



Métodos iterativos para la solución de sistemas de ecuaciones lineales

^aMaría Guadalupe Torres González, ^bVíctor Hugo Vázquez Guevara, ^cVelasco Luna Fernando

Facultad de Ciencias Físico-Matemáticas Benemérita Universidad Autónoma de Puebla, Puebla, Puebla, México.

e-mail: ^alupita1990_7@hotmail.com, ^bvvazquez@fcfm.buap.mx, ^cfvelasco@fcfm.buap.mx

Resumen

En este trabajo se estudiarán métodos iterativos y estocásticos para resolver sistemas de ecuaciones lineales de modo que una aproximación a su solución se determine en el menor tiempo posible.

Palabras clave: Sistemas de ecuaciones lineales, aproximación.

Introducción

En general, llamamos al problema que consiste en encontrar la solución (posiblemente múltiple) $x \in \mathbb{K}^m$ si existe, de la siguiente ecuación algebraica

$$Ax = b$$

un sistema lineal. La matriz $A \in \mathcal{M}_{n,m}(\mathbb{K})$ se llama "matriz de coeficientes del sistema", y $b \in \mathbb{K}^m$, vector de coeficientes independientes ambos datos del problema; el vector $x \in \mathbb{K}^m$ es desconocido. \mathbb{K} denota el campo \mathbb{R} ó \mathbb{C} .

Consideremos sistemas lineales con el mismo número de ecuaciones e incógnitas: $n = m$.

¿Cuándo aplicar métodos iterativos en lugar de métodos directos? si tomamos en cuenta que la cantidad de operaciones que se deben realizar en un método directo para llegar a la solución el valor es del orden n^3 donde n es el número de renglones en la matriz o el número de ecuaciones en el sistema. Si hablamos de un sistema de 100 o 200 ecuaciones entonces se requiere del orden de 1,000,000 u 8,000,000 de operaciones para resolver esos sistemas respectivamente. Es de considerar el trabajo computacional que deben hacer las máquinas

o el ser humano, si lo resolvemos a lápiz y papel, debemos entonces buscar alternativas para llegar a la solución de una manera eficiente.

Por ello, este trabajo está dedicado a presentar algunos métodos iterativos (deterministas y estocásticos) que nos ayuden a determinar la solución de sistemas de ecuaciones lineales, de tal manera que la solución del problema se determine de manera exacta o aproximada y que el trabajo computacional no sea muy costoso.

Metodología

Método de Jacobi

El primero de los métodos que consideramos es el que Carl Gustav Jacobi (1804-1851) desarrolló en 1845. Su mecánica es muy simple: supongamos que se desea resolver el sistema de tres ecuaciones lineales con tres incógnitas

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

Admitiendo que los coeficientes a_{11} , a_{22} y a_{33} son distintos de cero, se puede despejar de la primera ecuación

la incógnita x_1 , de la segunda x_2 y x_3 de la tercera,

$$x_1 = \frac{1}{a_{11}}(b_1 - a_{12}x_2 - a_{13}x_3)$$

$$x_2 = \frac{1}{a_{22}}(b_2 - a_{21}x_1 - a_{23}x_3)$$

$$x_3 = \frac{1}{a_{33}}(b_3 - a_{31}x_1 - a_{32}x_2)$$

la i -ésima ecuación en el sistema $Ax = b$ es

$$x_i = \frac{1}{a_{ii}}(b_i - \sum_{j \neq i} a_{ij}x_j); \quad i = 1, \dots, n$$

La generalización de esta idea es la base del método iterativo de Jacobi. La relación general de recurrencia para un sistema de tamaño $n \times n$ es:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j \neq i} a_{ij}x_j^{(k)} \right); \quad i = 1, \dots, n \quad (1)$$

en donde $x_i^{(k+1)}$ es la iteración $k + 1$ -ésima para la variable x_i y $x_i^{(0)}$ la mayoría de veces es igual al vector unitario o nulo.

Convergencia del método de Jacobi

Se presenta ahora un resultado para garantizar la convergencia del método de Jacobi

Teorema 1 *Sea A una matriz simétrica y definida positiva. El método iterativo de Jacobi para un sistema de ecuaciones con matriz de coeficientes A es convergente si y sólo si la matriz*

$$2D - A$$

es una matriz definida positiva,

donde D es la matriz diagonal formada con los elementos de la diagonal principal de la matriz A .

Método de Gauss-Seidel

En el método de Jacobi cada una de las componentes del vector solución en la iteración $k + 1$ se determina a partir de las de la iteración k . En el de Carl Friedrich Gauss (1777-1855) y Phillip Ludwig Seidel (1874) se modifica el de Jacobi utilizando en el cálculo de cada componente de la solución en una iteración el valor de aquellos ya calculados en esa misma iteración. Volviendo al sistema de tres ecuaciones que se consideró para introducir el método de Jacobi, son distintos de cero, el esquema iterativo del método de Gauss-Seidel es el siguiente:

$$x_1^{(k+1)} = \frac{1}{a_{11}}(b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)})$$

$$x_2^{(k+1)} = \frac{1}{a_{22}}(b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)})$$

$$x_3^{(k+1)} = \frac{1}{a_{33}}(b_3 - a_{31}x_1^{(k+1)} - a_{32}x_2^{(k+1)})$$

Si en el método de Jacobi las relaciones de recurrencia que conformaban su esquema iterativo se obtenían de despejar cada variable en su correspondiente ecuación, en el método de Gauss-Seidel esas relaciones surgen de hacer esto mismo, pero de una forma escalonada. En efecto, no es difícil comprobar que el método de Gauss-Seidel tiene una relación directa al escribir el sistema original en la forma

$$a_{11}x_1 = b_1 - a_{12}x_2 - a_{13}x_3$$

$$a_{21}x_1 + a_{22}x_2 = b_2 - a_{23}x_3$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

La relación de recurrencia general para un sistema $n \times n$ es la siguiente:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right);$$

$$i = 1, \dots, n$$

donde $x_i^{(0)}$ la mayoría de veces se escoge como el vector unitario o nulo.

Convergencia del método de Gauss-Seidel

A continuación se presenta un teorema importante para la convergencia del método de Gauss-Seidel.

Teorema 2 *El método iterativo de Gauss-Seidel es convergente para todo sistema de ecuaciones cuya matriz de coeficientes es simétrica y definida positiva.*

Sobre-relajación sucesiva (SOR)

Los dos métodos iterativos que se han estudiado hasta ahora, se pueden generalizar. En efecto, las relaciones de recurrencia de estos dos métodos se pueden escribir de la forma

$$x_i^{(k+1)} = x_i^{(k)} + r_i^{(k)}; \quad i = 1, \dots, n$$

En el caso del método de Jacobi, se tiene que

$$r_i^{(k)} = \frac{b_i - \sum_{j=1}^n a_{ij}x_j^{(k)}}{a_{ii}};$$

mientras que en el de Gauss-Seidel se observa que

$$r_i^{(k)} = \frac{b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i}^n a_{ij}x_j^{(k)}}{a_{ii}};$$

Visto así, estos dos procedimientos iterativos llegan a la solución a través de un número de pasos, en cada uno de los cuales se avanza una cantidad $r^{(k)}$ [5].

La idea de los métodos de relajación consiste, en cada iteración, aplicar la relación de recurrencia,

$$x_i^{(k+1)} = x_i^{(k)} + \omega r_i^{(k)}, \quad i = 1, \dots, n,$$

de tal forma que se mejoren las iteraciones del procedimiento avanzando un paso más amplio, $\omega > 1$, o más corto, $\omega < 1$. Al parámetro ω se le conoce como parámetro de relajación. El método de relajación más conocido es el SOR, (Successive Overrelaxation): resulta de aplicar esta idea sobre la base del método de Gauss-Seidel. Su relación de recurrencia es:

$$x_i^{(k+1)} = \frac{\omega}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} a_{ii} \right) + (1 - \omega) x_i^{(k)}, \quad i = 1, \dots, n.$$

donde $x_i^{(0)}$ la mayoría de veces se escoje como el vector unitario o nulo.

Una elección adecuada del valor de ω puede mejorar la convergencia del método. La idea que debe dirigir esa elección es que si la corrección que introduce cada iteración en la solución es excesiva, se puede disminuir con un factor $\omega < 1$. Por el contrario, si la corrección tiende a quedarse corta, se puede aumentar con un factor $\omega > 1$.

La elección del parámetro ω plantea dos problemas: en primer lugar, que ha de estudiarse el conjunto de valores del parámetro que hacen que el método SOR converja; en segundo, que hay que determinar el valor del parámetro que haga que la convergencia sea lo más rápida posible.

Convergencia del método de relajación sucesiva

Para matrices estrictamente diagonales, se tiene el siguiente resultado.

Proposición 1 *Para $0 < \omega \leq 1$, el método de relajación es convergente para matrices de diagonal estrictamente dominantes*

Mayor interés tiene el hecho de que cuando se tiene una matriz simétrica definida positiva, el método de relajación SOR converge para todos los valores permisibles del parámetro ω , es decir para $\omega \in (0, 2)$. En este caso la condición que antes se dio como necesaria se convierte en suficiente.

Teorema 3 (Ostrowski-Reich) *Para un sistema de ecuaciones con matriz simétrica y definida positiva, el método iterativo de relajación SOR converge si y sólo si el parámetro de relajación cumple que $0 < \omega < 2$.*

Número de operaciones de los metodos iterativos

Los métodos iterativos tienen un costo computacional; expresado en número de operaciones de $n^2 \times$ no. de iteraciones. Si el número de iteraciones es pequeño comparado con n , los métodos iterativos son preferibles a los directos ya que, el número de operaciones necesarias al utilizar métodos directos para resolver sistemas de ecuaciones lineales es del orden de n^3

Método iterativo de K. Sabelfeld

En esta sección se estudiará un análogo estocástico a los presentados con anterioridad. Sea H la matriz cuadrada de dimensión n , tal que $A = I - H$. Entonces el sistema de ecuaciones lineales $Ax = b$, puede ser escrito como

$$x = Hx + b,$$

en donde $x = (x_1, \dots, x_n)^T$, $b = (b_1, \dots, b_n)^T \in \mathbb{R}^n$ y $H = (h_{i,j})$, donde T representa la operación transpuesta.

Se presenta el siguiente resultado que nos dice, cuando la matriz $A = I - H$ es invertible.

Corolario 1 *La serie de Neumann*

$$\sum_{k=0}^{\infty} H^k$$

converge si $\rho(H) < 1$ y diverge si $\rho(H) > 1$, en donde $\rho(H)$ es el radio espectral de la matriz H

La solución del sistema $Ax = b$, se puede determinar por $x = A^{-1}b$. Si $A = I - H$ entonces A^{-1} se puede escribir como una serie de Neumann, suponiendo que tal serie es finita, es decir

$$A^{-1} = (I - H)^{-1} = \sum_{k=0}^{\infty} H^k$$

donde $A^0 = I$.

Por lo tanto, si H tiene radio espectral menor que uno, la solución de $x = Hx + b$ se puede calcular mediante el método de iteración de Richardson simple [5]

$$x^{(m+1)} = Hx^{(m)} + b; \quad x^{(0)} = b; \quad m = 0, 1, 2, \dots$$

Muestreo de columnas sin reemplazo.

A continuación, se presentará una versión estocástica del citado método de Richardson. Para esto, sea G un estimador insesgado para la matriz H que se define como una matriz aleatoria tal que $E[G] = H$, y sea $G^{(0)}, G^{(1)}, \dots, G^{(M-1)}$ una muestra aleatoria del estimador aleatorio G . El procedimiento iterativo se define por [6]

$$\xi^{(m+1)} = G^{(m)} \xi^{(m)} + b, \quad m = 0, 1, \dots, M - 1 \quad (2)$$

donde $\xi^{(0)} = b$. Dado que $G^{(m)}$, $m = 0, 1, \dots$ son independientes obtenemos de $\xi^{(m+1)} = G^{(m)}\xi^{(m)} + b$ que $E[\xi^M] = x^M$. Y M es el número de términos considerados en la serie de Neumann.

Consideremos el caso particular en el que se elige G como una matriz dispersa. Construiremos la matriz G en columnas: fijemos un entero arbitrario l menor que n , y elija un conjunto aleatorio J de l enteros elegidos uniformemente de 1 a n sin reemplazo, es decir, elegimos j_1 entero uniformemente entre $1, 2, \dots, n$ entonces j_2 entre el resto de $n - 1$ enteros, etc., el último es j_l y definimos las entradas de G por

$$G_{ik} = \begin{cases} \frac{n}{l} H_{ik} & \text{para } k \in J \\ 0 & \text{en otro caso} \end{cases}$$

para $i = 1, 2, \dots, n$

Por ejemplo, si H es una matriz de tamaño 4 y elegimos $J = \{1, 4\}$, entonces

$$G = \begin{pmatrix} 2h_{11} & 0 & 0 & 2h_{14} \\ 2h_{21} & 0 & 0 & 2h_{24} \\ 2h_{31} & 0 & 0 & 2h_{34} \\ 2h_{41} & 0 & 0 & 2h_{44} \end{pmatrix}$$

Por lo tanto la matriz aleatoria G tiene exactamente l columnas distintas de cero, y es obvio que para cualesquiera i, k tenemos

$$E[G_{ik}] = G_{ik} P\{k \in J\} = H_{ik}.$$

Notemos que para el cálculo de componentes del vector $\xi^{(m+1)}$ necesitamos solo l componentes del vector $\xi^{(m)}$ y para calcularlos solo necesitamos l componentes de $\xi^{(m-1)}$, etc. En consecuencia, necesitamos l^2 operaciones en cada etapa. Para hallar un valor de la aproximación $x^{(M)}$ necesitamos del orden de Ml^2 operaciones, donde M es la longitud de corte de la serie de Neumann.

Una forma de elegir el valor de M es establecer cierta tolerancia (ε) con respecto a la diferencia entre dos términos sucesivos de la serie de Neumann; es decir, M es el natural más pequeño tal que la norma de la diferencia de dos términos sucesivos es menor que ε .

Resultados y discusiones

Sea

$$A = \begin{pmatrix} 1 & 1/10 & -1/5 & 0 & 1/5 & -1/8 & 3/8 \\ 1/10 & 1 & 1/10 & -3/8 & 1/10 & -1/5 & 1/10 \\ -1/5 & 1/10 & 1 & 1/8 & 1/12 & 0 & 1/12 \\ 0 & -3/8 & 1/8 & 1 & -1/8 & 1/10 & -1/5 \\ 1/5 & 1/10 & 1/12 & -1/8 & 1 & 0 & 0 \\ -1/8 & -1/5 & 0 & 1/10 & 0 & 1 & -2/5 \\ 3/8 & 1/10 & 1/12 & -1/5 & 0 & -2/5 & 1 \end{pmatrix}$$

La matriz de un sistema de ecuaciones lineales y

$$b = \begin{pmatrix} 3/5 \\ 25/11 \\ -10/11 \\ 15/8 \\ 2/3 \\ -5/7 \\ 4/3 \end{pmatrix}$$

el vector solución. Resolviendo el sistema de ecuaciones en Matlab se obtiene que la solución exacta es:

$$x = \begin{pmatrix} -1.5631 \\ 4.0222 \\ -2.5183 \\ 4.3733 \\ 1.3336 \\ 0.5930 \\ 2.8390 \end{pmatrix}$$

Se resolverá el sistema con cada uno de los métodos mencionados anteriormente y se analizarán los resultados

Método de Jacobi

Notemos que A es simétrica, y veamos que es definida positiva. Haciendo los cálculos se obtiene que los valores propios de A son:

$$\lambda_1 = 0.3294, \lambda_2 = 0.6891, \lambda_3 = 0.7080, \lambda_4 = 1.0676, \lambda_5 = 1.1312, \lambda_6 = 1.2262 \text{ y } \lambda_7 = 1.8482$$

Como podemos observar, todos los valores propios son positivos. Por lo tanto, A es definida positiva.

Por el teorema 1, hay que verificar que $2D - A$ es definida positiva, es decir, que los valores propios de $2D - A$ sean positivos. En este caso, los valores propios de $2D - A$ son iguales a los de A . Por lo tanto, $2D - A$ es definida positiva.

Con ayuda de Matlab, y tomando $x = (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0)^T$ como vector inicial y una tolerancia de 0.000001 obtenemos que

$$x = \begin{pmatrix} -1.5631018 \\ 4.02221079 \\ -2.5183140 \\ 4.3733152 \\ 1.3335902 \\ 0.5930427 \\ 2.8390161 \end{pmatrix}$$

Por lo tanto, el número de operaciones que se realizan es: $49 \times 83 = 4067$.

Como podemos observar el número de operaciones

que se requiere es muy alto.

Método de Gauss-Seidel

Por el teorema 2 y los resultados obtenidos antes, la matriz A es simétrica y definida positiva.

Sea $x = (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0)^T$ el vector inicial y una tolerancia de 0.000001, la solución del sistema con este algoritmo es:

$$x = \begin{pmatrix} -1.56310 \\ 4.02221 \\ -2.51831 \\ 4.37331 \\ 1.33359 \\ 0.59304 \\ 2.83901 \end{pmatrix}$$

En la iteración 20, el método se aproxima a la solución del sistema. Por lo tanto, se necesitan $49 \times 20 = 980$ operaciones.

El número de operaciones es menor, pero aún falta analizar los otros métodos.

Método SOR

Por los métodos anteriores, se tiene que la matriz A es simétrica y definida positiva. Si tomamos $x = (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0)^T$ el vector inicial y $\omega = 1.5$ con 20 iteraciones obtenemos,

$$x = \begin{pmatrix} -1.5631063 \\ 4.0221967 \\ -2.5183090 \\ 4.3733057 \\ 1.3335896 \\ 0.5930396 \\ 2.8390146 \end{pmatrix}$$

se obtiene una buena aproximación a la solución. Así, se necesitan $49 \times 20 = 980$ operaciones.

El número de operaciones en este caso es igual al del método de Gauss-Seidel.

Método iterativo de K. Sabelfeld

En [6] los autores refieren que el valor de l deber ser mucho menor que "n" (la dimensión del sistema); es decir, el método estocástico presentado es eficaz si el sistema de ecuaciones es substancialmente grande. Por ejemplo, veamos que dicho método no es eficiente si el sistema de ecuaciones es pequeño,

Con ayuda de Matlab, se observa que $\rho(H) = 0.8483$. Sea $l = 2$ y hacemos 30 iteraciones del algoritmo en

Matlab, de donde se obtiene lo siguiente:

$$x = \begin{pmatrix} -1.3297 \\ 4.7252 \\ -0.2893 \\ 3.0811 \\ 2.0906 \\ 0.7143 \\ 2.7572 \end{pmatrix}$$

De acuerdo a lo anterior, el número de operaciones que se requieren es $l^2 M = (2)^2 \times 30 = 120$.

Concluimos que dicha situación ocurre ya que las observaciones de las matrices $G^{(0)}, G^{(1)}, \dots, G^{(M)}$ tienden a repetirse incluso para valores pequeños de M . Además se consideró a un sistema de ecuaciones de dimensión 50 cuya matriz y vector solución asociados son generados aleatoriamente apartir de la distribución normal estándar, de modo que se satisfagan las hipótesis del método de Sabesfeld: $\rho(A) = 0.688$.

En este caso, se considerará a $M = 10$, ya que el radio espectral de $H^{11} - H^{10}$ es del orden de 10^{-6} . Además se consideró a $l = 10$; es decir, el 20% de las columnas de la matriz A serán muestreadas en cada repetición.

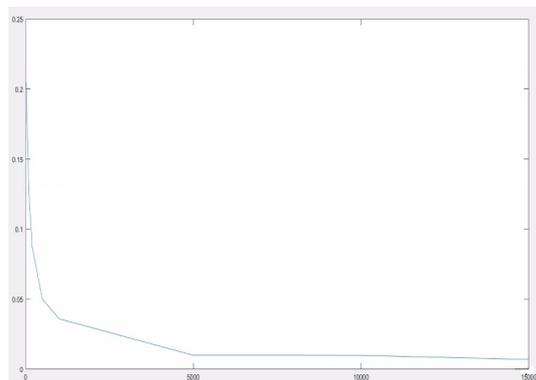


Figura 1:

En la la Figura 1 se observa la evolución del error absoluto medio de $A\bar{\xi}$ con respecto a b ; en donde $\bar{\xi}$ es el promedio de las aproximaciones generadas por el método de Sabelfeld, de manera más precisa, el promedio de

$$|A\bar{\xi} - b|$$

Conclusiones

Según los resultados obtenidos, con respecto al vector inicial, tolerancia, aproximación y el número de operaciones, el método de Gauss-Seidel y el método SOR son mejores que el método de Jacobi. Sin embargo, hace falta analizar el caso en que el orden de la matriz sea muy grande para decir que método es más eficiente.

Referencias

- [1] STANLEY I. GROSSMAN, JOSÉ JOB FLORES GODOY, *Álgebra lineal*, México, 2012.
- [2] GEORGE NAKOS, DAVID JOYNER, *Álgebra lineal con aplicaciones*, México, 1999.
- [3] DOLORES MARTÍNEZ, JAVIER PASTOR, *Introducción al Álgebra Lineal Numérica*, Universidad de valencia, Septiembre 2014.
- [4] DAVID S. WATKINS, *Fundamentals of Matrix Computations*, New York, 2002.
- [5] JOSÉ LUIS DE LA FUENTE O' CONNOR, *Técnicas de cálculo para sistemas de ecuaciones, programación lineal y programación entera*, Universidad Politécnica de Madrid, 2019.
- [6] K. SABELFELD, *"Stochastic algorithms in linear algebra-beyond the Markov chains and von Neumann-Ulam scheme"* in *Numerical Methods and Applications*, New York, NY, USA:Springer, pp. 14-28, 2011.