

Sistemas de ecuaciones lineales: métodos directos

Juan Hirschmann - jhirschmann@fi.uba.ar

1. Introducción

A grandes rasgos, un método directo se puede definir como un procedimiento empleado para resolver un sistema de ecuaciones lineales (SEL). Estos consisten en realizar operaciones permitidas entre filas y columnas despejando, de manera directa, las distintas variables.

De realizar ello, teóricamente se alcanzaría una solución exacta. Sin embargo, en la práctica los cálculos se realizan con precisión finita, debido a ello, se propagarán los errores de redondeo en menor o mayor grado. A continuación, se presentan 4 formas de resolver un SEL y una manera de refinar los resultados, a partir de ello, se observará de que manera los distintos métodos propagan el error de redondeo.

1.1. Solución trivial

Los métodos directos estudiados tienen como principal objetivo reducir un problema complejo a otro simple y con solución conocida. Como solución conocida se utilizará la de un sistema en donde la matriz que lo define es triangular superior, aunque bien podría ser triangular inferior y obtener un resultado similar:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & \vdots \\ \vdots & \ddots & \ddots & a_{2n} \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} \times \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}, \quad a_{ii} \neq 0 \quad \forall i \in [1, n] \quad (1)$$

A simple vista se observa que la solución del sistema es la siguiente:

$$x_i = \frac{b_i - \sum_{j>i}^n a_{ij} x_j}{a_{ii}} \quad (2)$$

1.2. Importancia del pivoteo

A fin de observar la importancia del pivoteo, y las consecuencias de no realizarlo, se puede considerar el siguiente SEL:

$$\left(\begin{array}{cc|c} \textcircled{a_{11}} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{array} \right) \xrightarrow{F_2 - m_{21} F_1} \left(\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ 0 & a_{22} - \frac{a_{21}}{a_{11}} a_{12} & b_2 - \frac{a_{21}}{a_{11}} b_1 \end{array} \right) \Rightarrow X_{sol} = \begin{pmatrix} b_1 - \frac{b_2 - \frac{a_{21}}{a_{11}} b_1}{a_{22} - \frac{a_{21}}{a_{11}} a_{12}} a_{12} \\ \frac{b_2 - \frac{a_{21}}{a_{11}} b_1}{a_{22} - \frac{a_{21}}{a_{11}} a_{12}} \end{pmatrix} \quad (3)$$

A modo demostrativo, se supone un caso extremo en donde el elemento pivote resulta mucho menor en módulo comparado con los demás elementos de la matriz, mientras que el resto de los elementos son de similar magnitud. De esta forma, se producen errores de redondeo que modifican a los resultados. Por este motivo, se pueden llegar a obtener resultados inexactos:

$$|a_{11}| \ll |a_{ij}|, |b_i| \Rightarrow \begin{cases} \left| \frac{a_{21}}{a_{11}} a_{12} \right| \gg |a_{22}| \\ \left| \frac{a_{21}}{a_{11}} b_1 \right| \gg |b_2| \end{cases} \quad X_{sol} \approx \begin{pmatrix} 0 \\ \frac{b_1}{a_{12}} \end{pmatrix} \quad (4)$$

Si, como caso contrario, el pivote fuese un elemento mucho mayor que el resto de los componentes, se obtiene que las inequaciones presentadas pierden validez y los errores de redondeo se propagarán en menor grado dependiendo del caso. Como solución al problema presentado, surgen las estrategias de pivoteo, que consisten en seleccionar con distintos criterios, pivotes grandes.

1.3. Normas matriciales

En primer lugar, se puede definir el concepto de norma de manera poco estricta como una función real que provee una medida para una variable multidimensional. Existen muchas normas matriciales con distintas propiedades, la menor de ellas es la norma espectral. Esta norma se calcula como la raíz del mayor autovalor de la matriz y, como la cota del error se expresa con normas, es la menor cota posible. Sin embargo, su cálculo tiene un costo computacional grande. Por este motivo, en la práctica se suele recurrir a distintas normas como las normas infinito o la norma 1, que si bien son mayores, proveen buenos resultados:

$$\|A\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|, \quad \|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \quad (5)$$

En otras palabras, las normas mencionadas consisten en obtener la mayor suma de elementos en valor absoluto según las filas o columnas respectivamente. En la resolución de ejercicios, se puede optar por una u otra pero siempre respetando la consistencia.

2. Ejemplo

Se propone hallar el vector solución del siguiente SEL utilizando las tres estrategias de pivoteo, expresando cada resultado con su cota de error. Además, se refinará la peor solución utilizando su descomposición LU.

$$\left(\begin{array}{ccc|c} 0,001 & 1 & 30 & 1 \\ 5 & 21 & 9 & 2 \\ 20 & 0,5 & 3 & 3 \end{array} \right) \quad (6)$$

Para todos los cálculos se utilizará una precisión de 6 dígitos significativos.

2.1. Resolución por eliminación de Gauss sin pivoteo

En principio, se observa que el primer pivote a tomar es mucho menor a cualquier otro elemento de la matriz. Como fue mencionado anteriormente, es probable que el resultado obtenido sin pivoteo sea inexacto. Esto no significa que la matriz se encuentre mal condicionada, si no que la estrategia utilizada no es óptima. Para comenzar la resolución, se toma el primer pivote, se calculan sus multiplicadores y se aplican las operaciones de filas:

$$\left(\begin{array}{ccc|c} \textcircled{0,001} & 1 & 30 & 1 \\ 5 & 21 & 9 & 2 \\ 20 & 0,5 & 3 & 3 \end{array} \right) \xrightarrow{\begin{array}{l} F_2 - \frac{5}{0,001} F_1 \\ F_3 - \frac{20}{0,001} F_1 \end{array}} \left(\begin{array}{ccc|c} \textcircled{0,001} & 1 & 30 & 1 \\ [5000] & -4979 & -149991 & -4998 \\ [20000] & -19999,5 & -599997 & -19997 \end{array} \right) \quad (7)$$

Como notación y para poder optimizar el uso de memoria en el código, los multiplicadores m_{ij} se almacenan en las posiciones a_{ij} que, tras realizar las operaciones entre filas, son nulos. Para poder diferenciarlos de los elementos de la matriz A, se notan de la siguiente manera: $[m_{ij}]$. Es importante reconocer que los multiplicadores no forman parte de la matriz A, si no que en esa posición hay un cero. A continuación, se accede al siguiente pivote:

$$\left(\begin{array}{ccc|c} 0,001 & 1 & 30 & 1 \\ [5000] & \textcircled{-4979} & -149991 & -4998 \\ [20000] & -19999,5 & -599997 & -19997 \end{array} \right) \xrightarrow{F_3 - \frac{-19999,5}{-4979} F_2} \left(\begin{array}{ccc|c} 0,001 & 1 & 30 & 1 \\ [5000] & \textcircled{-4979} & -149991 & -4998 \\ [20000] & [4,01667] & 2467,35 & 78,3167 \end{array} \right) \quad (8)$$

Una vez obtenida la matriz diagonal superior, se puede aplicar la solución conocida, obteniendo el siguiente vector solución:

$$X_{EG} = \begin{pmatrix} 0,1378 \\ 0,0488412 \\ 0,0317007 \end{pmatrix} \quad (9)$$

2.1.1. Resolución por eliminación de Gauss con pivoteo parcial

Por pivoteo parcial, se busca el pivote más grande dentro de la primera columna, una vez obtenido, se permutan las filas y el término independiente:

$$\left(\begin{array}{ccc|c} 0,001 & 1 & 30 & 1 \\ 5 & 21 & 9 & 2 \\ \textcircled{20} & 0,5 & 3 & 3 \end{array} \right) \xrightarrow{F_3 \times F_1} \left(\begin{array}{ccc|c} \textcircled{20} & 0,5 & 3 & 3 \\ 5 & 21 & 9 & 2 \\ 0,001 & 1 & 30 & 1 \end{array} \right) \quad (10)$$

Luego, se continua con la eliminación gaussiana:

$$\left(\begin{array}{ccc|c} \textcircled{20} & 0,5 & 3 & 3 \\ 5 & 21 & 9 & 2 \\ 0,001 & 1 & 30 & 1 \end{array} \right) \xrightarrow{\begin{array}{l} F_2 - \frac{5}{20} F_1 \\ F_3 - \frac{0,001}{20} F_1 \end{array}} \left(\begin{array}{ccc|c} \textcircled{20} & 0,5 & 3 & 3 \\ [0,25] & 20,875 & 8,25 & 1,25 \\ [0,00005] & 0,999975 & 29,9985 & 0,998500 \end{array} \right) \quad (11)$$

En este caso, se observa que no es necesario realizar permutaciones entre filas ya que $20,875 > 0,999975$.

$$\left(\begin{array}{ccc|c} 20 & 0,5 & 3 & 3 \\ [0,25] & \textcircled{20,875} & 8,25 & 1,25 \\ [0,00005] & 0,999975 & 29,9985 & 0,998500 \end{array} \right) \xrightarrow{F_3 - \frac{0,999975}{20,875} F_2} \left(\begin{array}{ccc|c} 20 & 0,5 & 3 & 3 \\ [0,25] & \textcircled{20,875} & 8,25 & 1,25 \\ [0,00005] & [0,0479030] & 29,6047 & 0,939971 \end{array} \right) \quad (12)$$

Utilizando la misma solución trivial se obtiene el siguiente vector solución:

$$X_{EGPP} = \begin{pmatrix} 0,144054 \\ 0,0473321 \\ 0,0317507 \end{pmatrix} \quad (13)$$

2.1.2. Resolución por eliminación de Gauss con pivoteo total

En primer lugar, se identifica el mayor elemento en la matriz para utilizar como pivote. Luego, se permutan las columnas, para mayor claridad se utilizarán las matrices de permutaciones:

$$\begin{pmatrix} 0,001 & 1 & \textcircled{30} \\ 5 & 21 & 9 \\ 20 & 0,5 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad (14)$$

$$\begin{pmatrix} 0,001 & 1 & \textcircled{30} \\ 5 & 21 & 9 \\ 20 & 0,5 & 3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Finalmente, al realizar los productos matriciales se obtiene el siguiente SEL con el vector incógnita permutado:¹

$$\begin{pmatrix} \textcircled{30} & 1 & 0,001 \\ 9 & 21 & 5 \\ 3 & 0,5 & 20 \end{pmatrix} \begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad (15)$$

A esta altura de la resolución, es relevante remarcar que el vector solución se encuentra permutado. Por ese motivo, al finalizar el ejercicio es necesario deshacer la permutación. A continuación, se continua de igual manera que en los dos métodos anteriores:

$$\left(\begin{array}{ccc|c} \textcircled{30} & 1 & 0,001 & 1 \\ 9 & 21 & 5 & 2 \\ 3 & 0,5 & 20 & 3 \end{array} \right) \xrightarrow{F_2 - \frac{9}{30}F_1} \left(\begin{array}{ccc|c} \textcircled{30} & 1 & 0,001 & 1 \\ [0,3] & 20,7 & 4,9997 & 1,7 \\ [0,1] & 0,4 & 19,9999 & 2,9 \end{array} \right) \quad (16)$$

Luego, se avanza al próximo pivote sin necesidad de realizar permutaciones, pues 20,7 es el mayor elemento de la submatriz:

$$\left(\begin{array}{ccc|c} 30 & 1 & 0,001 & 1 \\ [0,3] & \textcircled{20,7} & 4,9997 & 1,7 \\ [0,1] & 0,4 & 19,9999 & 2,9 \end{array} \right) \xrightarrow{F_3 - \frac{0,4}{20,7}F_2} \left(\begin{array}{ccc|c} 30 & 1 & 0,001 & 1 \\ [0,3] & \textcircled{20,7} & 4,9997 & 1,7 \\ [0,1] & [0,0193237] & 19,9033 & 2,84783 \end{array} \right) \quad (17)$$

Por último, se utiliza nuevamente la solución conocida, obteniendo el siguiente vector:

$$X_{EGPT} = \begin{pmatrix} 0,144054 \\ 0,0473320 \\ 0,0317508 \end{pmatrix} \quad (18)$$

2.2. Descomposición LU

Se puede demostrar que A es el producto de una matriz triangular inferior y otra triangular superior:

$$A = LU \quad L : \text{Triangular inferior} \quad U : \text{Triangular superior} \quad (19)$$

Además, al realizar la eliminación gaussiana se obtiene de manera colateral las matrices L y U. L será la matriz de diagonal 1 y, por debajo de la diagonal, los multiplicadores correspondientes. Por su parte, U será la matriz resultante de realizar la eliminación de Gauss:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{pmatrix} \quad U = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a'_{22} & a'_{23} \\ 0 & 0 & a''_{33} \end{pmatrix} \quad (20)$$

Por ejemplo, en la sección 2.1.1 tras finalizar la eliminación de Gauss sin pivoteo, se obtuvo el siguiente sistema:

$$\left(\begin{array}{ccc|c} 0,001 & 1 & 30 & 1 \\ [5000] & -4979 & -149991 & -4998 \\ [20000] & [4,01667] & 2467,35 & 78,3167 \end{array} \right) \quad (21)$$

¹Observar que si C_{ij} es la matriz que intercambia columna i por columna j entonces $C_{ij}^2 = I$, por lo que se conserva la equivalencia

En ese caso se puede observar que las matrices L y U son las siguientes:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 5000 & 1 & 0 \\ 20000 & 4,01667 & 1 \end{pmatrix} \quad U = \begin{pmatrix} 0,001 & 1 & 30 \\ 0 & -4979 & -149991 \\ 0 & 0 & 2467,35 \end{pmatrix} \quad (22)$$

Utilizando que $A = LU$, se puede descomponer el sistema $Ax = b$ en dos sistemas distintos. Dado que Ux es un vector en \mathbb{R}^n , se puede definir una nueva variable y , de forma tal que resolver $Ax = b$ sea equivalente a resolver los siguientes sistemas de ecuaciones:

$$Ax = b \equiv LUx = b \implies \begin{cases} Ux = y & , \quad y \in \mathbb{R}^n \\ Ly = b & , \quad Ly \in \mathbb{R}^n \end{cases} \quad (23)$$

Como particularidad de ello se puede notar que, al ser U y L triangulares, su solución es conocida. Además, evaluar el mismo sistema con distinto vector independiente implica unicamente reemplazar el vector b . Como ventaja de ello, no se debe realizar nuevamente el pivoteo si se quiere evaluar un mismo sistema reiteradas veces.

A modo demostrativo, se resuelve nuevamente $Ax = b$ utilizando su descomposición LU:

$$Ux = y \implies \begin{pmatrix} 0,001 & 1 & 30 \\ 0 & -4979 & -149991 \\ 0 & 0 & 2467,35 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \quad (24)$$

Aplicando la solución conocida, se obtiene el vector x :

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{y_1 - x_2 - 30x_3}{0,001} \\ \frac{y_2 + 149991x_3}{-4979} \\ \frac{y_3}{2467,35} \end{pmatrix} \quad (25)$$

En cuanto al segundo sistema de ecuaciones :

$$Ly = b \implies \begin{pmatrix} 1 & 0 & 0 \\ 5000 & 1 & 0 \\ 20000 & 4,01667 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad (26)$$

Como solución a este sistema resulta:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -4998 \\ 78,3167 \end{pmatrix} \quad (27)$$

Al reemplazar lo obtenido en la expresión 25, se obtiene que:

$$X_{EG} = \begin{pmatrix} 0,1378 \\ 0,0488412 \\ 0,0317007 \end{pmatrix} \quad (28)$$

Este resultado es idéntico al obtenido sin realizar la descomposición de manera explícita ya que ambos métodos son equivalentes.

En el caso de haber utilizado estrategias de pivoteo, es necesario permutar los vectores x y b según corresponda. Por ejemplo, en pivoteo parcial se permutaron las filas 1 y 3, por ese motivo es necesario hacer lo mismo en el vector b :

$$b = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \quad (29)$$

Por su parte, en la resolución por pivoteo total el vector b no cambia ya que, en este caso puntual, no se permutaron filas. Sin embargo, al haber realizado las permutaciones entre la columna 1 y 3, el vector x debe expresarse acorde:

$$x = \begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix} \quad (30)$$

Como ejercicio adicional, se sugiere corroborar la resolución del sistema mediante descomposición LU para los casos en donde se aplique pivoteo.

2.3. Refinamiento iterativo

Una de las grandes ventajas de los métodos directos es la inexistencia de errores de truncamiento. Sin embargo, la propagación del error de redondeo puede llegar a ser una fuente de error considerable en muchas aplicaciones prácticas. Por este motivo, es importante caracterizar y acotar el error para conocer que tipo de problema se presenta y actuar acorde.

Con esto en mente, se obtiene, tras aplicar un método directo, un vector solución, \tilde{x} . Naturalmente al tener error, \tilde{x} difiere de la solución exacta, x . De igual manera, el vector $A\tilde{x}$ difiere del vector independiente, b :

$$Ax = b, \quad x - \tilde{x} = \delta_x, \quad b - A\tilde{x} = R, \quad R: \text{Residuo} \quad (31)$$

Al ser δ_x una incógnita, pues de conocerla se sabría la solución exacta del problema, se puede definir un nuevo sistema de ecuaciones, en el que únicamente cambia el término independiente:

$$b - A\tilde{x} = A(x - \tilde{x}) = A\delta_x = R \quad (32)$$

Para el cual se obtiene una solución $\tilde{\delta}_x^{(1)}$, correspondiente al término de error entre iteraciones. Esta aproximación permite expresar los resultados de manera completa. Además, si se lo suma, es posible obtener una segunda solución, de mayor precisión, para la segunda *iteración*. Sin embargo, para expresar este nuevo resultado con su aproximación cota de error se debe iterar nuevamente.

$$\tilde{x}^{(2)} = \tilde{x}^{(1)} + \tilde{\delta}_x^{(1)} \quad (33)$$

En principio, este método parecería óptimo, pues se puede obtener una solución tan buena como se quiera y además se puede conocer la cota de error. No obstante, el residuo se obtiene a partir de la resta de dos vectores similares, esto produce que R sea un valor pequeño. Por este motivo, se produce una cancelación de términos insalvable. A raíz de ello, para obtener un resultado confiable hasta t dígitos significativos, se suele realizar el refinamiento iterativo con precisión $2t$.

2.3.1. Cálculo del error mediante descomposición LU

Al tener que evaluar el mismo SEL, pero con distinto término independiente, resulta conveniente utilizar la descomposición LU. Al realizar las distintas formas de eliminación de Gauss, se obtuvieron de manera colateral tres descomposiciones distintas para tres matrices distintas:²

$$\begin{aligned} A_{EG} &= \begin{pmatrix} 1 & 0 & 0 \\ 5000 & 1 & 0 \\ 20000 & 4,01667 & 1 \end{pmatrix} \begin{pmatrix} 0,001 & 1 & 30 \\ 0 & -4979 & -149991 \\ 0 & 0 & 2467,35 \end{pmatrix} \\ A_{EGPP} &= \begin{pmatrix} 1 & 0 & 0 \\ 0,25 & 1 & 0 \\ 0,00005 & 0,0479030 & 1 \end{pmatrix} \begin{pmatrix} 20 & 0,5 & 3 \\ 0 & 20,875 & 8,25 \\ 0 & 0 & 29,6047 \end{pmatrix} \\ A_{EGPT} &= \begin{pmatrix} 1 & 0 & 0 \\ 0,3 & 1 & 0 \\ 0,1 & 0,0193237 & 1 \end{pmatrix} \begin{pmatrix} 30 & 1 & 0,001 \\ 0 & 20,7 & 4,9997 \\ 0 & 0 & 19,9033 \end{pmatrix} \end{aligned} \quad (34)$$

Se propone realizar el cálculo del error para cada variante de la eliminación de Gauss. Para ello, es importante recordar que se debe duplicar la cantidad de dígitos significativos en el cálculo del residuo.

Para evitar la reiteración, se realiza de manera explícita el cálculo únicamente para la eliminación sin pivoteo pero se presentarán todos los resultados. Dado que es un caso óptimo para su uso, se resolverá utilizando la descomposición LU. En este caso, se deberá resolver el vector incógnita δ_x con el vector independiente R :

$$\begin{cases} U\delta_x = y \\ Ly = R \end{cases}, \quad R = b - A\tilde{x} = \begin{pmatrix} 1 - 1 \\ 2 - 1,9999715 \\ 3 - 2,8755227 \end{pmatrix} \quad (35)$$

Luego, se resuelve el sistema conociendo su solución :

$$Ly = R \implies \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 5000 & 1 & 0 & 2,85 \times 10^{-5} \\ 20000 & 4,01667 & 1 & 0,1244773 \end{array} \right) \Rightarrow y = \begin{pmatrix} 0 \\ 2,85 \times 10^{-5} \\ 0,124363 \end{pmatrix} \quad (36)$$

El segundo sistema, si bien es diagonal inferior, su solución es muy similar a la ya conocida:

$$U\delta_x = y \implies \begin{pmatrix} 0,001 & 1 & 30 \\ 0 & -4979 & -149991 \\ 0 & 0 & 2467,35 \end{pmatrix} \begin{pmatrix} \delta_{x_1} \\ \delta_{x_2} \\ \delta_{x_3} \end{pmatrix} = \begin{pmatrix} 0 \\ 2,85 \times 10^{-5} \\ 0,124363 \end{pmatrix} \quad (37)$$

²Si bien en todo caso se parte del mismo SEL, las permutaciones de realizadas alteran la matriz original. Esto permite la existencia de tres descomposiciones distintas sin violar la unicidad de la descomposición LU

Resolviendo el sistema, se obtuvo el vector δ_x :

$$\begin{pmatrix} \delta_{x_1} \\ \delta_{x_2} \\ \delta_{x_3} \end{pmatrix} = \begin{pmatrix} 6,29 \times 10^{-3} \\ -1,52 \times 10^{-3} \\ 5 \times 10^{-5} \end{pmatrix} \quad (38)$$

De esta forma, se pueden expresar los resultados de manera completa:

$$X_{EG} = \begin{pmatrix} 0,14 \pm 0,01 \\ 0,049 \pm 0,002 \\ 0,0317 \pm 0,0001 \end{pmatrix} \quad (39)$$

$$X_{EGPP} = \begin{pmatrix} 0,144054 \pm 0,000001 \\ 0,0473321 \pm 0,0000001 \\ 0,031751 \pm 0,000001 \end{pmatrix} \quad (40)$$

$$X_{EGPT} = \begin{pmatrix} 0,144054 \pm 0,000001 \\ 0,047332 \pm 0,000001 \\ 0,031751 \pm 0,000002 \end{pmatrix} \quad (41)$$

De manera esperada, se observa que el resultado de mayor incertidumbre es el obtenido sin pivotear. Además, se aprecia que en este caso, el pivoteo total no presentó un mayor exactitud comparado con el parcial. Sin embargo, se cumple con el principio de que siempre es mejor realizar pivoteo de algún tipo que no realizarlo.

2.3.2. Número de condición

Se buscará mejorar la solución obtenida mediante refinamiento iterativo. En primer lugar, se recomienda calcular el número de condición de la matriz. Caso contrario, el esfuerzo realizado puede ser en vano dado que una matriz muy mal condicionada no permite mejorar la solución. Utilizando la aproximación del $K[A]$ estudiada y los datos de la eliminación sin pivoteo:

$$K[A] \approx \frac{\|\delta_x\|}{\|x\|} 10^t = \frac{\sqrt{(2 \times 10^{-2})^2 + (-1,52 \times 10^{-3})^2 + (5 \times 10^{-5})^2}}{\sqrt{(0,14378)^2 + (0,0488412)^2 + (0,0317007)^2}} 10^6 \approx 4,1 \times 10^4 \quad (42)$$

Se observa que $K[A] \gg 1$, pero también $K[A] \ll 10^6$. A simple vista, este parecería ser un problema relativamente mal condicionado. Sin embargo, la estimación del $K[A]$ depende fuertemente de δ_x y, por lo tanto, del tipo de estrategia empleada. A modo de ejemplo, se calculó de manera exacta $K[A]$ dado que esta es una matriz pequeña y el cálculo es factible:

$$A = \begin{pmatrix} 0,001 & 1 & 30 \\ 5 & 21 & 9 \\ 20 & 0,5 & 30 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} -0,00473 & -0,000971 & 0,0502 \\ -0,0133 & 0,0485 & -0,0121 \\ 0,0338 & -0,00162 & 0,000403 \end{pmatrix} \implies \|A\|_1 \cdot \|A^{-1}\|_1 \approx 4,3 \quad (43)$$

Como motivo de lo observado, se puede señalar que la estimación de $K[A]$ depende proporcionalmente del vector error. Por su parte, este vector estará estrechamente ligado a la propagación del error de redondeo en la estrategia utilizada. Por este motivo, se recomienda utilizar la estimación de $K[A]$ con precaución, considerando el tipo de estrategia empleada.

Previo a realizar el refinamiento, se estimará la cantidad de dígitos significativos 'ganados' en la primera iteración:

$$q = t - \log(4 \times 10^4) = 1,4 \quad (44)$$

Es necesario aclarar que, para este cálculo se debe utilizar una estimación de $K[A]$ obtenida con la misma estrategia de pivoteo utilizada en la resolución y el refinamiento. Esto se debe a que la cantidad de dígitos 'ganados' depende de la calidad del refinamiento y la precisión de la solución y, por lo tanto, de la estrategia de pivoteo.

2.3.3. Cálculo

A modo de referencia y, conociendo de antemano que el problema se encuentra bien condicionado, se considerará exacta una solución obtenida a 6 dígitos significativos. También se recuerda el primer resultado obtenido:

$$X_{ref} = \begin{pmatrix} 0,144054 \\ 0,0473320 \\ 0,0317508 \end{pmatrix}, \quad X_{EG}^{(1)} = \begin{pmatrix} 0,1378 \\ 0,0488412 \\ 0,0317007 \end{pmatrix} \quad (45)$$

A continuación, se procederá a refinar la solución obtenida por eliminación sin pivoteo. En principio, se obtiene la nueva solución del sistema sumando a la original $\delta_x^{(1)}$, que ya fue calculado en la ecuación 38:

$$x^{(2)} = x^{(1)} + \delta_x^{(1)} = \begin{pmatrix} 0,14409 \\ 0,0473212 \\ 0,0317507 \end{pmatrix} \quad (46)$$

De comparar la solución 'exacta' con este resultado y el original, se observa que, efectivamente, se mejoró la solución alrededor de un dígito significativo. Luego, utilizando este nuevo vector solución se puede calcular el residuo para la segunda iteración:

$$R^{(2)} = b - A\tilde{x}^{(2)} = \begin{pmatrix} 1,078 \times 10^{-7} \\ 3 \times 10^{-7} \\ -7,587 \times 10^{-4} \end{pmatrix} \quad (47)$$

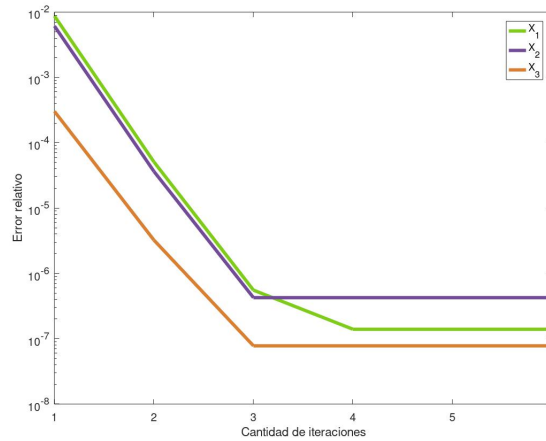
Al resolver el SEL nuevamente, se puede obtener el vector $\delta_x^{(2)}$ que, al sumarlo con la solución anterior se obtiene $x^{(3)}$. Para evitar ser reiterativo, el resultado se presenta con las cotas de error obtenidas calculando $\delta_x^{(3)}$:

$$x^{(3)} = \begin{pmatrix} 0,144054 \pm 0,000001 \\ 0,0473320 \pm 0,0000001 \\ 0,0317508 \pm 0,0000001 \end{pmatrix} \quad (48)$$

Se observa que la solución refinada es superior, en términos de error, a las soluciones con pivoteo pero sin refinar. Además, como fue mencionado, fue posible obtener una solución exacta hasta 6 dígitos significativos o medianamente significativos sin utilizar técnica de pivoteo. Sin embargo, para que ello sea posible, fue necesario realizar 3 iteraciones.

3. Análisis de resultados

A modo de síntesis, el mejor resultado obtenido fue mediante pivoteo parcial, marginalmente peor fue el pivoteo total y en último lugar sin pivoteo. Este fue un resultado esperado, dado que el pivoteo total no asegura ser mejor que el parcial y además que cualquier estrategia de pivoteo es mejor que ninguna. De todas formas, al refinar el resultado sin pivoteo sólo tres iteraciones fueron suficientes para alcanzar una mejor precisión que los otros resultados, esto demuestra las virtudes del refinamiento. Utilizando una solución exacta a 16 dígitos significativos como referencia, se graficó el error relativo en cada variable según la cantidad de iteraciones.



Se observa que, para las primeras dos iteraciones, el error relativo disminuye exponencialmente mientras que a partir de la tercera y cuarta este se mantiene constante. Esto se debe a que, el δ_x calculado es tan pequeño que al sumarse a la solución es redondeado. Como causa de ello, la nueva solución y, por lo tanto, las demás variables, se mantienen constantes. Por este motivo, elegir una condición adecuada para terminar el algoritmo resulta más complejo de lo esperado, dado que fijar una condición lo suficientemente chica podría llevar a infinitas iteraciones. Por lo general, se suele fijar condiciones sobre el error relativo y no el residuo.

4. Cuadro comparativo

No pivoteo		Pivoteo	
Ventajas	Desventajas	Ventajas	Desventajas
Veloz	Puede amplificar errores de redondeo	Disminuye la propagación de errores de redondeo	Replicar y considerar permutaciones en b y x según corresponda
Fácil implementación	Si algún pivote tomado es cero el método falla por <i>overflow</i>		Difícil implementación
Descomposición LU		Refinamiento iterativo	
Ventajas	Desventajas	Ventajas	Desventajas
La eliminación se realiza una única vez	No versátil	Puede mejorar los resultados hasta la precisión de trabajo	Su utilidad depende de la condición de la matriz
Útil para evaluar muchas veces un único SEL			Se necesita el doble de precisión para el residuo
			Se debe elegir un criterio de corte