[MN] Práctica 1 - Apunte

August 17, 2021

Métodos numéricos

Segundo cuatrimestre 2021

Apunte 1: Diferencias finitas y análisis de error

Cátedra: Pablo Dmitruk

---

1 Diferencias finitas

1.1 Aproximación asimétrica de primer orden para la primera derivada

Supongamos que tenemos una función suficientemente suave $f(x)$ cuyos valores conocemos sobre una cierta grilla equiespaciada con paso Δx, es decir, $x_k = x_0 + k\Delta x, \quad 0 \leq k \leq N-1$, con N la cantidad de puntos de la grilla. Llamemos f_k a $f(x_k)$.

Bajo estas condiciones, podemos estimar numéricamente su derivada primera utilizando, por ejemplo, el valor de f en el punto de interés x_j y un punto adicional. Naturalmente, la aproximación será, en general, más precisa si éste punto adicional se encuentra cerca de x_j.

Para encontrar una posible aproximación a la derivada, usando dos puntos, intentemos utilizar información conocida en x_j y x_{j+1}, es decir el punto de interés y el siguiente al mismo. Luego, utilizando el teorema de Taylor para escribir f_{j+1}, tenemos

$$f_{j+1} = f_j + \frac{1}{1!}f'_j h + \frac{1}{2!}f''(\xi)h^2 \qquad x_j \leq \xi \leq x_{j+1}. \tag{1}$$

donde $h = \Delta x = x_{j+1} - x_j$. Despejando la derivada de la primer ecuación obtenemos la *diferencia finita adelantada* (ya que usa información adelante del punto de interés) *de primer orden*,

$$f'_j = \frac{f_{j+1} - f_j}{h} - f''(\xi_\mathrm{f})h = \frac{f_{j+1} - f_j}{h} + \mathcal{O}(h) \qquad x_j \leq \xi_\mathrm{f} \leq x_{j+1}. \tag{2}$$

Utilizamos $\mathcal{O}(h)$ para denotar un término orden h. Finalmente, el aproximante buscado resulta

$$f'_j = \frac{f_{j+1} - f_j}{h},$$

y conlleva, en general, un error de orden h en la estimación de la derivada. De esta manera, en un entorno lo suficientemente chico alrededor de x_j, al disminuir h en un factor 2 el error disminuirá también en un factor 2. Por este motivo, se dice que el método es de primer orden. Noten, sin embargo, que el orden solo nos dice cómo se modifica (asintóticamente) el error a medida que la resolución de la grilla aumenta o disminuye, pero no nos permite calcular la magnitud del error (que dependerá, en este caso, de la magnitud de la derivada segunda en un entorno del punto).

De manera similar, utilzando expansiones de Taylor alrededor de x_{j+2}, ..., x_{j+p}, y también x_{j-1}, ..., x_{j-q} pueden obtenerse aproximaciones de mayor orden. Veamos como hacer esto para la derivada segunda.

1.2 Aproximación simétrica (centrada) de cuarto orden para la derivada segunda

Utilicemos ahora la misma idea para obtener un aproximador a la derivada segunda, utilizando información conocida en x_{j-2}, x_{j-1}, x_j, x_{j+1} y x_{j+2}. Escribiendo expresiones de Taylor análogas a las de (1) tenemos

$$f_{j-2} - f_j = -\frac{f_j^{(1)}}{1!}2h + \frac{f_j^{(2)}}{2!}(2h)^2 - \frac{f_j^{(3)}}{3!}(2h)^3 + \frac{f_j^{(4)}}{4!}(2h)^4 - \frac{f_j^{(5)}}{5!}(2h)^5 + \frac{1}{6!}f^{(6)}(\xi_{j-2})(2h)^6 \qquad x_{j-2} \leq \xi_{j-2} \leq x_{j-1}, \tag{3}$$

$$f_{j-1} - f_j = -\frac{f_j^{(1)}}{1!}h + \frac{f_j^{(2)}}{2!}h^2 - \frac{f_j^{(3)}}{3!}h^3 + \frac{f_j^{(4)}}{4!}h^4 - \frac{f_j^{(5)}}{5!}h^5 + \frac{1}{6!}f^{(6)}(\xi_{j-1})h^6 \qquad x_{j-1} \leq \xi_{j-1} \leq x_j, \tag{4}$$

$$f_{j+1} - f_j = \frac{f_j^{(1)}}{1!}h + \frac{f_j^{(2)}}{2!}h^2 + \frac{f_j^{(3)}}{3!}h^3 + \frac{f_j^{(4)}}{4!}h^4 + \frac{f_j^{(5)}}{5!}h^5 + \frac{1}{6!}f^{(6)}(\xi_{j+1})h^6 \qquad x_j \leq \xi_{j+1} \leq x_{j+1}, \tag{5}$$

$$f_{j+2} - f_j = \frac{f_j^{(1)}}{1!}2h + \frac{f_j^{(2)}}{2!}(2h)^2 + \frac{f_j^{(3)}}{3!}(2h)^3 + \frac{f_j^{(4)}}{4!}(2h)^4 + \frac{f_j^{(5)}}{5!}(2h)^5 + \frac{1}{6!}f^{(6)}(\xi_{j+2})(2h)^6 \qquad x_{j+1} \leq \xi_{j+2} \leq x_{j+2}, \tag{6}$$

donde utilizamos ahora la notación $f^{(i)}$ para denotar la i-ésima derivada.

Proponiendo ahora una combinación lineal de la forma $A(3) + B(4) + C(5) + D(6)$, y recordando que queremos obtener la mejor aproximación posible a la derivada segunda, esta combinación lineal debe verificar

$$\begin{aligned}
(-2A - B + C + 2D)\, f_j^{(1)} h &= 0, \\
\left(2A + \frac{B}{2} + \frac{C}{2} + 2D\right) f_j^{(2)} h^2 &= f_j^{(2)} h^2, \\
\left(-\frac{8}{6}A - \frac{1}{6}B + \frac{1}{6}C + \frac{8}{6}D\right) f_j^{(3)} h^3 &= 0, \\
\left(\frac{16}{24}A + \frac{1}{24}B + \frac{1}{24}C + \frac{16}{24}D\right) f_j^{(4)} h^4 &= 0,
\end{aligned}$$

que resolviendo para A, B, C y D, permite obtener

$$A = \frac{-1}{12} \qquad B = \frac{4}{3} \qquad C = \frac{4}{3} \qquad D = \frac{-1}{12}.$$

(pueden verificar este resultado acá). Notemos que esta solución anula, además, los coeficientes correspondientes a la derivada quinta, como podemos ver reemplazando

$$\left[-\frac{32}{120}\left(\frac{-1}{12}\right)-\frac{1}{120}\left(\frac{4}{3}\right)+\frac{1}{120}\left(\frac{4}{3}\right)+\frac{32}{120}\left(\frac{-1}{12}\right)\right]f_j^{(5)}h^5=0.$$

Como vieron en teóricas, este tipo de cancelaciones suceden frecuentemente para esquemas centrados.

Tenemos entonces que, por construcción,

$$(f_{j-2}-f_j)\left[\frac{-1}{12}\right]+(f_{j-1}-f_j)\left[\frac{4}{3}\right]+(f_{j+1}-f_j)\left[\frac{4}{3}\right]+(f_{j+2}-f_j)\left[\frac{-1}{6}\right]=f_j^{(2)}h^2+\mathcal{O}(h^6),$$

y despejando obtenemos

$$f_j''=\frac{-f_{j+2}+16f_{j+1}-30f_j+16f_{j-1}-f_{j-2}}{12h^2}+\mathcal{O}(h^4).$$

Vemos que, efectivamente, utilizando los valores de f sobre 5 puntos, pudimos encontrar un aproximante centrado a su derivada segunda que resulta de 4to orden.

1.3 Uso de aproximantes para construir aproximantes a derivadas superiores

Otra manera de obtener esquemas en diferencias finitas para derivadas superiores es usar recursivamente aproximantes conocidos. Por ejemplo, si conocemos la derivada primera de manera exacta en cada punto, podemos escribir una aproximación de primer orden a la derivada segunda de la siguiente manera

$$f_j''=\frac{f_{j+1}'-f_j'}{h}+\mathcal{O}(h).$$

Esto es, ni más ni menos, que la fórmula que vimos en (2), aplicada sobre la derivada segunda. Podemos ver fácilmente que no es necesario conocer f' de manera exacta, ya que hacerlo con orden h^2 alcanzará para que el error en toda la expresión anterior se mantenga de orden h.

Como demostrarán en el problema 1, el siguiente es un aproximante de segundo orden a la derivada primera

$$f_j'=\frac{f_{j+1}-f_{j-1}}{2h}+\mathcal{O}(h^2),$$

resulta en

$$f_j''=\frac{\left[\frac{f_{j+2}-f_j}{2h}+\mathcal{O}(h^2)\right]-\left[\frac{f_{j+1}-f_{j-1}}{2h}+\mathcal{O}(h^2)\right]}{h}+\mathcal{O}(h)=\frac{f_{j+2}-f_{j+1}-f_j+f_{j-1}}{2h^2}+\mathcal{O}(h).$$

Si bien este no es un aproximante óptimo considerando que utiliza valores de la función sobre 4 puntos, ilustra la idea de cómo construir esquemas de diferencias finitas para derivadas superiores a partir de aquellos conocidos para la primer derivada, siendo una alternativa a la construcción mediante expansiones de Taylor o, como vieron en teóricas, el polinomio interpolador de Lagrange.

1.4 Resolución matricial de problemas de contorno

En caso que nos interese resolver numéricamente un problema de valores de contorno lineal de primer orden, es decir,

$$\begin{aligned}\frac{\mathrm{d}u}{\mathrm{d}x} &= f(x)u + g(x), \qquad\qquad \text{para} \quad 0 < x < L \\ u(0) &= u_0,\end{aligned}$$

con f y g conocidas, una forma de hacerlo es integrando (analítica o numéricamente) el miembro derecho. Otra forma de realizar esta tarea, más relacionada con los contenidos del curso, es transformar esta ecuación diferencial en una ecuación de diferencias, usando los aproximantes que vimos en los problemas previos. Por ejemplo, si discretizamos nuestro dominio usando N puntos, $x_i = i\frac{L}{N-1}, \quad i = 0, \ldots, N-1$, y dado que conocemos u_0, podemos escribir una ecuación para u_1 usando diferencias finitas adelantadas como

$$\frac{u_1 - u_0}{h} = f_1 u_1 + g_1,$$

con $h = L/(N-1)$. Podemos despejar u_1 como

$$u_1 = -\frac{g_1 + u_0/h}{f_1 - 1/h}.$$

Siguiendo este procedimiento, podríamos hallar u_2 a partir de u_1 y así hasta u_{N-1}, aunque arrastrando un cierto error a cada paso.

Alternativamente, la misma idea puede utilizarse si reescribimos nuestro problema de contorno sobre una grilla discreta de la siguiente manera

$$D\mathbf{u} + \mathbf{b} = F\mathbf{u} + \mathbf{g},$$

con

$$D = \frac{1}{h}\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ -1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -1 & 1 \end{pmatrix}, \qquad F = \begin{pmatrix} f_1 & 0 & \ldots & 0 \\ 0 & f_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & f_{N-1} \end{pmatrix},$$

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{N-2} \\ u_{N-1} \end{pmatrix}, \qquad \mathbf{b} = \frac{1}{h}\begin{pmatrix} -u_0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \mathbf{g} = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_{N-2} \\ g_{N-1} \end{pmatrix}.$$

En particular, la matriz D es la representación matricial del operador de diferenciación y $\mathbf{b}$ contiene las condiciones de contorno (escaladas) del problema. Luego, $\mathbf{u}$ puede hallarse como la solución al sistema lineal

$$A\mathbf{u} = \mathbf{c},$$

con $A = D - F$ y $\mathbf{c} = \mathbf{g} - \mathbf{b}$. Para esto pueden utilizarse métodos clásicos de álgebra lineal (reducción gaussiana, metodos iterativos, matriz inversa, etc.).

Si bien cuando la cantidad de puntos de grilla es reducida esto funcionará aceptablemente, dado que las matrices involucradas poseen una gran cantidad de ceros (i.e. son ralas — *sparse* en inglés—), trabajar con las matrices completas se volverá muy ineficiente a resoluciones altas. Para suplir esta falencia, Numpy (y muchas otras bibliotecas de técnicas numéricas) implementan subrutinas especiales para matrices ralas. En prácticas posteriores veremos como emplearlas eficientemente.

1.5 Análisis de consistencia

Para estudiar la consistencia entre una ecuación diferencial y la correspondiente ecuación de diferencias queremos ver si, cuando disminuye el espaciamiento entre puntos, la ecuación de diferencias tiende a la ecuación diferencial. Naturalmente, la consistencia depende tanto de la ecuación diferencial bajo estudio como del esquema de diferencias propuesto.

Supongamos que nos interesa resolver la ecuación de advección 1D

$$\frac{\partial f(x,t)}{\partial t} + c\frac{\partial f(x,t)}{\partial x} = 0, \qquad 0 < t < T \quad \wedge \quad 0 < x < L.$$

Proponemos sendas grillas de N y J puntos $t^n = nT/(N-1) = n\Delta t$ y $x_j = jL/(J-1) = j\Delta x$. Noten que ahora el subíndice denota puntos del espacio y el superíndice puntos en tiempo (no potenciación). Proponemos entonces la aproximar nuestra ecuación con el esquema

$$\frac{f_j^{n+1} - f_j^{n-1}}{2\Delta t} + c\frac{f_{j+1}^n - f_{j-1}^n}{2\Delta x} = 0,$$

donde, naturalmente, $f_j^n = f(x_j, t^n)$.

Sea ahora F una solución analítica a la ecuación de advección. Reemplazando en el esquema de diferencias propuesto, obtenemos

$$\frac{F(x_j, t^{n+1}) - F(x_j, t^{n-1})}{2\Delta t} + c\frac{F(x_{j+1}, t^{n+1}) - F(x_{j-1}, t^n)}{2\Delta x} = \delta F,$$

donde δF es el ***error de consistencia*** (a veces llamado ***error de truncamiento***) y es una medida de qué tan bien o mal la solución real satisface nuestra ecuación en diferencias. Si la aproximación es buena, la solución analítica F debería satisfacerla exactamente para $\Delta t \to 0 \wedge \Delta x \to 0$, es decir, $\delta F \to 0$.

Podemos obtener un orden para δF en función de la resolución de la discretización empleando desarrollos de Taylor en t^n y x_j, es decir

$$\delta F = \frac{1}{2\Delta t}\left(\left[F + \frac{\partial F}{\partial t}\Delta t + \frac{1}{2}\frac{\partial^2 F}{\partial t^2}(\Delta t)^2 + \mathcal{O}(\Delta t)^3\right]_{x_j,t^n} - \left[F - \frac{\partial F}{\partial t}\Delta t + \frac{1}{2}\frac{\partial^2 F}{\partial t^2}(\Delta t)^2 + \mathcal{O}(\Delta t)^3\right]_{x_j,t^n}\right) +$$
$$\frac{c}{2\Delta x}\left(\left[F + \frac{\partial F}{\partial x}\Delta x + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}(\Delta x)^2 + \mathcal{O}(\Delta x)^3\right]_{x_j,t^n} - \left[F - \frac{\partial F}{\partial x}\Delta x + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}(\Delta x)^2 + \mathcal{O}(\Delta x)^3\right]_{x_j,t^n}\right).$$

Reagrupando tenemos

$$\delta F = \left[\frac{\partial F}{\partial t} + c\frac{\partial F}{\partial x}\right]_{x_j,t^n} + \mathcal{O}\left[(\Delta t)^2\right] + \mathcal{O}\left[(\Delta x)^2\right].$$

El primer término del miembro derecho se anula ya que F es solución a la ecuación de advección en todo punto, en particular, en (x_j, t^n). Obtuvimos entonces

$$\delta F = \mathcal{O}\left[(\Delta t)^2\right] + \mathcal{O}\left[(\Delta x)^2\right],$$

que tiende a 0 para $\Delta t \to 0$ y $\Delta x \to 0$ y, entonces, el esquema sugerido es consistente.

Noten que $\delta F \to 0$ no alcanza para garantizar que f converja a F sobre la grilla discreta. Además de consistente es necesario que el esquema sea estable para garantizar convergencia. Veremos más sobre esto en clases posteriores.