Método Gauss-Seidel

Marina Andretta/Franklina Toledo

ICMC-USP

20 de agosto de 2013

Baseado no livro Cálculo Numérico, de Neide B. Franco.

Introdução

Como vimos, queremos resolver o sistema linear

$$Ax = b.$$

Para isso, podemos reescrevê-lo na forma

$$x = Bx + g.$$

O Método Jacobi-Richardson

Supondo que $\det(D) \neq 0$, podemos transformar o sistema linear original em

$$
\begin{aligned}
Ax &= b & \Rightarrow \\
(L + D + R)x &= b & \Rightarrow \\
Dx &= -(L + R)x + b & \Rightarrow \\
x &= -D^{-1}(L + R)x + D^{-1}b,
\end{aligned}
$$

onde $B = -D^{-1}(L + R)$ e $g = D^{-1}b$.

O Método Jacobi-Richardson

O processo iterativo definido por

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

é chamado de *Método Jacobi-Richardson*.

O Método Jacobi-Richardson

Supondo que $det(D) \neq 0$ (ou seja, $a_{ii} \neq 0$, $i = 1, ..., n$) e dividindo cada linha de A pelo elemento da diagonal, temos

$$\begin{pmatrix} 1 & a_{12}/a_{11} & a_{13}/a_{11} \\ a_{21}/a_{22} & 1 & a_{23}/a_{22} \\ a_{31}/a_{33} & a_{32}/a_{33} & 1 \end{pmatrix} = \begin{pmatrix} 1 & a^*_{12} & a^*_{13} \\ a^*_{21} & 1 & a^*_{23} \\ a^*_{31} & a^*_{32} & 1 \end{pmatrix} =$$

$$\begin{pmatrix} 0 & 0 & 0 \\ a^*_{21} & 0 & 0 \\ a^*_{31} & a^*_{32} & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & a^*_{12} & a^*_{13} \\ 0 & 0 & a^*_{23} \\ 0 & 0 & 0 \end{pmatrix}.$$

Assim temos:

$$A^* = L^* + I + R^*.$$

Reescrevendo novamente o sistema, temos

$$
\begin{aligned}
Ax &= b & \Rightarrow \\
A^*x &= b^* & \Rightarrow \\
(L^* + I + R^*)x &= b^* & \Rightarrow \\
x &= -(L^* + R^*)x + b^*,
\end{aligned}
$$

onde $-(L^* + R^*) = B$ e $b^* = g$.

O processo iterativo fica

$$x^{(k+1)} = -(L^* + R^*)x^{(k)} + b^*.$$

Convergência

A sequência gerada pelo algoritmo converge?

Critério: A sequência será convergente se, para alguma norma de matrizes, $\|B\| < 1$.

O Método Gauss-Seidel

Transformando o sistema

$$(L^* + I + R^*)x = b^*$$

em

$$(L^* + I)x = -R^* x + b^*,$$

temos que

$$x = -(L^* + I)^{-1} R^* x + (L^* + I)^{-1} b^*.$$

O processo iterativo definido por

$$x^{(k+1)} = -(L^* + I)^{-1} R^* x^{(k)} + (L^* + I)^{-1} b^*$$

é chamado de *Método Gauss-Seidel*.

Observe que, multiplicando

$$x^{(k+1)} = -(L^* + I)^{-1} R^* x^{(k)} + (L^* + I)^{-1} b^*$$

por $(L^* + I)$, temos

$$(L^* + I)x^{(k+1)} = -R^* x^{(k)} + b^*.$$

Ou, ainda,

$$x^{(k+1)} = -L^* x^{(k+1)} - R^* x^{(k)} + b^*.$$

O Método Gauss-Seidel

Note que as componentes de $x^{(k+1)}$ podem ser calculadas sucessivamente, sem a necessidade de se calcular $(L^* + I)^{-1}$.

Observe também que, para o cálculo de uma componente de $x^{(k+1)}$, usamos o valor calculado mais recentemente das demais componentes. Por esse motivo, o Método Gauss-Seidel também é conhecido por *Método dos Deslocamentos Sucessivos*.

Esse método difere do Método Jacobi-Richardson por utilizar, no cálculo de uma componente de $x^{(k+1)}$, o valor mais recente das demais componentes.

Convergência

Já vimos que a sequência gerada pelo algoritmo converge se $\|B\| < 1$, para alguma norma de matriz.

A matriz B usada no Método Gauss-Seidel é dada por

$$B = -(L^* + I^*)^{-1} R^*.$$

Vamos deduzir um critério de convergência a partir da norma infinito de B.

Definição: uma norma de matriz é dita **induzida** por uma norma de vetor se ela é definida como

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Assim,

$$\|A\|_\infty = \max_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty}.$$

Convergência

Seja x o vetor tal que realiza o máximo na expressão de $\|B\|_\infty$. Considere

$$y = Bx.$$

Para o Método Gauss-Seidel, temos que

$$B = -(L^* + I)^{-1}R^*.$$

Logo,

$$y = -(L^* + I)^{-1}R^*x \Rightarrow$$

$$(L^* + I)y = -R^*x \Rightarrow$$

$$y = -L^*y - R^*x.$$

Dado que $y = -L^*y - R^*x$, podemos escrever

$$\begin{array}{rcccccccccccc}
y_1 & = & - & a^*_{12}x_2 & - & a^*_{13}x_3 & - & a^*_{14}x_4 & - & \ldots & - & a^*_{1n}x_n, \\
y_2 & = & - & a^*_{21}y_1 & - & a^*_{23}x_3 & - & a^*_{24}x_4 & - & \ldots & - & a^*_{2n}x_n, \\
y_3 & = & - & a^*_{31}y_1 & - & a^*_{32}y_2 & - & a^*_{34}x_4 & - & \ldots & - & a^*_{3n}x_n, \\
\vdots & & & & & & & & & & & \\
y_n & = & - & a^*_{n1}y_1 & - & a^*_{n2}y_2 & - & a^*_{n3}y_3 & - & \ldots & - & a^*_{n,n-1}y_n.
\end{array}$$

Queremos calcular $\|Bx\|_\infty$ e sabemos que $\|Bx\|_\infty = \|y\|_\infty$, ou seja,

$$\|Bx\|_\infty = \max_{1 \leq i \leq n} |y_i|.$$

Assim,

$$|y_1| = |-a_{12}^*x_2 - a_{13}^*x_3 - a_{14}^*x_4 - \ldots - a_{1n}^*x_n| = \left|\sum_{j=2}^{n} -a_{1j}^*x_j\right| \Rightarrow$$

$$|y_1| \leq \sum_{j=2}^{n} |-a_{1j}^*x_j| \leq \sum_{j=2}^{n} |a_{1j}^*||x_j| \Rightarrow$$

$$|y_1| \leq \sum_{j=2}^{n} |a_{1j}^*| \max_j |x_j| = \sum_{j=2}^{n} |a_{1j}^*| \|x\|_\infty .$$

Ou seja,

$$|y_1| \leq \beta_1 \|x\|_\infty, \quad \text{com} \quad \beta_1 = \sum_{j=2}^{n} |a_{1j}^*|.$$

Convergência

Fazendo os mesmos cálculos para $|y_2|$, temos

$$|y_2| = |-a^*_{21}y_1 - a^*_{23}x_3 - a^*_{24}x_4 - \ldots - a^*_{1n}x_n| = \left| -a^*_{21}y_1 + \sum_{j=3}^{n} -a^*_{2j}x_j \right| \Rightarrow$$

$$|y_2| \leq |-a^*_{21}y_1| + \sum_{j=3}^{n} |-a^*_{2j}x_j| \leq |a^*_{21}||y_1| + \sum_{j=3}^{n} |a^*_{2j}||x_j| \Rightarrow$$

$$|y_2| \leq |a^*_{21}|\beta_1\|x\|_\infty + \sum_{j=3}^{n} |a^*_{2j}| \max_j |x_j| \Rightarrow$$

$$|y_2| \leq |a^*_{21}|\beta_1\|x\|_\infty + \sum_{j=3}^{n} |a^*_{2j}|\|x\|_\infty.$$

Ou seja,

$$|y_2| \leq \beta_2 \|x\|_\infty, \quad \text{com} \quad \beta_2 = |a^*_{21}|\beta_1 + \sum_{j=3}^{n} |a^*_{2j}|.$$

Convergência

Para $|y_i|$, temos

$$|y_i| = |-a^*_{i1}y_1 - a^*_{i2}y_2 - \ldots - a^*_{i,i-1}y_{i-1} - a^*_{i,i+1}x_{i+1} - \ldots - a^*_{in}x_n| \Rightarrow$$

$$|y_i| = \left|\sum_{j=1}^{i-1} -a^*_{ij}y_j + \sum_{j=i+1}^{n} -a^*_{ij}x_j\right| \leq \sum_{j=1}^{i-1} |-a^*_{ij}y_j| + \sum_{j=i+1}^{n} |-a^*_{ij}x_j| \Rightarrow$$

$$|y_i| \leq \sum_{j=1}^{i-1} |-a^*_{ij}||y_j| + \sum_{j=i+1}^{n} |-a^*_{ij}||x_j| \Rightarrow$$

Convergência

$$|y_i| \leq \sum_{j=1}^{i-1} |-a_{ij}^*| \beta_j \|x\|_\infty + \sum_{j=i+1}^{n} |-a_{ij}^*| \max_j |x_j| \Rightarrow$$

$$|y_i| \leq \sum_{j=1}^{i-1} |-a_{ij}^*| \beta_j \|x\|_\infty + \sum_{j=i+1}^{n} |-a_{ij}^*| \|x\|_\infty .$$

Convergência

Portanto,

$$|y_i| \leq \beta_i \|x\|_\infty,$$

com

$$\beta_i = \sum_{j=1}^{i-1} |a^*_{ij}|\beta_j + \sum_{j=i+1}^{n} |a^*_{ij}|.$$

Convergência

Assim, temos que

$$\|Bx\|_\infty = \|y\|_\infty = \max_{1\le i\le n} |y_i| \le \max_{1\le i\le n} \beta_i \|x\|_\infty \Rightarrow$$

$$\frac{\|Bx\|_\infty}{\|x\|_\infty} \le \max_{1\le i\le n} \beta_i.$$

Como x foi escolhido de modo que

$$\|B\|_\infty = \frac{\|Bx\|_\infty}{\|x\|_\infty},$$

temos que

$$\|B\|_\infty \le \max_{1\le i\le n} \beta_i.$$

Convergência

Se

$$\max_{1 \leq i \leq n} \beta_i < 1,$$

temos que $\|B\|_\infty < 1$ e, portanto, estará satisfeita uma condição suficiente de convergência.

Este é o Critério de Sassenfeld.

O Método Gauss-Seidel converge se

- O Critério de **Sassenfeld** for satisfeito, ou seja, se

$$\max_{1 \leq i \leq n} \beta_i < 1,$$

com $\beta_i = \sum_{j=1}^{i-1} |a_{ij}^*| \beta_j + \sum_{j=i+1}^{n} |a_{ij}^*|$.

- O critério das linhas for satisfeito, ou seja, se

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

- A matriz dos coeficientes do sistema linear for estritamente diagonal dominante.

Exemplo

Resolva o sistema linear

$$\begin{cases} 5x_1 & + & x_2 & + & x_3 & = & 5, \\ 3x_1 & + & 4x_2 & + & x_3 & = & 6, \\ 3x_1 & + & 3x_2 & + & 6x_3 & = & 0 \end{cases}$$

utilizando o Método Gauss-Seidel, com ponto inicial $x^{(0)} = (0, 0, 0)^T$ e erro menor ou igual a 10^{-2}.

Exemplo

Primeiramente, vamos verificar se o Método Gauss-Seidel tem convergência garantida para a resolução do sistema.

- Como $|a_{22}| = 4 \not> 3 + 1 = |a_{21}| + |a_{23}|$, a matriz de coeficientes não é estritamente dominante.
- Como $\max_{1 \leq i \leq 3} \sum_{j=1, j \neq i}^{3} |a^*_{ij}| = \max\{0.4, 1.0, 1.0\} = 1.0 \not< 1$, o critério das linhas também não é satisfeito.
- Vamos, então, verificar se o Critério de Sassenfeld é satisfeito.

Exemplo

A matriz de coeficientes é dada por

$$A = \begin{pmatrix} 5 & 1 & 1 \\ 3 & 4 & 1 \\ 3 & 3 & 6 \end{pmatrix}.$$

Dividindo cada linha pelo elemento da diagonal, temos

$$A^* = \begin{pmatrix} 1.00 & 0.20 & 0.20 \\ 0.75 & 1.00 & 0.25 \\ 0.50 & 0.50 & 1.00 \end{pmatrix}.$$

Exemplo

Lembrando que

$$\beta_i = \sum_{j=1}^{i-1} |a^*_{ij}|\beta_j + \sum_{j=i+1}^{n} |a^*_{ij}|,$$

temos

$$\beta_1 = |a^*_{12}| + |a^*_{13}| = 0.20 + 0.20 = 0.40,$$

$$\beta_2 = |a^*_{21}|\beta_1 + |a^*_{23}| = 0.75(0.40) + 0.25 = 0.55,$$

$$\beta_3 = |a^*_{31}|\beta_1 + |a^*_{32}|\beta_2 = 0.50(0.40) + 0.50(0.55) = 0.475.$$

Exemplo

Assim,

$$\max\{\beta_1, \beta_2, \beta_3\} = \max\{0.40, 0.55, 0.475\} = 0.55 < 1.0.$$

Logo, segundo o Critério de Sassenfeld, o Método Gauss-Seidel converge.

Exemplo

Vamos, então, aplicar o Método Gauss-Seidel:

$$\begin{cases} x_1^{(k+1)} & = & - & 0.20x_2^k & - & 0.20x_3^k & + & 1.00, \\ x_2^{(k+1)} & = & - & 0.75x_1^{(k+1)} & + & 0.25x_3^k & + & 1.50, \\ x_3^{(k+1)} & = & - & 0.50x_1^{(k+1)} & - & 0.50x_2^{(k+1)} & + & 0.00. \end{cases}$$

A partir de $x^{(0)} = (0, 0, 0)^T$, temos

$$\begin{cases} x_1^{(k+1)} & = & - & 0.20(0.00) & - & 0.20(0.00) & + & 1.00 & = & 1.000, \\ x_2^{(k+1)} & = & - & 0.75(1.00) & + & 0.25(0.00) & + & 1.50 & = & 0.750, \\ x_3^{(k+1)} & = & - & 0.50(1.00) & - & 0.50(0.75) & + & 0.00 & = & -0.875. \end{cases}$$

Exemplo

Calculando o erro:

$$E_{abs} = \|x^{(1)} - x^{(0)}\|_\infty = 1.000 > 10^{-2}.$$

$$E_{rel} = \frac{\|x^{(1)} - x^{(0)}\|_\infty}{\|x^{(1)}\|_\infty} = \frac{1.000}{1.000} = 1.000 > 10^{-2}.$$

Exemplo

Como o erro não foi satisfeito, continuamos o processo iterativo:

k	0	1	2	3	4
$x_1^{(k)}$	0.0000	1.0000	1.0250	1.0513	1.0532
$x_2^{(k)}$	0.0000	0.7500	0.5125	0.5194	0.5138
$x_3^{(k)}$	0.0000	-0.8750	-0.7688	-0.7853	-0.7835
E_{abs}	-	1.0000	0.2375	0.0263	0.0056
E_{rel}	-	1.0000	0.2317	0.0250	0.0053

Exemplo

Como

$$E_{rel} = \frac{\|x^{(4)} - x^{(3)}\|_\infty}{\|x^{(4)}\|_\infty} = \frac{0.0056}{1.0532} \simeq 0.0053 < 10^{-2},$$

paramos o processo iterativo e devolvemos $x^{(4)}$ como solução aproximada do sistema.