término que contiene h^2 ,

$$\Rightarrow u(t_{n+1}) \approx u(t_n) + hf(t_n, u(t_n))$$

lo que nos lleva al mismo esquema de cálculo que en la forma geométrica,

$$u_{n+1} = u_n + hf(t_n, u_n)$$

2.1. Consistencia

La consistencia del método de Euler nos la dará el tamaño de los errores locales cometidos en cada paso. Como siempre, partimos del problema de Cauchy,

$$\begin{cases} u'(t) = f(t, u(t)) & t \in [t_0, t_0 + T] \\ u(0) = u_0 \end{cases}$$

Consideremos $y(t)$ una función. De la propia definición de derivada, si tomamos un h suficientemente pequeño,

$$y'(t) \approx \frac{y(t+h) - y(t)}{h}$$

Si tomamos el operador diferencial $\rightarrow \mathcal{L}(y(t)) = y'(t) - f(t, y(t)) \Rightarrow y$ es solución si cumple

$$\mathcal{L}(y(t)) = 0$$

Usando el desarrollo de Taylor siguiente para $h > 0$

$$\begin{aligned} y(t+h) &= y(t) + y'(t)h + y''(\xi)\frac{h^2}{2} \Rightarrow \mathcal{L}(y(t)) = y'(t) - f(t, y(t)) \\ &= \frac{y(t+h) - y(t)}{h} - f(t, y(t)) + \frac{h}{2}y''(\xi) \end{aligned}$$

Definimos

$$\textbf{Discretización de la Ecuación Diferencial} \quad \mathcal{D}(y(t), h) = \frac{y(t+h) - y(t)}{h} - f(t, y(t))$$

$$\textbf{Error de truncatura} \quad \mathcal{T}(y(t), h) = -\frac{h}{2}y''(\xi)$$

(error que se comete al reemplazar la ecuación diferencial por su discretización)

$$\Rightarrow \mathcal{L}(y(t)) = \mathcal{D}(y(t), h) - \mathcal{T}(y(t), h) \Rightarrow \boxed{\mathcal{L}(y(t)) = 0 \Leftrightarrow \mathcal{D}(y(t), h) = \mathcal{T}(y(t), h)}$$

Fijemos ahora una partición $\pi_h = \{t_0, t_1, \dots, t_N\}$ de $[t_0, t_0 + T]$ y supongamos que $u(t)$ es solución al problema de Cauchy. Sabemos que Euler explícito nos proporciona los valores $\{u_0, u_1, \dots, u_N\}$ que aproximan a $\{u(t_0), u(t_1), \dots, u(t_N)\}$.

Definición 3. Llamamos **Hipótesis de Localización** al proceso de suponer que $u(t_n) = u_n$ e.d. le aplicamos el método al valor exacto $u(t_n)$ para calcular u_{n+1} . Para identificar cuando hemos supuesto la hipótesis de localización, denotamos

$$\tilde{u}_{n+1} = \text{método aplicado a } u(t_n)$$

Por otro lado,

$$\mathcal{T}(u(t), h) = \mathcal{D}(u(t), h) = \frac{u(t+h) - u(t)}{h} - f(t, u(t))$$

y para un $n \in \{0, 1, \dots, N-1\}$ cualquiera, suponiendo cierta la hipótesis de localización

$$u(t_{n+1}) = u(t_n) + hf(t_n, u(t_n)) + h\mathcal{T}(u(t_n), h) = \tilde{u}_{n+1} + h\mathcal{T}(u(t_n), h)$$

Definición 4. Luego, al aproximar la solución u en la partición π_h con Euler explícito, el **error de truncatura en el punto t_{n+1} con paso h** es

$$\frac{u(t_{n+1}) - \tilde{u}_{n+1}}{h} = \mathcal{T}(u(t_n), h)$$

Definición 5. El método numérico se dice **consistente** cuando el error de truncatura global sobre cualquier solución $u(t)$ del P.C. lo suficientemente derivable cumple

$$\lim_{h \rightarrow 0} \left[\max_{t_* \in \pi_h} \mathcal{T}(u(t_*), h) \right] = 0$$

Si además, para $p \geq 0$, $\mathcal{T}(u(t_*), h) = O(h^p) \Rightarrow$ el método tiene **orden de consistencia p**

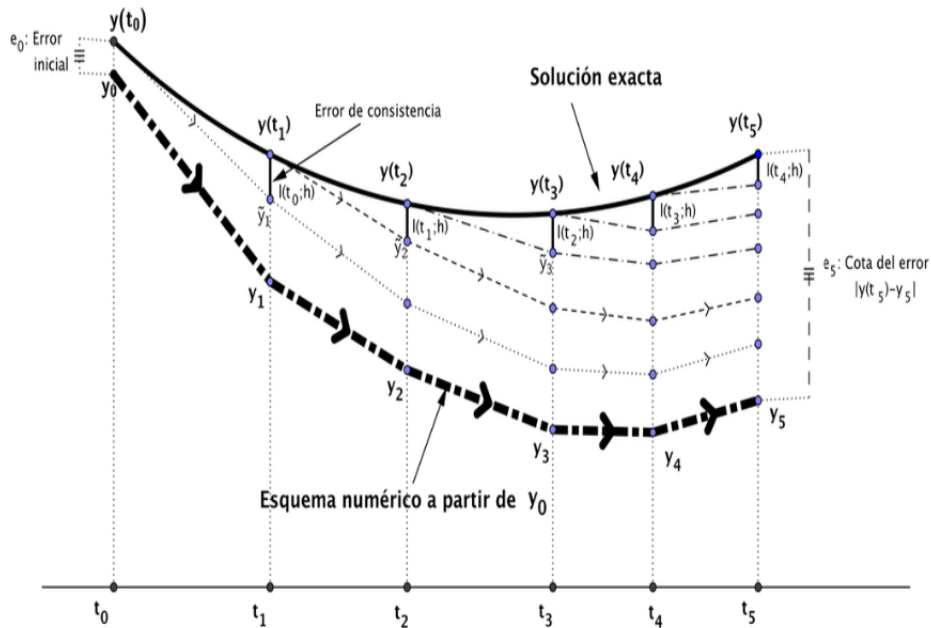
Definición 6. Llamamos **error local de consistencia en el punto t_{n+1} con paso h** para la solución $u(t)$ a

$$l(u(t_{n+1}), h) = h\mathcal{T}(u(t_{n+1}), h) = u(t_{n+1}) - \tilde{u}_{n+1}$$

es el error que se comete suponiendo cierta la hipótesis de localización. Se puede entender como lo que le falta al esquema para que la solución exacta lo cumpla.

Observación 4: $\mathcal{T}(u, h) = O(h^p) \Leftrightarrow l(u(t), h) = O(h^{p+1})$

Gráficamente:



Usando Taylor podemos obtener el orden de consistencia del método de Euler:
Como antes, si $u(t)$ es solución del P.C.

$$\mathcal{T}(u(t), h) = \mathcal{D}(u(t), h) = \frac{u(t+h) - u(t)}{h} - f(t, u(t))$$

entonces

$$u(t+h) = u(t) + hf(t, u(t)) + h\mathcal{T}(u(t), h) = u(t) + hf(t, u(t)) + l(u(t), h)$$

Con el desarrollo de Taylor tenemos

$$l(u(t), h) = \frac{h^2}{2} u''(\xi) \text{ para } \xi \in (t, t+h)$$

Pero u'' es la segunda derivada de la función que buscamos, por lo que es desconocida. Pero como es solución,

$$u'(t) = f(t, u(t)) \Rightarrow u''(t) = f_t(t, u(t)) + f_u(t, u(t)) u'(t)$$

Como $f \in C^1([t_0, t_0+T])$, si la gráfica de solución se encuentra en una región acotada $R \subset \mathbb{R}^2$ del plano, podemos obtener

$$M = \max_{t \in [t_0, t_0+T]} |f_t(t, u(t)) + f_u(t, u(t)) u'(t)| \Rightarrow \boxed{l(u(t), h) \leq \frac{h^2}{2} M = O(h^2)}$$

Y por tanto;

Teorema 2. El método de Euler explícito tiene un orden de consistencia 1.

2.2. Estabilidad

Datos de partida distintos en un esquema numérico de cálculo generarán valores finales distintos, observar estas diferencias se denomina el estudio de la estabilidad. Supongamos que damos un paso con dos valores iniciales distintos

$$\begin{cases} u_{n+1} = u_n + hf(t_n, u_n) \\ w_{n+1} = w_n + hf(t_n, w_n) \end{cases}$$

$$\text{buscamos acotar la diferencia } |u_{n+1} - w_{n+1}| = |u_n - w_n + h(f(t_n, u_n) - f(t_n, w_n))|$$

Definición 7. El **factor de amplificación** nos da la relación de la diferencia entre el paso n y el paso $n+1$. No hay un único factor de amplificación.

Por el Tma de los valores intermedios, $\exists \xi_n \in (u_n, w_n)$ tal que

$$u_{n+1} - w_{n+1} = (1 + h\partial_u f(t_n, \xi_n))(u_n - w_n) \Rightarrow |1 + h\partial_u f(t_n, \xi_n)| \text{ factor de amplificación}$$

Si $\partial_u f(t_n, \xi_n) < 0 \Rightarrow 1 + h\partial_u f(t_n, \xi_n) < 1 \Rightarrow$ tomando h tal que $|1 + h\partial_u f(t_n, \xi_n)| < 1 \Rightarrow$ la diferencia se reduce. Esta restricción sobre h se denomina **restricción de estabilidad**.

Definición 8. El esquema de Euler explícito tiene un carácter

Expansivo cuando $\partial_u f > 0$

Contractivo cuando $\partial_u f < 0$ (disipativo)

Desarrollando la diferencia de otra forma con Lipschitz, $|f(t_n, u_n) - f(t_n, w_n)| \leq L_f |u_n - w_n|$

$$\Rightarrow |u_{n+1} - w_{n+1}| \leq (1 + hL_f) |u_n - w_n| \text{ Estimación de estabilidad}$$

$$\Rightarrow (1 + hL_f) \text{ factor de amplificación}$$

La segunda forma nos va a interesar más porque relaciona la variación en el paso n -ésimo con la inicial iterando recursivamente:

$$|u_n - w_n| \leq (1 + hL_f)^n |u_0 - w_0|$$

Observación 5: Para estudiar la estabilidad hay que tener en cuenta ambas acotaciones.

Como $1 + hL_f > 1$, utilizando que

- $1 + x \leq e^x, \quad \forall x \geq 0$
- $t_n = nh \leq T$

$$|u_n - w_n| \leq (1 + hL_f)^n |u_0 - w_0| \leq e^{nhL_f} |u_0 - w_0| \leq e^{TL_f} |u_0 - w_0|$$

Obtenemos la **relación de estabilidad clásica** para el método de Euler explícito.

$$\boxed{|u_n - w_n| \leq e^{TL_f} |u_0 - w_0|}$$

$$\text{Si } |u_0 - w_0| \rightarrow 0 \Rightarrow |u_n - w_n| \rightarrow 0, \quad \forall n = 0, 1, \dots, N$$

Supongamos que el valor de L_f fuese muy grande. Como $Nh = T$, si queremos *controlar* el valor de TL_f para que no sea excesivamente grande, tendremos que tomar h suficientemente pequeño. Ésta sería en este caso la **restricción de estabilidad**.

Observación 6: Si lo relacionamos con la forma anterior, esta desigualdad nos sirve tanto si el esquema de Euler explícito tiene un caracter expansivo o contractivo.

2.3. Convergencia

Ahora que hemos estudiado la estabilidad del método, como podemos considerar u_n y \tilde{u}_n dos aproximaciones a un mismo punto que parten de dos puntos distintos,

$$\begin{cases} u_n = u_{n-1} + hf(t_{n-1}, u_{n-1}) \\ \tilde{u}_n = u(t_{n-1}) + hf(t_{n-1}, u(t_{n-1})) \end{cases}$$

podremos dar una cota para el error global de Euler Explícito $e_n = |u(t_n) - u_n|$

$$\begin{aligned} |e_n| &= |u(t_n) - u_n| \leq |u(t_n) - \tilde{u}_n| + |\tilde{u}_n - u_n| \\ &= l(u(t_n), h) + |u(t_{n-1}) + hf(t_{n-1}, u(t_{n-1})) - u_{n-1} - hf(t_{n-1}, u_{n-1})| \\ &\leq l(u(t_n), h) + (1 + hL_f)|e_{n-1}| \end{aligned}$$

Luego utilizando la cota que hemos dado en la sección de consistencia para el error local, llegamos a la relación de recurrencia,

$$\boxed{|e_n| \leq \frac{h^2}{2}M + (1 + hL_f)|e_{n-1}|} \quad \text{donde} \quad M = \max_{t \in [t_0, t_0+T]} |u''(t)|$$

Antes de desarrollar la expresión anterior, es importante destacar que el valor inicial u_0 no tiene porqué coincidir con el $u(t_0)$; puede darse el caso de que $u(t_0)$ provenga de un valor experimental o, por ejemplo, $u(t_0) = 1/3 \dots$. En general hay un error inicial e_0 que no podemos despreciar, veamos por qué

$$\begin{aligned} |e_n| &\leq \frac{h^2}{2}M + (1 + hL_f)|e_{n-1}| \\ &\leq \frac{h^2}{2}M + (1 + hL_f)\left(\frac{h^2}{2}M + (1 + hL_f)|e_{n-2}|\right) \\ &\leq \frac{h^2}{2}M + (1 + hL_f)\left(\frac{h^2}{2}M + (1 + hL_f)\left(\frac{h^2}{2}M + (1 + hL_f)|e_{n-3}|\right)\right) \\ &\leq \dots \leq (1 + L_f)^n e_0 + \frac{h^2}{2}M \left[(1 + L_f)^{n-1} + \dots + (1 + L_f) + 1 \right] \\ &= (1 + L_f)^n e_0 + \frac{h^2}{2}M \frac{(1 + L_f)^n - 1}{(1 + L_f) - 1} = (1 + L_f)^n e_0 + \frac{h^2}{2}M \frac{(1 + L_f)^n - 1}{hL_f} \\ &\leq e^{nhL_f} e_0 + \frac{h^2}{2}M \frac{e^{nhL_f} - 1}{hL_f} = e^{t_n L_f} e_0 + \frac{h}{2}M \frac{e^{t_n L_f} - 1}{L_f} = C(e_0 + h) \end{aligned}$$

$$\Rightarrow \boxed{\text{si } e_0 = O(h) \Rightarrow e_n = O(h)}$$

donde hemos usado la notación de t_* para un punto general de la partición π_h tal que $t_* = nh$. \Rightarrow El orden de convergencia de Euler explícito es 1. Si denotamos por $E(h)$ el error global sobre la función $u(t)$ en $[t_0, t_0 + T]$ y quisiésemos acotarlo por ε , como el orden de convergencia es 1,

$$E(h) = Ch < \varepsilon \Leftrightarrow h < \varepsilon C^{-1}$$

donde $C = \frac{M}{2}$ y M depende del tiempo y la función pendiente. Esta restricción sobre h se denomina **restricción por precisión**.

Dificultades y Mejoras Euler explícito

1. Tener un orden de convergencia uno no tiene fuerza suficiente para competir con las constantes que surgen en los problemas expansivos, ya que estas constantes crecen de manera exponencial.
2. En los problemas disipativos las constantes de la estimación de error podemos mejorarlas pero necesitamos restringir el valor de h de una forma que puede ser excesiva. El no hacerlo nos lleva a situaciones donde el cálculo se hace inestable.
3. En el mismo subintervalo $[t_n, t_{n+1}]$ la pendiente en el final $f(t_{n+1}, y_{n+1})$ puede cambiar mucho con respecto a $f(t_n, y_n)$ tanto en el caso expansivo como en el contractivo.

Como consecuencia, resulta que el método de Euler explícito es mejorable al menos en dos sentidos

1. Mayor precisión: mejorando el error local de consistencia de $O(h^2)$ a $O(h^{p+1})$ con $p > 1 \Rightarrow$ el error local de truncatura pasaría a $O(h^p)$ con $p > 1$.
2. Mejor estabilidad: mejorando la restricción de estabilidad sobre h , o incluso eliminando esta restricción.

Capítulo 3

Método de Euler implícito

Este método va a mejorar la estabilidad de Euler explícito, vamos a dar su definición estudiando a su vez su estabilidad para ver más claro el por qué.

Esquema Euler Implícito

Dados (t_0, u_0) y la partición π_h obtener

$$\begin{aligned}t_{n+1} &= t_n + h \\u_{n+1} &= u_n + hf(t_{n+1}, u_{n+1}), \quad n \in \{0, 1, \dots, N-1\}, \quad \text{donde } N = \frac{T}{h}\end{aligned}$$

Se interpreta geométicamente como avanzar en la dirección dada por la pendiente $f(t_{n+1}, u_{n+1})$ que hay en (t_{n+1}, u_{n+1}) . Esto supone una mejora si la pendiente decae fuertemente desde t_n a t_{n+1} , e.d. si $|f(t_n, u_n)| \gg |f(t_{n+1}, u_{n+1})|$.

Ejemplo: Supongamos que tenemos que resolver la edo

$$\begin{cases} u'(t) = -\lambda u(t) & 0 < t < T \\ u(0) = 1 \end{cases}$$

Supongamos $\lambda \gg 1$. Si usamos el método de Euler explícito en una partición uniforme se obtienen los valores

$$y_{n+1} = (1 - \lambda h)^{n+1}, \quad n = 0, 1, \dots, N-1, \quad \text{para } N = \frac{T}{h}$$

$$\Rightarrow f_u = -\lambda \Rightarrow |1 + hf_u| < 1 \Leftrightarrow h \in \left(0, \frac{2}{\lambda}\right)$$

Si $h \in \left(\frac{1}{\lambda}, \frac{2}{\lambda}\right) \Rightarrow$ hay decaimiento en la función estimada pero oscilatorio y sabemos que la gráfica de la solución es decreciente. Luego la restricción de estabilidad queda como

$$h \in \left(0, \frac{1}{\lambda}\right) \quad \text{que puede ser muy restrictiva si } \lambda \gg 1$$

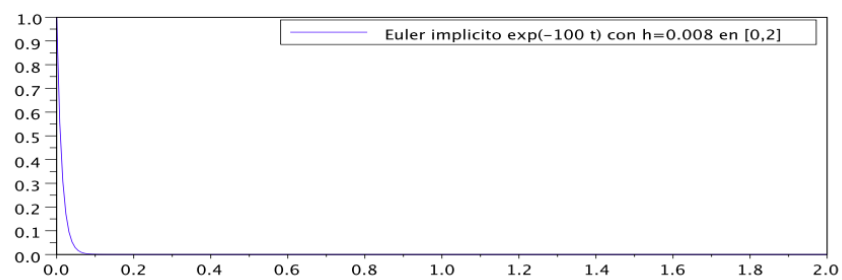
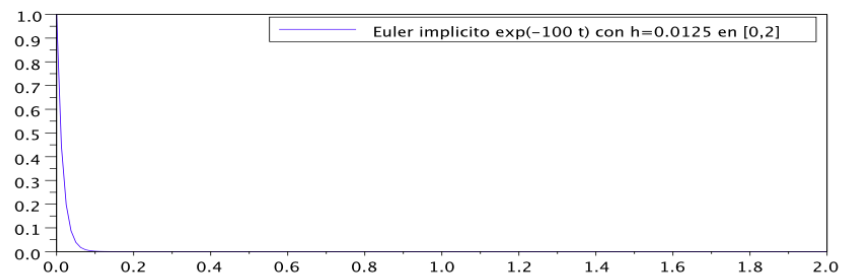
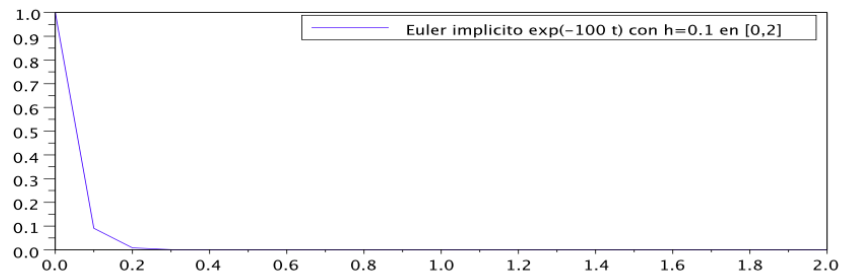
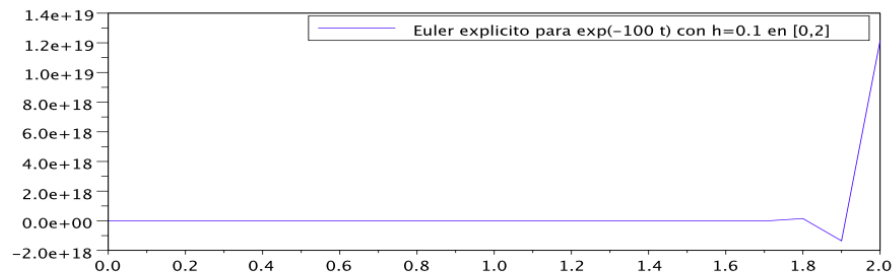
Como $h = \frac{T}{N}$, la restricción anterior queda como

$$N > T\lambda = T|\partial_u f(t, u)|$$

Por otro lado, si usamos el método de Euler implícito, obtenemos

$$y_{n+1} = (1 - \lambda h)^{-(n+1)}, \quad n = 0, 1, \dots, N-1, \quad \text{para } N = \frac{T}{h}$$

Y podemos comparar los resultados:



□

Aunque presente mejor estabilidad, el cálculo de u_{n+1} en cada paso es un poco más costoso que en caso de Euler explícito. Tenemos que resolver la ecuación

$$z = u_n + hf(t, z)$$

Computacionalmente se hace con cualquier proceso iterativo empezando por $z_0 = u_n + hf(t_{n+1}, u_n)$, ya que es bastante cercano al punto buscado. O también, podemos calcular el punto fijo de la función

$$G(z) := u_n + hf(t, z)$$

aplicando el método de Newton a la ecuación

$$F(z) = 0 \quad \text{donde} \quad F(z) = z - G(z)$$

Ahora, para ello necesitamos que G sea contractiva. Para un mismo $t_{n+1} \in \pi_h$ y L_f la constante de Lipschitz de f ,

$$|G(z) - G(w)| \leq L_f h |z - w|$$

$\Rightarrow G$ contractiva $\Leftrightarrow h < \frac{1}{L_f}$, lo que obliga a tener h muy pequeño si $L_f \gg 1$, pero en la práctica no suele ser una restricción muy severa.

3.1. Consistencia

Veamos el error local en un punto genérico fijo t_* . Supongamos cierta la hipótesis de localización, e.d. partimos del valor exacto $(t_*, u(t_*))$. Denotemos \bar{z} el punto *siguiente* en el desarrollo supuesta la hipótesis de localización,

$$\bar{z} = u(t_*) + hf(t_* + h, \bar{z})$$

$$\Rightarrow l(u(t_*), h) = u(t_* + h) - \bar{z} = u(t_* + h) - u(t_*) - hf(t_* + h, \bar{z})$$

Vamos a desarrollar $f(t_* + h, \bar{z})$ con Taylor en ambas variables con respecto al punto $(t_*, u(t_*))$ con incrementos $(h, hf(t_* + h, \bar{z}))$,

$$\begin{aligned} f(t_* + h, \bar{z}) &= f(t_* + h, u(t_*) + hf(t_* + h, \bar{z})) \\ &= f(t_*, u(t_*)) + hf_t(t_*, u(t_*)) + hf(t_* + h, \bar{z})f_u(t_*, u(t_*)) + O(h^2) \end{aligned}$$

Si ahora en la propia expresión de $f(t_* + h, \bar{z})$ escribimos la que nos acaba de salir, basta con agrupar términos para obtener:

$$f(t_* + h, \bar{z}) = u'(t_*) + hu''(t_*) + O(h^2)$$

Luego,

$$\begin{aligned} l(u(t_*), h) &= u(t_* + h) - u(t_*) - hf(t_* + h, \bar{z}) \\ &= u(t_* + h) - u(t_*) - h[u'(t_*) + hu''(t_*) + O(h^2)] \\ &= hu'(t_*) + \frac{h^2}{2}u''(t_*) + O(h^3) - h[u'(t_*) + hu''(t_*) + O(h^2)] \\ &= -\frac{h^2}{2}u''(t_*) + O(h^3) = O(h^2) \end{aligned}$$

Y por tanto,

$$l(u(t_*), h) = u(t_* + h) - \bar{z} = O(h^2)$$

3.2. Estabilidad

Como antes, tomamos el punto genérico t_* . Como en el caso de Euler explícito, vamos a comparar directamente los valores \bar{z} , z , desde sus definiciones:

$$\begin{aligned}\bar{z} - z &= u(t_*) + hf(t_* + h, \bar{z}) - [u_n + hf(t_{n+1}, z)] \\ &= u(t_*) - u_n + h(f(t_{n+1}, \bar{z}) - f(t_{n+1}, z)) \\ &= u(t_*) - u_n + hf_u(t_{n+1}, s)(\bar{z} - z)\end{aligned}$$

$$\Rightarrow u(t_*) - u_n = (1 - hf_u(t_{n+1}, s))(\bar{z} - z) \text{ para } s \text{ en el intervalo adecuado}$$

Tomando valores absolutos y usando la estimación de Lipschitz para la función f

$$\begin{aligned}|\bar{z} - z| &= \left| \frac{1}{1 - hf_u(t_{n+1}, s)} (u(t_*) - u_n) \right| \\ &\leq \left| \frac{1}{1 - hL_f} (u(t_*) - u_n) \right|\end{aligned}$$

Nos interesa que sea

$$\text{factor de amplificación } \frac{1}{1 - hL_f} \Leftrightarrow h < \frac{1}{L_f}$$

Si $\partial_u f < 0 \Rightarrow$ ya se cumple sin restricciones sobre h . (Por eso decimos que la estabilidad se mejora)

3.3. Convergencia

Para estudiar el error global, vemos la diferencia $u(t_{n+1}) - u_{n+1}$ donde

$$u_{n+1} = u_n + hf(t_{n+1}, u_{n+1})$$

Igual que con Euler explícito, tenemos que

$$\begin{aligned}|u(t_{n+1}) - u_{n+1}| &= u(t_{n+1}) - \bar{u}_{n+1} + \bar{u}_{n+1} - u_{n+1} \\ &= l(u(t_n), h) + \bar{u}_{n+1} - u_{n+1} = O(h^2) + \bar{u}_{n+1} - u_{n+1}\end{aligned}$$

Usando la estabilidad de la estimación añadiendo la restricción sobre h (porque no conocemos f_u), tenemos

$$e_{n+1} = |u(t_{n+1}) - u_{n+1}| \leq O(h^2) + \frac{1}{1 - hL_f} (u(t_n) - u_n)$$

$$\Rightarrow \boxed{e_{n+1} \leq O(h^2) + \frac{1}{1 - hL_f} e_n}$$

\Rightarrow Como en el caso de Euler explícito, hay convergencia de orden 1 con respecto a h , siendo además el error global decreciente por un factor de $(1 - hL_f)^{-1}$.

Dificultades y Mejoras de Euler Implícito

1. En los problemas lineales se encuentra la gran ventaja de usar este método implícito, ya que no plantea ninguna restricción sobre h en comparación con el uso del método explícito.
2. La restricción sobre h para resolver los problemas no lineales no es importante ya que las restricciones por precisión y por estabilidad son mas fuertes usualmente.
 - a) Si el sistema es expansivo los dos métodos dan similares cotas de error y suele ser más preciso Euler implícito.
 - b) Si el sistema es disipativo, ambos métodos dan cotas de error similares pero con Euler implícito la restricción de estabilidad sobre h no existe.
 - c) Si el problema es lineal, suele ser mejor usar Euler implícito en cualquier caso.

Capítulo 4

Método de Crank-Nicolson

Este método va a mejorar el orden de convergencia de Euler explícito.

Esquema Crank-Nicolson

Dados (t_0, u_0) y la partición π_h obtener

$$\begin{aligned} t_{n+1} &= t_n + h \\ u_{n+1} &= u_n + \frac{h}{2} [f(t_n, u_n) + f(t_{n+1}, u_{n+1})], \quad n \in \{0, 1, \dots, N-1\}, \quad \text{donde } N = \frac{T}{h} \end{aligned}$$

Se interpreta geométicamente como avanzar en la dirección promedio de las pendientes dada $f(t_{n+1}, u_{n+1})$ y $f(t_n, u_n)$, la del principio y la del final. También se puede interpretar como una aplicación de la fórmula de los trapecios a la expresión integral del problema de Cauchy.

Observación 7: El método de Heun consiste en linealizar esta ecuación

Esquema Heun

Dados (t_0, u_0) y la partición π_h obtener

$$\begin{aligned} t_{n+1} &= t_n + h \\ \tilde{u}_{n+1} &= u_n + hf(t_n, u_n) \\ u_{n+1} &= u_n + \frac{h}{2} [f(t_n, u_n) + f(t_{n+1}, \tilde{u}_{n+1})], \quad n \in \{0, 1, \dots, N-1\}, \quad \text{donde } N = \frac{T}{h} \end{aligned}$$

Tiene el mismo orden que el método de Crank-Nicolson pero una región de estabilidad más pequeña.

Siguiendo el mismo razonamiento que en el método anterior, tenemos que resolver la ecuación

$$z = u_n + \frac{h}{2} [f(t_n, u_n) + f(t_{n+1}, z)]$$

e.d. calcular el punto fijo de la función

$$G(z) := u_n + \frac{h}{2} [f(t_n, u_n) + f(t_{n+1}, z)]$$

Necesitamos que G sea contractiva, por lo que para un mismo $t_{n+1} \in \pi_h$,

$$|G(z) - G(w)| = \left| \frac{h}{2} [f(t_{n+1}, z) - f(t_{n+1}, w)] \right| \leq \frac{h}{2} |\partial_u f(t_{n+1}, \xi)| |z - w|$$

$$\Rightarrow G \text{ contractiva} \Leftrightarrow h < \frac{2}{L_f}, \text{ para } L_f = \max_{t, u} |f_u(t, u)|$$

Como en el caso de Euler Implícito, en la práctica no suele ser una restricción muy severa.

4.1. Consistencia. Estabilidad y Convergencia

De momento suponemos que el esquema permite el cálculo de u_{n+1} en cada paso. Para ver el error local en un punto generico fijo t_* suponemos, como siempre, la hipótesis de localización,

$$\bar{z} = u(t_*) + \frac{h}{2} [f(t_*, u(t_*)) + f(t_* + h, \bar{z})]$$

Entonces

$$\begin{aligned} l(u(t_*), h) &= u(t_* + h) - \bar{z} \\ &= u(t_* + h) - \left[u(t_*) + \frac{h}{2} [f(t_*, u(t_*)) + f(t_* + h, \bar{z})] \right] \\ &= u(t_* + h) - \left[u(t_*) + \frac{h}{2} u'(t_*) + \frac{h}{2} f(t_* + h, \bar{z}) \right] \end{aligned}$$

Desarrollando con Taylor el primer término

$$= \frac{h}{2} u'(t_*) + \frac{h}{2} u''(t_*) + O(h^3) - \frac{h}{2} f(t_* + h, \bar{z})$$

Por definición,

$$f(t_* + h, \bar{z}) = f\left(t_* + h, u(t_*) + \frac{h}{2} u'(t_*) + \frac{h}{2} f(t_* + h, \bar{z})\right)$$

Y si desarrollamos Taylor de la función f respecto al punto $(t_*, u(t_*))$ con los incrementos $(h, \frac{h}{2} u'(t_*) + \frac{h}{2} f(t_* + h, \bar{z}))$, se nos queda

$$f(t_* + h, \bar{z}) = f(t_*, u(t_*)) + h f_t(t_*, u(t_*)) + \frac{h}{2} f(t_*, u(t_*)) f_y(t_*, u(t_*)) + \frac{h}{2} f(t_* + h, \bar{z}) f_y(t_*, u(t_*)) + O(h^2)$$

y si reiteramos el desarrollo en la expresión de $f(t_* + h, \bar{z})$ ahora en el segundo miembro llegamos a

$$\frac{h}{2} f(t_* + h, \bar{z}) = \frac{h}{2} y'(t_*) + \frac{h^2}{2} y''(t_*) + O(h^3)$$

y como consecuencia,

$$\begin{aligned} l(y(t_*); h) &= \frac{h}{2} y'(t_*) + \frac{h}{2} y''(t_*) + O(h^3) - \frac{h}{2} f(t_* + h, \bar{z}) \\ &= \frac{h}{2} y'(t_*) + \frac{h}{2} y''(t_*) + O(h^3) - \frac{h}{2} y'(t_*) - \frac{h^2}{2} y''(t_*) - O(h^3) = O(h^3) \end{aligned}$$

luego hemos ganado una potencia en el orden del error local

$$l(y(t_*); h) = y(t_* + h) - \bar{z} = O(h^3) \Rightarrow \mathcal{T}(y(t_*); h) = O(h^2)$$

Esto implica un error global de orden 2 en h y hemos mejorado un orden.

Capítulo 5

Métodos de Runge-Kutta

Los métodos de Runge-Kutta se pueden interpretar geoméricamente como una extensión del método de Euler progresivo mejorando el orden de convergencia. Se ajusta mediante un promedio la mejor pendiente por la que avanzar en cada paso.

Esquema Runge-Kutta General

La forma más general de un método de Runge-Kutta consiste en tomar S muestras de pendientes k_1, k_2, \dots, k_S y promediarlas para generar una nueva dirección de avance.

1. Dado $u_0 \approx u(t_0)$
2. Construir las pendientes k_1, k_2, \dots, k_S y obtener la pendiente promedio. Dependiendo de cómo obtengamos las pendientes, el método es explícito o implícito.

$$k = \sum_{i=1}^S b_i k_i \quad \text{con } b_1 + b_2 + \dots + b_S = 1$$

3. $u_{n+1} = u_n + h k$

Definición 9. Al número de muestras S obtenidas se le llama **etapas del método**.

Observación 8:

1. Se dice que son métodos de un paso ya que se avanza desde t_n a t_{n+1}
2. Siempre son métodos de un paso explícitos con respecto al valor buscado u_{n+1} (aunque no lo sean con respecto a las pendientes k_1, k_2, \dots, k_S).
3. Es importante destacar que el número de muestras S no tiene porqué coincidir con el orden del método p , siempre es $p \leq S$. En los casos donde $S = p$ tenemos la situación óptima, pero esto se consigue sólo si $S \leq 4$.

5.1. Métodos de Runge-Kutta Explícitos

Desarrollemos con más detalle la forma general de un método Runge-Kutta. Supongamos que nos situamos en el punto u_n , en el eje x estaríamos en el punto t_n . Avanzamos hacia el punto avanzamos por la recta de pendiente k_1 una distancia $c_2 h$, llegaríamos al punto

$$u_n \rightsquigarrow u_{n+c_2} := u_n + c_2 h k_1$$

Calculamos ahora la pendiente en este nuevo punto

$$k_2 = f(t_n + c_2 h, u_{n+c_2}) = f(t_n + c_2 h, u_n + c_2 h k_1)$$

El método, por definición nos dará coeficientes $\alpha_{3,1}, \alpha_{3,2}$ tal que $\alpha_{3,1} + \alpha_{3,2} = 1$, y definiremos la pendiente promedio como

$$\alpha_{3,1} k_1 + \alpha_{3,2} k_2$$

Avanzamos en esa dirección desde u_n ahora una distancia $c_3 h$. La pendiente en este nuevo punto será

$$\begin{aligned} k_3 &= f(t_n + c_3 h, u_{n+c_3}) = f(t_n + c_3 h, u_n + c_3 h (\alpha_{3,1} k_1 + \alpha_{3,2} k_2)) \\ &= f(t_n + c_3 h, u_n + h (c_3 \alpha_{3,1} k_1 + c_3 \alpha_{3,2} k_2)) \\ &= f(t_n + c_3 h, u_n + h (a_{3,1} k_1 + a_{3,2} k_2)) \end{aligned}$$

y en general, cuando tenemos k_1, k_2, \dots, k_{i-1} construimos

$$u_{n+c_i} = u_n + h \sum_{j=1}^{i-1} a_{i,j} k_j$$

$$\text{donde } \begin{cases} a_{2,1} = c_2, & a_{i,j} = c_i \alpha_{i,j}, i \geq 2, j = 1, \dots, i-1 \\ a_{i1} + a_{i2} + \dots + a_{i,i-1} = c_i (\alpha_{i,1} + \alpha_{i,2} + \dots + \alpha_{i,i-1}) = c_i \end{cases}$$

Y la pendiente en este nuevo punto

$$k_i = f(t_n + c_i h, u_{n+c_i})$$

Resumiendo;

Cálculo de las pendientes Runge-Kutta explícito

tomando $c_1 = 0$, y sabiendo que $c_2 = a_{21}$

$$\begin{aligned} k_1 &= f(t_n, u_n) \\ k_2 &= f(t_n + c_2 h, u_n + h a_{21} k_1) \\ &\vdots \\ k_i &= f\left(t_n + c_i h, u_n + h \left(\sum_{j=1}^{i-1} a_{i,j} k_j\right)\right) \text{ para } i = 3, \dots, S \end{aligned}$$

Por último, habría que calcular el *promedio* de todas las pendientes los coeficientes $\{b_1, \dots, b_S\}$ también vienen dados por el método

$$k = \sum_{i=1}^S b_i k_i \text{ con } b_1 + b_2 + \dots + b_S = 1$$

Y entonces dar el paso definitivo

$$u_{n+1} = u_n + h k$$

5.1.1. Tablero de Butcher

Los coeficientes $\{a_{i,j}\}$, $\{c_i\}$, $\{b_i\}$ determinan completamente el método de Runge-Kutta y normalmente se recolectan en lo que se conoce como **matriz o tablero de Butcher**. En el caso explícito el tablero queda como

0	0	0	...	0	0
c_2	$a_{2,1}$	0	...	0	0
c_3	$a_{3,1}$	$a_{3,2}$	0	...	0
\vdots	\vdots	\vdots	\ddots	\ddots	0
c_S	$a_{S,1}$	$a_{S,2}$...	$a_{S,S-1}$	0
	b_1	b_2	...	b_{S-1}	b_S

Ejemplo:

Esquema de Heun de orden 2

Tomamos la pendiente inicial

$$k_1 = f(t_n, y_n)$$

y avanzamos en esa dirección hasta llegar al punto final del intervalo donde volvemos a tomar una nueva muestra para la pendiente:

$$k_2 = f(t_n + h, y_n + h k_1).$$

Con los dos valores k_1 y k_2 terminamos haciendo un promedio

$$k = \frac{1}{2}k_1 + \frac{1}{2}k_2$$

siendo esta la dirección de avance definitiva

$$y_{n+1} = y_n + h k.$$

Este método es de orden 2 con respecto a h . Su tablero de Butcher es

0	0	0
1	1	0
	1/2	1/2

□

Ejemplo:

Esquema de Heun de orden 3

Se avanza desde t_n hasta $t_n + h/3$ con la pendiente k_1 y se genera k_2 . Luego desde t_n hasta $t_n + 2h/3$ con la pendiente k_2 y se genera k_3 . Finalmente se toma el promedio para avanzar definitivamente a $t_n + h$

$$\begin{aligned} k_1 &= f(t_n, y_n) \\ k_2 &= f\left(t_n + \frac{1}{3}h, y_n + \frac{1}{3}h k_1\right) \\ k_3 &= f\left(t_n + \frac{2}{3}h, y_n + \frac{2}{3}h k_2\right) \\ y_{n+1} &= y_n + h\left(\frac{1}{4}k_1 + \frac{3}{4}k_3\right) \end{aligned}$$

Este método es de orden 3 con respecto a h . Su tablero de Butcher es

0	0	0	0
1/3	1/3	0	0
2/3	0	2/3	0
	0	1/4	3/4

□

Ejemplo:Regla del Punto Medio

Tomamos la pendiente inicial

$$k_1 = f(t_n, y_n)$$

y avanzamos en esa dirección hasta llegar al punto medio del intervalo donde volvemos a evaluar la pendiente como muestra.

$$k_2 = f\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_1\right)$$

Con los dos valores k_1 y k_2 terminamos haciendo un promedio que simplemente consiste en usar el valor k_2

$$k = 0k_1 + 1k_2 = k_2$$

siendo esta la dirección de avance definitiva

$$y_{n+1} = y_n + hk.$$

También se puede interpretar como que usamos el punto medio del intervalo para aproximar: tomamos

$$y_* = y_n + \frac{h}{2}f(t_n, y_n)$$

y escribimos

$$y_{n+1} = y_n + hf\left(t_n + \frac{h}{2}, y_*\right).$$

Este método es de orden 2 con respecto a h . Su tablero es

0	0	0
1/2	1/2	0
	0	1

□

Ejemplo:Esquema de Runge-Kutta Clásico de orden 4

Se avanza desde t_n hasta $t_n + h/2$ con la pendiente k_1 y desde t_n también hasta el mismo punto $t_n + h/2$ pero ahora con la pendiente k_2 . Luego avanzamos hasta $t_n + h$ con k_3 y finalmente se toma el promedio para avanzar definitivamente a $t_n + h$

$$k_1 = f(t_n, y_n)$$

$$k_2 = f\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_1\right)$$

$$k_3 = f\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_2\right)$$

$$k_4 = f(t_n + h, y_n + hk_3)$$

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

Este método es de orden 4 con respecto a h . Su tablero de Butcher es

0	0	0	0	0
1/2	1/2	0	0	0
1/2	0	1/2	0	0
1	0	0	1	0
	1/6	2/6	2/6	1/6

5.1.2. Consistencia

Por construcción, los métodos de Runge-Kutta explícitos son consistentes. Usando el desarrollo de Taylor en torno al punto $(t, y(t))$

$$k_i = f(t + c_i h, y(t) + h(a_{i1}k_1 + \dots + a_{is}k_s)) = f(t, y(t)) + O(h) = y'(t) + O(h)$$

luego

$$\begin{aligned} l(y(t), t) &= y(t+h) - y(t) - h k \\ &= y(t+h) - y(t) - h \sum_{i=1}^S b_i k_i = h y'(t) + O(h^2) - h \sum_{i=1}^S b_i k_i \\ &= h y'(t) + O(h^2) - h \sum_{i=1}^S (b_i y'(t) + O(h)) \\ &= h y'(t) + O(h^2) - h y'(t) \sum_{i=1}^S b_i - S h O(h) = O(h^2) \end{aligned}$$

Por lo tanto, el error local o de consistencia es siempre al menos $O(h^2)$.

La idea ahora es usar los parámetros y el número de etapas para aumentar el orden local o de consistencia del método, esto lo veremos más adelante.

5.1.3. Estabilidad

0-estabilidad

Consideramos el esquema general de Runge-Kutta, vamos a denotar a las pendientes como $k_i(t_n, y_n, h)$

$$y_{n+1} = y_n + h \sum_{i=1}^S b_i k_i(t_n, y_n, h), \quad 0 \leq n \leq N-1$$

Y supongamos que y_0 es el dato inicial. Denotamos su *perturbación* como δ_0 y sea $z_0 = y_0 + \delta_0$. Entonces

$$z_{n+1} = z_n + h \left(\sum_{i=1}^S b_i k_i(t_n, y_n, h) + \delta_n \right), \quad 0 \leq n \leq N-1$$

Definición 10. Sea el intervalo $[t_0, t_0 + T]$ y consideramos las particiones $\{\pi_h\}_h$ donde $h = \frac{T}{N}$. Decimos que un método es **0-estable** si su perturbación cumple la siguiente propiedad:

Si $|\delta_n^{(h)}| \leq \varepsilon$, para $\varepsilon > 0 \Rightarrow \exists C(T, f), h_0 > 0$ tal que para todo $h < h_0$ se cumple

$$|z_n - y_n| \leq C(\varepsilon + |z_0 - y_0|), \quad 1 \leq n \leq N$$

Vamos a ver que todos los métodos de Runge-Kutta explícitos son cero estables.

Teorema 3. La expresión definida por

$$\Phi_f(t, y, h) = \sum_{i=1}^S b_i k_i(t, y, h)$$

cumple la condición de Lipschitz con respecto a su segundo argumento de manera uniforme para $h < h_0$. E.d. $\exists h_0 > 0$ para el que $\exists M_{\Phi_f}$ (independiente de h) tal que para todo $h < h_0$

$$|\Phi_f(t, z, h) - \Phi_f(t, y, h)| \leq M_{\Phi_f} |z - y|$$

Ejemplo: La cero estabilidad para el método de Euler progresivo se obtiene de la condición de Lipschitz para f con respecto a la segunda variable. Esto se debe a que en este método la pendiente definitiva la da f en cada paso.

□

Ejemplo: Estabilidad del método de Heun.

$$\Phi(t, y, h) = \frac{1}{2} (k_1 + k_2) = \frac{1}{2} [f(t, y(t)) + f(t + h, y(t) + hf(t, y(t)))]$$

Como f cumple la condición de Lipschitz con respecto a su segunda variable, $\exists L_f$ tal que

$$|f(t, y) - f(t, z)| \leq L_f |y - z|$$

luego

$$\begin{aligned} |\Phi(t, y, h) - \Phi(t, z, h)| &\leq \frac{1}{2} |f(t, y) - f(t, z)| + \frac{1}{2} |f(t + h, y + hf(t, y)) - f(t + h, z + hf(t, z))| \\ &\leq \frac{L_f}{2} |y - z| + \frac{L_f}{2} |y + hf(t, y) - z - hf(t, z)| \\ &\leq L_f |y - z| + \frac{hL_f}{2} |f(t, y) - f(t, z)| \leq L_f |y - z| + \frac{hL_f^2}{2} |y - z| \\ &\leq \left(L_f + \frac{hL_f^2}{2} \right) |y - z| \end{aligned}$$

Entonces, para cualquier $h_0 > 0$, fijamos

$$M_{\Phi_f} = L_f + \frac{h_0 L_f^2}{2} \Rightarrow L_f + \frac{hL_f^2}{2} \leq M_{\Phi_f}, \quad \forall h \leq h_0$$

$$\Rightarrow \text{si } h \leq h_0, \quad |\Phi(t, y, h) - \Phi(t, z, h)| \leq \left(L_f + \frac{hL_f^2}{2} \right) |y - z| \leq M_{\Phi_f} |y - z|$$

□

Teorema 4. Todos los métodos de Runge-Kutta explícitos son cero estables

Observación 9: Supongamos que $\delta_n = 0$ para $n = 1, 2, \dots$

$$\begin{aligned} |z_{n+1} - y_{n+1}| &= \left| z_n + h \left(\sum_{i=1}^S b_i k_i(t_n, z_n, h) + \delta_n \right) - y_n - h \sum_{i=1}^S b_i k_i(t_n, y_n, h) \right| \\ &\leq |z_n - y_n| + h |\delta_n| + \sum_{i=1}^S h b_i |k_i(t_n, z_n, h) - k_i(t_n, y_n, h)| \\ &\leq |z_n - y_n| + h |\delta_n| + h M_{\Phi_f} |z - y| = |z_n - y_n| + h M_{\Phi_f} |z_n - y_n| \\ &= |z_n - y_n| (1 + h M_{\Phi_f}) \end{aligned}$$

Por tanto, sólo se tiene en cuenta la perturbación inicial, para $n = 1, 2, \dots$

$$|z_n - y_n| \leq (1 + h M_{\Phi_f})^n |\delta_0| \leq e^{TM_{\Phi_f}} |\delta_0|$$

Observación 10: Las constantes que garantizan la estabilidad dependen de T y de M_{Φ_f} de manera exponencial, es decir, aparece un $e^{TM_{\Phi_f}}$, por lo que pueden empeorar de manera muy importante.

5.1.4. Convergencia

Con la notación

$$\Phi_f(t, y, h) = \sum_{i=1}^S b_i k_i(t, y, h)$$

tenemos la misma estructura que con Euler explícito pero reemplazando la dirección de avance $f(t_n, y_n)$ por $\Phi_f(t_n, y_n, h)$. Por el Tma 3, Φ_f cumple la condición de Lipschitz global para $h < h_0$ y hemos visto también que el error de consistencia es al menos $O(h^2) \Rightarrow$ la misma demostración que en el caso de Euler explícito nos da la convergencia con orden al menos $O(h)$ de Runge-Kutta. En este caso además podemos tener ordenes mayores dependiendo de la elección del número de etapas S y de los coeficientes.

Teorema 5. Todos los métodos de Runge-Kutta explícitos son convergentes con orden al menos uno.

Además, si se cumplen estas dos condiciones:

- El método de Runge-Kutta es consistente con orden $O(h^p)$.
- El error inicial $|y(t_0) - y_0|$ tiende a cero con orden $O(h^p)$

Entonces converge con orden p .

Familias RKE de acuerdo al orden

Recordemos que un método convergente se dice que tiene orden de convergencia p si su error local tiene orden $p + 1$. Para ver el orden de convergencia de cada familia haremos siempre lo mismo:

- En la expresión del error $l(t, y(t))$ sustituimos el término $y(t + h) - y(t)$ por el desarrollo en Taylor de y en $(t + h)$.
- Escribimos el valor de k_i en términos de f
- Desarrollamos todas las expresiones de f con Taylor para varias variables.
- Forzar a 0 el mayor números de coeficientes de potencias de h posible.

1. Métodos RKE con $S = 1$

En el caso $S = 1$ reduce la matriz de Butcher a la situación trivial

$$\begin{array}{c|c} 0 & 0 \\ \hline c_2 & 1 \end{array}$$

que representa el método de Euler explícito, que sabemos que es de orden 1.

2. Métodos RKE con $S = 2$

En este caso tenemos como matriz de Butcher

$$\begin{array}{c|cc} 0 & 0 & 0 \\ c_2 & a_{2,1} & 0 \\ \hline b_1 & b_2 & \end{array} \quad \text{con} \quad \begin{cases} c_2 = a_{2,1} \\ b_1 + b_2 = 1 \end{cases}$$

Y por definición, $k = b_1 k_1 + b_2 k_2$, y

$$k_1 = f(t, y(t)) = y'(t), \quad k_2 = f(t + c_2 h, y(t) + h c_2 k_1)$$

El error local será

$$\begin{aligned} l(y(t), h) &= y(t+h) - y(t) - h[b_1 k_1 + b_2 k_2] \\ &= y(t+h) - y(t) - h[b_1 y'(t) + b_2 f(t + c_2 h, y(t) + h c_2 k_1)] \end{aligned}$$

Ahora, usando el desarrollo de Taylor: $y(t+h) - y(t) = h y'(t) + \frac{h^2}{2} y''(t) + \frac{h^3}{6} y'''(t) + O(h^4)$

$$\Rightarrow l(y(t), h) = h y'(t) + \frac{h^2}{2} y''(t) + \frac{h^3}{6} y'''(t) + O(h^4) - h[b_1 y'(t) + b_2 f(t + c_2 h, y(t) + h c_2 k_1)]$$

Desarrollando con Taylor k_2 en el punto $(t, y(t))$, llegamos a que

$$k_2 = y'(t) + hc_2 y''(t) + \frac{h^2 c_2^2}{2} (f_{tt} + 2f f_{ty} + f^2 f_{yy}) + O(h^3)$$

Entonces, si se sustituye en la expresión de $l(t, y(t))$, utilizando que $b_1 + b_2 = 1$, llegamos a que

$$l(t, y(t)) = \left(\frac{1}{2} - b_2 c_2 \right) y''(t) h^2 + h^3 \left(\frac{y'''(t)}{6} - \frac{b_2 c_2^2}{2} (f_{tt} + 2f f_{ty} + f^2 f_{yy}) \right) + O(h^4)$$

Si forzamos $c_2 b_2 = \frac{1}{2}$,

$$\Rightarrow l(t, y(t)) = h^3 \left(\frac{y'''(t)}{6} - \frac{b_2 c_2^2}{2} (f_{tt} + 2f f_{ty} + f^2 f_{yy}) \right) + O(h^4) \rightarrow 0 \text{ si } h \rightarrow 0$$

Luego podemos establecer que, siempre que $b_1 + b_2 = 1$,

$$\begin{aligned} \text{Si } c_2 b_2 &= \frac{1}{2} & \Rightarrow & \text{orden de convergencia 2} \\ \text{Si } c_2 b_2 &\neq \frac{1}{2} & \Rightarrow & \text{orden de convergencia 1} \end{aligned}$$

En el primer caso, como $c_2 \in [0, 1]$ por ser hc_2 la longitud que avanzamos dentro de $[t_n, t_{n+1}]$, podemos escribir

$$b_2 = \frac{1}{2c_2} \quad b_1 = 1 - \frac{1}{2c_2} \quad a_{2,1} = c_2$$

Luego la tenemos una familia infinita de métodos de Runge-Kutta explícitos de orden 2, con tablero de Butcher:

$$\begin{array}{c|cc} 0 & 0 & 0 \\ c_2 & c_2 & 0 \\ \hline 0 & 1 - \frac{1}{2c_2} & \frac{1}{2c_2} \end{array}$$

3. Métodos RKE con $S = 3$

Para el tablero de Butcher

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ c_2 & a_{21} & 0 & 0 \\ c_3 & a_{31} & a_{32} & 0 \\ \hline & b_1 & b_2 & b_3 \end{array}$$

Con un razonamiento simétrico, si establecemos las restricciones

$$\left\{ \begin{array}{l} b_1 + b_2 + b_3 = 1 \\ b_2 c_2 + b_3 c_3 = \frac{1}{2} \\ b_2 c_2^2 + b_3 c_3^2 = \frac{1}{3} \\ b_3 c_2 a_{3,2} = \frac{1}{6} \end{array} \right.$$

obtendríamos convergencia de orden 3. Como tenemos cuatro ecuaciones y seis incógnitas, aparecen dos familias infinitas de métodos RK explícitos de orden 3.

4. Métodos RKE con $S = 4$

Existen condiciones similares para obtener métodos de cuarto orden pero Eliseo no las quiere escribir. Hay un ejemplo clásico de RK explícito de orden 4 en la sección 5.1.1.

5.2. Métodos de Runge-Kutta Implícitos

En estos casos, cada pendiente k_j se calcula en función del resto de pendientes $\{k_i\}_{i=1}^S$, conocidas o no:

$$k_j = \sum_{i=1}^S \alpha_{j,i} k_i, \quad \text{donde} \quad \alpha_{j,1} + \alpha_{j,2} + \dots + \alpha_{j,S} = 1$$

Desde el punto u_n , avanzamos por la dirección k_j una longitud $c_j h$.

$$u_n \rightsquigarrow u_{n+c_j} := u_n + c_j h k_j = u_n + c_j h \sum_{i=1}^S \alpha_{j,i} k_i = u_n + h \sum_{i=1}^S (c_j \alpha_{j,i}) k_i$$

y denotaremos $a_{j,i} = c_j \alpha_{j,i}$. Ahora, queremos que k_j coincida con la pendiente en el punto que hemos calculado, u_{n+c_j} , e.d.

$$k_j = f(t_n + c_j h, u_{n+c_j}) = f\left(t_n + c_j h, u_n + h \sum_{i=1}^S (c_j \alpha_{j,i}) k_i\right)$$

\Rightarrow El cálculo de las pendientes k_j se hace mediante la resolución de una ecuación de punto fijo en general no lineal de dimensión S .

Ejemplo: Métodos Runge-Kutta semi-implícitos

En los casos en los que

$$k_i = \sum_{i=1}^j \alpha_{j,i} k_i$$

si obviamente empezamos con $j = 0$, al llegar al j -ésimo paso, el único coeficiente que no conocemos sería el j -ésimo, luego se trataría de resolver una ecuación de punto fijo que sólo involucra a k_j . Aquí el esfuerzo computacional es menor y **estos métodos se denominan métodos semi-implícitos**.

□

El tablero de Butcher en el caso implícito queda relleno completamente en general

c_1	$a_{1,1}$	$a_{1,2}$	\dots	$a_{1,S-1}$	$a_{1,S}$
c_2	$a_{2,1}$	$a_{2,2}$	\dots	$a_{2,S-1}$	$a_{2,S}$
c_3	$a_{3,1}$	$a_{3,2}$	$a_{3,3}$	\dots	$a_{3,S}$
\vdots	\vdots	\vdots	\ddots	\ddots	\vdots
c_S	$a_{S,1}$	$a_{S,2}$	\dots	$a_{S,S-1}$	$a_{S,S}$
	b_1	b_2	\dots	b_{S-1}	b_S

Observación 11: Los métodos Runge-Kutta implícitos tienen propiedades de estabilidad y de orden de convergencia distintas de las que tienen los Runge-Kutta explícitos y son muy usados para los problemas rígidos. A cambio de esta ventaja su dificultad de cómputo es mayor y por ello no vamos a profundizar más.

Capítulo 6

Métodos Multipaso