CIMAT

Centro de Investigación en Matemáticas, A.C.
Métodos Numéricos

Tarea 04

José Miguel Saavedra Aguilar

Resumen

En esta tarea se exploran diferentes técnicas para resolver un sistema lineal $Ax = b$, como son la factorización de Cholesky, el método de Jacobi y el de Gauss-Seidel. Así mismo, se utilizan dichos algoritmos para resolver el problema elíptico unidimensional.

1. Introducción

En muchos problemas de Métodos Numéricos y de sus aplicaciones, surge la necesidad de resolver un sistema de la forma $Bx = b$, donde B es simétrica. Si bien conocemos diferentes formas de resolver este sistema, podemos utilizar propiedades de la simetría para hacerlos aún más eficientes.

2. Metodología

2.1. Factorización de Cholesky

Sea $B \in \mathbb{R}^{n\times n}$ Simétrica Positiva Definida. La factorización de Cholesky consiste en encontrar $L \in \mathbb{R}^{n\times n}$ triangular inferior tal que $B = LL^\top$. Esta matriz está dada para $i = 1, \ldots, n$ por:

$$L_{i,j} = \frac{B_{i,j} - \sum_{k=1}^{j-1} L_{i,k} Lj,k}{L_{j,j}} \qquad j = 1, \ldots, i-1 \tag{1}$$

$$L_{i,i} = \sqrt{B_{i,i} - \sum_{k=1}^{i-1} L_{i,k}^2} \tag{2}$$

Es posible extender esta factorización para $B \in \mathbb{R}^{n\times n}$ Simétrica de forma que buscamos $L, D \in \mathbb{R}^{n\times n}$, con L triangular inferior y D diagonal tales que $B = LDL^\top$. Fijando Otro método es la factorización LDU, en el que L es una matriz triangular inferior, U es una matriz triangular superior, y D es una matriz diagonal, tales que $L_{i,i} = U_{i,i} = 1$ para $i = 1, \ldots, n$, y $A = LDU$. En este sistema, obtenemos las siguientes ecuaciones: Resolviendo para las entradas de L, D y U, tenemos que para $i = 1, \dots, n$:

$$L_{i,j} = \frac{B_{i,j} - \sum_{k=1}^{j-1} L_{i,k} D_{k,k} Lj,k}{D_{j,j}} \qquad j = 1, \dots, i-1 \tag{3}$$

$$D_{i,i} = B_{i,i} - \sum_{k=1}^{i-1} L_{i,k}^2 D_{k,k} \tag{4}$$

En el caso que B sea tridiagonal, notamos que para $i > 1$, $L_{i,1} = 0$, además:

$$\begin{aligned}
L_{i,j} &= \frac{0 - \sum_{k=1}^{j-1} L_{i,k} Lj,k}{L_{j,j}} && j = 1, \dots, i-2 \\
\implies L_{i,j} &= \frac{0 - \sum_{k=1}^{j-1} 0}{L_{j,j}} && j = 1, \dots, i-2 \\
\implies L_{i,j} &= 0 && j = 1, \dots, i-1
\end{aligned}$$

Por lo que $L_{i,i-1} = \frac{B_{i,i-1}}{Li-1,i-1}$ y $L_{i,i} = \sqrt{B_{i,i} - L_{i,i-1}^2}$.

2.2. Método de Jacobi

El método de Jacobi es un método iterativo para resolver el sistema $Ax = y$. Sea $A \in \mathbb{R}^{n \times n}$, descomponemos A en $A = D + A_L + A_U$ donde $D = \operatorname{diag}(A_{1,1}, \dots, A_{n,n})$, A_L y A_U dadas por:

$$A_L = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ A_{2,1} & 0 & 0 & \cdots & 0 \\ A_{3,1} & A_{3,2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{n,1} & A_{n,2} & A_{n,3} & \cdots & 0 \end{pmatrix} \qquad A_U = \begin{pmatrix} 0 & A_{1,2} & A_{1,3} & \cdots & A_{1,n} \\ 0 & 0 & A_{2,3} & \cdots & A_{2,n} \\ 0 & 0 & 0 & \cdots & A_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Supongamos que D es no singular, entonces si x es tal que $Ax = y$, se tiene

$$x = -D^{-1} (A_L + A_U) x + D^{-1} y$$

De forma que el método de Jacobi está definido en [1] por el esquema iterativo:

$$x_j^{i+1} = -\sum_{k=1}^{j-1} \frac{A_{j,k}}{A_{j,j}} x_k^i - \sum_{k=j+1}^{n} \frac{A_{j,k}}{A_{j,j}} x_k^i + \frac{y_j}{A_{j,j}} \qquad j = 1, \dots, n \tag{5}$$

Nos interesa saber cuándo es que el método de Jacobi es convergente, por lo que el siguiente teorema[1] nos será útil.

Teorema 1 (Convergencia del método de Jacobi)**.** *Si la matriz A satisface alguna de las siguientes condiciones:*

$$q_\infty := \max_{j=1,\dots,n} \sum_{k\neq j} \left| \frac{A_{j,k}}{A_{j,j}} \right| < 1$$

$$q_1 := \max_{k=1,\dots,n} \sum_{j\neq k} \left| \frac{A_{j,k}}{A_{j,j}} \right| < 1$$

$$q_2 := \sqrt{\sum_{j=1}^{n} \sum_{k\neq j} \left| \frac{A_{j,k}}{A_{j,j}} \right|^2} < 1$$

Entonces (5) *converge para todo $y \in \mathbb{C}^n$ y todo x_0 a la solución única de $Ax = y$. Además, para $p = 1, 2, \infty$*

$$\|x_i - x\|_p \leq \frac{q_p^i}{1 - q_p} \|x_1 - x_0\|_p$$

$$\|x_i - x\|_p \leq \frac{q_p}{1 - q_p} \|x_i - x_{i-1}\|_p$$

Es decir, tiene convergencia lineal.

2.3. Método de Gauss-Seidel

El método de Gauss-Seidel parte de la idea del método de Jacobi, con $A = D + A_L + A_U$. Supongamos que D es no singular, entonces si x es tal que $Ax = y$, se tiene

$$(A_L + D)x = (A_U)\, x + D^{-1}y$$

De forma que el método de Gauss-Seidel está definido en [1] por el esquema iterativo:

$$x_j^{i+1} = -\sum_{k=1}^{j-1} \frac{A_{j,k}}{A_{j,j}} x_k^{i+1} - \sum_{k=j+1}^{n} \frac{A_{j,k}}{A_{j,j}} x_k^i + \frac{y_j}{A_{j,j}} \qquad j = 1, \dots, n \tag{6}$$

En términos de convergencia, el teorema 1 no es válido para el método de Gauss-Seidel, pero tenemos otro resultado[1].

Teorema 2 (Criterio de Sassenfeld)**.** *Se definen p_j para $j = 1, \dots, n$*

$$p_1 := \sum_{k=2}^{n} \left| \frac{A_{1,k}}{A_{j,j}} \right|$$

$$p_j := \sum_{k=1}^{j-1} \left| \frac{A_{j,k}}{A_{j,j}} \right| p_k + \sum_{k=j+1}^{n} \left| \frac{A_{j,k}}{A_{j,j}} \right|$$

Si A cumple

$$p := \max_{j=1,\dots,n} p_j < 1$$

Entonces el método de Gauss-Seidel converge para todo $y \in \mathbb{C}^n$ y todo x_0 a la solución única de $Ax = y$. Además,

$$\|x_i - x\|_\infty \leq \frac{p^i}{1-p} \|x_1 - x_0\|_\infty$$
$$\|x_i - x\|_\infty \leq \frac{p}{1-p} \|x_i - x_{i-1}\|_\infty$$

Es decir, tiene convergencia lineal.

3. Resultados

3.1. Pseudocódigos

Algoritmo 1: Descomposición de Cholesky para A SPD

Entrada: A, n
Salida: L
para $i \in \{n, \dots, 1\}$ **hacer**
  para $j \in \{1, \dots, i-1\}$ **hacer**
    $L_{i,j} \leftarrow \frac{A_{i,j} - \sum_{k=1}^{j-1} L_{i,k} Lj,k}{L_{j,j}}$;
  fin
  $L_{i,i} \leftarrow \sqrt{A_{i,i} - \sum_{k=1}^{i-1} L_{i,k}^2}$
fin

Algoritmo 2: Descomposición de Cholesky para A SPD Tridiagonal

Entrada: A, n
Salida: L
para $i \in \{n, \dots, 1\}$ **hacer**
  $L_{i,i-1} \leftarrow \frac{A_{i,i-1}}{L_{i-1,i-1}}$;
  $L_{i,i} \leftarrow \sqrt{B_{i,i} - L_{i,i-1}^2}$
fin

Algoritmo 3: Método de Jacobi

Entrada: $A, y, n, maxIter, tol$
Salida: x
$x^0 \leftarrow$ resuelveDiagonal(diag(A), y);
$\epsilon \leftarrow \epsilon_0 > tol$;
$i \leftarrow 0$;
mientras $\epsilon < tol$ ***y*** $i < maxIter$ **hacer**
  para $j \in \{1, \ldots, n\}$ **hacer**
    $x_j^{i+1} \leftarrow -\sum_{k=1}^{j-1} \frac{A_{j,k}}{A_{j,j}} x_k^i - \sum_{k=j+1}^{n} \frac{A_{j,k}}{A_{j,j}} x_k^i + \frac{y_j}{A_{j,j}}$;
  fin
  $\epsilon \leftarrow \frac{\left\| x^{i+1} - x^i \right\|_2}{\left\| x^{i+1} \right\|_2}$;
  $i \leftarrow i + 1$
fin

Algoritmo 4: Método de Gauss-Seidel

Entrada: $A, y, n, maxIter, tol$
Salida: x
$x^0 \leftarrow$ resuelveDiagonal(diag(A), y);
$\epsilon \leftarrow \epsilon_0 > tol$;
$i \leftarrow 0$;
mientras $\epsilon < tol$ ***y*** $i < maxIter$ **hacer**
  para $j \in \{1, \ldots, n\}$ **hacer**
    $x_j^{i+1} \leftarrow -\sum_{k=1}^{j-1} \frac{A_{j,k}}{A_{j,j}} x_k^{i+1} - \sum_{k=j+1}^{n} \frac{A_{j,k}}{A_{j,j}} x_k^i + \frac{y_j}{A_{j,j}}$;
  fin
  $\epsilon \leftarrow \frac{\left\| x^{i+1} - x^i \right\|_2}{\left\| x^{i+1} \right\|_2}$;
  $i \leftarrow i + 1$
fin

3.2. Problema elíptico unidimensional

Consideramos el problema elíptico en una dimensión:

$$\begin{cases} \frac{\mathrm{d}^2 u}{\mathrm{d}x^2} = 2, & 0 \leq x \leq 1 \\ u(0) = 0, & u(1) = 2 \end{cases} \tag{7}$$

En un esquema de diferencias finitas por n nodos con aproximación de la segunda derivada por la fórmula:

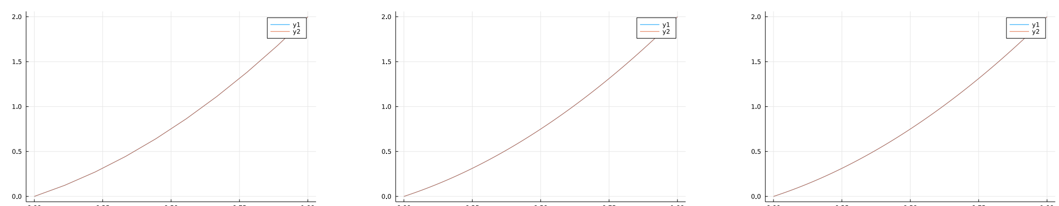
$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2}(x_i) \approx \frac{u(x_{i+1}) + u(x_{i-1}) - 2u(x_i)}{h^2} \qquad i = 2, \ldots, n-1$$
$$u(x_1) = 0 \qquad u(x_n) = 2$$

Donde $h = \frac{1}{n-1}$. Podemos reescribir este esquema como el sistema lineal $Ax = b$, donde $A \in \mathbb{R}^{n-2 \times n-2}$, $b \in \mathbb{R}^{n-2}$ están dados por:

$$A = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{pmatrix} \qquad b = \begin{pmatrix} -2 \\ -2 \\ -2 \\ \vdots \\ -2 \\ 2n(n-2) \end{pmatrix}$$

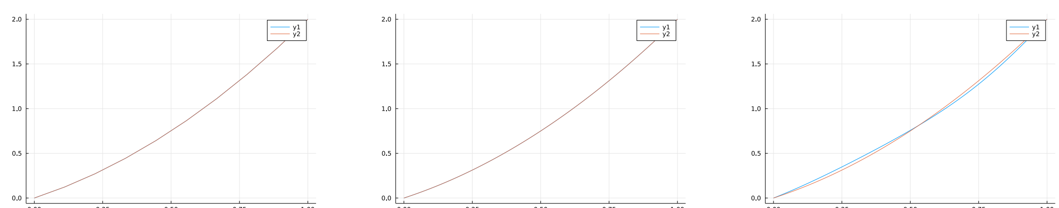
Sabemos que la solución exacta está dada por $u(x) = x^2 + x$, por lo que podemos comparar las soluciones obtenidas por el método de Cholesky, en particular 2, además de los métodos iterativos.

Nótese que A no cumple las condiciones de Convergencia del método de Jacobi pues $q_1 = q_\infty = 1$ y $q_2 = \sqrt{n-32}$, sin embargo cumplen el Criterio de Sassenfeld con $p = 1 - \frac{1}{2^{n-3}} < 1$, por lo que solo podemos garantizar convergencia de 4.
Se consideraron $n = 10, 50, 100$ nodos, y los resultados se muestran en las figuras 1, 2 y 3. Los erroes de aproximación están contenidos en los cuadros 1 y 2, el tiempo de máquina en 3 y las iteraciones en 4. Para los algoritmos iterativos se toma como criterio de paro $\frac{\|x_i - x_{i-1}\|}{\|x_i\|} < 10^{-10}$.



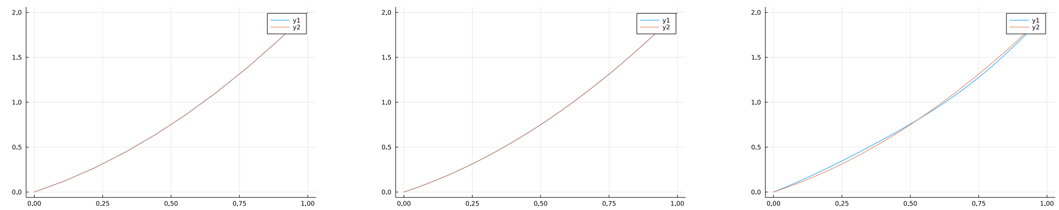
(a) $n = 10$ (b) $n = 1000$ (c) $n = 10000$

Figura 1: Solución de (7) con Factorización de Cholesky



(a) $n = 10$ (b) $n = 1000$ (c) $n = 10000$

Figura 2: Solución de (7) por el método de Jacobi



(a) $n = 10$ (b) $n = 1000$ (c) $n = 10000$

Figura 3: Solución de (7) por el método de Gauss-Seidel

$\|u - u_{aprox}\|$	**Cholesky**	**Gauss-Seidel**	**Jacobi**
10	1.1188630228E-16	1.6882851158E-05	1.6371505643E-08
50	3.2997982734E-14	1.9584430728E-07	4.6049765884E-07
100	3.7363556217E-14	9.9388670294E-05	2.6436889364E-06
1000	3.0986499738E-12	1.5617338353E-02	8.5763472228E-04
10000	6.7606426878E-10	2.4846533390E+00	2.6447325389E+00

Cuadro 1: Error absoluto de los métodos iterativos y de la solución por factorización de Cholseky

Tiempo (s)	**Cholesky**	**Gauss-Seidel**	**Jacobi**
10	0.000018	0.038535	0.033548
50	0.000033	0.040708	0.042869
100	0.000053	0.04633	0.07312
1000	0.001625	3.044957	12.44963
10000	0.147118	1135.811224	2879.959

Cuadro 2: Error relativo de los métodos iterativos y de la solución por factorización de Cholseky

$\frac{\|u-u_{aprox}\|}{\|u\|}$	**Cholesky**	**Gauss-Seidel**	**Jacobi**
10	3.3121791896E-17	2.3266506293E-06	4.8464699600E-09
50	4.5475006903E-15	5.7975947523E-08	6.3461861846E-08
100	3.6583480582E-15	9.7313635474E-06	2.5884940477E-07
1000	9.6349407483E-14	4.8560544415E-04	2.6667289957E-05
10000	6.6503951200E-12	2.4441354474E-02	2.6016041939E-02

Cuadro 3: Tiempo de CPU de los métodos iterativos y de la solución por factorización de Cholseky

Iteraciones	**Gauss-Seidel**	**Jacobi**
10	107	226
50	1918	5117
100	5285	18083
1000	150872	904744
10000	3285160	6256010

Cuadro 4: Iteraciones necesarias para los métodos iterativos

4. Conclusiones

En los cuadros podemos verificar que los métodos iterativos no son la mejor opción para el problema (7), en especial para $n \gg 1$, pues no podemos asegurar convergencia para Jacobi y para Gauss-Seidel, en el Criterio de Sassenfeld $p \xrightarrow[n\to\infty]{} 1$, por lo que toman muchas iteraciones y la aproximación es peor que los métodos analíticos.

Para este problema la factorización de Cholesky para matrices tridiagonales es muy efectiva, tanto por el punto de vista del tiempo de máquina, como por el error de aproximación. Esto se explica por la cantidad de operaciones realizadas en comparación con los métodos iterativos, así como por el problema que permite aproximar de manera exacta con el esquema de diferencias finitas que resolvemos.

Referencias

[1] R. Kress, *Numerical Analysis.* Springer New York, NY, 12 2012.