Condicionamiento y Estabilidad II

Alan Reyes-Figueroa

Métodos Numéricos II (Aula 05) 19.julio.2022

IEEE 754 (estándar):

Representación puntual de $\mathbb{R}$. En un sistema de punto flotante, las brechas entre los números representados adyacentes escalan en proporción a la magnitud de los números.
Tenemos base$=\beta=2$, precisión $=t$ ($t=24$ precisión simple, $t=53$ precisión doble).

La representación de un número es de la forma

$$x=\left(\frac{m}{\beta^t}\right)\beta^e,$$

donde m es un número entero en el rango $\beta^{t-1}\leq m\leq\beta^t-1$, y e es un entero arbitario. La cantidad $\frac{m}{\beta^t}$ se conoce entonces como la **mantisa**, mientras que e es el **exponente**.

Punto Flotante

La resolución de la máquina se resume tradicionalmente en un número conocido como el **épsilon de máquina**

$$\varepsilon_{maq} = \frac{1}{2}\beta^{1-t}. \tag{1}$$

Este número es la mitad de la distancia entre 1 y el siguiente número de punto flotante representable. En un sentido relativo, este es tan grande como los espacios entre los números de punto flotante. Es decir, ε_{maq} tiene la siguiente propiedad:

$$\forall x \in \mathbb{R}, \text{ existe } x' \text{ representable, tal que } |x - x'| < \varepsilon_{maq}|x|. \tag{2}$$

Denotamos por $fl : \mathbb{R} \to \mathbf{F}$ la función que da la aproximación más cercana de punto flotante a un número real. Esto es, $fl(x)$ es el equivalente a x redondeado en el sistema de punto flotante.

La desigualdad (2) se expresa en términos de *fl* como

$$\forall x \in \mathbb{R}, \text{ existe } \varepsilon \text{ con} |\varepsilon| < \varepsilon_{maq} \text{ tal que } fl(x) = x(1+\varepsilon). \tag{3}$$

Es decir, la diferencia entre un número real y su punto flotante más cercano, siempre menor que ε_{maq} (en términos relativos).

Operaciones de punto flotante: Denotamos las operaciones $+, -, \cdot$ y $\div$ por $x \oplus y = fl(x+y)$, $x \ominus y = fl(x-y)$, $x \odot y = fl(x \cdot y)$, $x \oslash y = fl(x/y)$.

Teorema (Axioma fundamental aritmética de punto flotante)

Para todo $x, y \in \mathbf{F}$, *existe* ε *con* $|\varepsilon| < \varepsilon_{maq}$ *tal que*

$$x \odot y = (x \cdot y)(1+\varepsilon). \tag{4}$$

Así, cada operación aritmética en punto flotante es exacta, hasta un error relativo máximo del tamaño de ε_{maq}.

Estabilidad

Hemos definido un problema matemático como una función $f : X \to Y$ desde un espacio vectorial X de datos a un espacio vectorial Y de soluciones.
Un algoritmo puede verse como otro mapa $\tilde{f} : X \to Y$.

Más precisamente, sea f es un problema, y dado un computador cuyo sistema de punto flotante satisface (4), un **algoritmo** para f (en el sentido amplio del término), y una implementación de este algoritmo en forma de programa informático. Dado un dato $\mathbf{x} \in X$, estos datos se redondean y se alimentan como entrada en el algoritmos $\tilde{f}$. Al correr el programa, el resultado es una colección de números de punto flotante que pertenecen al espacio vectorial Y. Denotamos este resultado por $\tilde{f}(\mathbf{x})$.

En el mínimo caso, $\tilde{f}(\mathbf{x})$ se verá afectado por errores de redondeo (pero existen otros posibles problemas que pueden afectar $\tilde{f}(\mathbf{x})$).

Así como $\tilde{f}$ es el análogo calculado de f, otras cantidades calculadas se marcarán por tildes. Por ejemplo, la solución calculada del sistema $A\mathbf{x} = \mathbf{b}$ se denotará por $\tilde{\mathbf{x}}$.

Estabilidad

Típicamente, $\tilde{f}$ no es continua, pero aún así un buen algoritmo debe aproximarse al problema asociado f. Consideramos el **error absoluto** de un cálculo, $||\tilde{f}(\mathbf{x}) - f(\mathbf{x})||$, o el **error relativo**, $\frac{||\tilde{f}(\mathbf{x})-f(\mathbf{x})||}{||f(\mathbf{x})||}$.

Definición

*Decimos que un algoritmo $\tilde{f}$ para un problema f es **preciso** si para cada $\mathbf{x} \in X$,*

$$\frac{||\tilde{f}(\mathbf{x}) - f(\mathbf{x})||}{||f(\mathbf{x})||} = O(\varepsilon_{maq}). \tag{5}$$

Sin embargo, si el problema f está mal condicionado, el objetivo de precisión definido por (5) es muy ambicioso. En ese caso, es mejor dar una definición alternativa para la exactitud de un algoritmo.

UVG UNIVERSIDAD DEL VALLE DE GUATEMALA

Definición

*Un algoritmo $\tilde{f}$ para un problema f es **estable** si para cada $\mathbf{x} \in X$,*

$$\frac{||\tilde{f}(\mathbf{x}) - f(\tilde{\mathbf{x}})||}{||f(\tilde{\mathbf{x}})||} = O(\varepsilon_{maq}), \tag{6}$$

para alguna $\tilde{\mathbf{x}}$ con $\frac{||\tilde{\mathbf{x}}-\mathbf{x}||}{||\tilde{\mathbf{x}}||} = O(\varepsilon_{maq})$.

Muchos algoritmos de álgebra lineal numérica satisfacen una condición que es a la vez más fuerte y simple que la estabilidad.

Definición

*Decimos que un algoritmo $\tilde{f}$ para un problema f es **estable hacia atrás** (backward stable) si para cada $\mathbf{x} \in X$,*

$$\tilde{f}(\mathbf{x}) = f(\tilde{\mathbf{x}}), \quad \textit{para algún } \tilde{\mathbf{x}} \textit{ con } \frac{||\tilde{\mathbf{x}}-\mathbf{x}||}{||\tilde{\mathbf{x}}||} = O(\varepsilon_{maq}). \tag{7}$$

Estabilidad

Obs! En cualquier aritmética de máquinas, el número ε_{maq} es una cantidad fija. Al hablar del límite $\varepsilon_{maq} \to 0$ estamos considerando una idealización de un computador. Las ecuaciones (5)-(7) hablan de la rapidez con la que la solución calculada del algoritmo $\tilde{f}$ tiende a la solución del problema f, a medida que la precisión de la máquina se mejora (de forma hipotética).