

# SME0305 - 2016

Roberto F. Ausas / Gustavo C. Buscaglia

ICMC - Ramal 736628, rfausas@gmail.com

ICMC - Ramal 738176, gustavo.buscaglia@gmail.com

---

## Métodos iterativos

Vamos considerar o problema:

Achar  $\mathbf{x}^* \in \mathbb{R}^n$  tal que

$$\mathbf{r}(\mathbf{x}^*) = 0$$

em que a função vetorial  $\mathbf{r} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  e pode ser linear ou não linear em  $\mathbf{x}$ .

Que é um método iterativo para resolver esse problema?

Um método iterativo para achar a solução  $\mathbf{x}^*$  do problema, é um método que, a partir de um vetor inicial  $\mathbf{x}^{(0)}$ , gera uma sequência de vetores  $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(k)}, \dots$ , que aproxima-se a  $\mathbf{x}^*$

Que significa se aproximar a  $\mathbf{x}^*$ ?

Matematicamente, o limite da sequência quando  $k$  tende para  $\infty$  tem que ser  $\mathbf{x}^*$ , i.e.

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}^*$$

Então, o método iterativo se diz convergente. De maneira equivalente, se o método for convergente

$$\| \mathbf{x}^{(k)} - \mathbf{x}^* \| \xrightarrow[k \rightarrow \infty]{} 0$$

em que  $\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}^*$  é o erro na iteração  $k$ .

Também podemos trabalhar com o residuo do problema na iteração  $k$ , que seria  $\mathbf{r}^{(k)} = \mathbf{r}(\mathbf{x}^{(k)})$ . O método se dirá convergente se

$$\| \mathbf{r}^{(k)} \| \xrightarrow[k \rightarrow \infty]{} 0$$

em que  $\| \cdot \|$  é alguma norma de  $\mathbb{R}^n$ .

Aqui vemos que precisamos trabalhar com normas de vetores, para medir que tão próximo nos encontramos da solução. Vamos aproveitar para definir algumas normas.

Uma norma é uma função  $\| \cdot \|: \mathbb{R}^n \rightarrow \mathbb{R}^+$  tal que

1.  $\| \mathbf{x} \| \geq 0 \forall \mathbf{x} \in \mathbb{R}^n$  e  $\| \mathbf{x} \| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ .
2.  $\| \alpha \mathbf{x} \| = |\alpha| \| \mathbf{x} \| \forall \alpha \in \mathbb{R}$ .
3.  $\| \mathbf{x} + \mathbf{y} \| \leq \| \mathbf{x} \| + \| \mathbf{y} \| \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ .

As mais usadas são:

- Norma Euclidiana:

$$\| \mathbf{x} \|_2 = \left( \sum_{i=1}^n x_i^2 \right)^{1/2}$$

- Norma do máximo:

$$\| \mathbf{x} \|_\infty = \max_{1 \leq i \leq n} |x_i|$$

- Norma 1:

$$\| \mathbf{x} \|_1 = \sum_{i=1}^n |x_i|$$

Como estamos em  $\mathbb{R}^n$  as normas são equivalentes, isto é:

$$\| \mathbf{x} \|_\infty \leq \| \mathbf{x} \|_2 \leq \sqrt{n} \| \mathbf{x} \|_\infty,$$

$$\| \mathbf{x} \|_\infty \leq \| \mathbf{x} \|_1 \leq n \| \mathbf{x} \|_\infty,$$

$$\| \mathbf{x} \|_2 \leq \| \mathbf{x} \|_1 \leq \sqrt{n} \| \mathbf{x} \|_2,$$

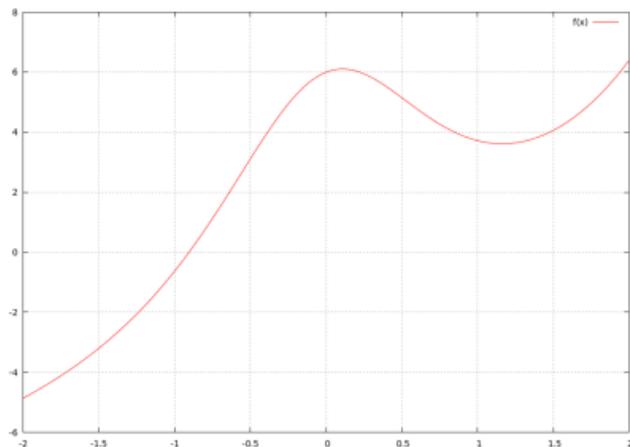
---

Vejamos um exemplo de processo iterativo para resolver um problema concreto.

Vamos calcular em que ponto se anula a função:

$$r(x) = x + e^x + \frac{10}{1+x^2} - 5$$

i.e., procuramos  $x^*$  tal que  $r(x^*) = 0$



Para resolver este problema, podemos utilizar a fórmula de iteração:

$$x^{(k+1)} = x^{(k)} - \frac{r(x^{(k)})}{r'(x^{(k)})}$$

Este é o famoso método de Newton. O código em Octave resulta:

```
func = inline ("x + exp(x) + 10/(1+x^2) - 5");  
derfunc = inline ("1 + exp(x) - 20*x/(1+x^2)^2");  
  
xold = 1.5;  
for k=1:100  
    xnew = xold - func(xold) / derfunc(xold);  
    xold = xnew  
    if(abs(func(xnew)) < 1.0e-10)  
        disp('Converged'); break ;  
    end  
end  
  
if(k == 100) disp('PANIC: Not Converged'); end
```

Agora vamos ver um problema mais difícil. Vamos considerar redes hidráulicas não lineares.

Vamos supor que a conductância dos canos é uma função da própria diferença de pressão nos extremos do cano, i.e., para o cano  $r$  unindo os nós  $i$  e  $j$

$$C_r = a_r + b_r |P_i - P_j|$$

Em Octave:

$$C(r) = a(r) + b(r) * \text{abs}(P(\text{conec}(r,1)) - P(\text{conec}(r,2)));$$

Notar que para conseguir avaliar as conductâncias, precisamos dos valores da pressão nos nós, que é justamente o que estamos procurando!

Neste caso, temos o problema de achar  $\mathbf{P}^*$  tal que:

$$\mathbf{r}(\mathbf{P}^*) = \tilde{\mathbf{A}}(\mathbf{P}^*) \mathbf{P}^* - \mathbf{b} = 0$$

Podemos propor um processo iterativo em que o vetor de pressões  $\mathbf{P}$  se atualizará a partir de um chute inicial.

Dado  $\mathbf{P}^{(k)}$ , na iteração  $k + 1$  do processo iterativo fazemos:

- Calcular as conductâncias usando  $\mathbf{P}^{(k)}$ ;
- Montar a matriz do sistema com essas conductâncias  $\tilde{\mathbf{A}}(\mathbf{P}^{(k)})$ ;
- Resolver o sistema para atualizar a pressão e calcular  $\mathbf{P}^{(k+1)}$ ;

Isto se conhece como método de Picard. A implementação em Octave é:

```

% Definicao do problema
...
[nv, nc, conec, C0, coords] = RedeHidraBairro(n,m,CH,CV);
...
MAX_IT = 20; TOLr = 1.0e-10;
Pold = zeros(nv,1); % Chute inicial
for k=1:MAX_IT % Loop de iteracoes
    C = ComputeConds(C0, nc, conec, Pold);
    [Atilde b] = BuildSystem(nv, nc, conec, C, QB, nB, ...
                            patm, natm);

    residual = Atilde*Pold - b;
    normres = norm(residual)
    if(normres < TOLr)
        disp('Converged');
        break ;
    end
    Pnew = Atilde \ b;
    Pold = Pnew;
end

```

Nesses exemplos já vimos vários dos ingredientes que aparecem num algoritmo iterativo:

- Um valor inicial ou um “chute”;
- Uma “receita” para calcular o novo elemento da sequência  $\mathbf{x}^{(k+1)}$  a partir do elemento anterior  $\mathbf{x}^{(k)}$ ;
- Um número máximo de iterações: *MAX\_IT*;
- Tolerâncias algorítmicas: *TOL<sub>r</sub>*. Também poderíamos ter utilizado uma tolerância de valor.

Se o processo for convergente, quanto mais iterações fizermos, mais preciso será o resultado com relação a  $\mathbf{x}^*$ . Então, o processo converge, a menos do erro de arredondamento, nesse valor exato. Para isto, precisamos escolher *MAX\_IT* e *TOL* apropriadamente.

A escolha do “chute” inicial também pode ser importante.

Para introduzir algumas ideias sobre como construir métodos iterativos, vamos propor um algoritmo geral da forma:

Dados  $\mathbf{x}^{(0)}$ ,  $TOL$ ,  $MAX\_IT$ ,  $k = 0$

Enquanto  $\|\mathbf{r}^{(k)}\| > TOL$  e  $k < MAX\_IT$

1. Achar uma direção de procura  $\mathbf{d}^{(k+1)}$
2. Determinar o escalar  $\beta_k$
3. Avançar:  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \beta_{k+1} \mathbf{d}^{(k+1)}$
4. Calcular o residual:  $\mathbf{r}^{(k+1)}$
5. Incrementar  $k$

Fim

e vamos trabalhar com problemas lineares. Então, dado o sistema

$$A \mathbf{x} = \mathbf{b}, \quad A \in \mathbb{R}^{n \times n}, \quad \mathbf{x}, \mathbf{b} \in \mathbb{R}^n$$

Neste caso, o residuo vem dado por

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

---

O ideal para um método iterativo seria ele converger na solução do sistema em apenas uma iteração. Isto significa pedir que:

$$\mathbf{x}^{(1)} = \mathbf{x}^*$$

Já que  $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \beta_1 \mathbf{d}^{(1)}$ , se tomarmos  $\beta_1 = 1$  e  $\mathbf{d}^{(1)} = \mathbf{x}^* - \mathbf{x}^{(0)}$ .

Claramente, para calcular  $\mathbf{d}^{(1)}$  precisamos conhecer a solução do sistema. Já que

$$\mathbf{d}^{(1)} = \mathbf{x}^* - \mathbf{x}^{(0)} = A^{-1} \mathbf{b} - A^{-1} (\mathbf{r}^{(0)} + \mathbf{b}) = -A^{-1} \mathbf{r}^{(0)}$$

isto significa que  $\mathbf{d}^{(1)}$  é solução de:

$$A \mathbf{d}^{(1)} = -\mathbf{r}^{(0)}$$

que envolve a matriz  $A$ .

Vamos supor que por algum motivo não queremos ou não podemos resolver o sistema envolvendo a matriz  $A$ , mas, que temos alguma forma **eficiente** de resolver o sistema

$$M \mathbf{d}^{(1)} = -\mathbf{r}^{(0)}$$

em que a matriz  $M$  é “parecida” com a matriz  $A$ . Nesse caso, ao resolver, acharemos um vetor  $\mathbf{d}^{(1)}$  tal que  $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \beta_1 \mathbf{d}^{(1)} \neq \mathbf{x}^{(*)}$ .

Se  $M$  for escolhida muito parecida com  $A$ ,  $\mathbf{x}^{(1)}$  resultará muito próximo de  $\mathbf{x}^*$ . Então, repetindo o processo iremos ficando cada vez mais e mais pertos de  $\mathbf{x}^*$ .

Como escolhemos a matriz  $M$ ?

Na iteração número  $k + 1$  teremos

$$\mathbf{r}^{(k+1)} = A\mathbf{x}^{(k+1)} - \mathbf{b} = A(\mathbf{x}^{(k)} + \beta_{k+1}\mathbf{d}^{(k+1)}) - \mathbf{b} = \mathbf{r}^{(k)} + A\beta_{k+1}(-M^{-1}\mathbf{r}^{(k)})$$

$$\mathbf{r}^{(k+1)} = \underbrace{(\mathbb{I} - \beta_{k+1}AM^{-1})}_S \mathbf{r}^{(k)}$$

Então, supondo que a matriz  $M$  e o parâmetro  $\beta$  estão fixos ao longo das iterações

$$\mathbf{r}^{(k+1)} = S\mathbf{r}^{(k)} = \dots = S^{k+1}\mathbf{r}^{(0)}$$

mas, como garantir que  $\|\mathbf{r}^{(k+1)}\|$  tende para zero quando  $k$  aumenta?

Tomando a norma

$$\|\mathbf{r}^{(k+1)}\| = \|S^{k+1}\mathbf{r}^{(0)}\|$$

O que precisamos, é que para todo vetor  $\mathbf{r}^{(0)}$ ,  $\|S^{k+1}\mathbf{r}^{(0)}\| \rightarrow 0$

Chamando de  $\lambda_i$ 's aos autovalores de  $S$ , pode-se provar que  $\| S^{k+1} \mathbf{r}^{(0)} \| \rightarrow 0$ , se o chamado raio espectral da matriz  $S$ , definido como:

$$\rho(S) = \max_{1 \leq i \leq n} \{|\lambda_1|, \dots, |\lambda_n|\}$$

é menor do que 1.

---

**Lembrete:** Autovalores e autovetores:

Seja  $A \in \mathbb{R}^{n \times n}$ ,  $\lambda \in \mathbb{R}$  é autovalor de  $A$  se  $A - \lambda \mathbb{I}$  é singular.

Portanto, se  $\lambda$  é autovalor de  $A$ , existe  $\mathbf{v} \in \mathbb{R}^n$ ,  $\mathbf{v} \neq \mathbf{0}$  tal que

$$(A - \lambda \mathbb{I}) \mathbf{v} = \mathbf{0}$$

ou seja:

$$A \mathbf{v} = \lambda \mathbf{v}$$

Uma forma de calcular os autovalores é através do polinômio característico: Se  $\lambda$  é autovalor da matriz  $A \in \mathbb{R}^{n \times n}$ , então:

$$P(\lambda) = \det(A - \lambda I) = 0$$

**Definição:** Uma matriz diz-se **diagonalizável** se ela for semelhante a uma matriz diagonal, ou seja, existe  $Q$  tal que

$$A = Q D Q^{-1}$$

em que  $D$  é uma matriz diagonal, i.e.,

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

Se  $A$  é simétrica, podemos achar uma  $Q$  ortogonal ( $Q^{-1} = Q^T$ ).

---

Agora voltemos ao método iterativo, para o qual temos o teorema:

**Teorema:** Para um método iterativo como o proposto, o método é convergente se e só se  $\rho(S) < 1$ .

Para entender isto, vejamos que acontece quando calculamos potências da matriz  $S$ . Para simplificar vamos supor que ela é diagonalizável. Então

$$S^2 = (Q D Q^{-1})(Q D Q^{-1}) = QD(Q^{-1} Q)D Q^{-1} = QD^2 Q^{-1}$$

similarmente

$$S^3 = S^2 S = (Q D^2 Q^{-1})(Q D Q^{-1}) = QD^2(Q^{-1} Q)D Q^{-1} = QD^3 Q^{-1}$$

e em geral  $S^k = QD^kQ^{-1}$ . Mas, a matriz  $D^k$  é:

$$D^k = \begin{bmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \lambda_n^k \end{bmatrix}$$

Se todos os  $\lambda_i$ ,  $i = 1, \dots, n$ , são menores do que 1, então  $\lambda_i^k \xrightarrow[k \rightarrow \infty]{} 0$   
e portanto  $S^k \xrightarrow[k \rightarrow \infty]{} 0$ .

Pode-se ver que estes processos iterativos podem ser escritos como:

$$\mathbf{x}^{(k+1)} = B \mathbf{x}^{(k)} + \mathbf{g}$$

Para uma certa matriz  $B$  e vetor  $\mathbf{g}$  e que

$$\mathbf{e}^{(k+1)} = B \mathbf{e}^{(k)}$$

em que,  $\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}^*$ . Fica como exercício para a lista verificar-lo.

Voltando ao algoritmo iterativo geral:

Dados  $\mathbf{x}^{(0)}$ ,  $TOL$ ,  $MAX\_IT$ ,  $k = 0$

Enquanto  $\|\mathbf{r}^{(k)}\| > TOL$  e  $k < MAX\_IT$

1. Resolver  $M \mathbf{d}^{(k+1)} = -\mathbf{r}^{(k)}$
2. Determinar o escalar  $\beta_{k+1}$
3. Avançar:  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \beta_{k+1} \mathbf{d}^{(k+1)}$
4. Calcular residual:  $\mathbf{r}^{(k+1)} = A \mathbf{x}^{(k+1)} - \mathbf{b}$
5. Incrementar  $k$

Fim

Diferentes métodos correspondem a diferentes escolhas para  $M$  e

$\beta_{k+1}$ .

Vamos estudar dois métodos básicos:

- Método de Jacobi;
- Método de Gauss-Seidel;

Primeiro, a matriz do sistema é decomposta aditivamente, i.e.:

$$A = L_A + D_A + R_A$$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \dots & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ a_{21} & 0 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ a_{41} & a_{42} & \dots & 0 \end{pmatrix} + \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} + \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ 0 & 0 & \dots & a_{2n} \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

O método de Jacobi consiste em tomar  $\beta_k = 1 \forall k$  e fazer a escolha mais simples para a matriz no algoritmo iterativo:

$$M = D_A$$

Podemos simplesmente utilizar o processo iterativo com essa escolha para  $M$  e  $\beta$ , mas, é interessante reescrever o método. Então,

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{d}^{(k+1)} \Rightarrow \mathbf{d}^{(k+1)} = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$$

sendo  $\mathbf{d}^{(k+1)}$  solução do sistema

$$D_A \mathbf{d}^{(k+1)} = -\mathbf{r}^{(k)} = -A \mathbf{x}^{(k)} + \mathbf{b}$$

Mas, já que  $A = L_A + D_A + R_A$ , resulta

$$D_A (\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = -(L_A + D_A + R_A) \mathbf{x}^{(k)} + \mathbf{b}$$

$$D_A \mathbf{x}^{(k+1)} = -(L_A + R_A) \mathbf{x}^{(k)} + \mathbf{b}$$

ou em notação de índices

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

Lembremos que o método é convergente se e só se  $\rho(S) < 1$ . Em geral, isto não é trivial de ser verificado para uma matriz arbitrária de  $n \times n$ , porém, pode-se demonstrar que:

**Teorema:** Se a matriz  $A$  é de diagonal estritamente dominante por linha, então, o método de Jacobi converge.

O outro método que vamos considerar consiste em tomar  $\beta_k = 1 \forall k$  e para matriz  $M$

$$M = D_A + L_A$$

Novamente, vamos reescrever o método para ver o que há por traz dessa escolha:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{d}^{(k+1)} \Rightarrow \mathbf{d}^{(k+1)} = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$$

sendo  $\mathbf{d}^{(k+1)}$  solução do sistema

$$(D_A + L_A) \mathbf{d}^{(k+1)} = -\mathbf{r}^{(k)} = -A \mathbf{x}^{(k)} + \mathbf{b}$$

Mas, já que  $A = L_A + D_A + R_A$ , resulta

$$(D_A + L_A) (\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = -(L_A + D_A + R_A) \mathbf{x}^{(k)} + \mathbf{b}$$

$$D_A \mathbf{x}^{(k+1)} + L_A \mathbf{x}^{(k+1)} = -R_A \mathbf{x}^{(k)} + \mathbf{b}$$

ou em notação de índices

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

Condições suficientes para que o método seja convergente:

**Teorema:** Se a matriz  $A$  é de diagonal estritamente dominante por linha ou ela é simétrica e definida positiva, então, o método de Gauss-Seidel converge.

Outras variantes parecidas são:

- Método SOR:

$$M = D_A + \omega L_A$$

- Método SSOR:

$$M = (D_A + \omega L_A) D_A^{-1} (D_A + \omega R_A)$$

em que  $\omega > 1$  é um parâmetro de sobrerelaxação.