ANÁLISIS DE ERROR

Máster de Modelización e Investigación Matemática, Estadística y Computación

Programación científica

Parte II

Curso 2014/2015

Contenido

Contenido

1.1 Introducción

- El tratamiento numérico de expresiones algebraicas matemáticamente correctas puede en ciertas ocasiones conducir a resultados erróneos.
- Esta paradoja encuentra su origen en el siguiente hecho: en el cálculo algebraico todas las operaciones simbolizadas se suponen hechas con una precisión infinita, mientras que en el cálculo numérico sobre el ordenador, el problema de la precisión aparece porque todas las operaciones son efectuadas con una precisión limitada.
- Es necesario analizar algunos de los errores asociados al planteamiento del problema que deseamos resolver, así como aquellas cuestiones que nos ayuden a controlar y evitar los errores que se producen en su implementación/resolución.

Contenido

1.2 Errores progresivos (forward) y regresivos (backward)

- Denotemos por $\hat{y}$ una aproximación de $y = f(x)$ computada con aritmética de precisión finita. ¿Cómo medir la "calidad" con la que $\hat{y}$ representa a y?
- Habitualmente existen dos forma de medir esta distancia: error forward (progresivo) y error backward (regresivo).
- El error forward es una medida de la diferencia entre la aproximación $\hat{y}$, y el valor real y. Esto es
  - Error forward absoluto: $|\hat{y} - y|$,
  - Error forward relativo: $|\hat{y} - y|/|y|$.

- El error forward es una forma “natural” de medir, pero por lo general no conocemos el verdadero valor de y, por lo que sólo podemos conseguir una cota superior de dicho error.
- El error backward considera que $\hat{y}$ es la solución exacta de un problema perturbado e intenta medir dicha perturbación, esto es $\hat{y} = f(x + \Delta x)$.
- El valor $|\Delta x|$ o $|\Delta x|/|x|$ es llamado error backward.

- **Ejemplo**: Se quiere computar $y = \sqrt{2}$, y se obtiene de manera aproximada que $\hat{y} = 1.4$, entonces
  - Error forward: $|\hat{y} - y| = |1.4 - 1.4142\dots| \approx 0.0142$.
  - Error backward: Si observamos que $\sqrt{1.96} = 1.4$, entonces $|\Delta x| = |2 - 1.96| = 0.04$.
- La siguiente figura ilustra la diferencia entre el error forward y el error backward en el estudio de la aproximación de $f(x)$.

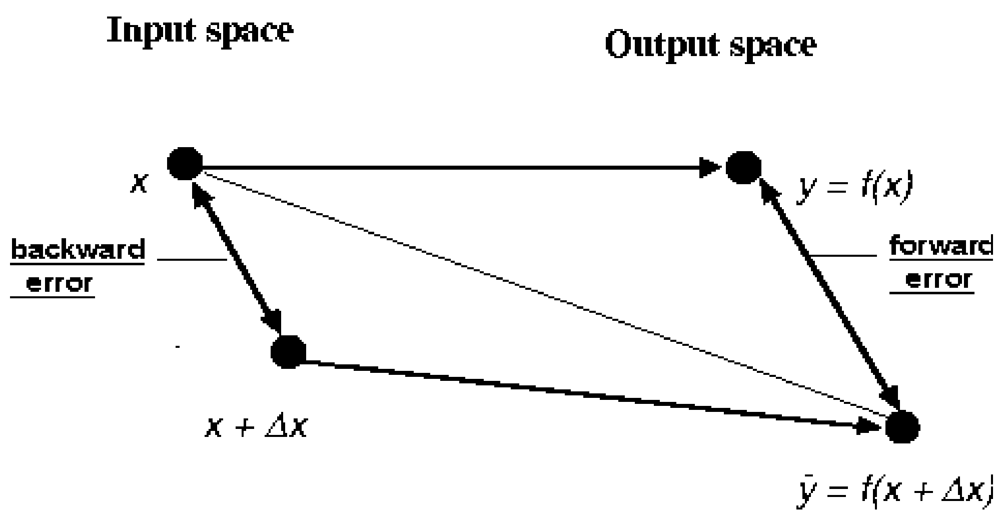

Figura 1: Error forward y backward.

Contenido

1.3 Condicionamiento y estabilidad

- Un *problema* matemático con entrada x y salida $y = f(x)$ se dice bien condicionado si “pequeños” cambios en x provocan “pequeños” cambios en y, por el contrario, se dice que está mal condicionado cuando “pequeños” cambios en x provocan “grandes” cambios en y.
- ¿Cómo puede medirse el condicionamiento?: utilizando el número de condición, κ. Este número mide la sensibilidad del problema a pequeñas perturbaciones de x.

$$\kappa = \frac{\|f(x+\Delta x) - f(x)\|}{\|\Delta x\|}.$$

- Conviene observar que el número condición puede calcularse de forma aproximada utilizando el primer término del desarrollo de f en serie de Taylor, esto es

$$f(x+\Delta x) \approx f(x) + f'(x)\Delta x \Rightarrow \frac{f(x+\Delta x) - f(x)}{\Delta x} \approx f'(x),$$

de donde $\kappa \approx \|f'(x)\|$.

- Si lo que se quiere es una idea de la relación entre los errores relativos, podemos obtener el número de condición relativo sin más que observar que

$$\frac{\|f(x+\Delta x) - f(x)\|}{\|f(x)\|} \approx \frac{\|\Delta x\|}{\|x\|} \cdot \frac{\|f'(x)\|\,\|x\|}{\|f(x)\|},$$

de donde

$$\left(\frac{\|f(x+\Delta x) - f(x)\|}{\|f(x)\|}\right) \Big/ \left(\frac{\|\Delta x\|}{\|x\|}\right) \approx \frac{\|f'(x)\|\,\|x\|}{\|f(x)\|}.$$

- **Ejemplos**:
  - $f(x) = \ln(x)$:
$$\kappa \approx \left|\frac{xf'(x)}{f(x)}\right| = \left|\frac{x(1/x)}{\ln(x)}\right| = \left|\frac{1}{\ln(x)}\right|.$$
  - $f(x) = x/2$:
$$\kappa \approx \left|\frac{1/2}{1/2}\right| = 1.$$
  - $f(x) = \sqrt{x}$:
$$\kappa \approx \left|\frac{1}{2}\right| = \frac{1}{2}.$$

- **Ejemplo (mal condicionado)**: Sea el polinomio $p(x) = (x+1)(x+2)(x+3)\cdots(x+19)(x+20)$, cuyas raíces son $x = -1, -2, \ldots, -19, -20$. Si sumamos al coeficiente de x^{19} la cantidad $2^{-23} = 1.1921 \times 10^{-7}$ se obtienen como raíces del nuevo polinomio
$$\begin{array}{lll} -1, & -6, & -10.0953 \pm 0.6435i, \\ -2, & -6.9997, & -11.7936 \pm 1.6523i, \\ -3, & -8.0073, & -13.9924 \pm 2.5188i, \\ -4, & -8.9173, & -16.7307 \pm 2.8126i, \\ -5, & -20.8469, & -19.5024 \pm 1.9403i. \end{array}$$
- Es evidente que no sólo se ha producido un cambio sustancial en algunas raíces, sino que algunas se han transformado en complejas.

- Un *algoritmo* se dice que es estable si el resultado que produce es "poco" sensible a los errores cometidos durante los cálculos. Es decir, el resultado que produce es la solución exacta de un problema "próximo" al de partida.
- La estabilidad de un algoritmo no garantiza una solución computada precisa.
- La exactitud depende del condicionamiento del problema y la estabilidad del algoritmo.
- Un algoritmo se dice que es forward estable si la solución computada que produce está próxima a la solución exacta, mientras que es backward estable cuando la solución que produce es la solución exacta de un problema próximo al original.

- Cuando el error backward, el error forward y el número de condición están definidos, podemos establecer la siguiente la relación

$$\texttt{error forward} \lesssim \texttt{número de condición} \cdot \texttt{error backward}.$$

- Una forma de interpretar esta regla es: "*the computed solution to an ill-conditioned problem can have a large forward error*".
- Así, cuando la solución computada tiene un error backward pequeño, este se puede ver amplificado a través del número de condición para obtener el error forward,
- La estabilidad backward implica la estabilidad forward, pero no al revés.

- **Ejemplo (algoritmo inestable)**: Supongamos que queremos calcular las raíces del polinomio ax^2+bx+c, utilizando para ello la bien conocida expresión

$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}.$$

¿Qué ocurre Si $b>0$ y $4ac<<b^2$?, ¿existe perdida de precisión?

- Si $b>0$ y $4ac<<b^2$, ocurrirá que $\sqrt{b^2-4ac}\simeq b$, teniéndose así en el cálculo de x_2 la diferencia de dos números casi iguales, por lo que se puede perder precisión. Para evitar este problema se recurre a calcular x_2 utilizando la expresión

$$x_2=\frac{(-b+\sqrt{b^2-4ac})(-b-\sqrt{b^2-4ac})}{(2a)(-b-\sqrt{b^2-4ac})}=\frac{-2c}{b+\sqrt{b^2-4ac}}.$$

1.3 Error en la resolución de un sistema de ecuaciones

- A continuación presentaremos algunos resultados sobre el análisis del error que se produce en la resolución de un sistema de ecuaciones lineales ($A\cdot x=b$) utilizando precisión finita.
- El estudio se planteará desde dos puntos de vista:
  - Número de condición en la resolución de un sistema de ecuaciones.
  - Error backward en la eliminación de Gauss.

1.3.2 Número de condición

- En la resolución de un sistema lineal $A\,x = b$ utilizando precisión finita, en general no obtendremos la solución exacta del sistema sino una solución aproximada que denotaremos por $\hat{x}$.
- Notación:

$$A x = b \;\rightarrow\; \begin{cases} \hat{x} & : \text{solución aproximada} \\ r = b - A\hat{x} & : \text{vector residuo} \\ e = x - \hat{x} & : \text{vector error absoluto} \end{cases}$$

- El residuo es una medida de hasta qué punto $\hat{x}$ satisface el sistema $A x = b$, pero ¿es fiable el residuo como medida del error?

- Sea el sistema lineal:

$$\begin{array}{rcrcl} 1.01\,x_1 & + & 0.99\,x_2 & = & 2 \\ 0.99\,x_1 & + & 1.01\,x_2 & = & 2 \end{array}$$

cuya solución exacta es $x_1 = x_2 = 1$.

$$\hat{x} = \begin{pmatrix} 1.01 \\ 1.01 \end{pmatrix},\; e = \begin{pmatrix} -0.01 \\ -0.01 \end{pmatrix},\; r = \begin{pmatrix} -0.02 \\ -0.02 \end{pmatrix},\; r \text{ “pequeño” y } e \text{ “pequeño”}$$

$$\hat{x} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad e = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad r = \begin{pmatrix} -0.02 \\ 0.02 \end{pmatrix}, \quad r \text{ “pequeño” y } e \text{ “grande”}$$

Definición 1.1

(Norma matricial) Una norma matricial es una aplicación:

$$\|\cdot\| : M_n(\mathbb{R}) \to \mathbf{R}^+ \cup \{0\}$$

que verifica las siguientes propiedades $\forall\ A, B \in M_n(\mathbb{R}),\ \forall\ \lambda \in \mathbb{R}$:

P1 $\|A\| = 0 \ \Leftrightarrow\ A = 0$,

P2 $\|A\| \geq 0$,

P3 $\|A + B\| \leq \|A\| + \|B\|$,

P4 $\|\lambda A\| = |\lambda| \|A\|$,

P5 $\|AB\| \leq \|A\| \|B\|$.

Definición 1.2

Una norma matricial $\|\cdot\|_M$ *en* $M_n(\mathbb{R})$ *y una norma vectorial* $\|\cdot\|_v$ *en* $\mathbb{R}^n$ *se dice que son compatibles si se verifica:*

$$\|Ax\|_v \leq \|A\|_M \|x\|_v,\ \forall\ A \in M_n(\mathbb{R}),\ \forall\ x \in \mathbb{R}^n.$$

Ejemplos de normas matriciales:

- $\|A\|_1 = \max\limits_{1 \leq j \leq n} \left\{ \sum\limits_{i=1}^{n} |a_{ij}| \right\}$ compatible con $\|x\|_1 = \sum\limits_{i=1}^{n} |x_i|$.
- $\|A\|_2 = \rho(A^t A)^{1/2}$ compatible con $\|x\|_2 = \left(\sum\limits_{i=1}^{n} |x_i|^2 \right)^{1/2}$.
- $\|A\|_\infty = \max\limits_{1 \leq i \leq n} \left\{ \sum\limits_{j=1}^{n} |a_{ij}| \right\}$ compatible con $\|x\|_\infty = \max\limits_{1 \leq j \leq n} |x_j|$.

Teorema 1.3

Sea A una matriz no singular y $Ax = b$ un sistema lineal de ecuaciones. Entonces:

$$\frac{1}{\kappa(A)}\frac{\|r\|}{\|b\|} \leq \frac{\|e\|}{\|x\|} \leq \kappa(A)\frac{\|r\|}{\|b\|}, \quad \kappa(A) \equiv \|A\|\|A^{-1}\|$$

donde las normas vectorial y matricial han de ser compatibles.

Al número real $\kappa(A)$ se le llama condicionamiento o número de condición de A.

- **Observaciones:**
  - $\kappa(A) \geq 1$:

$$1 = \rho(I) \leq \|I\| = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| = \kappa(A).$$

  - Si $\kappa(A) \approx 1$, el residuo relativo es una buena estimación del error relativo, ya que $\kappa(A)$ y $1/\kappa(A)$ serán muy próximos. Se dice en este caso que el sistema está bien condicionado.
  - Si $\kappa(A) \gg 1$, entones $\kappa(A)$ y $1/\kappa(A)$ serán muy diferentes. En este caso no podemos asegurar que el residuo relativo sea una buena estimación del error relativo. Se dice en este caso que el sistema está mal condicionado.

1.3.3 Error backward en la eliminación de Gauss

- Sea $Ax = b$ un sistema de ecuaciones lineales con $A = (a_{ij})_{1\leq i,j\leq n}$.
- Considerando que la solución computada con aritmética de punto flotante $\hat{x}$ es la solución exacta de un sistema perturbado ($(A+H)\hat{x} = b$ o $\hat{L}\hat{U} = A+E$), podemos plantearnos dos formas distintas de medir la perturbación, es decir cuantificar tanto E como H.
- Podemos medir las perturbaciones utilizando normas (*normwise*), apareciendo así acotaciones de $\|E\|$ y $\|H\|$, o acotando el valor absoluto de E y H elemento a elemento, lo que se conoce como error backward respecto de las componentes (*componentwise*).

- A continuación presentaremos algunos de los resultados más conocidos en el análisis del error backward para la eliminación gaussiana.
- Wilkinson (pivotaje parcial o total): La solución computada $\hat{x}$ satisface $(A+H)\hat{x} = b$, donde

$$\|H\|_\infty \leq \rho_n p(n) u \|A\|_\infty,$$

con $p(n)$ un polinomio de grado tres, u la unidad de redondeo y ρ_n el factor de crecimiento definido por

$$\rho_n = \rho_n(A) = \frac{\max\limits_{i,j,k} |a_{ij}^{(k)}|}{\max\limits_{i,j} |a_{ij}|},$$

donde los $a_{ij}^{(k)}$ son los elementos que aparecen en el paso k-ésimo de la eliminación, tras haber hecho ceros en las $k-1$ primeras columnas.

- Forsythe y Moler (pivotaje parcial): $\hat{L}\hat{U} = A + E$ con E verificando

$$\|E\|_\infty \leq n^2 \rho u \|A\|_\infty,$$

siendo

$$\rho = \frac{\max\limits_{i,j,k} |a_{ij}^{(k)}|}{\|A\|_\infty}.$$

Además, $(A+H)\hat{x} = b$, donde

$$\|H\|_\infty \leq 1.01(n^3 + 3n^2)\rho u \|A\|_\infty.$$

- De Boor y Pinkus (sin pivotaje): La solución computada $\hat{x}$ verifica $(A+H)\hat{x} = b$, con

$$|H| \leq \gamma_n(2+\gamma_n)|\hat{L}||\hat{U}|,$$

donde $\hat{L}$ y $\hat{U}$ son las matrices computadas por eliminación gaussiana y $\gamma_n = \dfrac{nu}{1-nu}$.

- Higham (sin pivotaje): Referencia el resultado de de Boor y Pinkus, comprobando previamente que $\hat{L}\hat{U} = A + E$, con E verificando

$$|E| \leq \gamma_n |\hat{L}||\hat{U}|.$$

Contenido

1.4 Bibliografía

- G.E. Forsythe, C.B. Moler, *Solución mediante computadoras de sistemas algebraicos lineales, Editorial universitaria de Buenos Aires (1973).*
- G.H. Golub, C.F. Van Loan, *Matrix computations* (3rd ed.), Johns Hopkins University Press, 1996.
- N.J. Higham, *Accuracy and Stability of Numerical Algorithms* (2nd ed.), Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.
- J.H. Wilkinson, *Rounding Errors in Algebraic Processes*, Notes on Applied Science 32 (1963), Her Majesty Stationery Office.