1. Algunas Ideas Generales sobre Métodos Numéricos

1.1 Introducción

En ciencia y tecnología son comunes los problemas para los que no es posible hallar una solución analítica. Es frecuente entonces reducir el problema a un caso particular, o simplificar el modelo de modo que pueda ser analizado. Hay, sin embargo, situaciones en que un modelo simplificado no es apropiado para describir los aspectos que son importantes en el comportamiento. Se recurre entonces a soluciones numéricas. La magnitud del trabajo es función de la precisión que se requiere. En los últimos 50 años, gracias a las computadoras digitales, las posibilidades para utilizar eficientemente los métodos numéricos han aumentado enormemente; y los puntos de vista con relación a ellos han ciertamente cambiado.

En la mayor parte de los métodos numéricos se aplican ideas relativamente simples. Una idea frecuente es la de *iteración*, es decir, la repetición de un proceso en forma tal que se obtienen cada vez mejores aproximaciones a la solución. Para ilustrar el uso de iteraciones considérese la solución de $x^3 = c$. En este caso x es la raíz cúbica de c. Esta ecuación puede reescribirse como:

$$x = \frac{1}{3}\left(2x + \frac{c}{x^2}\right)$$

Empezando con la aproximación inicial $x \approx x_0 \neq 0$, se puede iterar con:

$$x_{n+1} = \frac{1}{3}\left(2x_n + \frac{c}{x_n^2}\right)$$

Esta es una aplicación del conocido método de Newton para hallar raíces de una ecuación no lineal. Por ejemplo, para el caso $c = 2$ (es decir $x^3 = 2$) y con $x_0 = 1$ se obtienen:

$$x_1 = \frac{1}{3}\left(2 \cdot 1 + \frac{2}{(1)^2}\right) = 1.333$$

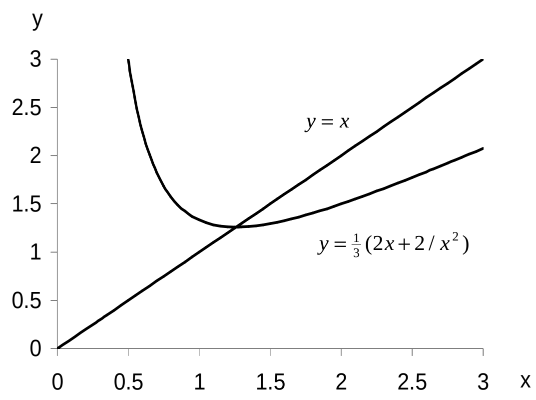
$$x_2 = \frac{1}{3}\left(2 \cdot 1.333 + \frac{2}{(1.333)^2}\right) = 1.263\,889$$

y así sucesivamente:

$$x_3 = 1.259\,933\,493\,450$$

$$x_4 = 1.259\,921\,050\,018$$

$$x_5 = 1.259\,921\,049\,895$$



Una interpretación geométrica de la iteración se muestra en la figura.

Puede en este caso probarse que el proceso converge siempre, para cualquier selección de x_0. Si x_n tiene t dígitos correctos, x_{n+1} tendrá por lo menos $2t - 1$ dígitos correctos.

Sin embargo, no todos los procesos iterativos funcionan. Por ejemplo, podría escribirse $x_{n+1} = 2/x_n^2$, lo que produce resultados alternados y obviamente no converge.

Otra idea frecuente es la de aproximar localmente una función complicada por una función lineal (o quizás parabólica u otra relativamente simple). Esto es lo que se hace al interpolar entre dos líneas de una tabla, o en procesos tales como el método de Newton – Raphson para mejorar la aproximación a una raíz de una función $f(x)=0$, la integración de una función por el método de los trapecios, la solución de una ecuación diferencial $y' = f(x,y)$ por el método de Euler, por citar sólo algunos de los métodos más conocidos.

En muchos casos se obtiene un conjunto de resultados en una sucesión de etapas, para cada una de las cuales se consideran como datos los resultados de la etapa anterior. Tales procesos se denominan de *recursión*. Son muy poderosos, pero deben ser utilizados con propiedad. La "Regla de Horner" para evaluar un polinomio tal como $p(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n$ proporciona un ejemplo simple de recursión. El polinomio $p(x)$ puede evaluarse realizando las operaciones:

$$
\begin{aligned}
p_0 &= 0 \\
p_1 &= p_0 x + a_0 \\
p_2 &= p_1 x + a_1 \\
&\vdots \\
p_n &= p_{n-1} x + a_n = p(x)
\end{aligned}
$$

La acumulación de errores en un proceso de este tipo puede ser importante.

El ejemplo siguiente ilustra también el uso de una recursión y el fenómeno conocido como "*inestabilidad numérica*". Supóngase que se requiere calcular, para $n = 0, 1, 2, \ldots$

$$y_n = \int_0^1 \frac{x^n}{x+5}\,dx$$

Puede observarse que los valores de y_n decrecen con n. Además:

$$y_n + 5y_{n-1} = \int_0^1 \frac{x^n}{x+5}\,dx + \int_0^1 \frac{5x^{n-1}}{x+5}\,dx = \int_0^1 \frac{x^{n-1}(x+5)}{x+5}\,dx = \int_0^1 x^{n-1}dx = \frac{1}{n}$$

y por lo tanto: $y_n = 1/n - 5y_{n-1}$. Esta expresión podría permitir determinar los sucesivos y_n a partir de un valor inicial, como y_0. Sabiendo que:

$$y_0 = \int_0^1 \frac{x^n}{x+5}\,dx = \left[Ln(x+5)\right]_0^1 = Ln\left(\frac{6}{5}\right) \approx 0.182$$

Se obtienen (en todos los cálculos de este ejemplo se han considerado sólo tres cifras significativas):

$$
\begin{aligned}
y_0 &\approx 0.182 \\
y_1 &= 1 - 5y_0 \approx 0.090 \\
y_2 &= \tfrac{1}{2} - 5y_1 \approx 0.050 \\
y_3 &= \tfrac{1}{3} - 5y_2 \approx 0.083 \\
y_4 &= \tfrac{1}{4} - 5y_3 \approx -0.165 \\
y_5 &= \tfrac{1}{5} - 5y_4 \approx 1.03 \quad \cdots
\end{aligned}
$$

¡Sorprendente que se obtenga $y_3 > y_2$!

¡Absurdo!

Los malos resultados se deben a que las aproximaciones y el uso de un número finito de dígitos introducen errores, que se "propagan" a etapas posteriores del cálculo. La forma en que estos errores se propagan (o disipan) es decisiva en la utilidad de un método numérico dado.

En el proceso utilizado, un pequeño error ε en y_o se multiplica por -5 en el cálculo de y_1. Sin tener en consideración los errores introducidos en los redondeos de este paso, se produce un error de $25\,\varepsilon$ en y_2. El resultado del paso k está afectado por el error inicial multiplicado por $(-5)^k$. A esto deben agregarse los efectos de los errores introducidos en todos los pasos intermedios. Si se hubieran utilizado más cifras decimales en los cálculos, los resultados absurdos habrían también aparecido, aunque un tanto más adelante en el proceso. La inestabilidad numérica puede evitarse seleccionando un algoritmo más adecuado. Así, utilizando la fórmula "en la otra dirección":

$$y_{n-1} = \frac{1}{5}\left(\frac{1}{n} - y_n\right)$$

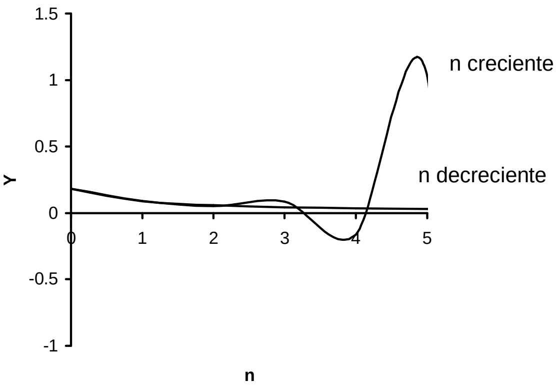
el error queda dividido por 5 en cada paso. Sabiendo que y_n decrece cuando n crece, pueden iniciarse los cálculos con algo tan pobre como $y_{10} = 0$, obteniéndose:

$$y_9 = \tfrac{1}{5}\left(\tfrac{1}{10} - 0\right) = 0.020$$

$$y_8 = \tfrac{1}{5}\left(\tfrac{1}{9} - y_9\right) \approx 0.019$$

$$y_7 = \tfrac{1}{5}\left(\tfrac{1}{8} - y_8\right) \approx 0.021$$

$$y_6 = \tfrac{1}{5}\left(\tfrac{1}{7} - y_7\right) \approx 0.025$$



y así sucesivamente:

$$y_5 \approx 0.028$$

$$y_4 \approx 0.034$$

$$y_3 \approx 0.043$$

$$y_2 \approx 0.058$$

$$y_1 \approx 0.088$$

$y_0 \approx 0.182$ ¡Correcto! (a pesar de la errada información inicial)

Sin embargo, no debe creerse que el utilizar fórmulas "al revés" es el remedio para todos los problemas numéricos. Cualquier proceso que se plantee no será siempre aplicable, ni en todos los casos el más efectivo.

1.2 Fuentes de Error

Los resultados numéricos están afectados por errores provenientes de diversas fuentes.

En primer lugar deben citarse errores en los datos, puesto que ellos son en general resultado de mediciones o estimaciones imperfectas. Es de esperar que los errores relativos en los resultados sean del mismo orden de magnitud (o menores) que aquellos de los datos. Sin embargo, éste no siempre es el caso: se dice entonces que el problema es "mal condicionado", es decir, la solución es muy sensible a pequeños errores en los datos. Dificultades de este tipo pueden también no ser debidas a la formulación del problema, sino a un mal condicionamiento del método numérico utilizado.

Un segundo grupo de errores es debido a simplificaciones en el modelo matemático del problema y a la truncación de expresiones (series por ejemplo), cuyo objetivo es evitar que la formulación se complique más allá de lo que razonablemente puede manejarse.

Más importantes desde el punto de vista de los métodos numéricos son los errores de truncación y redondeo. Éstos son función del procedimiento empleado y de las características de operación de la computadora. La mayor parte de las computadoras trabajan internamente con sistemas de numeración binarios, octales o hexadecimales y tienen dos “tipos” de aritmética: de punto fijo (o “enteros”) y de punto flotante (o “reales”). La aritmética de punto fijo es exacta, pero está limitada a números enteros y a un rango pequeño. En consecuencia, la mayor parte de las operaciones se efectúan con la aritmética de punto flotante. En la aritmética de punto flotante la representación interna de un número es de la forma: $a = m \cdot 10^q$, donde m es la *mantisa* y q el *exponente*. Sólo se almacenan t cifras (en base b) de la mantisa, y por lo tanto cualquier número puede ser representado con un error relativo que no excede $\frac{1}{2} b^{t-1}$ (habitualmente entre 10^{-6} y 10^{-15}). Para q se usa un número finito de posiciones de memoria y en consecuencia existe un “rango” aceptable (en general muy grande) para los números con punto flotante.

Las operaciones aritméticas en punto flotante tienen propiedades algo diferentes de aquellas correspondientes en la aritmética “exacta”. Así por ejemplo, la suma (o resta) no es estrictamente asociativa.

$$a = 0.1234567 \cdot 10^0$$
$$b = 0.123567 \cdot 10^4$$
$$c = -b$$

El esquema siguiente indica como se efectúa la suma en “punto flotante”:

$$b \rightarrow 0.1234567 \cdot 10^4$$
$$a \rightarrow 0.0000123 \cdot 10^4$$ (las cuatro cifras finales se recortan)
$$a + b \rightarrow 0.1234690 \cdot 10^4$$
$$c \rightarrow -0.1234567 \cdot 10^4$$
$$(a+b)+c \rightarrow 0.0000123 \cdot 10^4 = 0.1230000$$

mientras que $(b+c)+a = 0.1234567 \cdot 10^0$. El orden de las operaciones sí afecta los resultados.

Esto es válido también para operaciones de otro tipo. Por ejemplo, las raíces de $x^2 + 2bx + c = 0$ podrían obtenerse de: $x = -b \pm \sqrt{b^2 - c}$. Sin embargo el proceso alternativo (y teóricamente equivalente):

$$x_1 = -b - (signo\ \ b)\sqrt{b^2 - c}$$
$$x_2 = \frac{c}{x_1}$$

tiene mucho menos acumulación de error, especialmente cuando c es pequeño, porque evita la resta de dos números del mismo orden de magnitud. Considérese, por ejemplo, la ecuación: $x^2 - 64x + 1 = 0$. Trabajando con 5 cifras significativas:

$$x_1 = 32 + \sqrt{1023} \approx 32 + 31.984 = 63.984 \pm \tfrac{1}{2} \cdot 10^{-3}$$
$$x_2 = 32 - \sqrt{1023} \approx 32 - 31.984 = 0.016 \pm \tfrac{1}{2} \cdot 10^{-3}$$

El error relativo en x_2 es muy grande. La resta se ha hecho en forma exacta; la causa del error está más bien en el redondeo previo de la raíz cuadrada. Si en cambio se toma $x_2 = 1/x_1$ se obtiene:

$$x_2 = \frac{1}{63.984 \pm \frac{1}{2} \cdot 10^{-3}} = 0.015629 \pm 0.0000005$$

con un error relativo del mismo orden que el de x_1.

Finalmente, deben mencionarse errores humanos y errores de la computadora. Estos últimos son prácticamente inexistentes, los primeros son en cambio la causa de muchos resultados inesperados.