Sistemas de Ecuaciones Lineales IV

- **Propagación de errores:** Número de condición. Estimación a posteriori del error.

Propagación de errores en los datos (segundo miembro)

- Consideremos el siguiente ejemplo:

$$\begin{pmatrix} -0.10 & 1.00 \\ 0.11 & -1.00 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2.0 \\ 2.1 \end{pmatrix}. \qquad \text{Solución exacta:} \quad \boldsymbol{x} = \begin{pmatrix} 10.0 \\ -1.0 \end{pmatrix}.$$

- Supongamos que el segundo miembro $\boldsymbol{b}$ ha sido redondeado a un dígito decimal y que su valor redondeado a dos dígitos decimales es $\boldsymbol{b} = (-2.00, 2.14)^{\mathrm{t}}$. La solución de este nuevo sistema es:

$$\begin{pmatrix} -0.10 & 1.00 \\ 0.11 & -1.00 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2.00 \\ 2.14 \end{pmatrix}. \qquad \text{Solución exacta:} \quad \boldsymbol{x} = \begin{pmatrix} 14.0 \\ -0.6 \end{pmatrix}.$$

- **Notemos que un error de menos del 2% en los datos produjo un error del 40% en la solución!!!**

Propagación de errores en los datos (matriz)

- Consideremos de nuevo el mismo ejemplo:

$$\begin{pmatrix} -0.10 & 1.00 \\ 0.11 & -1.00 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2.0 \\ 2.1 \end{pmatrix}. \qquad \text{Solución exacta:} \quad \boldsymbol{x} = \begin{pmatrix} 10.0 \\ -1.0 \end{pmatrix}.$$

- Supongamos ahora que la entrada a_{11} de la matriz es en realidad -0.102 en lugar de -0.10 . La solución de este nuevo sistema es:

$$\begin{pmatrix} -0.102 & 1.00 \\ 0.11 & -1.00 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2.00 \\ 2.1 \end{pmatrix}. \qquad \text{Solución exacta:} \quad \boldsymbol{x} = \begin{pmatrix} 12.500 \\ -0.725 \end{pmatrix}.$$

- Esta vez un error del 2% en los datos produjo un error de alrededor del 25% en la solución.

  Otra vez, los errores en los datos se amplificaron, pese a que no se ha usado método numérico alguno!

Propagación de errores en el segundo miembro

- Supongamos que se quiere resolver el sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, pero que del segundo miembro $\boldsymbol{b}$, sólo se conoce un valor aproximado $\boldsymbol{b}' = \boldsymbol{b} + \delta\boldsymbol{b}$, sujeto a errores $\delta\boldsymbol{b}$.
- Sea $\boldsymbol{x}'$ la solución exacta del sistema $\boldsymbol{A}\boldsymbol{x}' = \boldsymbol{b}'$.
- Sea $\delta\boldsymbol{x} = \boldsymbol{x}' - \boldsymbol{x}$, el error de la solución aproximada $\boldsymbol{x}'$. Entonces,

$$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \left(\|\boldsymbol{A}\| \left\|\boldsymbol{A}^{-1}\right\|\right) \frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}.$$

Número de condición

- Se define el **número de condición** de la matriz $\boldsymbol{A}$ como

$$\mathrm{cond}(\boldsymbol{A}) := \|\boldsymbol{A}\| \left\|\boldsymbol{A}^{-1}\right\|.$$

- Dado que

  $\dfrac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|}$: **error relativo de la solución aproximada x'**,

  $\dfrac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}$: **error relativo del segundo miembro b'**,

  $$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \left(\|\boldsymbol{A}\| \left\|\boldsymbol{A}^{-1}\right\|\right) \frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|} = \mathrm{cond}(\boldsymbol{A}) \frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|},$$

  el número de condición $\mathrm{cond}(\boldsymbol{A})$ es una cota del factor de propagación entre el error relativo del dato b' y el error relativo de la solución aproximada x'.

- Para toda matriz no singular $\boldsymbol{A}$, $\mathrm{cond}(\boldsymbol{A}) \geq 1$.

Propagación de errores en los datos de un sistema

- El número de condición de la matriz de un sistema también rige la propagación de errores en la misma matriz.

- **Teorema.** Si $\boldsymbol{Ax} = \boldsymbol{b}$ y $(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = (\boldsymbol{b} + \delta\boldsymbol{b})$, con $\boldsymbol{A}$ no singular y $\delta\boldsymbol{A}$ suficientemente pequeño como para que $\operatorname{cond}(\boldsymbol{A})\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|} < 1$, entonces,

$$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \frac{\operatorname{cond}(\boldsymbol{A})}{1 - \operatorname{cond}(\boldsymbol{A})\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}}\left(\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|} + \frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}\right).$$

- La capacidad de propagar errores en los datos de un sistema $\boldsymbol{Ax} = \boldsymbol{b}$ depende esencialmente de la matriz $\boldsymbol{A}$ y no del segundo miembro $\boldsymbol{b}$.

- Si $\operatorname{cond}(\boldsymbol{A}) \approx 1$ decimos que el sistema está **bien condicionado**.

- Si $\operatorname{cond}(\boldsymbol{A}) \gg 1$ decimos que el sistema está **mal condicionado**.

Propagación de errores de redondeo

- El número de condición de la matriz también determina la capacidad de propagar los errores de redondeo cuando el sistema se resuelve computacionalmente.

- **Teorema.** Sea $\boldsymbol{x} \in \mathbb{R}^n$ la solución del sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ con $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ no singular y $\boldsymbol{b} \in \mathbb{R}^n$.

  Sea $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ la solución calculada mediante el método de **eliminación gaussiana con estrategia de pivoteo parcial**, en un computador con **constante de precisión** ε.

  Sea $p = \dfrac{1}{\|\boldsymbol{A}\|_\infty} \max\limits_{1\le i,j,k\le n} \left\{ \left| a_{ij}^{(k)} \right| \right\}$.

  Entonces, si el sistema no está demasiado mal condicionado de manera que $\varepsilon \operatorname{cond}_\infty(\boldsymbol{A}) < 1$, se tiene la siguiente acotación del error relativo de la solución calculada:

$$\frac{\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|_\infty}{\|\boldsymbol{x}\|_\infty} \le 2n^3 p\varepsilon \frac{\operatorname{cond}_\infty(\boldsymbol{A})}{1 - \varepsilon \operatorname{cond}_\infty(\boldsymbol{A})}.$$

Propagación de errores de redondeo (cont.)

- La cota del teorema nos dice que si el error de la solución calculada $\hat{\boldsymbol{x}}$ es muy grande, necesariamante se debe a alguna (o a varias) de estas razones:

  1. la **dimensión del sistema** n es muy grande;
  2. la **constante de precisión del computador** ε no es suficientemente pequeña;
  3. el **número de condición de la matriz** $\mathrm{cond}(\boldsymbol{A})$ es demasiado grande.

- ¿Quién es el culpable?

$$\boxed{\text{solución calculada mala}} \implies \left\{ \begin{array}{l} \boxed{\text{dimensión del problema excesivamente grande}} \\ \boxed{\text{mala precisión del computador (pocos dígitos)}} \\ \boxed{\textbf{sistema mal condicionado}} \end{array} \right.$$

- Teoremas análogos valen para el método de Cholesky, si la matriz $\boldsymbol{A}$ es simétrica y definida positiva, y para el algoritmo de Thomas, si la matriz $\boldsymbol{A}$ es tridiagonal y satisface las desigualdades $|b_1| > |c_1|$, $|b_n| > |a_n|$ y $|b_i| \geq |a_i| + |c_i|$, $i = 2, \ldots, n-1$.

Estimación a posteriori del error de una solución calculada

- Sea $\boldsymbol{x}$ la solución exacta del sistema $\boldsymbol{Ax} = \boldsymbol{b}$ y sea $\hat{\boldsymbol{x}} \approx \boldsymbol{x}$ una solución calculada.
- Sea $\boldsymbol{e} := \boldsymbol{x} - \hat{\boldsymbol{x}}$ el **error** de la solución calculada $\hat{\boldsymbol{x}}$.
- Sea $\boldsymbol{r} := \boldsymbol{b} - \boldsymbol{A}\hat{\boldsymbol{x}}$ el **residuo** de la solución calculada $\hat{\boldsymbol{x}}$.
- Si el residuo fuera nulo ($\boldsymbol{r} = 0$), la solución calculada coincidiría con la exacta.
- Si no, ¿cómo se puede estimar el error $\boldsymbol{e}$? Notemos que

$$\boldsymbol{Ae} = \boldsymbol{A}(\boldsymbol{x} - \hat{\boldsymbol{x}}) = \boldsymbol{Ax} - \boldsymbol{A}\hat{\boldsymbol{x}} = \boldsymbol{b} - \boldsymbol{A}\hat{\boldsymbol{x}} = \boldsymbol{r}.$$

- Por lo tanto, podemos proceder del siguiente modo:

1. Resolvemos computacionalmente
$$\boldsymbol{Ae} = \boldsymbol{r}.$$
2. La solución calculada de este sistema $\hat{\boldsymbol{e}}$ no va a coincidir exactamente con el error $\boldsymbol{e}$ pero va a ser una aproximación del mismo:
$$\hat{\boldsymbol{e}} \approx \boldsymbol{e} \implies \|\hat{\boldsymbol{e}}\| \approx \|\boldsymbol{e}\|.$$
3. $\|\hat{e}\|$ **es una estimación del tamaño del error** e**.**