

Sistemas Lineares: Métodos Iterativos

Prof. Afonso Paiva

Departamento de Matemática Aplicada e Estatística
Instituto de Ciências Matemáticas e de Computação
USP – São Carlos

Métodos Numéricos e Computacionais I – SME0305



Introdução

Seja $\mathbf{Ax} = \mathbf{b}$ um sistema linear de ordem n , com $\det(\mathbf{A}) \neq 0$.

Objetivo: queremos definir um processo iterativo, de modo que a sequência de vetores $\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots\}$ produzida por esse processo **convirja** para a solução \mathbf{x} , independentemente da escolha do chute inicial $\mathbf{x}^{(0)} \in \mathbb{R}^n$.

Introdução

Seja $\mathbf{Ax} = \mathbf{b}$ um sistema linear de ordem n , com $\det(\mathbf{A}) \neq 0$.

Objetivo: queremos definir um processo iterativo, de modo que a sequência de vetores $\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots\}$ produzida por esse processo **convirja** para a solução \mathbf{x} , independentemente da escolha do chute inicial $\mathbf{x}^{(0)} \in \mathbb{R}^n$.

Definição

Uma sequência de vetores $\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots\}$ converge para um vetor \mathbf{x} , se

$$\lim_{k \rightarrow \infty} \|\mathbf{x}^{(k)} - \mathbf{x}\| = 0.$$

Notação: $\mathbf{x}^{(k)} \rightarrow \mathbf{x}$.

Introdução

Ideia principal: vamos criar um processo recursivo através de um sistema equivalente a $\mathbf{Ax} = \mathbf{b}$.

Introdução

Ideia principal: vamos criar um processo recursivo através de um sistema equivalente a $\mathbf{Ax} = \mathbf{b}$.

- 1 Transformar $\mathbf{Ax} = \mathbf{b}$ em um sistema equivalente da forma:

$$\mathbf{x} = \mathbf{Cx} + \mathbf{g},$$

em que $\mathbf{C} \in \mathcal{M}(n, n)$ e $\mathbf{g} \in \mathbb{R}^n$ são conhecidos.

Introdução

Ideia principal: vamos criar um processo recursivo através de um sistema equivalente a $\mathbf{Ax} = \mathbf{b}$.

- 1 Transformar $\mathbf{Ax} = \mathbf{b}$ em um sistema equivalente da forma:

$$\mathbf{x} = \mathbf{Cx} + \mathbf{g},$$

em que $\mathbf{C} \in \mathcal{M}(n, n)$ e $\mathbf{g} \in \mathbb{R}^n$ são conhecidos.

- 2 Dado um chute inicial $\mathbf{x}^{(0)}$, obtemos uma sequência $\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \dots\}$ através do processo iterativo:

$$\mathbf{x}^{(k+1)} = \mathbf{Cx}^{(k)} + \mathbf{g}, \quad k = 0, 1, 2, \dots \quad (\star)$$

Introdução

Perguntas:

- Dado $\mathbf{Ax} = \mathbf{b}$ é possível obter um sistema equivalente $\mathbf{x} = \mathbf{Cx} + \mathbf{g}$?

Introdução

Perguntas:

- Dado $\mathbf{Ax} = \mathbf{b}$ é possível obter um sistema equivalente $\mathbf{x} = \mathbf{Cx} + \mathbf{g}$?
 - Sim. Por exemplo, basta tomar $\mathbf{C} = \mathbf{I} - \mathbf{A}$ e $\mathbf{g} = \mathbf{b}$.

Introdução

Perguntas:

- Dado $\mathbf{Ax} = \mathbf{b}$ é possível obter um sistema equivalente $\mathbf{x} = \mathbf{Cx} + \mathbf{g}$?
 - Sim. Por exemplo, basta tomar $\mathbf{C} = \mathbf{I} - \mathbf{A}$ e $\mathbf{g} = \mathbf{b}$.
- Se $\mathbf{x}^{(k)} \rightarrow \bar{\mathbf{x}}$ então $\bar{\mathbf{x}}$ é solução de $\mathbf{Ax} = \mathbf{b}$?

Introdução

Perguntas:

- Dado $\mathbf{Ax} = \mathbf{b}$ é possível obter um sistema equivalente $\mathbf{x} = \mathbf{Cx} + \mathbf{g}$?
 - Sim. Por exemplo, basta tomar $\mathbf{C} = \mathbf{I} - \mathbf{A}$ e $\mathbf{g} = \mathbf{b}$.
- Se $\mathbf{x}^{(k)} \rightarrow \bar{\mathbf{x}}$ então $\bar{\mathbf{x}}$ é solução de $\mathbf{Ax} = \mathbf{b}$?
 - Sim. Passando o limite em ambos lados da Equação $(*)$, temos que $\bar{\mathbf{x}} = \mathbf{C}\bar{\mathbf{x}} + \mathbf{g}$. Pela hipótese de equivalência, segue que $\bar{\mathbf{x}} = \mathbf{x}$.

Introdução

Perguntas:

- Dado $\mathbf{Ax} = \mathbf{b}$ é possível obter um sistema equivalente $\mathbf{x} = \mathbf{Cx} + \mathbf{g}$?
 - Sim. Por exemplo, basta tomar $\mathbf{C} = \mathbf{I} - \mathbf{A}$ e $\mathbf{g} = \mathbf{b}$.
- Se $\mathbf{x}^{(k)} \rightarrow \bar{\mathbf{x}}$ então $\bar{\mathbf{x}}$ é solução de $\mathbf{Ax} = \mathbf{b}$?
 - Sim. Passando o limite em ambos lados da Equação $(*)$, temos que $\bar{\mathbf{x}} = \mathbf{C}\bar{\mathbf{x}} + \mathbf{g}$. Pela hipótese de equivalência, segue que $\bar{\mathbf{x}} = \mathbf{x}$.
- Quando $\mathbf{x}^{(k)} \rightarrow \mathbf{x}$?

Introdução

Perguntas:

- Dado $\mathbf{Ax} = \mathbf{b}$ é possível obter um sistema equivalente $\mathbf{x} = \mathbf{Cx} + \mathbf{g}$?
 - Sim. Por exemplo, basta tomar $\mathbf{C} = \mathbf{I} - \mathbf{A}$ e $\mathbf{g} = \mathbf{b}$.
- Se $\mathbf{x}^{(k)} \rightarrow \bar{\mathbf{x}}$ então $\bar{\mathbf{x}}$ é solução de $\mathbf{Ax} = \mathbf{b}$?
 - Sim. Passando o limite em ambos lados da Equação $(*)$, temos que $\bar{\mathbf{x}} = \mathbf{C}\bar{\mathbf{x}} + \mathbf{g}$. Pela hipótese de equivalência, segue que $\bar{\mathbf{x}} = \mathbf{x}$.
- Quando $\mathbf{x}^{(k)} \rightarrow \mathbf{x}$?
- Quando terminar o processo iterativo $\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \dots\}$?

Convergência

Definição (raio espectral)

O **raio espectral** de uma matriz $\mathbf{A} \in \mathcal{M}(n, n)$ é definido como

$$\rho(\mathbf{A}) = \max_{i \in \{1, \dots, n\}} \{|\lambda_i|\},$$

onde λ_i são os autovalores de \mathbf{A} .

Convergência

Definição (raio espectral)

O **raio espectral** de uma matriz $\mathbf{A} \in \mathcal{M}(n, n)$ é definido como

$$\rho(\mathbf{A}) = \max_{i \in \{1, \dots, n\}} \{|\lambda_i|\},$$

onde λ_i são os autovalores de \mathbf{A} .

Teorema (critério geral de convergência)

Seja $\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots\}$ sequência gerada pelo processo iterativo (\star) .

- 1 Se $\|\mathbf{C}\|_M < 1$, onde $\|\cdot\|_M$ é uma norma consistente, então a sequência converge.
- 2 $\mathbf{x}^{(k)} \rightarrow \mathbf{x}$ se e somente se $\rho(\mathbf{C}) < 1$.

Critérios de Parada

Voltando a pergunta de quando devemos parar a sequência $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$?

Critérios de Parada

Voltando a pergunta de quando devemos parar a sequência $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$?

Dados $\varepsilon > 0$ e $k_{max} \in \mathbb{N}$, temos:

1 Erro absoluto:

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| < \varepsilon;$$

Critérios de Parada

Voltando a pergunta de quando devemos parar a sequência $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$?

Dados $\varepsilon > 0$ e $k_{max} \in \mathbb{N}$, temos:

1 Erro absoluto:

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| < \varepsilon;$$

2 Erro relativo:

$$\frac{\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|}{\|\mathbf{x}^{(k+1)}\|} < \varepsilon;$$

Critérios de Parada

Voltando a pergunta de quando devemos parar a sequência $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$?

Dados $\varepsilon > 0$ e $k_{max} \in \mathbb{N}$, temos:

1 Erro absoluto:

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| < \varepsilon;$$

2 Erro relativo:

$$\frac{\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|}{\|\mathbf{x}^{(k+1)}\|} < \varepsilon;$$

3 Teste de resíduo:

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}^{(k)}\| < \varepsilon;$$

Critérios de Parada

Voltando a pergunta de quando devemos parar a sequência $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$?

Dados $\varepsilon > 0$ e $k_{max} \in \mathbb{N}$, temos:

1 Erro absoluto:

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| < \varepsilon;$$

2 Erro relativo:

$$\frac{\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|}{\|\mathbf{x}^{(k+1)}\|} < \varepsilon;$$

3 Teste de resíduo:

$$\|\mathbf{b} - \mathbf{Ax}^{(k)}\| < \varepsilon;$$

4 Número máximo de iterações:

$$k = k_{max}.$$

Método de Gauss-Jacobi

Dado $\mathbf{Ax} = \mathbf{b}$ e supondo sem perda de generalidade que $a_{ii} \neq 0, i = 1, \dots, n$, temos:

$$\left\{ \begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n & = & b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n & = & b_3 \\ \vdots & & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots + a_{nn}x_n & = & b_n \end{array} \right.$$

Método de Gauss-Jacobi

Dado $\mathbf{Ax} = \mathbf{b}$ e supondo sem perda de generalidade que $a_{ii} \neq 0, i = 1, \dots, n$, temos:

$$\left\{ \begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n & = & b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n & = & b_3 \\ \vdots & & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots + a_{nn}x_n & = & b_n \end{array} \right.$$

A forma como o Método de Gauss-Jacobi transforma $\mathbf{Ax} = \mathbf{b}$ em $\mathbf{x} = \mathbf{Cx} + \mathbf{g}$ é feita isolando cada coordenada x_i do vetor \mathbf{x} na i -ésima equação do sistema.

Método de Gauss-Jacobi

Logo,

$$\left\{ \begin{array}{lcl} x_1 & = & (b_1 - a_{12}x_2 - a_{13}x_3 - \cdots - a_{1n}x_n) / a_{11} \\ x_2 & = & (b_2 - a_{21}x_1 - a_{23}x_3 - \cdots - a_{2n}x_n) / a_{22} \\ x_3 & = & (b_3 - a_{31}x_1 - a_{32}x_2 - \cdots - a_{3n}x_n) / a_{33} \\ \vdots & \vdots & \vdots \\ x_n & = & (b_n - a_{n1}x_1 - a_{n2}x_2 - \cdots - a_{n,n-1}x_{n-1}) / a_{nn} \end{array} \right.$$

Método de Gauss-Jacobi

Logo,

$$\begin{cases} x_1 = (b_1 - a_{12}x_2 - a_{13}x_3 - \cdots - a_{1n}x_n)/a_{11} \\ x_2 = (b_2 - a_{21}x_1 - a_{23}x_3 - \cdots - a_{2n}x_n)/a_{22} \\ x_3 = (b_3 - a_{31}x_1 - a_{32}x_2 - \cdots - a_{3n}x_n)/a_{33} \\ \vdots \\ x_n = (b_n - a_{n1}x_1 - a_{n2}x_2 - \cdots - a_{n,n-1}x_{n-1})/a_{nn} \end{cases}$$

Desta forma temos o sistema equivalente $\mathbf{x} = \mathbf{C}\mathbf{x} + \mathbf{g}$, em que:

$$\mathbf{C} = \begin{bmatrix} 0 & -a_{12}/a_{11} & -a_{13}/a_{11} & \cdots & -a_{1n}/a_{11} \\ -a_{21}/a_{22} & 0 & -a_{23}/a_{22} & \cdots & -a_{2n}/a_{22} \\ -a_{31}/a_{33} & -a_{32}/a_{33} & 0 & \cdots & -a_{3n}/a_{33} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{n1}/a_{nn} & -a_{n2}/a_{nn} & \cdots & -a_{n,n-1}/a_{nn} & 0 \end{bmatrix} \quad \text{e} \quad \mathbf{g} = \begin{bmatrix} b_1/a_{11} \\ b_2/a_{22} \\ b_3/a_{33} \\ \vdots \\ b_n/a_{nn} \end{bmatrix}$$

Método de Gauss-Jacobi

Portanto, dado o chute inicial $\mathbf{x}^{(0)}$, o processo iterativo é dado por:

$$\begin{cases} x_1^{(k+1)} = (b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \cdots - a_{1n}x_n^{(k)})/a_{11} \\ x_2^{(k+1)} = (b_2 - a_{21}x_1^{(k)} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)})/a_{22} \\ x_3^{(k+1)} = (b_3 - a_{31}x_1^{(k)} - a_{32}x_2^{(k)} - \cdots - a_{3n}x_n^{(k)})/a_{33} \\ \vdots \\ x_n^{(k+1)} = (b_n - a_{n1}x_1^{(k)} - a_{n2}x_2^{(k)} - \cdots - a_{n,n-1}x_{n-1}^{(k)})/a_{nn} \end{cases}$$

Desta forma temos o sistema equivalente $\mathbf{x}^{(k+1)} = \mathbf{C}\mathbf{x}^{(k)} + \mathbf{g}$, em que:

$$\mathbf{C} = \begin{bmatrix} 0 & -a_{12}/a_{11} & -a_{13}/a_{11} & \cdots & -a_{1n}/a_{11} \\ -a_{21}/a_{22} & 0 & -a_{23}/a_{22} & \cdots & -a_{2n}/a_{22} \\ -a_{31}/a_{33} & -a_{32}/a_{33} & 0 & \cdots & -a_{3n}/a_{33} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{n1}/a_{nn} & -a_{n2}/a_{nn} & \cdots & -a_{n,n-1}/a_{nn} & 0 \end{bmatrix} \quad \text{e} \quad \mathbf{g} = \begin{bmatrix} b_1/a_{11} \\ b_2/a_{22} \\ b_3/a_{33} \\ \vdots \\ b_n/a_{nn} \end{bmatrix}$$

Método de Gauss-Jacobi

Forma Matricial

Vamos mostrar como obter $\mathbf{x}^{(k+1)} = \mathbf{C}\mathbf{x}^{(k)} + \mathbf{g}$ a partir de $\mathbf{A}\mathbf{x} = \mathbf{b}$. Seja \mathbf{D} uma matriz diagonal formada pela diagonal de \mathbf{A} . Assim,

$$\mathbf{A}\mathbf{x} = \mathbf{b} \Leftrightarrow (\mathbf{A} - \mathbf{D} + \mathbf{D})\mathbf{x} = \mathbf{b} \Leftrightarrow (\mathbf{A} - \mathbf{D})\mathbf{x} + \mathbf{D}\mathbf{x} = \mathbf{b}$$

Método de Gauss-Jacobi

Forma Matricial

Vamos mostrar como obter $\mathbf{x}^{(k+1)} = \mathbf{C}\mathbf{x}^{(k)} + \mathbf{g}$ a partir de $\mathbf{A}\mathbf{x} = \mathbf{b}$. Seja \mathbf{D} uma matriz diagonal formada pela diagonal de \mathbf{A} . Assim,

$$\mathbf{A}\mathbf{x} = \mathbf{b} \Leftrightarrow (\mathbf{A} - \mathbf{D} + \mathbf{D})\mathbf{x} = \mathbf{b} \Leftrightarrow (\mathbf{A} - \mathbf{D})\mathbf{x} + \mathbf{D}\mathbf{x} = \mathbf{b}$$

Dessa forma,

$$(\mathbf{A} - \mathbf{D})\mathbf{x}^{(k)} + \mathbf{D}\mathbf{x}^{(k+1)} = \mathbf{b} \Leftrightarrow \mathbf{D}\mathbf{x}^{(k+1)} = (\mathbf{D} - \mathbf{A})\mathbf{x}^{(k)} + \mathbf{b}$$

Método de Gauss-Jacobi

Forma Matricial

Vamos mostrar como obter $\mathbf{x}^{(k+1)} = \mathbf{C}\mathbf{x}^{(k)} + \mathbf{g}$ a partir de $\mathbf{A}\mathbf{x} = \mathbf{b}$. Seja \mathbf{D} uma matriz diagonal formada pela diagonal de \mathbf{A} . Assim,

$$\mathbf{A}\mathbf{x} = \mathbf{b} \Leftrightarrow (\mathbf{A} - \mathbf{D} + \mathbf{D})\mathbf{x} = \mathbf{b} \Leftrightarrow (\mathbf{A} - \mathbf{D})\mathbf{x} + \mathbf{D}\mathbf{x} = \mathbf{b}$$

Dessa forma,

$$(\mathbf{A} - \mathbf{D})\mathbf{x}^{(k)} + \mathbf{D}\mathbf{x}^{(k+1)} = \mathbf{b} \Leftrightarrow \mathbf{D}\mathbf{x}^{(k+1)} = (\mathbf{D} - \mathbf{A})\mathbf{x}^{(k)} + \mathbf{b}$$

Portanto,

$$\mathbf{x}^{(k+1)} = \underbrace{(\mathbf{I} - \mathbf{D}^{-1}\mathbf{A})}_{\mathbf{C}} \mathbf{x}^{(k)} + \underbrace{\mathbf{D}^{-1}\mathbf{b}}_{\mathbf{g}}.$$

MATLAB – Método de Gauss-Jacobi

```
function [x,k]=gauss_jacobi(A,b,x0,tol)
n = size(A,1);
D = diag(diag(A));
C = eye(n)-D\A;
g = D\b;
kmax = 10000; k = 0;

while (norm(b-A*x0)>tol && k<kmax)
    k = k+1;
    x0 = C*x0+g;
end
if (k == kmax)
    disp('Erro: o metodo nao converge.');
```

return;

```
end
x = x0;
```

Método de Gauss-Jacobi

Crítérios de Convergência

O Método de Gauss-Jacobi **converge** para a solução de $\mathbf{Ax} = \mathbf{b}$, independentemente da escolha de $\mathbf{x}^{(0)}$, se satisfazer um dos critérios:

Método de Gauss-Jacobi

CrITÉRIOS de Convergência

O Método de Gauss-Jacobi **converge** para a solução de $\mathbf{Ax} = \mathbf{b}$, independentemente da escolha de $\mathbf{x}^{(0)}$, se satisfazer um dos critérios:

1 Critério das linhas:

$$\alpha = \max_{1 \leq k \leq n} \{\alpha_k\} < 1, \quad \text{com} \quad \alpha_k = \frac{\sum_{j=1, j \neq k}^n |a_{kj}|}{|a_{kk}|}$$

Método de Gauss-Jacobi

Crítérios de Convergência

O Método de Gauss-Jacobi **converge** para a solução de $\mathbf{Ax} = \mathbf{b}$, independentemente da escolha de $\mathbf{x}^{(0)}$, se satisfazer um dos critérios:

1 Critério das linhas:

$$\alpha = \max_{1 \leq k \leq n} \{\alpha_k\} < 1, \quad \text{com} \quad \alpha_k = \frac{\sum_{j=1, j \neq k}^n |a_{kj}|}{|a_{kk}|}$$

2 Critério das colunas:

$$\alpha = \max_{1 \leq k \leq n} \{\alpha_k\} < 1, \quad \text{com} \quad \alpha_k = \frac{\sum_{i=1, i \neq k}^n |a_{ik}|}{|a_{kk}|}$$

Método de Gauss-Jacobi

Crítérios de Convergência

Observações:

- Uma matriz que satisfaz o critério das linhas é dita **estritamente diagonal dominante**;

Método de Gauss-Jacobi

Critérios de Convergência

Observações:

- Uma matriz que satisfaz o critério das linhas é dita **estritamente diagonal dominante**;
- Quanto menor o valor de α , mais rápida será a convergência.

Método de Gauss-Jacobi

Exercício 1

Considere o sistema linear:

$$\begin{cases} 8x_1 + x_2 - x_3 &= 8 \\ x_1 - 7x_2 + 2x_3 &= -4 \\ 2x_1 + x_2 + 9x_3 &= 12 \end{cases}.$$

- 1 Determine o Método de Gauss-Jacobi para resolver o sistema acima;
- 2 O Método de Gauss-Jacobi converge?
- 3 Dado o chute inicial $\mathbf{x}^{(0)} = (0, 0, 0)^\top$, calcule $\mathbf{x}^{(1)}$.

Método de Gauss-Jacobi

Exercício 2

Considere o sistema linear:

$$\begin{cases} x_1 + 3x_2 + x_3 &= -2 \\ 5x_1 + 2x_2 + 2x_3 &= 3 \\ 6x_2 + 8x_3 &= -6 \end{cases} .$$

Teria como usar o Método de Gauss-Jacobi para resolver o sistema acima analisando a convergência do método?

Método de Gauss-Seidel

Como acelerar a convergência do Método de Gauss-Jacobi?

- No cálculo de $x_i^{(k+1)}$ usar os valores atualizados $x_1^{(k+1)}, \dots, x_{i-1}^{(k+1)}$ e os valores restantes $x_{i+1}^{(k)}, \dots, x_n^{(k)}$.

Método de Gauss-Seidel

Como acelerar a convergência do Método de Gauss-Jacobi?

- No cálculo de $x_i^{(k+1)}$ usar os valores atualizados $x_1^{(k+1)}, \dots, x_{i-1}^{(k+1)}$ e os valores restantes $x_{i+1}^{(k)}, \dots, x_n^{(k)}$.

$$\begin{cases} x_1^{(k+1)} = (b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - a_{14}x_4^{(k)} - \dots - a_{1n}x_n^{(k)})/a_{11} \\ x_2^{(k+1)} = (b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} - a_{24}x_4^{(k)} - \dots - a_{2n}x_n^{(k)})/a_{22} \\ x_3^{(k+1)} = (b_3 - a_{31}x_1^{(k+1)} - a_{32}x_2^{(k+1)} - a_{34}x_4^{(k)} - \dots - a_{3n}x_n^{(k)})/a_{33} \\ \vdots \\ x_n^{(k+1)} = (b_n - a_{n1}x_1^{(k+1)} - a_{n2}x_2^{(k+1)} - \dots - a_{n,n-1}x_{n-1}^{(k+1)})/a_{nn} \end{cases}$$

Método de Gauss-Seidel

Forma Matricial

Vamos mostrar como obter $\mathbf{x}^{(k+1)} = \mathbf{C}\mathbf{x}^{(k)} + \mathbf{g}$ a partir de $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Considere $\mathbf{A} = \mathbf{L} + \mathbf{R}$, em que \mathbf{L} é a matriz triangular inferior de \mathbf{A} e \mathbf{R} é a matriz triangular superior de \mathbf{A} sem a diagonal. Assim,

$$\mathbf{A}\mathbf{x} = \mathbf{b} \Leftrightarrow (\mathbf{L} + \mathbf{R})\mathbf{x} = \mathbf{b} \Leftrightarrow \mathbf{L}\mathbf{x} + \mathbf{R}\mathbf{x} = \mathbf{b}$$

Método de Gauss-Seidel

Forma Matricial

Vamos mostrar como obter $\mathbf{x}^{(k+1)} = \mathbf{C}\mathbf{x}^{(k)} + \mathbf{g}$ a partir de $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Considere $\mathbf{A} = \mathbf{L} + \mathbf{R}$, em que \mathbf{L} é a matriz triangular inferior de \mathbf{A} e \mathbf{R} é a matriz triangular superior de \mathbf{A} sem a diagonal. Assim,

$$\mathbf{A}\mathbf{x} = \mathbf{b} \Leftrightarrow (\mathbf{L} + \mathbf{R})\mathbf{x} = \mathbf{b} \Leftrightarrow \mathbf{L}\mathbf{x} + \mathbf{R}\mathbf{x} = \mathbf{b}$$

Dessa forma,

$$\mathbf{L}\mathbf{x}^{(k+1)} + \mathbf{R}\mathbf{x}^{(k)} = \mathbf{b} \Leftrightarrow \mathbf{L}\mathbf{x}^{(k+1)} = -\mathbf{R}\mathbf{x}^{(k)} + \mathbf{b}$$

Método de Gauss-Seidel

Forma Matricial

Vamos mostrar como obter $\mathbf{x}^{(k+1)} = \mathbf{C}\mathbf{x}^{(k)} + \mathbf{g}$ a partir de $\mathbf{A}\mathbf{x} = \mathbf{b}$. Considere $\mathbf{A} = \mathbf{L} + \mathbf{R}$, em que \mathbf{L} é a matriz triangular inferior de \mathbf{A} e \mathbf{R} é a matriz triangular superior de \mathbf{A} sem a diagonal. Assim,

$$\mathbf{A}\mathbf{x} = \mathbf{b} \Leftrightarrow (\mathbf{L} + \mathbf{R})\mathbf{x} = \mathbf{b} \Leftrightarrow \mathbf{L}\mathbf{x} + \mathbf{R}\mathbf{x} = \mathbf{b}$$

Dessa forma,

$$\mathbf{L}\mathbf{x}^{(k+1)} + \mathbf{R}\mathbf{x}^{(k)} = \mathbf{b} \Leftrightarrow \mathbf{L}\mathbf{x}^{(k+1)} = -\mathbf{R}\mathbf{x}^{(k)} + \mathbf{b}$$

Portanto,

$$\mathbf{x}^{(k+1)} = \underbrace{(-\mathbf{L}^{-1}\mathbf{R})}_{\mathbf{C}} \mathbf{x}^{(k)} + \underbrace{\mathbf{L}^{-1}\mathbf{b}}_{\mathbf{g}}.$$

MATLAB – Método de Gauss-Seidel

```
function [x,k]=gauss_seidel(A,b,x0,tol)
L = tril(A); R = triu(A,1);
C = -L\R;
g = L\b;
kmax = 10000; k = 0;

while (norm(b-A*x0)>tol && k<kmax)
    k = k+1;
    x0 = C*x0+g;
end
if (k == kmax)
    disp('Erro: o metodo nao converge. ');
    return;
end
x = x0;
```

Método de Gauss-Seidel

Critério de Sassenfeld

O Método de Gauss-Seidel **converge** para a solução de $\mathbf{Ax} = \mathbf{b}$, independentemente da escolha de $\mathbf{x}^{(0)}$, se satisfazer:

$$\beta = \max_{1 \leq i \leq n} \{\beta_i\} < 1, \quad \text{com}$$

$$\beta_1 = \frac{\sum_{j=2}^n |a_{1j}|}{|a_{11}|} \quad \text{e} \quad \beta_i = \frac{\sum_{j=1}^{i-1} |a_{ij}| \beta_j + \sum_{j=i+1}^n |a_{ij}|}{|a_{ii}|}$$

Obs.: quanto menor o valor de β , mais rápida será a convergência.

Método de Gauss-Seidel

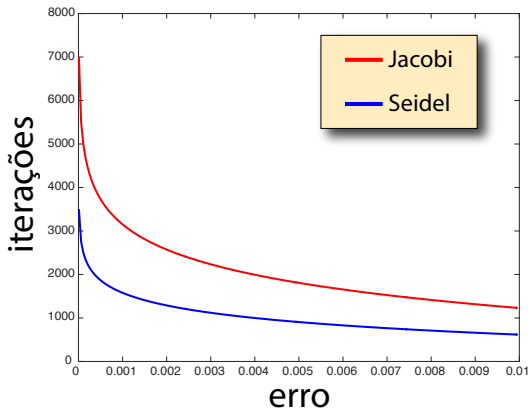
Exercício 3

Considere o sistema linear:

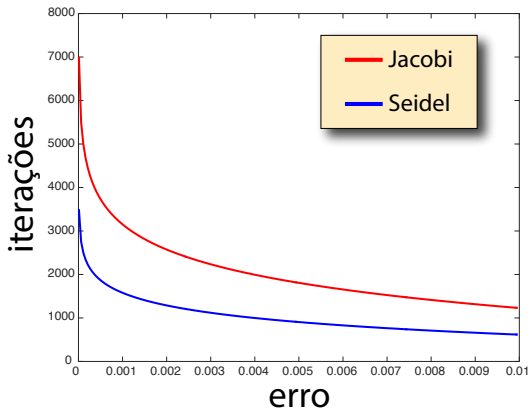
$$\begin{cases} x_1 + 0.5x_2 + 0.1x_3 &= 0.2 \\ 0.3x_1 + 2x_2 + 0.2x_3 &= 0.7 \\ -0.5x_1 + x_2 + 7x_3 &= 1 \end{cases}.$$

- 1 O Método de Gauss-Seidel converge?
- 2 Dado o chute inicial $\mathbf{x}^{(0)} = (0, 0, 0)^\top$, calcule $\mathbf{x}^{(1)}$.

Gauss-Jacobi \times Gauss-Seidel



Gauss-Jacobi × Gauss-Seidel

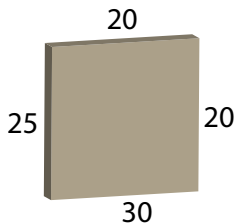


- Método de Gauss-Seidel converge mais **rápido**;
- Método de Gauss-Jacobi é **paralelizável**.

Aplicação

Distribuição de Temperatura

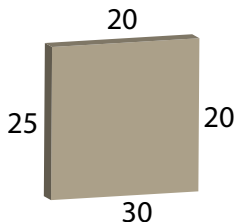
Problema: dada uma placa \mathcal{R} sujeita a 3 temperaturas (em Celsius) distintas na fronteira $\partial\mathcal{R}$, como calcular a temperatura de equilíbrio no interior da placa?



Aplicação

Distribuição de Temperatura

Problema: dada uma placa \mathcal{R} sujeita a 3 temperaturas (em Celsius) distintas na fronteira $\partial\mathcal{R}$, como calcular a temperatura de equilíbrio no interior da placa?

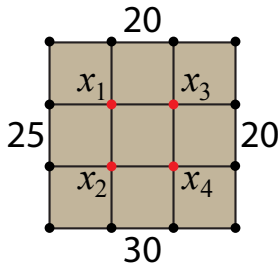


Propriedade do Valor Médio: a temperatura de equilíbrio em um ponto P é o valor médio da temperatura de sua vizinhança.

Aplicação

Distribuição de Temperatura

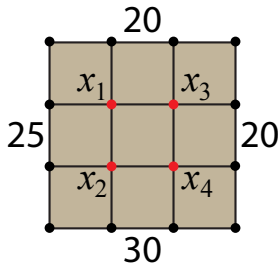
Suponha que \mathcal{R} já alcançou a temperatura de equilíbrio. Vamos discretizar \mathcal{R} por uma grade (*grid*):



Aplicação

Distribuição de Temperatura

Suponha que \mathcal{R} já alcançou a temperatura de equilíbrio. Vamos discretizar \mathcal{R} por uma grade (*grid*):

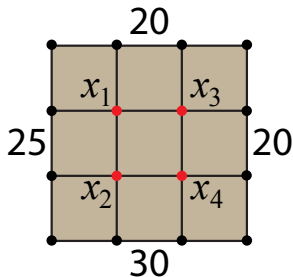


Propriedade do Valor Médio: a temperatura em um ponto $P \notin \partial\mathcal{R}$ é o valor médio da temperatura dos seus 4 pontos mais próximos.

Aplicação

Distribuição de Temperatura

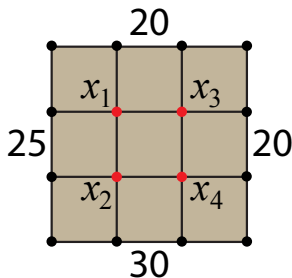
Qual o valor da temperatura em x_1 , x_2 , x_3 e x_4 ?



Aplicação

Distribuição de Temperatura

Qual o valor da temperatura em x_1 , x_2 , x_3 e x_4 ?



$$x_1 = \frac{20 + 25 + x_2 + x_3}{4}$$

$$x_2 = \frac{x_1 + 25 + 30 + x_4}{4}$$

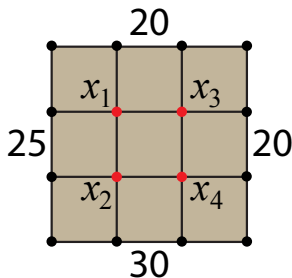
$$x_3 = \frac{20 + x_1 + x_4 + 20}{4}$$

$$x_4 = \frac{x_3 + x_2 + 30 + 20}{4}$$

Aplicação

Distribuição de Temperatura

Qual o valor da temperatura em x_1 , x_2 , x_3 e x_4 ?



$$x_1 = \frac{20 + 25 + x_2 + x_3}{4}$$

$$x_2 = \frac{x_1 + 25 + 30 + x_4}{4}$$

$$x_3 = \frac{20 + x_1 + x_4 + 20}{4}$$

$$x_4 = \frac{x_3 + x_2 + 30 + 20}{4}$$

$$\begin{bmatrix} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 45 \\ 55 \\ 40 \\ 50 \end{bmatrix}$$

Aplicação

Distribuição de Temperatura

Exercício 4

Dada uma placa quadrada de lado $[0, 1] \times [0, 1]$ metros, já com os valores de temperatura prescritos na fronteira. Faça uma função em MATLAB que calcule e visualize a distribuição de temperaturas nesta placa usando o Método de Gauss-Seidel e um grid de resolução $n \times n$.

Formas Quadráticas

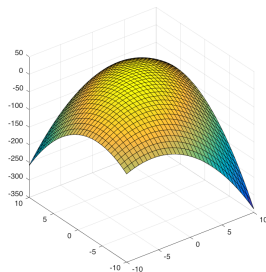
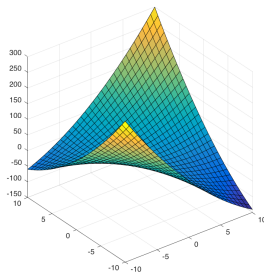
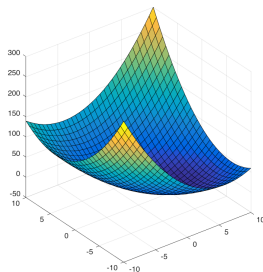
Sejam $\mathbf{A} \in \mathcal{M}(n, n)$, $\mathbf{b} \in \mathbb{R}^n$ e $c \in \mathbb{R}$. Uma **forma quadrática** é uma função $F : \mathbb{R}^n \rightarrow \mathbb{R}$ escrita da seguinte maneira:

$$F(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top \mathbf{A} \mathbf{x} - \mathbf{x}^\top \mathbf{b} + c,$$

Formas Quadráticas

Sejam $\mathbf{A} \in \mathcal{M}(n, n)$, $\mathbf{b} \in \mathbb{R}^n$ e $c \in \mathbb{R}$. Uma **forma quadrática** é uma função $F : \mathbb{R}^n \rightarrow \mathbb{R}$ escrita da seguinte maneira:

$$F(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top \mathbf{A} \mathbf{x} - \mathbf{x}^\top \mathbf{b} + c,$$



Formas Quadráticas

Proposição

Se \mathbf{A} é SPD então $F(\mathbf{x})$ é **minimizada** pela solução de $\mathbf{Ax} = \mathbf{b}$.

Formas Quadráticas

Proposição

Se \mathbf{A} é SPD então $F(\mathbf{x})$ é **minimizada** pela solução de $\mathbf{Ax} = \mathbf{b}$.

Esboço da prova:

- 1 Mostrar que \mathbf{x} é ponto de crítico, isto é,

$$\nabla F(\mathbf{x}) = \mathbf{Ax} - \mathbf{b} = \bar{\mathbf{0}} \Rightarrow \mathbf{Ax} = \mathbf{b}.$$

- 2 Mostrar que \mathbf{x} é ponto de crítico. A matriz **Hessiana**

$$\mathbf{H} = \left[\frac{\partial^2 F(\mathbf{x})}{\partial x_i \partial x_j} \right] = \mathbf{A},$$

como \mathbf{A} é SPD $\Rightarrow \mathbf{x}$ é ponto de mínimo.

Método dos Gradientes

Entrada: $\mathbf{Ax} = \mathbf{b}$, com \mathbf{A} SPD.

Método dos Gradientes

Entrada: $\mathbf{Ax} = \mathbf{b}$, com \mathbf{A} SPD.

Objetivo: construir um processo iterativo (sequência) $\mathbf{x}^{(k)}$ que aproxime \mathbf{x} .

Método dos Gradientes

Entrada: $\mathbf{Ax} = \mathbf{b}$, com \mathbf{A} SPD.

Objetivo: construir um processo iterativo (sequência) $\mathbf{x}^{(k)}$ que aproxime \mathbf{x} .

Ideia: vamos “descer” o parabolóide no sentido contrário de $\nabla F(\mathbf{x})$, isto é,

$$-\nabla F(\mathbf{x}^{(k)}) = \mathbf{b} - \mathbf{Ax}^{(k)} = \underbrace{\mathbf{r}^{(k)}}_{\text{resíduo}}$$

Método dos Gradientes

Entrada: $\mathbf{Ax} = \mathbf{b}$, com \mathbf{A} SPD.

Objetivo: construir um processo iterativo (sequência) $\mathbf{x}^{(k)}$ que aproxime \mathbf{x} .

Ideia: vamos “descer” o parabolóide no sentido contrário de $\nabla F(\mathbf{x})$, isto é,

$$-\nabla F(\mathbf{x}^{(k)}) = \mathbf{b} - \mathbf{Ax}^{(k)} = \underbrace{\mathbf{r}^{(k)}}_{\text{resíduo}}$$

Dado um chute inicial $\mathbf{x}^{(0)}$, obtemos $\mathbf{x}^{(1)}$ da seguinte forma

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - \alpha \nabla F(\mathbf{x}^{(0)}) = \underbrace{\mathbf{x}^{(0)} + \alpha \mathbf{r}^{(0)}}_{\text{reta}}$$

Método dos Gradientes

Entrada: $\mathbf{Ax} = \mathbf{b}$, com \mathbf{A} SPD.

Objetivo: construir um processo iterativo (sequência) $\mathbf{x}^{(k)}$ que aproxime \mathbf{x} .

Ideia: vamos “descer” o parabolóide no sentido contrário de $\nabla F(\mathbf{x})$, isto é,

$$-\nabla F(\mathbf{x}^{(k)}) = \mathbf{b} - \mathbf{Ax}^{(k)} = \underbrace{\mathbf{r}^{(k)}}_{\text{resíduo}}$$

Dado um chute inicial $\mathbf{x}^{(0)}$, obtemos $\mathbf{x}^{(1)}$ da seguinte forma

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - \alpha \nabla F(\mathbf{x}^{(0)}) = \underbrace{\mathbf{x}^{(0)} + \alpha \mathbf{r}^{(0)}}_{\text{reta}}$$

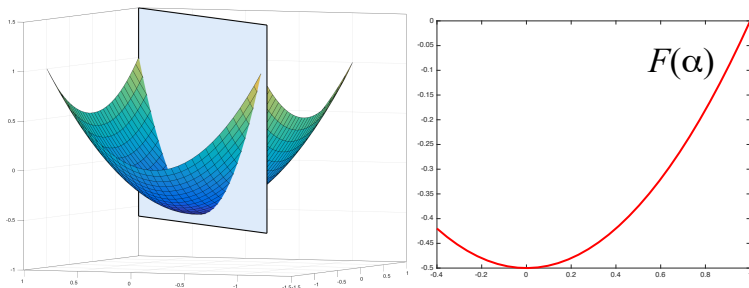
Vamos “caminhar” nessa reta, mas qual o tamanho do passo α ???

Método dos Gradientes

Resposta: o valor de α tem que minimizar $F(\mathbf{x})$ ao longo da reta $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha \mathbf{r}^{(0)}$, ou seja, **minimizar** $F(\mathbf{x}^{(1)})$. Logo,

Método dos Gradientes

Resposta: o valor de α tem que minimizar $F(\mathbf{x})$ ao longo da reta $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha \mathbf{r}^{(0)}$, ou seja, **minimizar** $F(\mathbf{x}^{(1)})$. Logo,



$$\frac{\partial F}{\partial \alpha}(\mathbf{x}^{(1)}) \underbrace{=}_{\text{r. cadeia}} \nabla F(\mathbf{x}^{(1)}) \cdot \frac{\partial \mathbf{x}^{(1)}}{\partial \alpha} = \underbrace{\nabla F(\mathbf{x}^{(1)})}_{-\mathbf{r}^{(1)}} \cdot \mathbf{r}^{(0)} = 0$$

Método dos Gradientes

Portanto, $\mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} = 0$ (ortogonais). Segue que,

Método dos Gradientes

Portanto, $\mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} = 0$ (ortogonais). Segue que,

$$0 = \mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} = \left[\mathbf{b} - \mathbf{A}\mathbf{x}^{(1)} \right] \cdot \mathbf{r}^{(0)}$$

Método dos Gradientes

Portanto, $\mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} = 0$ (ortogonais). Segue que,

$$\begin{aligned} 0 = \mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} &= \left[\mathbf{b} - \mathbf{A}\mathbf{x}^{(1)} \right] \cdot \mathbf{r}^{(0)} \\ &= \left[\mathbf{b} - \mathbf{A}(\mathbf{x}^{(0)} + \alpha \mathbf{r}^{(0)}) \right] \cdot \mathbf{r}^{(0)} \end{aligned}$$

Método dos Gradientes

Portanto, $\mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} = 0$ (ortogonais). Segue que,

$$\begin{aligned} 0 = \mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} &= \left[\mathbf{b} - \mathbf{A}\mathbf{x}^{(1)} \right] \cdot \mathbf{r}^{(0)} \\ &= \left[\mathbf{b} - \mathbf{A}(\mathbf{x}^{(0)} + \alpha \mathbf{r}^{(0)}) \right] \cdot \mathbf{r}^{(0)} \\ &= \left[(\mathbf{b} - \mathbf{A}\mathbf{x}^{(0)}) - \alpha \mathbf{A}\mathbf{r}^{(0)} \right] \cdot \mathbf{r}^{(0)} \end{aligned}$$

Método dos Gradientes

Portanto, $\mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} = 0$ (ortogonais). Segue que,

$$\begin{aligned} 0 = \mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} &= [\mathbf{b} - \mathbf{A}\mathbf{x}^{(1)}] \cdot \mathbf{r}^{(0)} \\ &= [\mathbf{b} - \mathbf{A}(\mathbf{x}^{(0)} + \alpha\mathbf{r}^{(0)})] \cdot \mathbf{r}^{(0)} \\ &= [(\mathbf{b} - \mathbf{A}\mathbf{x}^{(0)}) - \alpha\mathbf{A}\mathbf{r}^{(0)}] \cdot \mathbf{r}^{(0)} \\ &= [\mathbf{r}^{(0)} - \alpha\mathbf{A}\mathbf{r}^{(0)}] \cdot \mathbf{r}^{(0)} \end{aligned}$$

Método dos Gradientes

Portanto, $\mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} = 0$ (ortogonais). Segue que,

$$\begin{aligned} 0 = \mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} &= [\mathbf{b} - \mathbf{A}\mathbf{x}^{(1)}] \cdot \mathbf{r}^{(0)} \\ &= [\mathbf{b} - \mathbf{A}(\mathbf{x}^{(0)} + \alpha\mathbf{r}^{(0)})] \cdot \mathbf{r}^{(0)} \\ &= [(\mathbf{b} - \mathbf{A}\mathbf{x}^{(0)}) - \alpha\mathbf{A}\mathbf{r}^{(0)}] \cdot \mathbf{r}^{(0)} \\ &= [\mathbf{r}^{(0)} - \alpha\mathbf{A}\mathbf{r}^{(0)}] \cdot \mathbf{r}^{(0)} \\ &= \mathbf{r}^{(0)} \cdot \mathbf{r}^{(0)} - \alpha(\mathbf{r}^{(0)} \cdot \mathbf{A}\mathbf{r}^{(0)}) \end{aligned}$$

Método dos Gradientes

Portanto, $\mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} = 0$ (ortogonais). Segue que,

$$\begin{aligned} 0 = \mathbf{r}^{(1)} \cdot \mathbf{r}^{(0)} &= [\mathbf{b} - \mathbf{A}\mathbf{x}^{(1)}] \cdot \mathbf{r}^{(0)} \\ &= [\mathbf{b} - \mathbf{A}(\mathbf{x}^{(0)} + \alpha\mathbf{r}^{(0)})] \cdot \mathbf{r}^{(0)} \\ &= [(\mathbf{b} - \mathbf{A}\mathbf{x}^{(0)}) - \alpha\mathbf{A}\mathbf{r}^{(0)}] \cdot \mathbf{r}^{(0)} \\ &= [\mathbf{r}^{(0)} - \alpha\mathbf{A}\mathbf{r}^{(0)}] \cdot \mathbf{r}^{(0)} \\ &= \mathbf{r}^{(0)} \cdot \mathbf{r}^{(0)} - \alpha(\mathbf{r}^{(0)} \cdot \mathbf{A}\mathbf{r}^{(0)}) \end{aligned}$$

$$\alpha = \frac{\mathbf{r}^{(0)} \cdot \mathbf{r}^{(0)}}{\mathbf{r}^{(0)} \cdot \mathbf{A}\mathbf{r}^{(0)}}$$

Método dos Gradientes

O processo iterativo é definido como:

$$\mathbf{1} \quad \mathbf{r}^{(k)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)};$$

$$\mathbf{2} \quad \alpha^{(k)} = \frac{\mathbf{r}^{(k)} \cdot \mathbf{r}^{(k)}}{\mathbf{r}^{(k)} \cdot \mathbf{A}\mathbf{r}^{(k)}};$$

$$\mathbf{3} \quad \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{r}^{(k)}.$$

Método dos Gradientes

O processo iterativo é definido como:

- 1 $\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)};$
- 2 $\alpha^{(k)} = \frac{\mathbf{r}^{(k)} \cdot \mathbf{r}^{(k)}}{\mathbf{r}^{(k)} \cdot \mathbf{A}\mathbf{r}^{(k)}};$
- 3 $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{r}^{(k)}.$

Exercício 5

Utilize o Método dos Gradiente com $\mathbf{x}^{(0)} = (-2, 2)^\top$ para calcular uma aproximação da solução do sistema abaixo:

$$\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ -8 \end{bmatrix}.$$

Note que a solução exata é $\mathbf{x} = (2, -2)^\top$.

MATLAB – Método dos Gradientes

```
function [x,k] = gradientes(A,b,x0,tol)
% A: matriz SPD

kmax = 1000;
for k=1:kmax
    r = b - A*x0;
    if norm(r)<tol
        x = x0;
        k = k-1;
        return;
    end
    alpha = (r'*r)/(r'*A*r);
    x0 = x0 + alpha*r;
end

disp('Erro: o metodo nao converge.');
```