Cuarta relación de problemas | Técnicas Numéricas
Profesor Francisco R. Villatoro | 27 de Noviembre de 1998

SOLUCIONES

Soluciones de los ejercicios de la tercera relación de problemas.

1. Demuestra que la inversa de una matriz triangular superior (inferior) es una matriz triangular superior (inferior).

   Solución. Sea la matriz triangular inferior

$$L = \begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix}, \qquad |L| = \prod_{i=1}^{n} l_{ii}$$

   y su inversa

$$L^{-1} = \frac{1}{|L|} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix},$$

   donde $b_{ij} = (-1)^{i+j}\, L_{(ij)}^{\top}$, donde $L_{(ij)}$ es el determinante de la matriz traspuesta sin la fila i y la columna j. Como el determinante es una suma de productos cada uno de los cuales contiene un elemento de cada fila y de cada columna, debe ser

$$b_{ij} = 0, \qquad i > j,$$

   con lo cual

$$L^{-1} = \frac{1}{|L|} \begin{pmatrix} b_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix},$$

   la inversa es triangular inferior, como queríamos probar. Para las triangulares superiores el resultado es del todo similar.

2. Una matriz A de $n \times n$ es diagonalmente dominante (estrictamente) por filas si

$$|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}| > 0, \qquad i = 1, 2, \ldots n.$$

Demuestre que A es invertible, es decir, existe su inversa A^{-1}.

Solución. Para una matriz diagonalmente dominante por filas

$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}| = \max_{1\le i\le n} \left(|a_{ii}| + \sum_{j=1, j\ne i}^{n} |a_{ij}| \right) < 2 \max_{1\le i\le n} |a_{ii}|.$$

Además $a_{ii} \ne 0$, por lo que

$$A = D\,B, \qquad D = (d_{ii}) = (a_{ii}),$$

donde D tiene inversa, y

$$b_{ii} = 1, \qquad i = 1, 2, \ldots, n,$$

y para $1 \le i \ne j \le n$,

$$d_{ii}\, b_{ij} = a_{ij}, \qquad b_{ij} = \frac{a_{ij}}{d_{ii}}.$$

Ahora vamos a demostrar que la matriz $I - B$ es invertible. Ya que

$$\|I - B\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |\delta_{ij} - b_{ij}|,$$

donde δ_{ij} es la delta de Kronecker, y

$$\|I - B\|_\infty = \max_{1\le i\le n} \sum_{j=1, j\ne i}^{n} |b_{ij}| = \max_{1\le i\le n} \sum_{j=1, j\ne i}^{n} \left| \frac{a_{ij}}{a_{ii}} \right| < 1 = \frac{1}{\|I^{-1}\|_\infty},$$

por lo cual B es invertible y como también lo es D, A tiene como inversa

$$A^{-1} = D^{-1}\, B^{-1}.$$

3. Determina el número total de operaciones de división, producto y suma requeridas por el procedimiento de resolución de un sistema lineal mediante factorización LU de Crout. Calcula también el número total de operaciones, su orden de magnitud y compara con el del procedimiento de Gauss-Jordan.

   Solución. El procedimiento de factorización LU requiere para cada $k = 1, 2, \ldots, n$ el cálculo de los $n - k + 1$ últimos elementos de la k-ésima fila de U, lo que requiere $k - 1$ productos y sumas, y el cálculo de los $n - k$ últimos elementos de la k-ésima columna de L, lo que también requiere 1 división y $k - 1$ productos y sumas. De esta forma el número total de

a) divisiones
$$\sum_{k=1}^{n-1} = \frac{n\,(n-1)}{2} = O\left(\frac{n^2}{2}\right),$$

b) sumas y productos
$$\sum_{k=1}^{n}(n-k+1)\,(k-1) + \sum_{k=1}^{n}(n-k)\,(k-1)$$
$$= \frac{n\,(n-1)\,(n+1)}{6} + \frac{n\,(n-1)\,(n-2)}{3} = O\left(\frac{n^3}{3}\right).$$

El número de operaciones requeridas para la resolución del sistema triangular superior U es de n divisiones y $n\,(n-1)/2$ productos y sumas, y el del sistema triangular superior L es de sólo $n\,(n-1)/2$ productos y sumas.

Por lo tanto, el número total de operaciones para el procedimiento de factorización LU es de $O(2\,n^3/3)$, que es el mismo que para el procedimiento de eliminación de Gauss-Jordan.

4. Dado el sistema lineal
$$1{,}01\,x + 0{,}99\,y = 2, \qquad 0{,}99\,x + 1{,}01\,y = 2.$$

Calcule

a) La solución exacta del problema

b) La solución usando solamente dos cifras decimales y redondeo

c) La inversa de la matriz de coeficientes A^{-1} con dos cifras decimales y redondeo

d) El número de condición de A para el apartado (d).

e) El residuo obtenido en (b).

Solución.

a) Por inspección visual del sistema la solución es $x = y = 1$.

b) Operando por sustitución

$$x = \frac{2 - 0{,}99\, y}{1{,}01} = 1{,}98 - 0{,}98\, y,$$

$$\begin{aligned} 0{,}99\, x + 1{,}01\, y &= 0{,}99\,(1{,}98 - 0{,}98\, y) + 1{,}01\, y \\ &= 1{,}96 - 0{,}97\, y + 1{,}01\, y = 1{,}96 + 0{,}04\, y = 2, \end{aligned}$$

con lo que

$$y = 1, \qquad x = 1{,}98 - 0{,}98\, y = 1.$$

c) El determinante es

$$|A| = 1{,}01 \times 1{,}01 + 0{,}99 \times 0{,}99 = 1{,}02 - 0{,}98 = 0{,}04,$$

y la matriz inversa

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} 1{,}01 & -0{,}99 \\ -0{,}99 & 1{,}01 \end{pmatrix} = \begin{pmatrix} 25{,}25 & -24{,}75 \\ -24{,}75 & 25{,}25 \end{pmatrix}$$

d) Utilizando normas infinito, el número de condición es

$$\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}| = 2,$$

$$\|A^{-1}\|_\infty = 50, \qquad \kappa(A) = \|A\|_\infty \, \|A^{-1}\|_\infty = 100$$

e) El residuo es

$$r = b - A\,x = \begin{pmatrix} 2 \\ 2 \end{pmatrix} - \begin{pmatrix} 1{,}01 & -0{,}99 \\ -0{,}99 & 1{,}01 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

5. Demuestre que si $|A| \neq 0$, $|B| = 0$ entonces

$$\|A^{-1}\| \ge \frac{1}{\|A - B\|}.$$

Solución. Si $|B| = 0$ existe un vector $x \neq 0$ tal que $B\,x = 0$, luego

$$A\,x = A\,x - B\,x = (A - B)\,x, \qquad \|A\,x\| \le \|A - B\|\,\|x\|,$$

y como $x = A^{-1} A\, x$,

$$\|x\| \leq \|A^{-1}\| \, \|A\, x\| \leq \|A^{-1}\| \, \|A - B\| \, \|x\|.$$

Como $x \neq 0$, entonces

$$\|A^{-1}\| \geq \frac{1}{\|A - B\|}.$$

6. Demuestre que

   a) si $\|I - B\| < 1$, entonces $|B| \neq 0$,

   b) si $\|C\| < 1$, entonces $|I - C| \neq 0$,

   c) si $|A| \neq 0$ y $|B|$ es tal que

   $$\|A^{-1}\| < \frac{1}{\|A - B\|},$$

   entonces $|B| \neq 0$.

   <u>Solución</u>.

   a) En el ejercicio anterior escogemos $A = I$, entonces si suponemos que $|B| = 0$,

   $$1 \leq \|A - B\|,$$

   que contradice la hipótesis $\|I - B\| < 1$, luego $|B| \neq 0$.

   b) Definiendo $C = I - B$, la hipótesis se escribe

   $$\|I - B\| = \|C\| < 1,$$

   y usando el apartado anterior,

   $$|B| = |I - C| \neq 0.$$

   c) De la hipótesis

   $$1 > \|A^{-1}\| \, \|A - B\| \geq \|A^{-1}\,(A - B)\| = \|I - A^{-1}\, B\|.$$

   Definiendo $C = I - A^{-1}\, B$,

   $$\|C\| < 1, \qquad \|C\| \leq \|C\| \, \|C\| = \|C\|^2 < 1, \qquad \|C^n\| \leq \|C\|^n,$$

con lo que
$$\lim_{n\to\infty} \|C^n\| = 0.$$
Por otro lado, aplicando la definición de norma subordinada a los autovectores de A,
$$\|A\| \geq \rho(A), \qquad \|A^n\| \geq \rho(A^n) = \rho^n(A),$$
y como
$$\lim_{n\to\infty} \|C^n\| = \lim_{n\to\infty} \rho^n(C) = 0,$$
entonces
$$\rho(C) < 1, \qquad |\lambda_C| < 1.$$
Como
$$1 > \|I - A^{-1}\, B\| = \|I - D\| \geq \rho(I - D) = \max_i |1 - \lambda_{Di}| \geq 0,$$
por lo que los autovalores de D no pueden ser cero (ya que en ese caso el máximo sería la unidad y se violaría la desigualdad de la izquierda). Tampoco pueden ser negativos, por la misma razón. Luego son positivos y
$$|D| = \prod_{i=1}^{n} \lambda_{Di} \neq 0,$$
y finalmente,
$$D = A^{-1}\, B, \qquad B = A\, D, \qquad |B| = |A|\, |D| \neq 0.$$

7. Resuelva el siguiente sistema de ecuaciones lineales con aritmética de cuatro dígitos (mantisa de 4 dígitos decimales),
$$0{,}1410 \times 10^{-2}\, x + 0{,}4004 \times 10^{-1}\, y = 0{,}1142 \times 10^{-1},$$
$$0{,}2000 \times 10^{0}\, x + 0{,}4912 \times 10^{1}\, y = 0{,}1428 \times 10^{1},$$
por medio de los siguientes métodos:
   a) regla de Cramer,
   b) eliminación de Gauss y en el orden en que se dan las ecuaciones,
   c) eliminación de Gauss e intercambiando el orden de las ecuaciones,

d) eliminación de Gauss con pivotaje completo.

Explique sus resultados y calcule el número de condición de la matriz de coeficientes. ¿Es esta matriz simétrica?, ¿y definida positiva? ¿Cuáles son su determinante, sus autovalores y sus autovectores?

Solución. Solución. Es fácil comprobar que la solución exacta de este sistema es x=1, y=0.25.

a) Aplicando la regla de Cramer, determinamos el determinante del sistema

$$|A| = a_{11}\,a_{22} - a_{12}\,a_{21} = 0{,}6926 \times 10^{-2} - 0{,}8008 \times 10^{-2}$$
$$= -0{,}1082 \times 10^{-2},$$

y las dos soluciones

$$x = \frac{a_{22}\,b_1 - a_{12}\,b_2}{|A|} = \frac{0{,}5610 \times 10^{-1} - 0{,}5712 \times 10^{-1}}{-1{,}082 \times 10^{-2}}$$
$$= \frac{-0{,}1020 \times 10^{-2}}{-1{,}082 \times 10^{-2}} = 0{,}9427 \times 10^{0},$$

$$y = \frac{a_{11}\,b_2 - a_{21}\,b_1}{|A|} = \frac{0{,}2013 \times 10^{-2} - 0{,}2284 \times 10^{-2}}{-1{,}082 \times 10^{-2}}$$
$$= \frac{-0{,}2710 \times 10^{-3}}{-1{,}082 \times 10^{-2}} = 0{,}2505 \times 10^{0}.$$

b) Aplicando eliminación de Gauss en el orden dado, obtenemos como primer (y único) multiplicador

$$m = \frac{a_{11}}{a_{21}} = 0{,}7050 \times 10^{-2},$$

y para la segunda ecuación

$$\begin{aligned}
(m\,a_{22} - a_{12})\,y &= (m\,b_2 - b_1),\\
(0{,}3463 \times 10^{-1} - 0{,}4004 \times 10^{-1})\,y &= 0{,}1007 \times 10^{-1} - 0{,}1142 \times 10^{-1}\\
-0{,}5410 \times 10^{-2}\,y &= -0{,}1350 \times 10^{-2}\\
y &= 0{,}2495\,10^{0},
\end{aligned}$$

con lo que despejando la otra variable

$$x = \frac{b_1 - a_{12}\, y}{a_{11}} = \frac{0{,}1142 \times 10^{-1} - 0{,}9990 \times 10^{-2}}{a_{11}}$$
$$= \frac{0{,}1430 \times 10^{-2}}{0{,}1410 \times 10^{-2}} = 0{,}1014 \times 10^{1}.$$

c) Aplicando eliminación de Gauss intercambiando las ecuaciones, obtenemos como primer (y único) multiplicador

$$m = \frac{a_{21}}{a_{11}} = 0{,}4072 \times 10^{-1},$$

y para la segunda (antes la primera) ecuación

$$\begin{aligned} (m\, a_{12} - a_{22})\, y &= (m\, b_1 - b_2), \\ 0{,}7657 \times 10^{0}\, y &= 0{,}1910 \times 10^{0} \\ y &= 0{,}2494\, 10^{0}, \end{aligned}$$

con lo que despejando la otra variable

$$x = \frac{b_2 - a_{22}\, y}{a_{22}} = \frac{0{,}2029 \times 10^{0}}{0{,}2000 \times 10^{0}} = 0{,}1015 \times 10^{1}.$$

d) Aplicando eliminación de Gauss con pivotaje completo tenemos que resolver el sistema

$$0{,}4912 \times 10^{1}\, y + 0{,}2000 \times 10^{0}\, x = 0{,}1428 \times 10^{1},$$
$$0{,}4004 \times 10^{-1}\, y + 0{,}1410 \times 10^{-2}\, x = 0{,}1142 \times 10^{-1},$$

por lo que obtenemos como primer (y único) multiplicador

$$m = \frac{0{,}4912 \times 10^{1}}{0{,}4004 \times 10^{-1}} = 0{,}1227 \times 10^{3},$$

y para la segunda (antes la primera) ecuación

$$\begin{aligned} (m\, 0{,}1410 \times 10^{-2} - 0{,}2000 \times 10^{0})\, x &= (m\, 0{,}1142 \times 10^{-1} - 0{,}1428 \times 10^{1}) \\ -0{,}2699 \times 10^{-}1\, x &= -0{,}2677 \times 10^{-}1 \\ x &= 0{,}9918\, 10^{0}, \end{aligned}$$

con lo que despejando la otra variable

$$\begin{aligned} 0{,}4912 \times 10^1\, y &= 0{,}1428 \times 10^1 - 0{,}2\, x \\ 0{,}4912 \times 10^1\, y &= 0{,}1230 \times 10^1 \\ y &= 0{,}2504 \times 10^0. \end{aligned}$$

e) El determinante de la matriz de coeficientes A es

$$|A| = -0{,}00108208 = -0{,}108208 \times 10^-2,$$

y su polinomio característico es

$$|A - \lambda\, I| = -0{,}00108208 - 4{,}91341\, \lambda + \lambda^2,$$

cuyas raíces son $\lambda \in \{4{,}91363, -0{,}00022022\}$ y sus autovectores

$$v \in \{(-0{,}00815083, -0{,}999967)^\top, (-0{,}999172, 0{,}0406811)^\top\},$$

respectivamente. Como el determinante de la matriz es muy pequeño y sus autovalores son de tamaño muy diferente, es de esperar que esta matriz esté mal condicionada. Esta matriz son es simétrica, obviamente, y tampoco es definida positiva porque uno de sus autovalores es negativo.

f) Calcularemos el número de condición en las tres normas más importantes de A,

$$\|A\|_1 = 4{,}95204, \qquad \|A\|_\infty = 5{,}112, \qquad \|A\|_2 = 4{,}91623,$$

y para la inversa

$$A^{-1} = \begin{pmatrix} -4539{,}41 & 37{,}0028 \\ 184{,}829 & -1{,}30305 \end{pmatrix},$$

sus normas son

$$\|A^{-1}\|_1 = 4724{,}23, \qquad \|A^{-1}\|_\infty = 4576{,}41, \qquad \|A^{-1}\|_2 = 4543{,}32,$$

por lo que los tres números de condición son

$$\kappa_1(A) = 23394{,}58, \qquad \kappa_\infty(A) = 23394{,}61, \qquad \kappa_2(A) = 22336{,}01,$$

que indica que la matriz está mal condicionada.

g) Conclusiones: Aunque el error relativo en la variable y es del mismo orden $\approx 0{,}5 \times 10^{-3}$ para todos los métodos, el error relativo para la regla de Cramer (0.057) es una cuatro veces mayor que el error para los métodos de Gauss (b, 0.014) y (c, 0.015), que es prácticamente igual, y que a su vez son unas dos veces más grande que el error cuando se usa pivotaje completo (0.0082). Es decir, debido al mal condicionamiento de la matriz de coeficientes los métodos que usan más operaciones aritméticas (Cramer) tiene mayor error y los que minimizan posibles diferencias cancelativas (pivotaje completo) provocan menor error.

8. Calcule el rango, los autovalores y los autovectores de la matriz

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & -1 & -1 \\ 6 & 2 & 0 \end{pmatrix}.$$

¿Cuál es la solución de $A\,x = 0$? ¿Por qué?

Solución. Aplicando eliminación de Gauss

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & -1 & -1 \\ 6 & 2 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & 5 & 3 \\ 0 & 10 & 6 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & 5 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

observamos que el rango de esta matriz es 2. Los autovalores de esta matriz son las raíces de $\lambda - 9\,\lambda^3$, es decir, $\lambda \in \{-3, 0, 3\}$ y sus autovectores son

$$v \in \{(-2, 3, 2)^\top, (1, -3, 5)^\top, (1, 0, 2)^\top\}.$$

Por el teorema del rango, al ser el rango de la matriz de coeficientes menor que la dimensión de la matriz, e igual al rango de la matriz ampliada, este sistema indeterminado tiene infinitas soluciones de la forma

$$x_1 = -\frac{1}{3}\,x_2, \qquad x_3 = -\frac{5}{3}\,x_2.$$

9. Resuelva el sistema

$$\begin{aligned} 2\,x + 2\,y + 3\,z &= 1 \\ x + y + z &= 2 \\ 2\,x + y + 2\,z &= 3 \end{aligned}$$

con el método de eliminación de Gauss y calcule las matrices de permutación que le sean necesarias. Explique lo que hace y por qué lo hace.

Solución. Primero haremos ceros en la primera columna, para lo sustituiremos la segunda fila por la multiplicación de la primera fila por $-1/2$ y su suma a la segunda fila anterior, y sustituiremos la tercera fila por la anterior tercera fila menos la primera fila, es decir, utilizando matrices de permutación

$$\begin{pmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|c} 2 & 2 & 3 & 1 \\ 1 & 1 & 1 & 2 \\ 2 & 1 & 2 & 3 \end{array}\right) = \left(\begin{array}{ccc|c} 2 & 2 & 3 & 1 \\ 0 & 0 & -1/2 & 3/2 \\ 0 & -1 & -1 & 2 \end{array}\right).$$

Seguidamente podemos permutar las filas segunda y tercera,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \left(\begin{array}{ccc|c} 2 & 2 & 3 & 1 \\ 0 & 0 & -1/2 & 3/2 \\ 0 & -1 & -1 & 2 \end{array}\right). = \left(\begin{array}{ccc|c} 2 & 2 & 3 & 1 \\ 0 & -1 & -1 & 2 \\ 0 & 0 & -1/2 & 3/2 \end{array}\right).$$

Ahora podemos resolver el sistema triangular superior obtenido hacia atrás

$$z = -3, \qquad y = -(2-3) = 1, \qquad x = 1/2\,(1+9-2) = 4.$$

10. Resuelva

$$\begin{aligned} 0{,}780\,x + 0{,}563\,y &= 0{,}217 \\ 0{,}913\,x + 0{,}659\,y &= 0{,}254 \end{aligned}$$

con tres cifras decimales. En un ordenador (1) se ha obtenido el resultado

$$x_{c1} = \begin{pmatrix} 0{,}341 \\ -0{,}087 \end{pmatrix},$$

mientras que un segundo ordenador (2) se ha obtenido

$$x_{c2} = \begin{pmatrix} 0{,}999 \\ -1{,}001 \end{pmatrix}.$$

Calcule el residuo $r = b - A\,x$ para las soluciones x_{c1} y x_{c2}. ¿Cuál es el ordenador que le da mejor resultados? ¿Por qué? ¿Sugiere usted el uso del residuo como una indicación de la exactitud de la solución calculada? ¿Por qué? ¿Cuál es el número de condición de la matriz de coeficientes? ¿Es un problema bien condicionado?

Solución. Operando con aritmética exacta es fácil comprobar que la solución es $x = 1$ e $y = -1$. Intentemos resolver este sistema con tres dígitos decimales mediante el procedimiento de eliminación de Gauss. El primer multiplicador es

$$m = \frac{0{,}780}{0{,}913} = 0{,}854,$$

y multiplicando por la segunda fila y restando la primera,

$$0{,}854 \times 0{,}659 - 0{,}563 = 0{,}563 - 0{,}563 = 0,$$

con lo que este pivote es nulo y la solución no puede ser determinada por este procedimiento. Utilizando pivotaje parcial, intercambiamos las filas primera y segunda, y obtenemos como nuevo multiplicador

$$m' = \frac{0{,}913}{0{,}780} = 1{,}171,$$

y multiplicando por la segunda fila y restando la primera,

$$1{,}171 \times 0{,}563 - 0{,}659 = 0{,}659 - 0{,}659 = 0,$$

y volvemos a tener un pivote nulo. Podemos utilizar reescalado y escribir el sistema como

$$\begin{aligned} x + 0{,}722\, y &= 0{,}278 \\ x + 0{,}722\, y &= 0{,}278 \end{aligned}$$

y observamos que este sistema es incompatible cuando se utiliza aritmética de tres dígitos decimales.

Supongamos que dos ordenadores han obtenido las soluciones x_{c1} y x_{c2}, el residuo para estas soluciones (calculado con más de tres dígitos) es

$$r_1 = (10^{-6}, 0)^\top, \qquad r_2 = (0{,}001343, 0{,}001572)^\top,$$

aunque el segundo ordenador da resultados precisos hasta el último dígito y el primero da una solución sin ningún dígito significativo correcto, el residuo para el primer ordenador es prácticamente cero, mientras que para el segundo es del mismo orden que el error en la solución. Por tanto, el residuo no es un indicativo adecuado de la exactitud de la solución de un problema. Un residuo pequeño no necesariamente indica una solución precisa, como indica la desigualdad

$$\|r\| \leq \|A\| \, \|e\|,$$

donde e es el error. Un error pequeño, si $\|A\|$ es grande puede conducir a un residuo grande, y a la inversa, un error grande, si $\|A\|$ es pequeña puede conducir a un residuo pequeño.

El número de condición de esta matriz es (omitimos los detalles de su cálculo)

$$\kappa_1(A) = \kappa_\infty(A) = 2{,}66 \times 10^6, \qquad \kappa_2(A) = 2{,}19 \times 10^6,$$

lo que indica que este problema está extremadamente mal condicionado. Hemos de notar que para problemas bien condicionados el residuo no es un indicativo del error demasiado malo.

11. Calcule la solución exacta de $A\,x = b$ donde

$$A = \begin{pmatrix} 1{,}6384 & 0{,}8065 \\ 0{,}8321 & 0{,}4096 \end{pmatrix}, b = \begin{pmatrix} 0{,}8319 \\ 0{,}4225 \end{pmatrix}.$$

Calcule el vector x_c tal que $r = A\,x_c - b$ es exactamente igual a

$$e = \begin{pmatrix} -10^{-8} \\ 10^{8} \end{pmatrix}.$$

Calcule el número de condición de A usando norma infinito. Si el ordenador representa b exactamente (sin error), calcule el error relativo de A tal que el error relativo de la solución sea menor o igual que 10^{-8} en la norma infinito. Si el ordenador no comete error al representar A, calcule el error relativo de b tal que el error relativo de la solución sea menor o igual que 10^{-8} en la norma infinito. Repita los resultados anteriores en norma 1.

Solución.

a) Para calcular la solución exacta podemos utilizar eliminación de Gauss, o cualquier otro procedimiento con aritmética exacta. Calcularemos directamente la inversa (ya que más tarde tendremos que determinar el número de condicionamiento),

$$|A| = -10^{-8}, \qquad A^{-1} = \frac{1}{|A|} \begin{pmatrix} 0{,}4096 & -0{,}8065 \\ -0{,}8321 & 0{,}4096 \end{pmatrix},$$

con lo que la solución exacta es

$$\begin{aligned} x &= A^{-1}\, b = 10^8 \begin{pmatrix} 0{,}4096 & -0{,}8065 \\ -0{,}8321 & 0{,}4096 \end{pmatrix} \begin{pmatrix} 0{,}8319 \\ 0{,}4225 \end{pmatrix} \\ &= 10^8 \begin{pmatrix} -10^{-8} \\ 10^{-8} \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \end{aligned}$$

b) El vector x_c pedido es

$$\begin{aligned} x_c &= A^{-1}\,(r+e) = \begin{pmatrix} 0{,}4096 & -0{,}8065 \\ -0{,}8321 & 0{,}4096 \end{pmatrix} \begin{pmatrix} 83189999 \\ 42250001 \end{pmatrix} \\ &= \begin{pmatrix} -2{,}2161 \\ 3{,}4684 \end{pmatrix}. \end{aligned}$$

c) El número de condición de la matriz de coeficientes es

$$\kappa(A) = \|A\|\,\|A^{-1}\|,$$

por lo que en norma infinito

$$\|A\|_\infty = 2{,}44, \quad \|A^{-1}\|_\infty = 2{,}47 \times 10^8, \qquad \kappa_\infty(A) = 6{,}04 \times 10^8.$$

y en norma uno

$$\|A\|_1 = 2{,}47, \quad \|A^{-1}\|_1 = 2{,}44 \times 10^8, \qquad \kappa_1(A) = 6{,}04 \times 10^8.$$

d) Si suponemos que A tiene un error $\delta\, A$ y que b es exacto,

$$(A + \delta\, A)\,(x + \delta\, x) = b,$$

y la ecuación del error es

$$A\,\delta\, x = -\delta\, A\,(x + \delta\, x),$$

con lo que aplicando normas y la desigualdad triangular

$$\|\delta\, x\| \leq \frac{\|\delta\, A\|}{\|A\|}\, (\|x\| + \|\delta\, x\|),$$

de donde

$$\frac{\|\delta\, x\|}{\|x\|} \leq \frac{e_r(A)}{1 - e_r(A)}, \qquad e_r(A) = \frac{\|\delta\, A\|}{\|A\|},$$

por lo que

$$\frac{e_r(A)}{1 - e_r(A)} \leq 10^{-8}, \qquad e_r(A) \leq \frac{10^{-8}}{1 + 10^{-8}} \approx 10^{-8}$$

(tanto en norma 1 como ∞) garantiza que

$$\frac{\|\delta\, x\|}{\|x\|} \leq 10^{-8}.$$

e) Si suponemos que b tiene un error $\delta\, b$ y que A es exacto,

$$A\,(x + \delta\, x) = (b + \delta\, b),$$

de donde la ecuación del error es

$$A\,\delta\, x = \delta\, b,$$

y como

$$A\, x = b, \quad \Rightarrow \quad \|x\| \geq \frac{\|b\|}{\|A\|},$$

$$\delta\, x = A^{-1}\,\delta\, b, \quad \Rightarrow \quad \|\delta\, x\| \leq \|\delta\, b\| \|A^{1}\|,$$

entonces

$$\frac{\|\delta\, x\|}{\|x\|} \leq \frac{\|A\|}{\|b\|}\, \|\delta\, x\| \leq \|A\|\, \|A^{-1}\|\, \frac{\|\delta\, b\|}{\|b\|} = \kappa(A)\, \frac{\|\delta\, b\|}{\|b\|},$$

y (tanto en norma uno como infinito),

$$\frac{\|\delta\, b\|}{\|b\|} \leq \frac{10^{-8}}{\kappa(A)} \approx 1{,}66 \times 10^{-17},$$

garantiza que

$$\frac{\|\delta\, x\|}{\|x\|} \leq 10^{-8}.$$

12. Si $r = A\,x_c - b$, $x_c = x_e + \Delta\,x$, $A\,x_e = b$, y $R = A\,C - I$ donde C es una aproximación a la inversa de A, demuetre que

$$\frac{\|C\|\,\|r\|}{\|R\|-1} \leq \|\Delta\,x\| \leq \frac{\|C\|\,\|r\|}{1+\|R\|}.$$

Solución. Como

$$r = A\,x_c - b = A\,\Delta\,x, \qquad \|\Delta\,x\| \geq \frac{\|r\|}{\|A\|},$$

y para demostrar la primera desigualdad del enunciado, basta demostrar que

$$\frac{1}{\|A\|} \geq \frac{\|C\|}{\|R\|-1},$$

lo que es falso en general, ya que

$$\|R\| = \|A\,C - I\| \leq \|A\|\,\|C\| + 1, \qquad \|A\| \geq \frac{\|R\|-1}{\|C\|},$$

es la desigualdad opuesta.

Por otro lado, como

$$A^{-1}\,r = \Delta\,x, \qquad \|\Delta\,x\| \leq \|A^{-1}\|\,\|r\|,$$

y para demostrar la segunda desigualdad del enunciado, basta probar que

$$\|A^{-1}\| \leq \frac{\|C\|}{1+\|R\|},$$

pero esto es falso en general, ya que

$$C = A^{-1}\,(I+R), \qquad \|A^{-1}\| \geq \frac{\|C\|}{1+\|R\|}.$$

Sin embargo, como

$$A = (I+R)\,C^{-1}, \qquad \|A\| \leq \|C^{-1}\|\,(1+\|R\|),$$

$$A^{-1}\,R = C - A^{-1}, \qquad |\|C\| - \|A^{-1}\|| \leq \|A^{-1}\|\,\|R\|,$$

$$-\|A^{-1}\|\,\|R\| \leq \|C\| - \|A^{-1}\|, \qquad \|A^{-1}\| \leq \frac{\|C\|}{1-\|R\|}$$

y por tanto

$$\frac{\|r\|}{\|C^{-1}\|\,(1+\|R\|)} \leq \|\Delta\, x\| \leq \frac{\|r\|\,\|C\|}{1-\|R\|}.$$

Nota: si en el enunciado de un ejercicio le piden demostrar algo falso, basta dar un contraejemplo simple para resolver el ejercicio.

13. Dada la matriz

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix},$$

calcule su factorización LU por medio de los métodos de Doolittle y Crout. Calcule también su descomposición de Cholesky.

Solución. Operando paso a paso, omitiremos los detalles, fácilmente se obtiene que la descomposición de Doolittle es

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 \\ 0 & \frac{2}{7} & 1 & 0 \\ 0 & 0 & \frac{7}{26} & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & \frac{7}{2} & 1 & 0 \\ 0 & 0 & \frac{26}{7} & 1 \\ 0 & 0 & 0 & \frac{45}{26} \end{pmatrix}.$$

Como la matriz es simétrica $A = A^\top$, la descomposición de Crout $A = L'\,U' = (L\,U)^\top = U^\top\,L^\top$, por lo que

$$L' = U^\top, \qquad U' = L^\top.$$

A partir de la factorización de Crout es fácil calcular la descomposición de Cholesky modificada ya que al ser A simétrica $A = L\,D\,L^\top$, y D es la diagonal de $U = D\,L^\top$. Escribiendo $\tilde{L} = L\,D^{1/2}$ obtenemos fácilmente

$$\tilde{L} = \begin{pmatrix} \sqrt{2} & 0 & 0 & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{14}}{2} & 0 & 0 \\ 0 & \frac{\sqrt{14}}{7} & \frac{\sqrt{182}}{7} & 0 \\ 0 & 0 & \frac{\sqrt{182}}{26} & \frac{\sqrt{1170}}{26} \end{pmatrix}.$$

14. Resuelva el sistema

$$3{,}9\,x_1 + 1{,}6\,x_2 = 5{,}5, \qquad 6{,}8\,x_1 + 2{,}9\,x_2 = 9{,}7,$$

con aritmética de dos cifras y el método de Crout. Si es posible mejore el resultado utilizando el residuo.

<u>Solución</u>. Utilizaremos aritmética de dos cifras significativas (mantisa de dos dígitos) y aplicaremos el procedimiento de Crout, con lo que obtnemos

$$A = \begin{pmatrix} 3{,}9 & 1{,}6 \\ 6{,}8 & 2{,}9 \end{pmatrix} = \begin{pmatrix} 3{,}9 & 0 \\ 6{,}8 & 0{,}1 \end{pmatrix} \begin{pmatrix} 1 & 0{,}41 \\ 0 & 1 \end{pmatrix} = L\,U,$$

y ahora resolvemos los sistemas $L\,y = b$ y $U\,x = y$,

$$\begin{pmatrix} 3{,}9 & 0 \\ 6{,}8 & 0{,}1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 5{,}5 \\ 9{,}7 \end{pmatrix},$$

$$y_1 = \frac{5{,}5}{3{,}9} = 1{,}4, \qquad y_2 = \frac{9{,}7 - 6{,}8 \times 1{,}4}{0{,}1} = 2;$$

$$\begin{pmatrix} 1 & 0{,}41 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1{,}4 \\ 2 \end{pmatrix},$$

$$x_2 = 2, \qquad x_1 = 1{,}4 - 2 \times 0{,}41 = 0{,}58.$$

Calculemos el residuo para ver si podemos mejorar la solución que hemos obtenido (también usaremos aritmética de dos dígitos)

$$\begin{aligned} r = b - A\,x &= \begin{pmatrix} 5{,}5 \\ 9{,}7 \end{pmatrix} - \begin{pmatrix} 3{,}9 & 1{,}6 \\ 6{,}8 & 2{,}9 \end{pmatrix} \begin{pmatrix} 0{,}58 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} 5{,}5 \\ 9{,}7 \end{pmatrix} - \begin{pmatrix} 2{,}3 + 3{,}2 \\ 3{,}9 + 5{,}8 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

con lo que no podemos mejorar nuestro resultado. Se debería haber calculado el residuo con mayor precisión y entonces sí podría ser utilizado para mejorar el resultado obtenido.