

CAPÍTULO 5

EJERCICIOS RESUELTOS: MÉTODOS ITERATIVOS PARA ECUACIONES LINEALES

Ejercicios resueltos¹

1. **Examen 21/Junio/1994.** Para la inversión de una matriz cuadrada A de orden $n \times n$, cuya inversa existe, se ha definido la siguiente iteración

$$B_{i+1} = B_i(2I - AB_i), \quad i = 1, 2, \dots$$

donde I es la matriz identidad (o unidad) de orden $n \times n$, y B_1 es una matriz cualquiera de orden $n \times n$. Defina la matriz $C_i = I - AB_i$ y deduzca las propiedades que debe satisfacer dicha matriz para que la iteración anterior converja a la inversa A^{-1} . ¿Cuáles son las desventajas de este método comparado con otros que usted conozca para calcular la inversa de una matriz? Ayuda: Deduzca cotas superiores e inferiores para la norma de la matriz B_{i+1} .

Solución. Lo primero que hay que estudiar en una iteración es su punto fijo

$$B = B(2I - AB), \quad I = AB, \quad B = A^{-1},$$

que en este caso es la inversa de A , como debe ser. Ahora tendremos que estudiar la convergencia de la iteración, es decir, si

$$\lim_{i \rightarrow \infty} B_i = B = A^{-1}.$$

¹© Francisco R. Villatoro, Carmen M. García, Juan I. Ramos. Estas notas están protegidas por derechos de copyright y pueden ser distribuidas libremente sólo con propósitos educativos sin ánimo de lucro. *These notes are copyright-protected but may be freely distributed for instructional nonprofit purposes.*

Utilizando la matriz C_i , podemos escribir la iteración como

$$B_{i+1} = B_i (I + I - A B_i) = B_i (I + C_i),$$

y como

$$\begin{aligned} C_i &= I - A B_i = I - A B_{i-1} (I + C_{i-1}) \\ &= I - A B_{i-1} - A B_{i-1} C_{i-1} \\ &= C_{i-1} - A B_{i-1} C_{i-1} = (I - A B_{i-1}) C_{i-1} \\ &= C_{i-1} C_{i-1} = C_{i-1}^2 = C_{i-2}^4 = C_{i-3}^8, \end{aligned}$$

obtenemos finalmente

$$C_i = C_1^{2^{i-1}}, \quad C_1 = I - A B_1.$$

De esta forma, siguiendo la ayuda y aplicando normas a la iteración

$$\begin{aligned} \|B_{i+1}\| &\leq \|B_i\| \|I + C_i\| \leq \|B_i\| (\|I\| + \|C_i\|) \\ &= \|B_i\| (1 + \|C_i\|) \leq \|B_i\| (1 + \|C_1\|^{2^{i-1}}) \\ &= \|B_1\| \prod_{j=0}^i (1 + \|C_1\|^{2^{j-1}}), \end{aligned}$$

que es poco útil para la demostración de convergencia.

Podemos reescribir la iteración

$$B_{i+1} = B_i + B_i C_i, \quad B_{i+1} - B_i = B_i C_i,$$

y aplicando normas

$$\|B_{i+1}\| - \|B_i\| \leq \|B_i\| \|C_i\| = \|B_i\| \|C_1^{2^{i-1}}\| \leq \|B_i\| \|C_1\|^{2^{i-1}}.$$

Por otro lado, escribiendo la iteración como

$$B_{i+1} = B_i + B_i C_i, \quad B_{i+1} - B_i C_i = B_i,$$

y aplicando normas

$$\|B_i\| = \|B_{i+1} - B_i C_i\| \leq \|B_{i+1}\| + \|B_i C_i\| \leq \|B_{i+1}\| + \|B_i\| \|C_i\|,$$

con lo que

$$\|B_{i+1}\| \geq \|B_i\| (1 - \|C_i\|) \geq \|B_i\| (1 - \|C_1\|^{2^{i-1}}).$$

Finalmente hemos logrado acotar la iteración tanto superior como inferiormente de la forma

$$\|B_i\| \left(1 - \|C_1\|^{2^{i-1}}\right) \leq \|B_{i+1}\| \leq \|B_i\| \left(1 + \|C_1\|^{2^{i-1}}\right).$$

Por tanto para que el método converja es necesario que

$$\lim_{i \rightarrow \infty} \|C_1\|^{2^{i-1}} = 0,$$

es decir, el radio espectral de C_1 (y el valor absoluto de todos los autovalores de C_1) deben ser menores que la unidad. En este caso, también se cumple que

$$\lim_{i \rightarrow \infty} C_i = 0 = \lim_{i \rightarrow \infty} I - AB_i,$$

con lo que

$$\lim_{i \rightarrow \infty} B_i = A^{-1}.$$

La desventaja fundamental de este método es que B_1 se tiene que escoger de tal manera que los autovalores de C_1 sean menores que la unidad. Otra desventaja es que requiere muchas multiplicaciones de matrices (dos multiplicaciones y una suma en cada iteración).

2. ¿Para qué valores de α , la iteración

$$x_{i+1} = \alpha x_i + \beta, \quad \text{con } x_0 = a, \quad \beta \neq 0,$$

converge? Es decir, ¿para qué valores de α ,

$$\lim_{i \rightarrow \infty} x_i,$$

existe?

Solución. Iterando obtenemos

$$\begin{aligned} x_n &= \alpha x_{n-1} + \beta = \alpha^2 x_{n-2} + \beta \alpha + \beta = \dots \\ &= a \alpha^n + \beta \alpha^{n-1} + \dots + \beta \alpha + \beta. \end{aligned}$$

Definamos

$$\begin{aligned} s &= \beta \alpha^{n-1} + \dots + \beta \alpha + \beta, \\ \alpha s &= \beta \alpha^n + \dots + \beta \alpha, \end{aligned}$$

con lo que restando

$$(\alpha - 1) s = \alpha^n \beta - \beta = \beta (\alpha^n - 1),$$

y finalmente

$$x_n = a \alpha^n + \beta \frac{\alpha^n - 1}{\alpha - 1}.$$

Esta secuencia convergerá si $|\alpha| < 1$ al valor

$$\lim_{i \rightarrow \infty} x_n = \frac{-\beta}{\alpha - 1} = \frac{\beta}{1 - \alpha},$$

que corresponde al punto fijo de la iteración (sucesión), ya que si ésta converge a \bar{x} entonces

$$\bar{x} = \alpha \bar{x} + \beta, \quad \bar{x} = \frac{\beta}{1 - \alpha},$$

que es el punto fijo.

3. Dados

$$A = \begin{pmatrix} 10 & 3 & 1 \\ 2 & -10 & 3 \\ 1 & 3 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} 14 \\ -5 \\ 14 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

resuelva el sistema $Ax = b$ por el método de Gauss-Jacobi y determine la tasa de convergencia usando las normas ∞ y 1. Haga lo mismo con el método de Gauss-Seidel y deduzca cuál de ellos converge más rápidamente.

Solución.

(a) El método de Gauss-Jacobi se basa en la descomposición $A = L + D + U$ y tiene la iteración

$$x^{(k+1)} = D^{-1} (b - (L + U) x^{(k)}).$$

Operando, obtenemos la secuencia Operando, obtenemos la secuencia

$$x^{(1)} = \begin{pmatrix} 7/5 \\ 1/2 \\ 7/5 \end{pmatrix}, \quad x^{(2)} = \begin{pmatrix} 111/100 \\ 6/5 \\ 111/100 \end{pmatrix},$$

$$x^{(3)} = \begin{pmatrix} 0.929 \\ 1.055 \\ 0.929 \end{pmatrix}, \quad x^{(4)} = \begin{pmatrix} 0.9906 \\ 0.9645 \\ 0.9906 \end{pmatrix},$$

$$x^{(5)} = \begin{pmatrix} 1.01159 \\ 0.9953 \\ 1.01159 \end{pmatrix}, \quad x^{(6)} = \begin{pmatrix} 1.00025 \\ 1.0058 \\ 1.00025 \end{pmatrix},$$

que claramente converge a la solución exacta $(1, 1, 1)^\top$.

(b) El método de Gauss-Seidel se basa en la iteración

$$x^{(k+1)} = (L + D)^{-1} (b - U x^{(k)}).$$

Operando, obtenemos la secuencia

$$x^{(1)} = \begin{pmatrix} 7/5 \\ 39/50 \\ 513/500 \end{pmatrix}, \quad x^{(2)} = \begin{pmatrix} 1.0634 \\ 1.02048 \\ 0.987516 \end{pmatrix},$$

$$x^{(3)} = \begin{pmatrix} 0.995104 \\ 0.995276 \\ 1.00191 \end{pmatrix}, \quad x^{(4)} = \begin{pmatrix} 1.00123 \\ 1.00082 \\ 0.999632 \end{pmatrix},$$

$$x^{(5)} = \begin{pmatrix} 0.999792 \\ 0.999848 \\ 1.00007 \end{pmatrix}, \quad x^{(6)} = \begin{pmatrix} 1.00004 \\ 1.00003 \\ 0.999988 \end{pmatrix},$$

que claramente converge, y más rápidamente que el Gauss-Jacobi, a la solución exacta $(1, 1, 1)^\top$.

(c) La tasa de convergencia del método de Gauss-Jacobi viene dada por la norma de

$$J = D^{-1}(L + U) = \begin{pmatrix} 0 & 3/10 & 1/10 \\ -1/5 & 0 & -3/10 \\ 1/10 & 3/10 & 0 \end{pmatrix},$$

cuyas normas son $\|J\|_1 = 0.6$ y $\|J\|_\infty = 0.5$. La tasa de convergencia del método de Gauss-Seidel viene dada por la norma de

$$S = (L + D)^{-1}U = \begin{pmatrix} 0 & 3/10 & 1/10 \\ 0 & 3/50 & -7/25 \\ 0 & -6/125 & 37/500 \end{pmatrix},$$

cuyas normas son $\|J\|_1 = 227/500 = 0.454$ y $\|J\|_\infty = 2/5 = 0.4$. Estos resultados confirman la convergencia más rápida observada para el método de Gauss-Seidel.

4. **Examen 4/Septiembre/1996.** Considere el siguiente sistema de ecuaciones

$$x + ay = 1,$$

$$x + y + z = 1,$$

$$by + z = 1.$$

- Determine los valores de a y b para que el sistema tenga solución única
- Determine los valores de a y b para asegurar la convergencia del método de Gauss-Jacobi para la resolución de dicho sistema
- Determine los valores de a y b para asegurar la convergencia del método de Gauss-Seidel
- Estudiar la convergencia de los métodos anteriores y el método de Cholesky para los valores de a y b para los que la matriz de los coeficientes del sistema es simétrica.

Solución.

- El sistema se puede escribir como $Ax = b$ donde

$$A = \begin{pmatrix} 1 & a & 0 \\ 1 & 1 & 1 \\ 0 & b & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

El sistema tendrá solución única cuando A no sea singular,

$$|A| = 1 - a - b \neq 0, \quad a + b \neq 1.$$

(b) El método de Gauss-Jacobi convergerá si su tasa de convergencia

$$\|J\| = \|D^{-1}(L + U)\| = \|L + U\|,$$

es menor que la unidad para alguna matriz. No sirven las norma uno e infinito, ya que para ellas el método no converge,

$$\|J\|_1 = \max\{1, |a| + |b|\}, \quad \|J\|_\infty = \max\{2, |a|, |b|\}.$$

Sin embargo, pueden existir normas para las que converja. La convergencia quedará asegurada si el radio espectral, que siempre es menor o igual que cualquier norma, es menor que la unidad. El radio espectral para el método de Gauss-Jacobi es

$$\rho(J) = \max |\lambda_J|,$$

y sus autovalores son

$$|J - \lambda_J I| = \begin{vmatrix} -\lambda_J & a & 0 \\ 1 & -\lambda_J & 1 \\ 0 & b & -\lambda_J \end{vmatrix} = 0,$$

es decir,

$$-\lambda_J^3 + a\lambda_J + b\lambda_J = 0,$$

que tiene como raíces

$$\lambda_J = 0, \quad \lambda_J = \pm\sqrt{a+b},$$

para $a+b \in \mathbb{R}^+ \cup 0$. Por tanto, la convergencia del método de Gauss-Jacobi requiere que

$$\rho(J) = \sqrt{a+b} < 1.$$

Note que $a+b < 1$ está incluido en dicho conjunto.

(c) El método de Gauss-Seidel convergerá si su matriz de convergencia

$$S = (L + D)^{-1}U = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ b & -b & 1 \end{pmatrix} \begin{pmatrix} 0 & a & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & a & 0 \\ 0 & -a & 1 \\ 0 & ab & -b \end{pmatrix},$$

tiene radio espectral menor que la unidad. Sus autovalores son

$$|S - \lambda_S I| = \begin{vmatrix} -\lambda_S & a & 0 \\ 0 & -a - \lambda_S & 1 \\ 0 & ab & -b - \lambda_S \end{vmatrix} = 0,$$

es decir,

$$\lambda_S (ab - (a + \lambda_S)(b + \lambda_S)) = 0,$$

que tiene como soluciones

$$\lambda_S = 0, \quad \lambda_S^2 + \lambda_S(a + b),$$

es decir,

$$\lambda_S = -(a + b), \quad \lambda_S = 0.$$

Por tanto, el método de Gauss-Seidel converge para

$$\rho(S) = |a + b| < 1.$$

Note que las condiciones de convergencia de Gauss-Jacobi y Gauss-Seidel son disjuntas y que hay valores de (a, b) tales que uno puede converger y el otro no, y viceversa.

- (d) La matriz A es simétrica para $a = b = 1$, por lo que el método de Gauss-Jacobi no converge,

$$\rho(J) = \sqrt{a + a} = \sqrt{2} > 1,$$

el método de Gauss-Seidel converge cuando

$$\rho(S) = 2|a| < 1.$$

En cuanto al método de Cholesky, no se trata de un método iterativo si no uno directo, por lo que no se puede estudiar su convergencia. Además el método de Cholesky requiere que la matriz A sea hermitiana y positiva definida. Si $a = b = 1$, la matriz es simétrica por lo que hay que estudiar si es definida positiva. Podemos determinar sus autovalores,

$$|C - \lambda_C I| = \begin{vmatrix} 1 - \lambda_C & 1 & 0 \\ 1 & 1 - \lambda_C & 1 \\ 0 & 1 & 1 - \lambda_C \end{vmatrix} = 0,$$

$$-\lambda_C^3 + 3\lambda_C^2 - \lambda_C - 1 = 0,$$

una de cuyas raíces es $\lambda_C = 1$, y por la regla de Ruffini,

$$-\lambda_C^2 + 2\lambda_C + 1 = 0$$

por lo que las otras dos raíces son

$$\lambda_C = 1 \pm \sqrt{2},$$

una de las cuales es negativa, por lo que no es definida positiva.

Otra posibilidad para comprobar que no es definida positiva es utilizar el signo de los menores principales

$$C_1 = |1| = 1 > 0,$$

$$C_2 = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0,$$

$$C_3 = |C| = 1 - 2 = -1 < 0,$$

que claramente indican que no es definida, y por tanto, el método de Cholesky no es aplicable.

5. Considere el siguiente sistema de ecuaciones

$$x + ay = a$$

$$ax + y + bz = b$$

$$by + z = c$$

- Determine los valores de a y b para que el sistema tenga solución única
- Determine los valores de a y b para asegurar la convergencia del método de Gauss-Jacobi para la resolución de dicho sistema
- Determine los valores de a y b para asegurar la convergencia del método de Gauss-Seidel
- Estudiar la convergencia de los métodos anteriores y el método de Cholesky para los valores de a y b para los que la matriz de los coeficientes del sistema es simétrica.

Solución.

(a) El sistema se puede escribir como $Ax = b$ donde

$$A = \begin{pmatrix} 1 & a & 0 \\ a & 1 & b \\ 0 & b & 1 \end{pmatrix}, \quad b = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

El sistema tendrá solución única cuando A no sea singular,

$$|A| = 1 - a^2 - b^2 \neq 0, \quad a^2 + b^2 \neq 1.$$

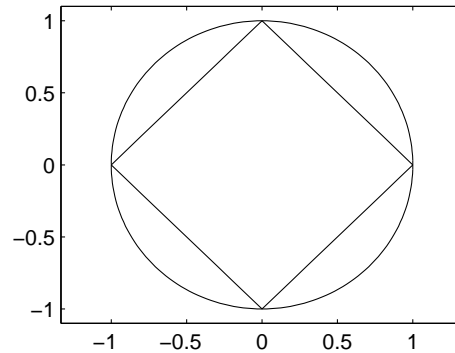
(b) El método de Gauss-Jacobi convergerá si su tasa de convergencia

$$\|J\| = \|D^{-1}(L + U)\| = \|L + U\|,$$

es menor que la unidad para alguna norma. Las norma uno e infinito nos dan la condición

$$\|J\|_1 = \|J\|_\infty = \max\{|a|, |b|, |a| + |b|\} = |a| + |b| < 1,$$

es decir, (a, b) incluido en el rombo simétrico que aparece en la siguiente figura.



Sin embargo, la condición anterior es suficiente, pero no necesaria. Una condición suficiente y necesaria se obtiene usando el radio espectral,

$$\rho(J) = \max |\lambda_J|.$$

Los autovalores de J son

$$|J - \lambda_J I| = \lambda_J (a^2 + b^2 - \lambda_J^2) = 0,$$

es decir, $\lambda_J \in \{0, \pm\sqrt{a^2 + b^2}\}$, y la convergencia del método de Gauss-Jacobi requiere que el punto (a, b) se encuentre en el disco unidad S_1 (ver figura).

(c) El método de Gauss-Seidel convergerá si su matriz de convergencia

$$S = (L + D)^{-1} U = \begin{pmatrix} 0 & a & 0 \\ 0 & -a^2 & b \\ 0 & a^2 b & -b^2 \end{pmatrix},$$

tiene radio espectral menor que la unidad. Sus autovalores cumplen

$$|S - \lambda_S I| = -\lambda_S^2(a^2 + b^2 + \lambda_S) = 0,$$

es decir, $\lambda_S \in \{0, 0, -(a^2 + b^2)\}$, por lo que la condición de convergencia de Gauss-Seidel es la misma que la de Jacobi, (a, b) debe pertenecer al disco unidad.

(d) La matriz A es simétrica para todo (a, b) . En cuanto al método de Cholesky, no se trata de un método iterativo sino uno directo, por lo que no se puede estudiar su convergencia. Además el método de Cholesky requiere que la matriz A sea hermítica y positiva definida. La condición para que sea definida positiva es que sus autovalores sean positivos,

$$|A - \lambda I| = 1 - a^2 - b^2 - (3 - a^2 - b^2)\lambda + 3\lambda^2 - \lambda^3 = 0,$$

que por inspección nos conduce a $\lambda = 1$ como primera raíz y a

$$-1 + a^2 + b^2 + 2\lambda - \lambda^2 = 0, \quad \lambda_{\pm} = 1 \pm \sqrt{a^2 + b^2},$$

que serán dos raíces positivas sólo si (a, b) pertenece al disco unidad. Sólo en dicho caso será aplicable el método de Cholesky.

Otra posibilidad para estudiar cuando es definida positiva es utilizar el signo de los menores principales

$$\begin{aligned} A_1 &= |1| = 1 > 0, \\ A_2 &= \begin{vmatrix} 1 & a \\ a & 1 \end{vmatrix} = 1 - a^2 > 0, \quad a^2 < 1 \\ A_3 &= |A| = 1 - a^2 - b^2 > 0, \quad a^2 + b^2 < 1, \end{aligned}$$

que muestra de nuevo la condición previa..

6. **Examen 6/Julio/1997.** Sea B una matriz real cuadrada de orden $n \times n$ tal que $Bx = 0$ y $|B| = 0$.

- (a) ¿Cuál es la relación entre b_{ii} y b_{ij} , donde $B = (b_{ij})$?
- (b) Si $B = A - \lambda I$ donde λ son los autovalores de A , ¿cuál es la relación entre λ y los coeficientes o elementos de la matriz A ?
- (c) ¿Cuál es la relación entre $|\lambda|_{\max}$ y $|\lambda|_{\min}$ y los elementos de A ?
- (d) Sea A una matriz cuadrada de orden $n \times n$ tal que $a_{ii} = 1$. Deduzca las condiciones que deben satisfacer los coeficientes de esta matriz para que el método iterativo de Gauss-Jacobi converja cuando se utiliza para resolver el sistema $Ax = b$.

Solución.

- (a) La ecuación $Bx = 0$ tiene infinitas soluciones ya que $|B| = 0$. Sea x^* una de estas soluciones y sea x_i^* tal que

$$|x_i^*| = \max_{1 \leq j \leq n} |x_j^*|,$$

entonces $Bx^* = 0$ implica que

$$0 = \sum_{j=1}^n b_{ij} x_j^* = b_{ii} x_i^* + \sum_{j=1, j \neq i}^n b_{ij} x_j^*,$$

por lo que

$$|b_{ii}| |x_i^*| \leq \sum_{j=1, j \neq i}^n |b_{ij}| |x_j^*|,$$

y, finalmente, observamos que B no es diagonalmente dominante por filas

$$|b_{ii}| \leq \sum_{j=1, j \neq i}^n |b_{ij}| \frac{|x_j^*|}{|x_i^*|} \leq \sum_{j=1, j \neq i}^n |b_{ij}|.$$

- (b) Si $B = A - \lambda I$ entonces

$$b_{ii} = a_{ii} - \lambda, \quad b_{ij} = a_{ij},$$

y por la condición del apartado anterior

$$|a_{ii} - \lambda| \leq \sum_{j=1, j \neq i}^n |a_{ij}|.$$

- (c) Aplicando la desigualdad triangular inversa

$$||\lambda| - |a_{ii}|| \leq |\lambda - a_{ii}|,$$

y el resultado del apartado anterior tenemos que se cumplen las dos desigualdades

$$|\lambda| \leq |a_{ii}| + \sum_{j=1, j \neq i}^n |a_{ij}| = \sum_{j=1}^n |a_{ij}|,$$

$$|\lambda| \geq |a_{ii}| - \sum_{j=1, j \neq i}^n |a_{ij}|,$$

por lo que en los casos límites podemos obtener una cota superior para el mayor autovalor y una inferior para el menor,

$$|\lambda|_{\max} \leq \max_i \sum_{j=1}^n |a_{ij}|,$$

$$|\lambda|_{\min} \geq \min_i \left(|a_{ii}| - \sum_{j=1, j \neq i}^n |a_{ij}| \right),$$

y por tanto el número de condición de A cumple que

$$\kappa(A) \geq \frac{|\lambda|_{\max}}{|\lambda|_{\min}} \geq \frac{\max_i \sum_{j=1}^n |a_{ij}|}{\min_i \left(2|a_{ii}| - \sum_{j=1}^n |a_{ij}| \right)}.$$

- (d) Sea A una matriz cuadrada de orden $n \times n$ tal que $a_{ii} = 1$. Deduzca las condiciones que deben satisfacer los coeficientes de esta matriz para que el método iterativo de Gauss-Jacobi converja cuando se utiliza para resolver el sistema $Ax = b$.

Descomponiendo la matriz $A = L + D + U$, el método de Gauss-Jacobi tiene como iteración

$$x^{(k+1)} = D^{-1} (b - (L + U)x^{(k)}),$$

por lo que para que converja, el radio espectral de su matriz de iteración debe ser menor que la unidad

$$\rho(J) = \rho(D^{-1}(L + U)) = |\lambda_J|_{\max} < 1.$$

Ahora bien, por ser $D = I$, $A = I + N$ y los autovalores de A

$$\lambda_A = 1 + \lambda_J,$$

por lo que la condición de convergencia es $\rho(A) < 2$, pero por el apartado anterior del ejercicio

$$\rho(A) \leq \max_i \left(1 + \sum_{j=1, i \neq j}^n |a_{ij}| \right) < 2,$$

luego

$$\sum_{j=1, i \neq j}^n |a_{ij}| < 1.$$

De forma del todo similar, ya que los autovalores son invariantes ante trasposición,

$$\sum_{i=1, j \neq i}^n |a_{ij}| < 1.$$

7. Comprobar que la matriz que determina el sistema

$$\begin{aligned} 10x_1 - x_2 &= 9, \\ -x_1 + 10x_2 - 2x_3 &= 7, \\ -2x_2 + 10x_3 &= 6, \end{aligned}$$

es definida positiva. ¿Qué parámetro de relajación w escogería para obtener una convergencia más rápida? Escribir las 3 primeras iteraciones del método de relajación con esa w tomando como valores iniciales $x = 0$. Comparar con las 3 primeras iteraciones de Gauss-Seidel.

Solución. El sistema se puede escribir como $Ax = b$ donde

$$A = \begin{pmatrix} 10 & -1 & 0 \\ -1 & 10 & -2 \\ 0 & -2 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} 9 \\ 7 \\ 6 \end{pmatrix}.$$

Para comprobar que esta matriz es definida positiva podemos calcular sus autovalores

$$|A - \lambda I| = (\lambda - 10)(-\lambda^2 + 20\lambda - 95) = 0,$$

que son

$$\lambda = 10, \quad \lambda = 10 \pm \sqrt{5},$$

que claramente son positivos. También podríamos haber aplicado la regla de los menores principales

$$A_1 = |10| = 10 > 0, \quad A_2 = \begin{vmatrix} 10 & -1 \\ -1 & 10 \end{vmatrix} = 99 > 0,$$

y $A_3 = |A| = 950 > 0$.

Como no se dice si debemos aplicar el método de relajación al método de Gauss-Jacobi (que converge porque la matriz es diagonalmente dominante) o al de Gauss-Seidel (que converge por lo anterior y además porque la matriz de coeficientes es simétrica y definida positiva). Vamos a escoger este último. Descomponiendo la matriz $A = L + D + U$, tenemos que el método de Gauss-Seidel es

$$x^{(k+1)} = (L + D)^{-1}(b - Ux^{(k)}),$$

e introduciendo un parámetro de relajación

$$x^{(k+1)} = w(L + D)^{-1}(b - Ux^{(k)}) + (1 - w)x^{(k)},$$

con lo que el error $e^{(k)} = x - x^{(k)}$ sigue la ecuación

$$e^{(k+1)} = ((1 - w)I - w(L + D)^{-1}U)e^{(k)} = Ne^{(k)},$$

y la convergencia del método queda garantizada si la tasa de convergencia es menor que la unidad

$$\rho(N) < 1.$$

Como

$$(L + D)^{-1} = \begin{pmatrix} 1/10 & 0 & 0 \\ 1/100 & 1/10 & 0 \\ 1/500 & 1/50 & 1/10 \end{pmatrix},$$

$$(L + D)^{-1}U = \begin{pmatrix} 0 & -1/10 & 0 \\ 0 & -1/100 & -1/5 \\ 0 & -1/500 & -1/25 \end{pmatrix},$$

tenemos que

$$N = (1 - w)I - w(L + D)^{-1}U = \begin{pmatrix} 1 - w & w/10 & 0 \\ 0 & 1 - 99w/100 & w/5 \\ 0 & w/500 & 1 - 24w/25 \end{pmatrix},$$

y por tanto su polinomio característico es

$$|N - \lambda I| = (1 - 19w/20 - \lambda)(1 - w - \lambda)^2,$$

y el radio espectral será menor que la unidad si

$$|1 - w| < 1, \quad |1 - 19w/20| < 1,$$

y por tanto $0 < w < 2$ (de la primera desigualdad) garantiza la convergencia del método. Para obtener la convergencia más rápida debemos buscar el valor w^* que minimiza el radio

espectral. Dibujando gráficamente el valor del radio espectral obtenemos que

$$\rho(N) = \begin{cases} 1 - 19w/20 & 0 < w \leq w^* \\ w - 1 & w^* \leq w < 2 \end{cases},$$

por lo que el valor óptimo es

$$1 - 19w^*/20 = w^* - 1, \quad w^* = \frac{40}{39} \approx 1.026.$$

Seguidamente vamos a realizar tres iteraciones del método de Gauss-Seidel con relajación con w^* ,

$$x_w^{(1)} = \begin{pmatrix} 12/13 \\ 158/195 \\ 758/975 \end{pmatrix} = \begin{pmatrix} 0.923077 \\ 0.810256 \\ 0.777436 \end{pmatrix},$$

$$x_w^{(2)} = \begin{pmatrix} 0.982512 \\ 0.957265 \\ 0.791059 \end{pmatrix}, \quad x_w^{(3)} = \begin{pmatrix} 0.996065 \\ 0.957798 \\ 0.79157 \end{pmatrix},$$

y tres iteraciones del método de Gauss-Seidel sin relajación ($w = 1$),

$$x^{(1)} = \begin{pmatrix} 9/10 \\ 79/100 \\ 379/500 \end{pmatrix} = \begin{pmatrix} 0.9 \\ 0.79 \\ 0.758 \end{pmatrix},$$

$$x^{(2)} = \begin{pmatrix} 0.979 \\ 0.9495 \\ 0.7899 \end{pmatrix} \quad x^{(3)} = \begin{pmatrix} 0.99495 \\ 0.957475 \\ 0.791495 \end{pmatrix}.$$

Al ser el sistema de 3×3 la solución exacta se puede calcular por un método directo dando

$$x = \begin{pmatrix} 0.995789 \\ 0.957895 \\ 0.791579 \end{pmatrix}.$$

Comparando los errores obtenidos

$$\|x - x_w^{(3)}\|_\infty = 0.000276, \quad \|x - x^{(3)}\|_\infty = 0.000839,$$

$$\|x - x_w^{(3)}\|_1 = 0.000382, \quad \|x - x^{(3)}\|_1 = 0.001343,$$

se observa que el método con relajación tiene una convergencia más rápida (menor error), aunque las diferencias entre los dos métodos son pequeñas debido a que el valor óptimo del parámetro de relajación w^* es muy próximo a la unidad.

8. Comprobar que la matriz que determina el sistema

$$10x_1 - 3x_2 = 2$$

$$-3x_1 + 10x_2 - 2x_3 = 3$$

$$-2x_2 + 10x_3 = 5,$$

es definida positiva. ¿Qué parámetro de relajación w escogería para obtener una convergencia más rápida? Escribir las 3 primeras iteraciones del método de relajación con esa w tomando como valores iniciales $x = 0$. Comparar con las 3 primeras iteraciones de Gauss-Seidel.

Solución. El sistema se puede escribir como $Ax = b$ donde

$$A = \begin{pmatrix} 10 & -3 & 0 \\ -3 & 10 & -2 \\ 0 & -2 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}.$$

Para comprobar que esta matriz es definida positiva podemos calcular sus autovalores

$$|A - \lambda I| = (\lambda - 10)(-\lambda^2 + 20\lambda - 87) = 0,$$

que son

$$\lambda = 10, \quad \lambda = 10 \pm \sqrt{13},$$

que claramente son positivos. También podríamos haber aplicado la regla de los menores principales

$$A_1 = |10| = 10 > 0, \quad A_2 = \begin{vmatrix} 10 & -3 \\ -3 & 10 \end{vmatrix} = 91 > 0,$$

y $A_3 = |A| = 870 > 0$.

Como no se dice si debemos aplicar el método de relajación al método de Gauss-Jacobi (que converge porque la matriz es diagonalmente dominante) o al de Gauss-Seidel (que converge por lo anterior y además porque la matriz de coeficientes es simétrica y definida positiva), vamos a escoger este último, que es más rápido. Descomponiendo la matriz $A = L + D + U$, tenemos que el método de Gauss-Seidel es

$$x^{(k+1)} = (L + D)^{-1} (b - U x^{(k)}),$$

e introduciendo un parámetro de relajación

$$x^{(k+1)} = w (L + D)^{-1} (b - U x^{(k)}) + (1 - w) x^{(k)},$$

con lo que el error $e^{(k)} = x - x^{(k)}$ sigue la ecuación

$$e^{(k+1)} = ((1 - w) I - w (L + D)^{-1} U) e^{(k)} = N e^{(k)},$$

y la convergencia del método queda garantizada si la tasa de convergencia es menor que la unidad

$$\rho(N) < 1.$$

Como

$$(L + D)^{-1} = \begin{pmatrix} 1/10 & 0 & 0 \\ 3/100 & 1/10 & 0 \\ 3/500 & 1/50 & 1/10 \end{pmatrix},$$

$$(L + D)^{-1} U = \begin{pmatrix} 0 & -3/10 & 0 \\ 0 & -9/100 & -1/5 \\ 0 & -9/500 & -1/25 \end{pmatrix},$$

tenemos que

$$N = (1 - w) I - w (L + D)^{-1} U = \begin{pmatrix} 1 - w & 3w/10 & 0 \\ 0 & 1 - 91w/100 & w/5 \\ 0 & 9w/500 & 1 - 24w/25 \end{pmatrix},$$

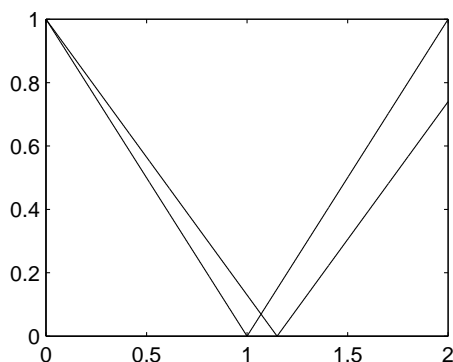
y por tanto su polinomio característico es

$$|N - \lambda I| = (1 - 87w/100 - \lambda)(1 - w - \lambda)^2,$$

y el radio espectral será menor que la unidad si

$$|1 - w| < 1, \quad |1 - 87w/100| < 1,$$

y por tanto $0 < w < 2$ (de la primera desigualdad) garantiza la convergencia del método. Para obtener la convergencia más rápida debemos buscar el valor w^* que minimiza el radio espectral. Dibujando gráficamente el valor del radio espectral obtenemos



por lo que el valor óptimo es

$$1 - 87w^*/100 = w^* - 1, \quad w^* = \frac{200}{187} \approx 1.069.$$

Seguidamente vamos a realizar tres iteraciones del método de Gauss-Seidel con relajación con w^* con $x_w^{(0)} = 0$,

$$x_w^{(1)} = \begin{pmatrix} 40/187 \\ 72/187 \\ 52/85 \end{pmatrix} = \begin{pmatrix} 0.322571 \\ 0.52618 \\ 0.60282 \end{pmatrix},$$

$$x_w^{(2)} = \begin{pmatrix} 0.360307 \\ 0.528041 \\ 0.605776 \end{pmatrix}, \quad x_w^{(3)} = \begin{pmatrix} 0.358281 \\ 0.528723 \\ 0.605733 \end{pmatrix},$$

y tres iteraciones del método de Gauss-Seidel sin relajación ($w = 1$),

$$x_w^{(1)} = \begin{pmatrix} 1/5 \\ 9/25 \\ 143/250 \end{pmatrix} = \begin{pmatrix} 0.308 \\ 0.5068 \\ 0.60136 \end{pmatrix},$$

$$x_w^{(2)} = \begin{pmatrix} 0.35204 \\ 0.525884 \\ 0.605177 \end{pmatrix}, \quad x_w^{(3)} = \begin{pmatrix} 0.357765 \\ 0.528365 \\ 0.605673 \end{pmatrix},$$

Al ser el sistema de 3×3 la solución exacta se puede calcular por un método directo dando

$$x = \begin{pmatrix} 0.358621 \\ 0.528736 \\ 0.605747 \end{pmatrix}.$$

Comparando los errores obtenidos

$$\|x - x_w^{(3)}\|_\infty = 0.0017, \quad \|x - x^{(3)}\|_\infty = 0.0066,$$

$$\|x - x_w^{(3)}\|_1 = 0.0024, \quad \|x - x^{(3)}\|_1 = 0.0100,$$

se observa que el método con relajación tiene una convergencia más rápida (menor error), aunque las diferencias entre los dos métodos son pequeñas debido a que el valor óptimo del parámetro de relajación w^* es muy próximo a la unidad.

9. ¿Puede converger un método iterativo para la resolución de un sistema de ecuaciones lineales de la forma

$$x^{(k)} = B x^{(k-1)} + c,$$

siendo B una matriz singular ($|B| = 0$)?

Solución. El punto fijo de esta iteración es

$$x = Bx + c, \quad x = (I - B)^{-1} c,$$

que requiere que $(I - B)^{-1}$ exista, pero no es necesario que exista B^{-1} . Para estudiar la convergencia de este método, debemos estudiar la ecuación del error

$$e^{(k)} = B e^{(k-1)},$$

que nos indica que el método convergerá si

$$\rho(B) < 1.$$

Que B sea singular (por ejemplo, por tener un autovalor nulo) no implica que no se pueda cumplir la condición de convergencia $\rho(B) < 1$. Luego la respuesta a la pregunta es afirmativa.

10. **Examen 13/Marzo/1993.** Dados

$$A = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 6 \\ 2 \end{pmatrix},$$

determine

- el vector x tal que $Ax = b$ por factorización de Cholesky,
- A^{-1} a partir de la factorización de Cholesky,
- la convergencia del método de Gauss-Jacobi realizando cuatro iteraciones con el vector inicial $x^{(0)} = 0$,
- la convergencia del método de Gauss-Seidel realizando cuatro iteraciones con el vector inicial $x^{(0)} = 0$,
- el polinomio característico de A por el método de LeVerrier y acote sus raíces (es decir, los autovalores de A),
- la descomposición LU (es decir, $A = LU$ donde L es una matriz triangular inferior con unos en la diagonal principal y U es una matriz triangular superior).

Solución.

- Para aplicar el método directo de Cholesky es necesario que la matriz sea simétrica y definida positiva. Para comprobar este último punto aplicaremos la regla de los menores principales

$$A_1 = |4| = 4 > 0, \quad A_2 = \begin{vmatrix} 4 & -1 \\ -1 & 4 \end{vmatrix} = 15 > 0,$$

y $A_3 = |A| = 56 > 0$, luego es definida positiva.

La factorización de Cholesky $A = L L^\top$ se determina fácilmente igualando fila a fila las matrices

$$\begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{pmatrix} = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix},$$

es decir,

$$\begin{aligned} l_{11}^2 &= 4, & l_{11} &= 2, \\ l_{11} l_{21} &= -1, & l_{21} &= -1/2, \\ l_{11} l_{31} &= 0, & l_{31} &= 0, \\ l_{21}^2 + l_{22}^2 &= 4, & l_{22}^2 &= 4 - 1/4 = 15/4, \\ l_{21} l_{31} + l_{22} l_{32} &= -1, & l_{32} &= -2/\sqrt{15}, \\ l_{31}^2 + l_{32}^2 + l_{33}^2 &= 4, & l_{33}^2 &= 4 - 4/15 = 56/15, \end{aligned}$$

por lo que

$$L = \begin{pmatrix} 2 & 0 & 0 \\ -1/2 & \sqrt{15}/2 & 0 \\ 0 & -2/\sqrt{15} & \sqrt{56/15} \end{pmatrix}.$$

Ahora para resolver el sistema lineal $Ax = L L^\top x = b$, debemos resolver los dos sistemas triangulares $Ly = b$ y $L^\top x = y$. El primero de ellos tiene como solución

$$y = (1, 13/\sqrt{15}, \sqrt{56/15})^\top,$$

y el segundo,

$$x = (1, 2, 1)^\top,$$

que es la solución buscada.

(b) $A^{-1} = (L^\top)^{-1} L^{-1} = (L^{-1})^\top L^{-1}$ y la inversa de la matriz L es fácil de determinar

$$L^{-1} = \frac{1}{\sqrt{56}} \begin{pmatrix} \sqrt{56}/2 & 0 & 0 \\ \sqrt{56/15}/2 & 2\sqrt{56/15} & 0 \\ 1/\sqrt{15} & 4/\sqrt{15} & \sqrt{15} \end{pmatrix}$$

$$= \begin{pmatrix} 1/2 & 0 & 0 \\ \sqrt{15}/30 & 2/\sqrt{15} & 0 \\ \sqrt{210}/420 & \sqrt{210}/105 & \sqrt{56}/15 \end{pmatrix},$$

con lo que

$$A^{-1} = (L^{-1})^{\top} L^{-1} = \begin{pmatrix} 15/56 & 1/14 & 1/56 \\ 1/14 & 2/7 & 1/14 \\ 1/56 & 1/14 & 15/56 \end{pmatrix}.$$

- (c) Como la matriz A es diagonalmente dominante, la convergencia del método de Gauss-Jacobi

$$x^{(k+1)} = D^{-1} (b - (L + U) x^{(k)}),$$

queda garantizada. Realizaremos cuatro iteraciones con el vector inicial $x^{(0)} = 0$, que dan lugar a

$$x^{(1)} = \begin{pmatrix} 1/2 \\ 3/2 \\ 1/2 \end{pmatrix}, \quad x^{(2)} = \begin{pmatrix} 7/8 \\ 7/4 \\ 7/8 \end{pmatrix},$$

$$x^{(3)} = \begin{pmatrix} 0.9375 \\ 1.9375 \\ 0.9375 \end{pmatrix}, \quad x^{(4)} = \begin{pmatrix} 0.984375 \\ 1.96875 \\ 0.984375 \end{pmatrix},$$

que claramente converge a la solución exacta determinada en la pregunta anterior.

- (d) Como la matriz A es simétrica y definida positiva, la convergencia del método de Gauss-Seidel

$$x^{(k+1)} = (L + D)^{-1} (b - U x^{(k)}),$$

queda garantizada. Realizaremos cuatro iteraciones con el vector inicial $x^{(0)} = 0$,

$$x^{(1)} = \begin{pmatrix} 1/2 \\ 13/8 \\ 29/32 \end{pmatrix}, \quad x^{(2)} = \begin{pmatrix} 0.90625 \\ 1.95312 \\ 0.988281 \end{pmatrix},$$

$$x^{(3)} = \begin{pmatrix} 0.988281 \\ 1.99414 \\ 0.998535 \end{pmatrix}, \quad x^{(4)} = \begin{pmatrix} 0.998535 \\ 1.99927 \\ 0.999817 \end{pmatrix},$$

que converge más rápidamente que el Gauss-Jacobi.

(e) El método de LeVerrier nos da el polinomio característico

$$p(\lambda) = (-1)^n (\lambda^n - p_1 \lambda^{n-1} - p_2 \lambda^{n-2} - \dots - p_{n-1} \lambda - p_n) = 0,$$

mediante el cálculo de los coeficientes

$$\begin{aligned} p_1 &= \text{traza}(B_1), & B_1 &= A, \\ p_2 &= \frac{1}{2} \text{traza}(B_2), & B_2 &= A(B_1 - p_1 I), \\ & \dots & \dots & \\ p_n &= \frac{1}{n} \text{traza}(B_n), & B_n &= A(B_{n-1} - p_{n-1} I). \end{aligned}$$

El método de LeVerrier también permite determinar la inversa

$$A^{-1} = \frac{1}{p_n} (B_{n-1} - p_{n-1} I),$$

ya que $B_n - p_n I = 0$.

Aplicando el método de LeVerrier en nuestro caso, obtenemos

$$\begin{aligned} B_1 &= \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix}, & p_1 &= 12, \\ B_2 &= \begin{pmatrix} -31 & 4 & 0 \\ 4 & -30 & 4 \\ 1 & 4 & -31 \end{pmatrix}, & p_2 &= -92/2 = -46, \\ B_3 &= \begin{pmatrix} 56 & 0 & 0 \\ 0 & 56 & 0 \\ 0 & 0 & 56 \end{pmatrix}, & p_3 &= 168/3 = 56, \end{aligned}$$

por lo que

$$p(\lambda) = (-1)^3 (\lambda^3 - 12\lambda^2 + 46\lambda - 56) = 0.$$

Para acotar las raíces podemos tantear algunos valores del polinomio característico, aunque este no es el procedimiento más eficiente y/o adecuado,

$$p(1) = -21, \quad p(2) = -4, \quad p(3) = 1,$$

$$p(4) = 0, \quad p(5) = -1, \quad p(6) = 4,$$

con lo que tenemos, por el teorema de Bolzano, una raíz en $[2, 3]$ (≈ 5.59), otra en $[5, 6]$ (≈ 5.41) y la última es exactamente 4.

- (f) Vamos a realizar una descomposición LU de tipo Doolittle (es decir, $A = LU$ donde L es una matriz triangular inferior con unos en la diagonal principal y U es una matriz triangular superior). Esta factorización se determina fácilmente igualando fila y columna, alternativamente, en el producto de matrices siguiente

$$\begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{21} & u_{31} \\ 0 & u_{22} & u_{32} \\ 0 & 0 & u_{33} \end{pmatrix} = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix},$$

es decir,

$$u_{11} = 4, \quad u_{12} = -1, \quad u_{13} = 0,$$

$$l_{21} = -1/4, \quad l_{31} = 0,$$

$$l_{11} l_{31} = 0, \quad l_{31} = 0,$$

$$u_{22} = 4 - l_{21} u_{12} = 15/4, \quad u_{23} = -1,$$

$$l_{32} = -1/u_{22} = -4/15, \quad u_{33} = 4 - l_{32} u_{23} = 56/15,$$

por lo que

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -1/4 & 1 & 0 \\ 0 & -4/15 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 4 & -1 & 0 \\ 0 & 15/4 & -1 \\ 0 & 0 & 56/15 \end{pmatrix}.$$

11. **Examen 9/Junio/1993.** Dado el sistema $Ax = b$, donde

$$A = \begin{pmatrix} 10 & 3 & 1 \\ 2 & -10 & 3 \\ 1 & 3 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} 14 \\ -5 \\ 14 \end{pmatrix},$$

resuélvalo por medio de los siguientes métodos

- un método directo que, a su juicio, sea el más preciso,
- justifique el uso del método que utilizó en (a),
- ¿puede resolver este sistema por el método de Gauss-Jacobi? ¿Por qué? Si lo puede hacer, haga sólo dos iteraciones y determine la tasa de convergencia,
- ¿puede resolver este sistema por el método de Gauss-Seidel? ¿Por qué? Si lo puede hacer, haga sólo dos iteraciones y determine la tasa de convergencia,
- utilice un parámetro de relajación w y determine para qué valores de dicho parámetro un método de relajación converge. Para un valor $w \neq 0$ y $w \neq 1$, si para dicho valor el método converge, haga dos iteraciones y determine la tasa de convergencia del método de relajación que ha utilizado.

Solución. Dado el sistema $Ax = b$, donde

$$A = \begin{pmatrix} 10 & 3 & 1 \\ 2 & -10 & 3 \\ 1 & 3 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} 14 \\ -5 \\ 14 \end{pmatrix},$$

resuélvalo por medio de los siguientes métodos

- Dado que la matriz no es simétrica, vamos a utilizar el método de Gauss-Jorda con pivotaje total o completo. Para el primer pivote no hay que reordenar la matriz, y tenemos

$$\left(\begin{array}{ccc|c} 10 & 3 & 1 & 14 \\ 2 & -10 & 3 & -5 \\ 1 & 3 & 10 & 14 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 10 & 3 & 1 & 14 \\ 0 & 53 & -14 & 39 \\ 0 & -27 & -99 & -126 \end{array} \right);$$

ahora debemos intercambiar las filas 2 y 3, y las columnas 2 y 3,

$$\left(\begin{array}{ccc|c} 10 & 3 & 1 & 14 \\ 0 & -99 & -27 & -126 \\ 0 & -14 & 53 & 39 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 10 & 3 & 1 & 14 \\ 0 & -99 & -27 & -126 \\ 0 & 0 & 5625 & 5625 \end{array} \right);$$

y ahora resolviendo el sistema triangular superior obtenido recordando que las incógnitas están en el orden 2, 3, 1, tenemos

$$x_2 = 1, \quad x_3 = (-126 + 27)(-99) = 1, \quad x_1 = 1.$$

- (b) La ventaja del método de Gauss-Jordan con pivotaje completo es que minimiza posibles diferencias cancelativas y divisiones por números cercanos pequeños, que son susceptibles a propagación de errores.
- (c) Al ser la matriz de coeficientes A diagonalmente dominante por filas, el método iterativo de Gauss-Jacobi está garantizado que converge,

$$x^{(k)} = D^{-1} (b - (L + U) x^{(k-1)}).$$

Vamos a realizar dos iteraciones a partir de $x^{(0)} = 0$,

$$x^{(1)} = (7/5, 1/2, 7/5)^\top, \quad x^{(2)} = (111/100, 6/5, 111/100)^\top.$$

La tasa de convergencia del método de Gauss-Jacobi es el radio espectral de su matriz de convergencia

$$\rho(D^{-1} (L + U)),$$

donde

$$J = D^{-1} (L + U) = \frac{1}{10} \begin{pmatrix} 0 & 3 & 1 \\ -2 & 0 & -3 \\ 1 & 3 & 0 \end{pmatrix}$$

cuyo polinomio característico es

$$|J - \lambda I| = -3/200 - 7\lambda/50 - \lambda^3 = (1 + 10\lambda)(-3 + 2\lambda - 20\lambda^2) = 0,$$

que resolviendo la ecuación cuadrática (2 raíces complejas conjugadas) nos da para el radio espectral, es decir, la tasa de convergencia del método de Gauss-Jacobi

$$\rho(J) = \max_i |\lambda_{Ji}| = \max\{1/10, \sqrt{15}/10\} = \sqrt{15}/10 \approx 0.387.$$

La tasa de convergencia también se puede estimar utilizando los dos últimos (en nuestro caso los únicos) iterados del método de la forma

$$\|e_2\| \leq \|J\| \|e_1\|, \quad \|J\| \approx \frac{\|e_2\|_1}{\|e_1\|_1} = \frac{21}{65} \approx 0.323,$$

donde $e_i = x - x_i$.

- (d) Al ser la matriz de coeficientes A diagonalmente dominante por filas, el método iterativo de Gauss-Seidel está garantizado que converge (de hecho, si Gauss-Jacobi converge entonces también lo hace Gauss-Seidel y más rápido),

$$x^{(k)} = (L + D)^{-1} (b - U x^{(k-1)}).$$

Vamos a realizar dos iteraciones a partir de $x^{(0)} = 0$,

$$x^{(1)} = (7/5, 39/50, 513/500)^\top, \quad x^{(2)} = (1.0634, 1.02048, 0.987516)^\top.$$

La tasa de convergencia del método de Gauss-Jacobi es el radio espectral de su matriz de convergencia

$$\rho((L + D)^{-1} U),$$

donde

$$S = (L + D)^{-1} U = \begin{pmatrix} 0 & 3/10 & 1/10 \\ 0 & 3/50 & -7/25 \\ 0 & -6/125 & 37/500 \end{pmatrix}$$

cuyo polinomio característico es

$$|J - \lambda I| = 9\lambda/1000 + 67\lambda^2/500 - \lambda^3,$$

que resolviendo la ecuación cuadrática (2 raíces reales) nos da para el radio espectral, es decir, la tasa de convergencia del método de Gauss-Seidel

$$\rho(S) = \max_i |\lambda_{S_i}| = \max\{0, -0.0491, 0.183\} \approx 0.183.$$

La tasa de convergencia también se puede estimar utilizando los dos últimos (en nuestro caso los únicos) iterados del método de la forma

$$\|e_2\| \leq \|S\| \|e_1\|, \quad \|S\| \approx \frac{\|e_2\|_1}{\|e_1\|_1} = \frac{24091}{161500} \approx 0.149,$$

donde $e_i = x - x_i$.

(e) Vamos a aplicar el método de relajación al método iterativo de Gauss-Seidel,

$$x^{(k)} = w(L + D)^{-1}(b - Ux^{(k-1)}) + (1 - w)x^{(k-1)}.$$

Para estudiar la convergencia de este método tenemos que calcular su tasa de convergencia

$$\rho(SOR) = \rho((1 - w)I - w(L + D)^{-1}U),$$

donde

$$SOR = \begin{pmatrix} 1 - w & -3w/10 & -w/10 \\ 0 & 1 - 53w/50 & 7w/25 \\ 0 & 6w/125 & 1 - 537w/500 \end{pmatrix}$$

cuyo polinomio característico es

$$|SOR - \lambda I| = (w - 1 + \lambda)(-1 + 2\lambda - 1\lambda^2 + \frac{2134}{1000}(1 - \lambda)w - \frac{1125}{1000}w^2),$$

que resolviendo la ecuación cuadrática nos dan los autovalores

$$\lambda_{SOR} = 1 - w, \quad \lambda_{SOR} = 1 - (1067 \pm \sqrt{13489})w/1000,$$

es decir,

$$\lambda_{SOR} = 1 - w, \quad \lambda_{SOR} = 1 - 1.183w, \quad \lambda_{SOR} = 1 - 0.9509w,$$

y la tasa de convergencia del método de Gauss-Seidel con relajación

$$\rho(SOR) = \max_i |\lambda_{SORi}| = \max_{0 \leq w \leq 2} \{|1 - w|, |1 - 1.183w|, |1 - 0.9509w|\},$$

será mínima para un w^* y valdrá (como se puede comprobar fácilmente dibujando las funciones en w) en la región en la que converge

$$1 > \rho(SOR) = \begin{cases} 1 - 0.9509w, & 0 < w \leq w^*, \\ -1 + 1.183w, & w^* \leq w < 1.690, \end{cases}$$

dando como valor óptimo

$$1 - 0.9509w^* = -1 + 1.183w^*, \quad w^* = 0.937,$$

es decir, realizaremos una subrelajación en lugar de una sobrerrelajación con una tasa de convergencia de

$$\rho(SOR) \approx 0.1088.$$

Vamos a realizar dos iteraciones a partir de $x^{(0)} = 0$,

$$x^{(1)} = (1.312, 0.73102, 0.9616)^\top, \quad x^{(2)} = (1.0988, 0.9882, 0.9882)^\top.$$

12. **Examen 6/Julio/1995.** Para la resolución del sistema $Bx = b$, donde $x, b \in \mathbb{R}^2$, se propone el siguiente método iterativo

$$x^{(k+1)} = b + Ax^{(k)}, \quad k \geq 0,$$

donde

$$A = \begin{pmatrix} \lambda & c \\ 0 & -\lambda \end{pmatrix}, \quad \lambda, c \in \mathbb{R}.$$

- ¿Para qué valores de λ y c existe solución única?
- Sea x_{PF} el punto fijo de la iteración. Calcule $x_{PF} - x^{(k)}$ para aquellos valores de λ y c para los que existe solución única. Suponga que $|\lambda| < 1$.
- Para las condiciones del apartado (b), ¿cómo se comporta $\|x_{PF} - x^{(k)}\|_\infty$ y $\|x^{(k+1)} - x^{(k)}\|_\infty$, cuando $k \rightarrow \infty$? ¿Es la convergencia al punto fijo independiente de c ? Justifique todos los resultados.

Solución.

- Para el punto fijo $(I - A)x_{PF} = b$,

$$\begin{pmatrix} 1 - \lambda & -c \\ 0 & 1 + \lambda \end{pmatrix} x_{PF} = Bx_{PF} = b,$$

y para que haya solución $|B| = 1 - \lambda^2 \neq 0$, por lo que $\lambda \neq \pm 1$.

- Calculemos las potencias de A ,

$$A^2 = AA = \begin{pmatrix} \lambda & c \\ 0 & -\lambda \end{pmatrix} \begin{pmatrix} \lambda & c \\ 0 & -\lambda \end{pmatrix} = \begin{pmatrix} \lambda^2 & 0 \\ 0 & \lambda^2 \end{pmatrix},$$

$$A^3 = A^2A = \begin{pmatrix} \lambda^2 & 0 \\ 0 & \lambda^2 \end{pmatrix} \begin{pmatrix} \lambda & c \\ 0 & -\lambda \end{pmatrix} = \begin{pmatrix} \lambda^3 & \lambda^2 c \\ 0 & -\lambda^3 \end{pmatrix},$$

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

y finalmente

$$A^k = \begin{pmatrix} \lambda^k & 0 \\ 0 & \lambda^k \end{pmatrix}, \quad \text{si } k \text{ es par,}$$

$$A^k = \begin{pmatrix} \lambda^k & \lambda^{k-1} c \\ 0 & -\lambda^k \end{pmatrix}, \quad \text{si } k \text{ es impar.}$$

Para estudiar la convergencia de la iteración acotaremos el error

$$e^{(k)} = x_{PF} - x^{(k+1)} = A e^{(k-1)} = A^k e^{(0)},$$

aplicando normas infinito

$$\|e^{(k)}\|_\infty = \|A^k e^{(0)}\|_\infty \leq \|A^k\|_\infty \|e^{(0)}\|_\infty,$$

y utilizando las expresiones que hemos derivado previamente

$$\|A^k\|_\infty = \max\{|\lambda^k|, |\lambda^k|\} = |\lambda^k|, \quad \text{si } k \text{ es par,}$$

$$\begin{aligned} \|A^k\|_\infty &= \max\{|\lambda^k| + |\lambda^{k-1} c|, |-\lambda^k|\} \\ &= |\lambda^k| \max\left\{1 + \left|\frac{c}{\lambda}\right|, 1\right\} \quad \text{si } k \text{ es impar.} \end{aligned}$$

Si c y λ tienen el mismo signo y $c/\lambda \gg 1$ está claro que $\|A^k\|_\infty$ oscila entre $|\lambda|^k$ para k par y $|\lambda|^k (1 + |c|/|\lambda|) \gg |\lambda|^k$, para k impar. Estas oscilaciones también ocurren si $|c|/|\lambda| \gg 1$. Para fijar ideas suponga que $c = \alpha + \beta \lambda^\gamma$, donde α , β y γ son independientes de λ . Para este valor tenemos

$$\lambda^k + \lambda^{k-1} c = \lambda^k + \alpha \lambda^{k-1} + \beta \lambda^{k-1+\beta},$$

y si α y β son $O(1)$, está claro que para $|\lambda| < 1$, $|\lambda|^{k-1} > |\lambda|^k$, mientras que el tercer término es mayor que el segundo si $\beta < 0$. Una vez que $k - 1 + \beta > 1$, el tercer término $|\lambda|^{k-1+\beta} < 1$. Esto significa que la convergencia al punto fijo depende de c y de su relación con λ . Más concretamente, depende de c/λ .

(c) Queremos estudiar

$$\begin{aligned}x^{(k+1)} - x^{(k)} &= A(x^{(k)} - x^{(k-1)}) = A^2(x^{(k-1)} - x^{(k-2)}) \\ &= A^k(x^{(1)} - x^{(0)}),\end{aligned}$$

por lo que

$$\|x^{(k+1)} - x^{(k)}\| \leq \|A^k\|_{\infty} \|x^{(1)} - x^{(0)}\|_{\infty},$$

por lo que $\|x^{(k+1)} - x^{(k)}\|$ oscila de la misma manera que $\|e^{(k)}\|_{\infty}$ en el apartado anterior.

13. **Examen 4/Septiembre/1997.** Para la resolución del sistema lineal en $\mathbb{R}^{n \times n}$ de

$$Ax = b, \quad |A| \neq 0,$$

se propone el método iterativo

$$x^{(k+1)} = Bx^{(k)} + c.$$

Se pide:

- (a) Calcular o determinar B y c
- (b) Las condiciones para que el método converja si B y c se mantienen fijos durante el proceso iterativo (método estacionario)
- (c) Las condiciones para que el método converja si B y c cambian durante el proceso iterativo (método no estacionario), es decir,

$$x^{(k+1)} = B_k x^{(k)} + c_k.$$

- (d) Utilice un método de relajación para la iteración del apartado (c) y determine las condiciones para las que dicho método converja.

Solución.

- (a) La iteración

$$x^{(k+1)} = B_k x^{(k)} + c_k,$$

cuando converja se convertirá en

$$x = Bx + c, \quad \Rightarrow \quad A^{-1}b = BA^{-1}b + c,$$

luego

$$c = (I - B)A^{-1}b = Mb,$$

y, por otro lado,

$$M A = (I - B), \quad \Rightarrow \quad B = I - M A.$$

Para estudiar la convergencia del método estudiaremos el error

$$e^{(k)} = x^{(k)} - x, \quad e^{(k)} = B e^{(k-1)} = B^k e^{(0)},$$

y aplicando normas

$$\|e^{(k)}\| \leq \|B\|^k \|e^{(0)}\|,$$

y para la convergencia $\|B\| < 1$. Como $\|B\| \geq |\lambda_B|$, los autovalores tienen que ser de módulo menor que la unidad.

(b) Para un método iterativo no estacionario

$$x^{(k+1)} = B_k x^{(k)} + M_k b, \quad x = B_k x + M_k b,$$

el error toma la forma

$$e^{(k)} = B_k e^{(k-1)} = \prod_{i=0}^{k-1} B_i e^{(0)},$$

y aplicando normas

$$\|e^{(k)}\| \leq \prod_{i=0}^{k-1} \|B_i\| \|e^{(0)}\|,$$

la condición de convergencia es

$$\prod_{i=0}^{k-1} \|B_i\| < 1.$$

(c) Un método no estacionario de relajación se obtiene de la expresión

$$x^{(k+1)} (1 + w_k) = B_k x^{(k)} (1 + w_k) + M_k b (1 + w_k),$$

$$x^{(k+1)} = -w_k x^{(k+1)} + (1 + w_k) B_k x^{(k)} + (1 + w_k) M_k b,$$

y toma la siguiente forma (cerca de la convergencia $x^{(k)} \approx x^{(k+1)}$),

$$x^{(k+1)} = -w_k x^{(k)} + (1 + w_k) B_k x^{(k)} + (1 + w_k) M_k b,$$

que se puede escribir

$$x^{(k+1)} = ((1 + w_k) B_k - w_k I) x^{(k)} + (1 + w_k) M_k b,$$

con lo que el error sigue la ecuación

$$e^{(k+1)} = P_k e^{(k)}, \quad P_k \equiv (1 + w_k) B_k - w_k I,$$

$$e^{(k)} = \prod_{i=0}^{k-1} P_i e^{(0)},$$

y tomando normas

$$\|e^{(k)}\| \leq \prod_{i=0}^{k-1} \|P_i\| \|e^{(0)}\|,$$

la iteración convergerá si

$$\begin{aligned} 1 > \prod_{i=0}^{k-1} \|P_i\| &= \prod_{i=0}^{k-1} \|(1 + w_k) B_k - w_k I\| \\ &= (1 + w_k) \prod_{i=0}^{k-1} \left\| B_k - \frac{w_k}{(1 + w_k)} I \right\|. \end{aligned}$$

14. Dado el sistema $Ax = b$, donde

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},$$

resuélvalo por medio de los siguientes métodos

- por factorización de Cholesky.
- ¿Puede resolver este sistema por el método de Gauss-Seidel? ¿Por qué? Si lo puede hacer, haga sólo dos iteraciones a partir de la solución nula y determine la tasa numérica de convergencia. Además calcula la tasa exacta de convergencia. ¿Cuántas iteraciones necesitará para alcanzar un error absoluto de 10^{-5} .
- ¿Puede resolver este sistema por el método del descenso más rápido? ¿Por qué? Si lo puede hacer, haga sólo tres iteraciones y determine la tasa numérica de convergencia.
- ¿Puede resolver este sistema por el método del gradiente conjugado? ¿Por qué? Si lo puede hacer, haga sólo tres iteraciones a partir de la solución nula y determine la tasa numérica de convergencia. ¿Cuántas iteraciones necesitará para alcanzar un error absoluto de 10^{-5} .
- Desarrolle un método de relajación basado en el método del gradiente conjugado. Determine para qué valores del parámetro w el método de relajación converge. ¿Cómo determinaría el w óptimo? Itere tres veces el método que ha obtenido (con el w óptimo).

Solución.

- (a) Se puede aplicar la factorizando de Cholesky $L^\top L$ porque la matriz es simétrica y definida positiva (demuestrolo). Operando (realice los detalles) se obtiene como solución

$$x = (-5/13, 8/13, 12/13)^\top.$$

- (b) El método de Gauss-Seidel es aplicable porque la matriz es simétrica y definida positiva. Dos iteraciones conducen a

$$x^{(0)} = (0, 0, 0)^\top, \quad x^{(1)} = (1/3, 4/9, 20/27)^\top,$$

$$x^{(2)} = (-17/81, 136/243, 644/729)^\top,$$

y la tasa de convergencia numérica la podemos calcular como (en norma infinito)

$$\frac{\|x^{(2)} - x\|}{\|x^{(1)} - x\|} = \frac{46}{189} = 0.24,$$

que se parece poco a la tasa de convergencia exacta

$$\rho((L + D)^{-1}U) = 0.413.$$

NOTA: calculando más iteraciones nos acercamos a la tasa teórica, por ejemplo,

$$\frac{\|x^{(10)} - x\|}{\|x^{(9)} - x\|} = 0.413.$$

Para alcanzar (en norma infinito) un error absoluto menor que 10^{-5} se requieren 13 iteraciones.

- (c) El sistema se puede resolver al ser la matriz A simétrica y definida positiva. El método consiste en, para la iteración k -ésima

$$r^{(k)} = b - Ax^{(k)}, \quad t_k = \frac{\langle r^{(k)}, r^{(k)} \rangle}{\langle r^{(k)}, Ar^{(k)} \rangle}, \quad x^{(k+1)} = x^{(k)} + t_k r^{(k)},$$

lo que conduce a

$$x^{(0)} = (0, 0, 0)^\top, \quad x^{(1)} = (7/34, 7/17, 21/34)^\top,$$

$$x^{(2)} = (-0.170987, 0.317547, 0.806082)^\top,$$

$$x^{(3)} = (-0.154596, 0.450727, 0.90545)^\top.$$

La tasa numérica del error nos da en este caso

$$\frac{\|x^{(3)} - x\|}{\|x^{(2)} - x\|} = 0.77, \quad \frac{\|x^{(10)} - x\|}{\|x^{(9)} - x\|} = 0.66,$$

que indica que este método es más lento que Gauss-Seidel.

- (d) El sistema se puede resolver al ser la matriz A simétrica y definida positiva. El método consiste en, para la iteración k -ésima

$$t_k = \frac{\langle r^{(k)}, r^{(k)} \rangle}{\langle p^{(k)}, Ap^{(k)} \rangle}, \quad x^{(k+1)} = x^{(k)} + t_k p^{(k)},$$

$$r^{(k+1)} = r^{(k)} - t_k Ap^{(k)},$$

$$s_k = \frac{\langle r^{(k+1)}, r^{(k+1)} \rangle}{\langle r^{(k)}, r^{(k)} \rangle}, \quad p^{(k+1)} = r^{(k+1)} + s_k p^{(k)},$$

con $p^{(0)} = r^{(0)} = b - Ax^{(0)}$ y con la condición de fin $\|r^{(k+1)}\| \approx 0$. Este método nos da entonces

$$x^{(0)} = (0, 0, 0)^\top, \quad x^{(1)} = (7/34, 7/17, 21/34)^\top,$$

$$x^{(2)} = (-14/67, 26/67, 66/67)^\top,$$

$$x^{(3)} = (-5/13, 8/13, 12/13)^\top = x,$$

esta última solución con residuo cero, como es de esperar. En tres iteraciones hemos obtenido la solución exacta, luego es un método más rápido que los dos anteriores. Para obtener un error menor que 10^{-5} es necesario realizar las tres iteraciones.

- (e) Sea $x^{(k+1),CG}$ la solución en cada iteración del método del gradiente conjugado, entonces definimos el método de relajación

$$x^{(k+1)} = (1 - w)x^{(k)} + wx^{(k+1),CG}.$$

Este método no es consistente, aunque converge, pero a un vector que no es solución del problema lineal a resolver, salvo que $w = 1$. Es fácil ver que el método converge a wx .

15. 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 la descomposición de Cholesky y la descomposición de Cholesky modificada ($L^\top DL$).

Solución. Este ejercicio es caso particular del próximo, luego no detallaremos su solución.

16. Dada un sistema tridiagonal $\langle a_i, b_i, c_i \rangle x = \langle d_i \rangle$,

$$b_1x_1 + c_1x_2 = d_1,$$

$$a_2x_1 + b_2x_2 + c_2x_3 = d_2,$$

$$a_3x_2 + b_3x_3 + c_3x_4 = d_3,$$

\vdots

$$a_{n-1}x_{n-1} + b_nx_n = d_n,$$

escribe el algoritmo de factorización de Doolittle y de Crout para dicho sistema general (Algoritmos de Thomas).

Solución. El método de Doolittle consiste en factorizar $A = LU$ con L de diagonal unitaria,

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & & 0 \\ 0 & l_{32} & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & l_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & 0 & \cdots & 0 \\ 0 & u_{22} & u_{23} & & 0 \\ 0 & 0 & u_{33} & \ddots & \vdots \\ \vdots & & & \ddots & u_{n-1,n} \\ 0 & \cdots & 0 & & u_{nn} \end{pmatrix}$$

$$u_{11} = b_1, \quad u_{12} = c_1, \quad l_{21} = a_2/u_{11},$$

$$u_{kk} = b_k - l_{k,k-1}u_{k-1,k}, \quad u_{k,k+1} = c_k, \quad l_{k+1,k} = a_{k+1}/u_{kk},$$

$$k = 2, 3, \dots, n.$$

Y como solución del sistema, $Ly = d$, $Ux = y$, obtenemos

$$y_1 = d_1, \quad y_k = d_k - l_{kk-1}y_{k-1}, \quad k = 2, \dots, n,$$

$$x_n = y_n/u_{nn}, \quad x_k = (y_k - u_{k,k+1}y_{k+1})/u_{kk}, \quad k = n-1, \dots, 1.$$

Para la factorización de Crout el procedimiento es del todo similar,

$$\begin{pmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ l_{21} & l_{22} & 0 & & 0 \\ 0 & l_{32} & l_{33} & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & & l_{n,n-1} & l_{nn} \end{pmatrix} \begin{pmatrix} 1 & u_{12} & 0 & \cdots & 0 \\ 0 & 1 & u_{23} & & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & & \ddots & u_{n-1,n} \\ 0 & \cdots & & 0 & 1 \end{pmatrix}$$

$$l_{11} = b_1, \quad l_{21} = a_2, \quad u_{12} = c_1/l_{11},$$

$$l_{kk} = b_k - l_{k,k-1}u_{k-1,k}, \quad l_{k+1,k} = a_k, \quad u_{k,k+1} = c_k/l_{kk},$$

$$k = 2, 3, \dots, n.$$

Y como solución del sistema, $Ly = d$, $Ux = y$, obtenemos

$$y_1 = d_1/l_{11}, \quad y_k = (d_k - l_{kk-1}y_{k-1})/l_{kk}, \quad k = 2, \dots, n,$$

$$x_n = y_n, \quad x_k = y_k - u_{k,k+1}y_{k+1}, \quad k = n-1, \dots, 1.$$