# Métodos numéricos

Pablo Dmitruk

2do cuatrimestre 2021 - DF/FCEN/UBA

# Error de truncamiento en un esquema en diferencias: consistencia

Consideremos la ecuación de difusión : $$\frac{\partial f}{\partial t} = \kappa \frac{\partial^2 f}{\partial x^2}$$

con f = temperatura en una barra, $\boldsymbol{\kappa}$ = conductividad térmica

$0 \leq x \leq L$ $\quad f(x, t=0)$ condición inicial

\+ cond. contorno sobre $f(x=0,t)$ $y$ $f(x=L,t)$

Discretizamos la parte temporal y la espacial:

$$f_j^n = f(x_j, t^n) \Longrightarrow \begin{cases} x_j = j.\triangle x = j.h & j = 0, 1, \ldots, N \\ t^n = n.\triangle t = n.k & n = 0, 1, \ldots, M \end{cases}$$

Para la parte espacial utilizamos un esquema centrado: $$\left(\frac{\partial^2 f}{\partial x^2}\right)_j^n = \frac{f_{j+1}^n + f_{j-1}^n - 2f_j^n}{\triangle x^2}$$

Para la parte temporal utilizamos un esquema adelantado: $$\left(\frac{\partial f}{\partial t}\right)_j^n = \frac{f_j^{n+1} - f_j^n}{\triangle t}$$

Entonces tenemos, $$\frac{f_j^{n+1} - f_j^n}{\triangle t} = \kappa \frac{f_{j+1}^n + f_{j-1}^n - 2f_j^n}{\triangle x^2}$$

que sería la ecuación de difusión en su forma discreta (con este esquema particular)

Supongamos ahora que $F(x,t)$ es la solución exacta de la ecuación diferencial

si reemplazamos $f_j^n = F(x_j, t^n)$ en la ecuación discreta, no va a satisfacerse en general .

(es decir, la solución de la ec. diferencial no es necesariamente solución de la ec. discreta)

Llamamos $$E = \frac{F(x_j, t^{n+1}) - F(x_j, t^n)}{\triangle t} - \kappa \frac{F(x_{j+1}, t^n) + F(x_{j-1}, t^n) - 2F(x_j, t^n)}{\triangle x^2}$$

error de truncamiento de la aproximación en diferencias finitas

Hacemos Taylor en los desplazamientos temporales y espaciales,

$$E = \frac{F(x_j, t^n) + \Delta t(\frac{\partial F}{\partial t})_j^n + \frac{\Delta t^2}{2}(\frac{\partial^2 F}{\partial t^2})_j^n + \ldots - F(x_j, t^n)}{\triangle t}$$
$$- \kappa \frac{F(x_j, t^n) + \Delta x(\frac{\partial F}{\partial x})_j^n + \frac{\Delta x^2}{2}(\frac{\partial^2 F}{\partial x^2})_j^n + \ldots + F(x_j, t^n) - \Delta x(\frac{\partial F}{\partial x})_j^n + \frac{\Delta x^2}{2}(\frac{\partial^2 F}{\partial x^2})_j^n + \ldots - 2F(x_j, t^n)}{\triangle x^2}$$

$$E = \left(\frac{\partial F}{\partial t}\right)_j^n + \frac{\triangle t}{2}\left(\frac{\partial^2 F}{\partial x^2}\right)_j^n + \ldots - \kappa\left(\frac{\partial^2 F}{\partial x^2}\right)_j^n - \kappa\frac{\triangle x^2}{12}\left(\frac{\partial^4 F}{\partial x^4}\right)_j^n + O(\triangle t^2, \triangle x^4)$$

el 1er y el 3er término se cancelan, ya que F satisface la ec. diferencial, con lo cual queda:

$$E = \frac{\triangle t}{2}\left(\frac{\partial^2 F}{\partial t^2}\right)_j^n - \kappa\frac{\triangle x^2}{12}\left(\frac{\partial^4 F}{\partial x^4}\right)_j^n + O(\triangle t^2, \triangle x^4)$$

$$E = \frac{\triangle t}{2}\left(\frac{\partial^2 F}{\partial t^2}\right)_j^n - \kappa\frac{\triangle x^2}{12}\left(\frac{\partial^4 F}{\partial x^4}\right)_j^n + O(\triangle t^2, \triangle x^4)$$

vemos que (si las derivadas están acotadas) $E \to 0 \;\; \text{si} \;\; \Delta t \to 0\,,\; \Delta x \to 0$

Decimos que un esquema (de discretización) es *consistente* si el error de truncamiento tiende a 0 cuando hacemos tender a 0 el espaciado (de grilla) temporal y espacial.

No es obvio que esto pase siempre !

Ejemplo de un cambio “inocente” en la ec. discreta:

$$\frac{f_j^{n+1} - f_j^n}{\triangle t} = \kappa \frac{f_{j+1}^n + f_{j-1}^n - 2 f_j^n}{\triangle x^2}$$ lo modificamos por $$\frac{f_j^{n+1} - f_j^{n-1}}{2 \triangle t} = \kappa \frac{f_{j+1}^n + f_{j-1}^n - (f_j^{n+1} + f_j^{n-1})}{\triangle x^2}$$

Reemplazando la solución de la ec. diferencial F en este nuevo esquema discreto y haciendo Taylor queda:

$$E = \left(\frac{\partial F}{\partial t}\right)_j^n - \kappa \left(\frac{\partial^2 F}{\partial x^2}\right)_j^n + \kappa \left(\frac{\triangle t}{\triangle x}\right)^2 \left(\frac{\partial^2 F}{\partial t^2}\right)_j^n + O(\triangle t^2, \triangle x^2, \left(\frac{\Delta t}{\Delta x}\right)^2 \Delta t^2)$$

Los dos 1ros términos se cancelan cómo en el caso anterior, pero el siguiente término no necesariamente tiende a 0 cuando el espaciado temporal y espacial tienden a 0, ya que si los espaciados son proporcionales entre sí entonces ese término sobrevive.

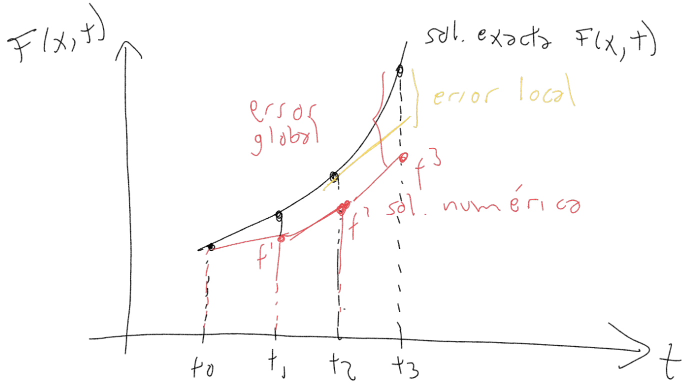
**Este esquema no es consistente con la ec. de difusión !!**

Alcanza con que el esquema sea consistente ? → Nop

Si resolvemos la ecuación discreta numéricamente (a menos del error de redondeo que vamos a ignorar) lo que nos interesa es la diferencia entre esa solución numérica y la exacta. A esa diferencia la vamos a llamar *error de discretización*

$$D = f_j^n - F(x_j, t^n)$$

Podemos distinguir entre *error de discretización local y global*




El error local se define calculándolo cómo la diferencia entre la solución exacta y la numérica partiendo del último instante cómo condición inicial

Lo que quisiéramos es que nuestra *solución numérica se aproxime a la solución exacta* a medida que disminuimos los espaciados temporal y espacial → en ese caso decimos que la solución numérica *converge* a la solución de la ecuación diferencial.

Es decir, decimos que el esquema numérico es *convergente*

Vinculado con el concepto de *estabilidad* de un esquema numérico:

→ a medida que iteramos (en el tiempo) esperamos que la solución numérica se mantenga acotada

Es decir, queremos que el (módulo) del error de discretización $|f_j^n - F(j\Delta x, n\Delta t)|$

se mantenga acotado, cuando incrementamos "n", dejando fijos $\Delta x, \Delta t$

En general hay 3 formas de analizar estabilidad:

1) Método directo: $max_{(j)}|f_j^{n+1}|$ se mantiene acotado

2) Método de la *energía*: $\sum_j |f_j^{n+1}|^2$ se mantiene acotado

3) Método de von Neumann:

proponemos $f_j^n = \hat{f}^{\,n} e^{ikj\Delta x}$

y definimos el "factor de amplificación" $\lambda$ , $\hat{f}^{\,n+1} = \lambda \hat{f}^{\,n}$

es decir, el cociente de la solución numérica entre la iteración en un paso y el paso anterior

(también se lo llama el método del autovalor)

# Esquemas temporales: integración de ecuaciones diferenciales ordinarias

$$\frac{dq}{dt} = f[q(t), t] \qquad\qquad q(t), f \in \mathbb{C}^N$$

$$q(t=0) = q_0 \quad , t \in [0, T]$$ → *problema de valores iniciales*

Ejemplo: ecuación de decaimiento $\quad \frac{dq}{dt} = Aq \quad , A \in \mathbb{R} < 0 \quad , \quad q(0) = q_0 \in \mathbb{R}$

Ejemplo: ec. de Newton para una masa en un resorte $\quad m\frac{d^2x}{dt^2} = -kx$

$$v = \frac{dx}{dt} \rightarrow m\frac{dv}{dt} = -kx \qquad \rightarrow \frac{d}{dt}\begin{pmatrix} x \\ v \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{pmatrix}\begin{pmatrix} x \\ v \end{pmatrix} \qquad q = (x, v) \in \mathbb{R}^2$$

En gral cualquier sistema de orden más alto se puede llevar a un sist de 1er orden
(Lagrange → Hamilton en Meca Clásica)

y cómo veremos después una ec. en derivadas parciales se puede llevar a un sist. de ec. ordinarias en el tiempo una vez que hayamos discretizado en el espacio (o reemplazado por modos normales).

## Métodos no-iterativos

Supongamos que integramos la ec. diferencial en el tiempo desde $(n-m)\Delta t$ hasta $(n+1)\Delta t$

con $n \geq m$ enteros positivos

$$q[(n+1) \triangle t] - q[(n-m) \triangle t] = q^{n+1} - q^{n-m} = \int_{(n-m)\Delta t}^{(n+1)\Delta t} f(q,t)\, dt$$

aproximamos la integral utilizando los valores de f en los niveles de tiempo discretos

$$\int_{(n-m)\Delta t}^{(n+1)\Delta t} f(q,t)dt \simeq [(n+1) \triangle t - (n-m) \triangle t]. [\beta f^{n+1} + \alpha_n f^n + \alpha_{n-1} f^{n-1} + \alpha_{n-2} f^{n-2} + \ldots \alpha_{n-l} f^{n-l}]$$

$l \geq 0$ entero

con $f^n = f[q(n\Delta t), n\Delta t]$

$$\frac{q^{n+1}-q^{n-m}}{(1+m)\triangle t} = \beta f^{n+1} + \alpha_n f^n + \alpha_{n-1} f^{n-1} + \ldots + \alpha_{n-l} f^{n-l}$$

Si $\beta = 0$ el método se llama *explícito* y el valor de $q^{n+1}$ se puede obtener de los valores en instantes anteriores.

Si $\beta \neq 0$ el método se llama *implícito* y hay que "despejar" $q^{n+1}$ cómo argumento de f de la ecuación en diferencias

La ecuación de arriba define una *familia de método*s, en donde tenemos la libertad de elegir los valores de "m", "l" y los coeficientes (bajo ciertas condiciones)

Para obtener las condiciones sobre los coeficientes, hacemos Taylor alrededor del instante $t = n\Delta t$

$$q'[1-(\beta+\alpha_n+\alpha_{n-1}+\ldots+\alpha_{n-l})]+\ \Delta t\, q''\left[\frac{1}{2}\left(\frac{1-m^2}{1+m}\right)-\beta+\alpha_{n-1}+2\alpha_{n-2}+\ldots+l\alpha_{n-l}\right]$$

$$+\frac{\Delta t^2}{2}q'''\left[\frac{1}{3}\left(\frac{1+m^3}{1+m}\right)-\beta-\alpha_{n-1}-4\alpha_{n-2}+\ldots-l^2\alpha_{n-l}\right]+\frac{\Delta t^3}{6}q''''\left[\frac{1}{4}\left(\frac{1+m^4}{1+m}\right)-\beta+\alpha_{n-1}+8\alpha_{n-2}+\ldots+l^3\alpha_{n-l}\right]+\ldots=\epsilon$$

donde usamos $q'=\dfrac{dq}{dt}=f\,;\ q''=f';\ldots$

y $\epsilon$ es el error de truncamiento

De pedir que $\epsilon\to 0 \ \ \text{cuando} \ \ \Delta t\to 0$

obtenemos la *condición de consistencia*

$$1=\beta+\alpha_n+\alpha_{n-1}+\ldots+\alpha_{n-l}$$

y en ese caso obtenemos

$$\epsilon=\Delta t\, q''\left[\frac{1}{2}\left(\frac{1-m^2}{1+m}\right)-\beta+\alpha_{n-1}+2\alpha_{n-2}+\ldots+l\alpha_{n-l}\right]+O[(\Delta t)^2]$$

y el esquema tiene al menos precisión de 1er orden

Tenemos todavía la posibilidad de elegir el valor de “l” y “l+1” coeficientes

Por ejemplo podríamos elegir las cosas para que el primer término del error sea 0 y obtendríamos un esquema de orden $O[(\Delta t)^2]$

En general podemos llegar hasta un orden de precisión “l+2” eligiendo convenientemente los coeficientes. En el caso de un esquema explícito queda fijo $\beta = 0$ y por lo tanto el orden de precisión puede ser hasta “l+1”

Métodos explícitos: $\beta = 0$

*Método de Euler o Adelantado:* $m = 0 \ , \ l = 0 \ \rightarrow \ \alpha_n = 1$ y el error es $\epsilon = \dfrac{\Delta t}{2} q'' + O(\Delta t^2)$

el esquema queda $$\frac{q^{n+1} - q^n}{\triangle t} = f^n + O(\Delta t) \qquad q^{n+1} = q^n + \Delta t f^n + O(\Delta t^2)$$

$\rightarrow$ error *local* $O(\Delta t^2)$ , error *global* $O(\Delta t)$

Voy de 0 a T en N pasos

$$\Delta t \, N = T \rightarrow N = \frac{T}{\Delta t}$$

$$N(\Delta t)^2 = \frac{T}{\Delta t}(\Delta t)^2 = T(\Delta t)$$

*Métodos de Adams-Bashforth:* $m = 0 \ , \ l > 0$

Si $l = 1$ , $$\frac{q^{n+1} - q^n}{\triangle t} = \alpha_n f^n + \alpha_{n-1} f^{n-1}$$

La cond. consistencia es $\alpha_n + \alpha_{n-1} = 1$ y el error $\epsilon = \triangle t \, q'' \left( \alpha_{n-1} + \dfrac{1}{2} \right) + O(\triangle t^2)$

Si elegimos $\alpha_{n-1} = -\dfrac{1}{2}$ el *esquema tiene precisión de segundo orden*. Además $\alpha_n = \dfrac{3}{2}$

y queda $$\frac{q^{n+1} - q^n}{\triangle t} = \frac{3}{2} f^n - \frac{1}{2} f^{n-1} + O(\triangle t^2)$$

AB2

(en un programa guardamos el valor previo de f, para no evaluarla dos veces por iteración)

(además hay que hacer algo especial en el 1er paso)

Con l=2 obtenemos un esquema de tercer orden de precisión y resulta $\alpha_n = \frac{23}{12}$ , $\alpha_{n-1} = -\frac{4}{3}$ , $\alpha_{n-2} = \frac{5}{12}$

$$\frac{q^{n+1} - q^n}{\triangle t} = \frac{23}{12} f^n - \frac{4}{3} f^{n-1} + \frac{5}{12} f^{n-2} + O(\triangle t^3)$$ AB3

Con l=3,

$$\frac{q^{n+1} - q^n}{\triangle t} = \frac{55}{24} f^n - \frac{59}{24} f^{n-1} + \frac{37}{24} f^{n-2} - \frac{9}{24} + O(\triangle t^4)$$ AB4

*Método "salto de rana" (leapfrog):* $m = 1$ , $l = 0$

$$\frac{q^{n+1} - q^{n-1}}{2 \triangle t} = f^n$$ con error $\epsilon = \frac{\triangle t^2}{6} q'''$

Métodos implícitos: $\beta \neq 0$

Con $m = 0\ ,\ l = 0$ $\quad \dfrac{q^{n+1} - q^n}{\triangle t} = \beta f^{n+1} + \alpha_n f^n \qquad \beta + \alpha_n = 1 \qquad \epsilon = \Delta t\, q'' \left(\frac{1}{2} - \beta\right) + O(\Delta t^2)$

Si $\beta = 1\ ,\ \alpha_n = 0$ obtenemos el *Método atrasado implícito* $\dfrac{q^{n+1} - q^n}{\triangle t} = f^{n+1} + O(\Delta t)$

de precisión de primer orden

Si tomamos $\beta = \dfrac{1}{2}\ ,\ \alpha_n = \dfrac{1}{2}$ obtenemos el *Método trapezoidal* $\dfrac{q^{n+1} - q^n}{\triangle t} = \dfrac{1}{2} f^{n+1} + \dfrac{1}{2} f^n + O(\Delta t^2)$

que es de segundo orden

Con $m = 0\ ,\ l > 0$ se obtienen los llamados *Métodos de Adams-Moulton*

Si $l = 1\ ,\ \dfrac{q^{n+1} - q^n}{\triangle t} = \dfrac{5}{12} f^{n+1} + \dfrac{8}{12} f^n - \dfrac{1}{12} f^{n-1} + O(\triangle t^3)$ AM3

Si $l = 2\ ,\ \dfrac{q^{n+1} - q^n}{\triangle t} = \dfrac{9}{24} f^{n+1} + \dfrac{19}{24} f^n - \dfrac{5}{24} f^{n-1} + \dfrac{1}{24} f^{n-2} + O(\triangle t^4)$ AM4

# Métodos iterativos

La idea es obtener $q^{n+1}$ a través de un procedimiento iterativo, en pasos, que involucran múltiples evaluaciones de la función "f"

Si son dos pasos el 1er paso suele llamarse "*predictor*" y el 2do paso se llama "*corrector*"

En general tienen la ventaja de ganar mayor precisión y de que es posible tomar pasos de tiempo mayores que con los esquemas no-iterativos. La desventaja es que al involucrar múltiples evaluaciones de "f" son más costosos computacionalmente.

Ejemplo: en el esq implícito atrasado que vimos podemos reemplazar

$$f^{n+1} = f[q^{n+1}, (n+1)\Delta t] \;\text{ por }\; f^{(n+1)*} = f[q^{(n+1)*}, (n+1)\Delta t] \;\text{ con }\; q^{(n+1)*} = q^n + \Delta t\, f^n$$

obtenido del método de Euler

$$\frac{q^{n+1}-q^n}{\triangle t} = \beta^* f^{(n+1)*} + \alpha f^n$$

Si $\beta^* = 1\,,\ \alpha = 0\,,\qquad q^{n+1} = q^n + \Delta t\ f[q^n + \Delta t f^n, (n+1)\Delta t]$

se llama el *Método de Euler-atrasado o Matsuno,*
con precisión de primer orden

Si $\beta^* = \frac{1}{2}\,,\ \alpha = \frac{1}{2}\,,\qquad q^{n+1} = q^n + \frac{\Delta t}{2}\ f[q^n + \Delta t f^n, (n+1)\Delta t] + \frac{\Delta t}{2}\ f^n$

se llama el *Método de Heun*,
con precisión de segundo orden

una variante muy utilizada es usar un "*medio paso*" en el método Matsuno, se obtiene el llamado *Método de Runge-Kutta de segundo orden* RK2

$$q^{n+1} = q^n + \Delta t\ f[q^n + \frac{\Delta t}{2} f^n, (n + \frac{1}{2})\Delta t]$$

y también el conocido *Runge-Kutta de orden 4* RK4 cuyo esquema es:

$$q^{n+1} = q^n + \frac{\Delta t}{6}(K_1 + 2K_2 + 2K_3 + K_4)$$

$$K_1 = f[q^n, n\Delta t]$$

$$K_2 = f\left[q^n + K_1\frac{\Delta t}{2}, \left(n + \frac{1}{2}\right)\Delta t\right]$$

$$K_3 = f\left[q^n + K_2\frac{\Delta t}{2}, \left(n + \frac{1}{2}\right)\Delta t\right]$$

$$K_4 = f[q^n + K_3\Delta t, (n+1)\Delta t]$$

el del teorema de la “pelota no dobla” → aerodinámica
(en simultáneo con el ruso Yukovski, claro)

(ver Métodos de Runge-Kutta de orden arbitrario)

Carl Runge, Martin Kutta → Alemania, 1900

Veamos cómo ejemplo que RK2 es efectivamente de segundo orden:

$$q^{n+1} = q^n + \Delta t \; f[q^n + \frac{\Delta t}{2} f^n, t^n + \frac{\Delta t}{2}]$$

$$q' = \frac{dq}{dt} = f(q,t) \;\; \rightarrow \;\; q' = f \;\; , \;\; q'' = f' = \frac{df}{dt} = \frac{\partial f}{\partial q}\frac{dq}{dt} + \frac{\partial f}{\partial t} = f_q q' + f_t = f_q f + f_t$$

Hacemos Taylor para $q(t) \;\; \rightarrow \;\; q^{n+1} = q^n + \frac{dq}{dt}_n \Delta t + \frac{d^2 q}{dt^2}_n \frac{\Delta t^2}{2} + O(\Delta t^3)$

$$q^{n+1} = q^n + f^n \Delta t + (f_q f + f_t)\frac{\Delta t^2}{2} + O(\Delta t^3)$$

Por otro lado haciendo Taylor en la 1er ecuación

$$q^{n+1} = q^n + \Delta t \; f^n + \Delta t \frac{\partial f}{\partial q}\frac{\Delta t}{2} f + \Delta t \frac{\partial f}{\partial t}\frac{\Delta t}{2} + O(\Delta t^3)$$

$$q^{n+1} = q^n + \Delta t \; f^n + \frac{\Delta t^2}{2}(f_q f + f_t) + O(\Delta t^3)$$

coinciden a $O(\Delta t^3)$

es decir el error en RK2 al ir de $q^n$ a $q^{n+1}$ es $O(\Delta t^3)$ → error en $\frac{dq}{dt} = \frac{q^{n+1} - q^n}{\Delta t}$ es $O(\Delta t^2)$

Si voy de 0 a T en N pasos, $\Delta t \, N = T \to N = \dfrac{T}{\Delta t}$

y el *error global* es $N(\Delta t)^3 = \dfrac{T}{\Delta t}(\Delta t)^3 = T(\Delta t)^2$

entonces el error global en RK2 es $O(\Delta t^2)$