

Notas de Aulas de Algoritmos Numéricos

Resolução de Sistemas Lineares via Métodos Iterativos

1 Introdução

Para se resolver sistemas lineares há duas grandes categorias de métodos: os métodos diretos e os métodos iterativos. Os métodos diretos são aqueles que envolvem uma sequência finita e pré-definida de operações e levam à solução exata do problema (a menos de erros de arredondamento da máquina).

O método de eliminação de Gauss é um exemplo de um método direto.

Os métodos iterativos são aqueles que calculam a solução via sucessivas aproximações. Não se pode prever quantas operações serão realizadas para se chegar à solução (e, eventualmente, dependendo do problema, pode nem haver convergência).

2 Métodos iterativos

A ideia envolvida nos métodos iterativos é partir de uma aproximação inicial (um vetor chute inicial) e ir obtendo melhores soluções caminhando segundo alguma estratégia.

De uma forma bem geral, pode-se dizer que o processo iterativo consiste em gerar uma nova solução fornecendo um incremento δ a uma solução já existente, isto é, consiste em repetir a seguinte atualização:

$$x^{Nova} = x^{Antiga} + \delta.$$

A forma de calcular este incremento δ varia em cada processo iterativo mas todos buscam obter um valor que seja uma aproximação para o erro contido no vetor existente.

Se, em uma dada iteração k , temos uma solução aproximada x^k , pode se obter uma nova aproximação (um vetor x^{k+1}) somando “algo” relacionado com o erro contido naquela aproximação x^k .

Observe que o erro contido em uma aproximação em x^k é dado por:

$$erro^k = x_{exato} - x^k$$

Multiplicando por A

$$A(erro^k) = Ax_{exato} - Ax^k = b - Ax^k$$

Assim, o erro poderia ser obtido via

$$erro^k = A^{-1}[b - Ax^k].$$

Se somássemos, ao vetor x^k , o erro contido em x^k chegaríamos na solução exata em um único passo já que

$$x_{exato} = x^k + erro^k$$

No entanto, para calcular este $erro^k$ seria necessário calcular A^{-1} . Como o cálculo de A^{-1} demandaria um esforço ainda maior (computacionalmente) do que

resolver via eliminação de Gauss então não faz sentido calculá-la. O que é feito, como alternativa, é calcular uma aproximação para este $erro^k$ e, em vez de somar o $erro^k$, soma-se apenas um vetor δ^k que seja uma dada aproximação do erro.

Assim o processo iterativo fica

$$x^{k+1} = x^k + \delta^k$$

onde

$$\delta^k = (A^{-1})_{approx} * [b - Ax^k].$$

e $(A^{-1})_{approx}$ consiste em uma aproximação para a matriz (A^{-1}) .

Desta forma, a atualização de x^{k+1} é dada por,

$$x^{k+1} = x^k + ((A^{-1})_{approx} * [b - Ax^k])$$

Os diversos métodos iterativos diferem entre si na forma de obter uma aproximação para a matriz (A^{-1}) , isto é, diferem entre si no cálculo de $(A^{-1})_{approx}$, resultando em diferentes formas de obter o δ^k .

Para ver, alguma alternativas de como obter $(A^{-1})_{approx}$, vamos olhar como A pode ser expressa.

Uma matriz A pode ser desmembrada como uma soma três partes. Denominando de L, D e R a parte triangular da esquerda (L de Left), a parte diagonal (D) e a parte triangular da direita (R de Right), respectivamente, então $A = L + D + R$.

$$A = \begin{bmatrix} 0 & 0 & \dots & 0 \\ a_{21} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & 0 \end{bmatrix} + \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} + \begin{bmatrix} 0 & a_{12} & \dots & a_{1n} \\ 0 & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

O método de Gauss Jacobi usa: $(A^{-1})_{approx} = D^{-1}$.

O método de Gauss Seidel usa : $(A^{-1})_{approx} = (L + D)^{-1}$.

2.1 Resolução via método de Gauss Jacobi

Usando $(A^{-1})_{approx} = D^{-1}$, o processo iterativo é dado por:

$$x^{k+1} = x^k + D^{-1}[b - Ax^k]$$

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

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

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

Ou seja, a atualização do vetor é:

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

Escrevendo cada membro da equação matricial acima temos:

$$D^{-1} = \begin{bmatrix} 1/a_{11} & 0 & \dots & 0 \\ 0 & 1/a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/a_{nn} \end{bmatrix}$$

$$(L + R) = \begin{bmatrix} 0 & a_{12} & \dots & a_{1n} \\ a_{21} & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & 0 \end{bmatrix}$$

$$D^{-1}b = \begin{bmatrix} 1/a_{11} & 0 & \dots & 0 \\ 0 & 1/a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/a_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} b_1/a_{11} \\ b_2/a_{22} \\ \vdots \\ b_n/a_{nn} \end{bmatrix}$$

$D^{-1}(L + R)$ é dada por

$$\begin{bmatrix} 1/a_{11} & 0 & \dots & 0 \\ 0 & 1/a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/a_{nn} \end{bmatrix} \begin{bmatrix} 0 & a_{12} & \dots & a_{1n} \\ a_{21} & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & 0 \end{bmatrix} = \begin{bmatrix} 0/a_{11} & a_{12}/a_{11} & \dots & a_{1n}/a_{11} \\ a_{21}/a_{22} & 0/a_{22} & \dots & a_{2n}/a_{22} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}/a_{nn} & a_{n2}/a_{nn} & \dots & 0/a_{nn} \end{bmatrix}$$

Portanto,

$$\begin{aligned} D^{-1}b - D^{-1}(L + R)x^k &= \begin{bmatrix} b_1/a_{11} \\ b_2/a_{22} \\ \vdots \\ b_n/a_{nn} \end{bmatrix} - \begin{bmatrix} 0/a_{11} & a_{12}/a_{11} & \dots & a_{1n}/a_{11} \\ a_{21}/a_{22} & 0/a_{22} & \dots & a_{2n}/a_{22} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}/a_{nn} & a_{n2}/a_{nn} & \dots & 0/a_{nn} \end{bmatrix} \begin{bmatrix} x_1^k \\ x_2^k \\ \vdots \\ x_n^k \end{bmatrix} \\ &= \begin{bmatrix} b_1/a_{11} \\ b_2/a_{22} \\ \vdots \\ b_n/a_{nn} \end{bmatrix} - \begin{bmatrix} (0 + a_{12}x_2^k + a_{13}x_3^k + \dots + a_{1n}x_n^k)/a_{11} \\ (a_{21}x_1^k + 0 + a_{23}x_3^k + \dots + a_{2n}x_n^k)/a_{22} \\ \vdots \\ (a_{n1}x_1^k + a_{n2}x_2^k + \dots + a_{nn-1}x_{n-1}^k + 0)/a_{nn} \end{bmatrix} \end{aligned}$$

Assim, a atualização de uma componente i é dada por:

$$x_i^{k+1} = [b_i - (a_{i1}x_1^k + a_{i2}x_2^k + \dots + a_{i,i-1}x_{i-1}^k + a_{i,i+1}x_{i+1}^k + \dots + a_{i,n}x_n^k)]/a_{ii}$$

Escrito de forma mais compacta, tem -se:

$$x_i^{k+1} = [b_i - (\sum_{j=1}^{j=(i-1)} (a_{ij} * x_j^k) + \sum_{j=(i+1)}^{j=n} (a_{ij} * x_j^k))]/a_{ii}$$

2.2 Resolução via método de Gauss Seidel

No método de Gauss Seidel, como $(A^{-1})_{approx} = (L + D)^{-1}$, a atualização é dada por:

$$\begin{aligned}x^{k+1} &= x^k + (L + D)^{-1}[b - Ax^k] \\x^{k+1} &= x^k + (L + D)^{-1}[b - (L + D + R)x^k] \\x^{k+1} &= x^k + (L + D)^{-1}b - Ix^k - (L + D)^{-1}Rx^k \\x^{k+1} &= (L + D)^{-1}b - (L + D)^{-1}Rx^k\end{aligned}$$

Multiplicando a expressão acima por $(L + D)$ a expressão pode escrita via:

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

Ou seja, temos que

$$\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1^{k+1} \\ x_2^{k+1} \\ \vdots \\ x_n^{k+1} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} - \begin{bmatrix} 0 & a_{12} & \dots & a_{1n} \\ 0 & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} x_1^k \\ x_2^k \\ \vdots \\ x_n^k \end{bmatrix}$$

Componente a componente

$$\sum_{j=1}^{j=i-1} a_{ij} * x_j^{k+1} + a_{ii} * x_i^{k+1} = b_i - \sum_{j=i+1}^{j=n} a_{ij} * x_j^k$$

ou

$$x_i^{k+1} = (b_i - \sum_{j=1}^{j=i-1} a_{ij} * x_j^{k+1} - \sum_{j=i+1}^{j=n} a_{ij} * x_j^k) / a_{ii}$$

2.3 Critério de parada do processo iterativo

A ideia central é parar quando os últimos dois vetores gerados pelo processo iterativo estão “suficientemente” próximos (dada uma tolerância e uma métrica).

Deve-se verificar a distância (relativa) entre os últimos dois vetores gerados ou seja, deve-se calcular a distância relativa entre os vetores x^k e x^{k+1} , relativamente ao vetor mais atual:

$$DiffRel = \frac{\|x^{k+1} - x^k\|}{\|x^{k+1}\|}.$$

Para isso é preciso adotar uma métrica (uma forma de medir o tamanho do vetor). Existem várias normas vetoriais que permitem calcular o tamanho de um vetor (como, por exemplo, a norma euclidiana).

A norma do máximo (também chamada de norma infinito) é bastante usada computacionalmente pois é barata (exige pouco cálculos). A norma do máximo é definida por

$$\|v\|_{max} = \max(|v_i|), i = 1, n$$

Assim, adotando esta norma, deve-se parar o processo iterativo quando a distância relativa entre eles for menor que uma dada tolerância (tol):

$$DiffRel = \frac{\|x^{k+1} - x^k\|_{max}}{\|x^{k+1}\|_{max}} \leq tol.$$

3 Análise de convergência

Para se fazer uma análise da convergência da sequência de vetores gerados pelos processos iterativos de Gauss Jacobi e Gauss Seidel é importante observar como se comporta o erro contido na aproximação $x^{(k)}$. É importante, também, expressar o processo iterativo em função da matriz de iteração.

3.1 Estudo do erro e o processo iterativo

O erro ($erro^{(k)}$) contido em uma aproximação $x^{(k)}$ é:

$$erro^{(k)} = x_{exato} - x^{(k)}$$

Para garantir a convergência temos garantir que

$$\lim_{k \rightarrow \infty} (x_{exato} - x^{(k)}) = 0$$

Para entender como se comporta o erro é preciso ver que os dois métodos iterativos podem ser escritos através de uma atualização que envolve uma matriz M , denominada de matriz de iteração, e um vetor g . Os processos iterativos podem ser escritos via:

$$x^{k+1} = Mx^k + g.$$

Para **Gauss Jacobi** a matriz de iteração é $M = -D^{-1}(L + R)$ e o vetor g é $g = D^{-1}b$ já que o processo iterativo, neste método, é dado por

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

Para **Gauss Seidel**, a matriz é $M = -(D + L)^{-1}R$ e o vetor g é $g = (D + L)^{-1}b$ já que, em Gauss Seidel, a atualização do vetor é

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

Além disso, a solução exata pode ser escrita em função da matriz de iteração, isto é:

$$x_{exato} = Mx_{exato} + g$$

pois são válidas as seguintes igualdades:

para **Gauss Jacobi**:

$$\begin{aligned}
 Ax_{exato} &= b \\
 (L + D + R)x_{exato} &= b \\
 Dx_{exato} &= b - (L + R)x_{exato} \\
 x_{exato} &= D^{-1}b - D^{-1}(L + R)x_{exato} \\
 x_{exato} &= g + Mx_{exato}
 \end{aligned}$$

(2) para **Gauss Seidel**:

$$\begin{aligned}
 Ax_{exato} &= b \\
 (L + D + R)x_{exato} &= b \\
 (L + D)x_{exato} &= b - Rx_{exato} \\
 x_{exato} &= (L + D)^{-1}b - (L + D)^{-1}Rx_{exato} \\
 x_{exato} &= g + Mx_{exato}
 \end{aligned}$$

O erro $erro^{(k)}$ contido na aproximação da iteração (k) está relacionado com o erro existente na iteração anterior ($erro^{(k-1)}$) através de:

$$\begin{aligned}
 erro^{(k)} &= x_{exato} - x^{(k)} = (Mx_{exato} + g) - (Mx^{(k-1)} + g) \\
 erro^{(k)} &= (Mx_{exato}) - (Mx^{(k-1)}) \\
 erro^{(k)} &= M(x_{exato} - x^{(k-1)}) \\
 erro^{(k)} &= M(erro^{(k-1)})
 \end{aligned}$$

De forma análoga, pode-se escrever que erro contido na aproximação $erro^{(k-1)}$ está relacionado com o erro existente na iteração anterior ($erro^{(k-2)}$) através de:

$$\begin{aligned}
 erro^{(k-1)} &= x_{exato} - x^{(k-1)} = (Mx_{exato} + g) - (Mx^{(k-2)} + g) \\
 erro^{(k-1)} &= (Mx_{exato}) - (Mx^{(k-2)}) \\
 erro^{(k-1)} &= M(x_{exato} - x^{(k-2)}) \\
 erro^{(k-1)} &= M(erro^{(k-2)})
 \end{aligned}$$

Assim,

$$\begin{aligned}
 erro^{(k)} &= M(erro^{(k-1)}) = M(M(erro^{(k-2)})) \\
 erro^{(k)} &= M^2(erro^{(k-2)})
 \end{aligned}$$

Pode-se, ainda, escrever que

$$erro^{(k-2)} = x_{exato} - x^{(k-2)} = (Mx_{exato} + g) - (Mx^{(k-3)} + g) = M(erro^{(k-3)})$$

e, portanto,

$$erro^{(k)} = M^3(erro^{(k-3)})$$

Continuando desta forma, expressa-se $erro^{(k)}$ em função do erro inicial $erro^{(0)}$

$$erro^{(k)} = M^k(erro^{(0)})$$

É possível mostrar que se o raio espectral de M for menor que 1 (o que equivale a exigir que todos os autovalores de M devem ser, em módulo, menores que 1) então

$$\lim_{k \rightarrow \infty} (erro^{(k)}) = \lim_{k \rightarrow \infty} M^k(erro^{(0)}) = 0$$

o que equivale a dizer que a sequência gerada pelo processo iterativo $x^{k+1} = Mx^k + g$ converge para a solução exata. A determinação do raio espectral de M requer um esforço computacional maior do que o esforço de se obter a solução de $Ax = b$ por isso, usualmente, outros critérios são empregados para se garantir a convergência.

3.2 A convergência e o critério das linhas

Para que se possa garantir que a sequência de vetores $x^{(k)}$ seja convergente, um critério será estabelecido em função da norma de $erro^{(k)}$.

Para que haja convergência é necessário que

$$\lim_{k \rightarrow \infty} (erro^{(k)}) = 0$$

Como

$$e^{(k)} = M^k(erro^{(0)})$$

Tomando a norma

$$\|e^{(k)}\| = \|M^k(erro^{(0)})\|$$

Usando propriedades de norma matriciais, sabe-se que

$$\|M^k(erro^{(0)})\| \leq \|M\| \cdot \|M^{k-1}(erro^{(0)})\| \leq \|M\| \|M\| \|M^{k-2}(erro^{(0)})\|$$

e

$$\leq \|M\| \cdot \|M\| \cdot \|M\| \cdot \|M^{k-3}(erro^{(0)})\|.$$

Continuando, desta forma, é possível escrever que

$$\|M^k(erro^{(0)})\| \leq (\|M\|)^k \cdot \|(erro^{(0)})\|$$

Se

$$\|M\| < 1$$

então

$$\lim_{k \rightarrow \infty} (\|M\|)^k \cdot \|(erro^{(0)})\| = 0$$

e, por consequência,

$$\lim_{k \rightarrow \infty} \|(erro^{(k)})\| = 0$$

Demonstração: lembrando que

$$0 \leq \|(erro^{(k)})\| \leq (\|M\|)^k \cdot \|(erro^{(0)})\|$$

então, pelo teorema do sanduíche, é possível garantir $\lim_{k \rightarrow \infty} ||(erro^{(k)})|| = 0$.

A condição $||M|| < 1$ garante a convergência sob qualquer norma. Se a norma do máximo ($||\cdot||_{max}$) (ou também chamada de norma infinito ($||\cdot||_{\infty}$) for empregada, a condição (suficiente) para garantir a convergência é simples de ser verificada.

Lembrando que no método de **Gauss Jacobi**, a matriz de iteração $M = D^{-1}(L + R)$ é dada por:

$$\begin{bmatrix} 1/a_{11} & 0 & \dots & 0 \\ 0 & 1/a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/a_{nn} \end{bmatrix} \begin{bmatrix} 0 & a_{12} & \dots & a_{1n} \\ a_{21} & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & 0 \end{bmatrix} = \begin{bmatrix} 0/a_{11} & a_{12}/a_{11} & \dots & a_{1n}/a_{11} \\ a_{21}/a_{22} & 0/a_{22} & \dots & a_{2n}/a_{22} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}/a_{nn} & a_{n2}/a_{nn} & \dots & 0/a_{nn} \end{bmatrix}$$

Assim, exigir que $||M||_{\infty} = ||M||_{max} = \max(\sum_{j=1}^{j=n} |m_{i,j}|) < 1$, equivale a exigir, que para toda linha i , a seguinte desigualdade seja verdadeira:

$$\left(\frac{|a(i, 1)|}{|a(i, i)|} + \dots + \frac{|a(i, i-1)|}{|a(i, i)|} + 0 + \frac{|a(i, i+1)|}{|a(i, i)|} + \dots + \frac{|a(i, n)|}{|a(i, i)|} \right) < 1$$

Portanto, exigir que $||M||_{max} < 1$, para o método de **Gauss Jacobi**, é equivalente a exigir que matriz A seja diagonalmente dominante. O critério é conhecido como o critério das linhas.

Para o método de **Gauss Seidel**, onde a matriz de iteração é $M = (D + L)^{-1}R$, também é possível estabelecer um critério (um deles é conhecido como o critério de Sassenfeld) mas não o definiremos aqui.

O que se pode garantir é que se a matriz A for diagonalmente dominante, a sequencia gerada pelo método de Gauss-Seidel também será convergente para a solução exata.

Algumas vezes, quando o problema linear $Ax = b$ não possuir uma matriz que seja diagonalmente dominante pode-se efetuar trocas de linhas para que em uma nova configuração ele tenha uma matriz A diagonalmente dominante.