# Cálculo Numérico

## Notas de aula - Processos Iterativos Matriciais Lineares

Prof. Yuri Dumaresq Sobral

Departamento de Matemática
Universidade de Brasília

2025

- Já aprendemos a resolver sistemas de equações lineares usando a técnica da fatoração LU.
- Esta técnica, porém, requer manipulações não triviais de matrizes, que podem ser computacionalmente caras para sistemas muito grandes : a eliminação gaussiana requer $\mathcal{O}(N^3)$ operações! CARO!
- Vamos construir, agora, métodos que tenham como base processos iterativos!
- Esperamos que estes métodos sejam mais fáceis de programar, e mais eficientes quando aplicados a sistemas muito grandes.
- Antes, porém, precisamos estudar com um pouco mais de detalhe os processos iterativos construídos com matrizes! Vamos nos concentrar inicialmente apenas nos processos lineares. (Lembrete: eles governam o erro de processos iterativos nas vizinhanças de seus pontos fixos!)

## Relembrando:

- Considere uma função $g : IR \to IR$ tal que possamos construir, com base nela, o seguinte processo iterativo escalar:

$$x_{n+1} = g(x_n), \qquad n = 0, 1, 2, 3, \ldots$$

- O processo iterativo escalar linear mais geral que podemos construir é tomando $g(x) = ax + b$, isto é:

$$x_{n+1} = ax_n + b.$$

- Estudamos detalhadamente a estabilidade do seu ponto fixo

$$x^* = \frac{b}{1-a}.$$

- Se $|a| < 1$, então temos que $x^*$ é ponto fixo assintoticamente estável do processo iterativo linear.
- Se $|a| > 1$, então temos que $x^*$ é ponto fixo instável do processo iterativo linear.
- Se $a = 0$, então temos que $x^* = b$ é ponto fixo estável (trivial) do processo iterativo linear.

- Considere, agora, uma matriz $T$ quadrada de ordem $m$, isto é, de tamanho $m \times m$. Considere os vetor $\mathbf{x}, \mathbf{b} \in I\!R^m$.
- O processo iterativo matricial linear mais geral que podemos construir é

$$\mathbf{x}_{n+1} = T\mathbf{x}_n + \mathbf{b}$$

- Vamos considerar o caso mais simples com $\mathbf{b} = \mathbf{0}$, isto é, $\mathbf{x}_{n+1} = T\mathbf{x}_n$, e vamos considerar $\mathbf{x}_0$ como valor inicial do processo. Neste caso, o processo iterativo se dá da seguinte maneira:

$$\mathbf{x}_1 = T\mathbf{x}_0, \qquad \mathbf{x}_2 = T\mathbf{x}_1 = T(T\mathbf{x}_0) = T^2\mathbf{x}_0$$

$$\mathbf{x}_3 = T\mathbf{x}_2 = T^3\mathbf{x}_0, \qquad \dots \qquad \mathbf{x}_n = T\mathbf{x}_{n-1} = T^n\mathbf{x}_0.$$

- Portanto, o processo iterativo é determinado por potências da matriz $T$! Muito parecido ao processo iterativo escalar linear $x_{n+1} = ax_n$, cuja solução é $x_n = a^n x_0$, com $a \in I\!R$.

- Apesar da semelhança com o caso escalar, não é possível encontrar um padrão óbvio em potências de matrizes. Vamos ver isto num exemplo.
- Exemplo: Considere $\mathbf{x}_{n+1} = T\mathbf{x}_n$ com $m = 2$ e

$$T = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix}, \qquad \mathbf{x}_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}, \qquad \mathbf{x}_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Então, o processo iterativo é dado por:

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} \Leftrightarrow \begin{cases} x_{n+1} = 0.6x_n + 0.2y_n \\ y_{n+1} = 0.2x_n + 0.6y_n \end{cases}$$

Assim, o processo será dado por:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix} \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} =$$

$$\cdots = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix}^2 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 0.4 & 0.24 \\ 0.24 & 0.4 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

$$\begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix}^3 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} =$$

$$\cdots = \begin{pmatrix} 0.288 & 0.224 \\ 0.224 & 0.288 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

$$\vdots$$

$$\begin{pmatrix} x_{10} \\ y_{10} \end{pmatrix} = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix}^{10} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 0.05373\ldots & 0.05363\ldots \\ 0.05363\ldots & 0.05373\ldots \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

$$\vdots$$

$$\begin{pmatrix} x_{100} \\ y_{100} \end{pmatrix} = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix}^{100} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \approx \begin{pmatrix} 1.02 \cdot 10^{-10} & 1.02 \cdot 10^{-10} \\ 1.02 \cdot 10^{-10} & 1.02 \cdot 10^{-10} \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

$$\vdots$$

Conclusão: Parece que $\mathbf{x}_n \to \mathbf{0}$ com $n \to \infty$. E parece que $T^n \to \mathbb{O}$ (matriz nula, isto é, $\{T^n\}_{ij} \to 0$), com $n \to \infty$.

- Vemos, portanto, que é muito difícil determinar o comportamento do sistema a partir das potências das matrizes.
- Será que não seria possível encontrar uma estrutura similar àquela dos processos iterativos escalares lineares?
- A idéia, portanto, seria tentar escrevermos o resultado do processo iterativo matricial linear como

$$\mathbf{x}_n = \lambda^n \mathbf{v}, \qquad \mathbf{v} \neq \mathbf{0},$$

  com algum $\lambda \in I\!R$ e algum vetor $\mathbf{v} \in I\!R^m$. Mas será que isto funciona?
- Vamos testar:

$$\mathbf{x}_{n+1} = T\mathbf{x}_n \quad \Leftrightarrow \quad \lambda^{n+1}\mathbf{v} = T\lambda^n\mathbf{v} \quad \Leftrightarrow \quad \lambda\mathbf{v} = T\mathbf{v}$$

- Portanto, se encontrarmos $\lambda$, $\mathbf{v}$ que satisfaçam $T\mathbf{v} = \lambda\mathbf{v}$, a solução do processo iterativo será $\mathbf{x}_n = \lambda^n\mathbf{v}$ e conseguiremos determinar o que acontece com $n \to \infty$ de maneira mais trivial.

- Os números $\lambda$ que satisfazem $T\mathbf{v} = \lambda\mathbf{v}$ são chamados de autovalores da matriz $T$.
- Os vetores $\mathbf{v}$ que que satisfazem $T\mathbf{v} = \lambda\mathbf{v}$ são chamados de autovetores da matriz $T$.
- Os autovalores são encontrados a partir do polinômio característico de $T$:

$$T\mathbf{v} = \lambda\mathbf{v} \;\Leftrightarrow\; T\mathbf{v} - \lambda\mathbf{v} = \mathbf{0} \;\Leftrightarrow\; (T - \lambda I)\,\mathbf{v} = \mathbf{0}$$

$$\text{Para } \mathbf{v} \neq \mathbf{0}, \quad \text{impomos} \quad \det\left(T - \lambda I\right) = 0.$$

- Como o polinômio característico tem grau $m$ e todos seus coeficientes são $I\!R$, uma matriz $T$ tem $m$ autovalores complexos, com alguns podendo ser valores repetidos (multiplicidade!).
- Então, para cada par $\lambda$, $\mathbf{v}$, o processo iterativo terá seu comportamento determinado por $\mathbf{x}_n = \lambda^n\mathbf{v}$. Mas e o caso geral? Vamos apenas considerar um caso geral mais simples:

- Se a matriz $T$ for tal que seus $m$ autovetores $\mathbf{v}_1 \ldots \mathbf{v}_m$ forem base de $I\!R^m$, isto é, se eles forem linearmente independentes e se qualquer vetor de $I\!R^m$ puder ser escrito como uma combinação linear deles, então

$$\mathbf{x}_n = c_1\lambda_1^n\mathbf{v}_1 + c_2\lambda_2^n\mathbf{v}_2 + \cdots + c_m\lambda_m^n\mathbf{v}_m.$$

  e as constantes $c_1 \ldots c_m \in I\!R$ serão unicamente determinadas pela valor de $\mathbf{x}_0$.
- Vamos voltar ao exemplo anterior e reescrevê-lo de acordo com o que aprendemos agora:

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix}.$$

  Então, para determinarmos os autovalores da matriz $T$:

$$\det\left(\begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right) = 0 \Leftrightarrow \det\begin{pmatrix} 0.6-\lambda & 0.2 \\ 0.2 & 0.6-\lambda \end{pmatrix} = 0$$

$$\Leftrightarrow (0.6-\lambda)^2-0.04=0 \;\Leftrightarrow\; \lambda^2-1.2\lambda-0.32=0 \;\Leftrightarrow \lambda_1=0.8,\; \lambda_2=0.4$$

- O autovetor associado ao autovalor $\lambda_1 = 0.8$ será dado por:

$$(T-\lambda_1 I)\,\mathbf{v}_1=\mathbf{0} \;\Leftrightarrow\; (T-0.8I)\,\mathbf{v}_1=\mathbf{0} \;\Leftrightarrow$$

$$\begin{pmatrix} 0.6-0.8 & 0.2 \\ 0.2 & 0.6-0.8 \end{pmatrix}\begin{pmatrix} v_{1x} \\ v_{1y} \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -0.2 & 0.2 \\ 0.2 & -0.2 \end{pmatrix}\begin{pmatrix} v_{1x} \\ v_{1y} \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow\; v_{1x}=v_{1y} \;\Leftrightarrow\; \mathbf{v}_1=\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

- O autovetor associado ao autovalor $\lambda_2 = 0.4$ será dado por:

$$\mathbf{v}_2=\begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

- Desta forma, podemos finalmente escrever a solução geral do processo iterativo:

$$\mathbf{x}_n = c_1 \cdot 0.8^n \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 \cdot 0.4^n \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

- Conhecendo o valor de $\mathbf{x}_0$, podemos calcular $c_1$ e $c_2$:

$$\mathbf{x}_n = 0.8^n \, \frac{x_0 + y_0}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0.4^n \, \frac{y_0 - x_0}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

- E, agora, podemos ver claramente que tanto $0.8^n$ quanto $0.4^n$ tendem a 0 quando $n \to \infty$, isto é, $\mathbf{x}_n \to \mathbf{0} \quad n \to \infty$.
- Vemos, também, que a convergência para o ponto fixo assintoticamente estável $\mathbf{x}^* = \mathbf{0}$ é dominada pelo maior autovalor $\lambda_1 = 0.8$, pois, com $n \to \infty$:

$$\mathbf{x}_n = 0.8^n \left( \underbrace{\frac{x_0 + y_0}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{\textit{constante}} + \underbrace{0.5^n \, \frac{y_0 - x_0}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix}}_{\to \mathbf{0}} \right) \approx 0.8^n \, \frac{x_0 + y_0}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \to \mathbf{0}.$$

- Finalmente, vemos que quando $\mathbf{x}_n \to \mathbf{0}$, seu módulo $|\mathbf{x}_n| \to 0$. Estas condições são equivalentes. Então, vamos usar com certa frequência esta condição ao invés da condição vetorial.
- Com a intuição que ganhamos com os resultados deste exemplo, podemos agora sedimentar os seguintes conceitos fundamentais:
- O ponto fixo $\mathbf{x}^* = \mathbf{0}$ será assintoticamente estável se e somente se todos os autovalores $\lambda_i$ da matrix $T$ tiverem módulo menor que um, isto é, $|\lambda_i| < 1$. Se algum tiver módulo maior (ou igual) que 1, o ponto fixo do processo iterativo será instável (estável).
- Se uma matriz $T$ for tal que todos seus autovalores, $|\lambda_i| < 1$, então $T^n \to \mathbb{O}$ com $n \to \infty$, isto é, todos os seus elementos $\{T^n\}_{ij} \to 0$ com $n \to \infty$. Dizemos que $T$ é uma matriz convergente.

- O autovalor $\lambda_i$ (complexo ou real) de maior módulo da matriz $T$, é chamado de raio espectral de $T$, denotado por $\rho(T)$.

$$\rho(T) = \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_m|\}$$

- Portanto, uma matriz $T$ será convergente se e somente se seu raio espectral $\rho(T) < 1$.
- O raio espectral de $T$ determina quão rápido o processo vai convergir (ou divergir) para seu ponto fixo, pois:

$$\begin{aligned}
|\mathbf{x}_n| &= |c_1\lambda_1^n\mathbf{v}_1 + c_2\lambda_2^n\mathbf{v}_2 + \cdots + c_m\lambda_m^n\mathbf{v}_m| \\
&\leq |c_1\lambda_1^n\mathbf{v}_1| + |c_2\lambda_2^n\mathbf{v}_2| + \cdots + |c_m\lambda_m^n\mathbf{v}_m| \quad \text{(desigualdade triangular)} \\
&\leq |\lambda_1^n||c_1\mathbf{v}_1| + |\lambda_2^n||c_2\mathbf{v}_2| + \cdots + |\lambda_m^n||c_m\mathbf{v}_m| \\
&\leq |\lambda_1^n||c_1||\mathbf{v}_1| + |\lambda_2^n||c_2||\mathbf{v}_2| + \cdots + |\lambda_m^n||c_m||\mathbf{v}_m|
\end{aligned}$$

Como $\rho(T) = \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_m|\}$, então:

$$\begin{aligned} |\mathbf{x}_n| &\leq |\lambda_1^n||c_1||\mathbf{v}_1| + |\lambda_2^n||c_2||\mathbf{v}_2| + \cdots + |\lambda_m^n||c_m||\mathbf{v}_m| \\ &\leq \rho(T)^n \Big( \underbrace{|c_1||\mathbf{v}_1| + |c_2||\mathbf{v}_2| + \cdots + |c_m||\mathbf{v}_m|}_{K \in I\!R} \Big) = K\rho(T)^n \end{aligned}$$

- Portanto, se $\rho(T) < 1$, temos que $|\mathbf{x}_n| \leq K\rho(T)^n \to 0$, quando $n \to \infty$, $\mathbf{x}^* = \mathbf{0}$ é um ponto fixo assintóticamente estável e $T$ é uma matriz convergente. Caso contrário, se $\rho(T) \geq 1$, $\mathbf{x}^* = \mathbf{0}$ é um ponto fixo instável ou apenas estável.
- Quanto menor for $\rho(T)$, mais rapidamente o processo iterativo converge.
- Portanto, vamos querer construir processos iterativos com matrizes com o menor $\rho(T)$ possível! (lembrar da escolha de $|g'(x^*)|$ no caso de processos iterativos escalares!)
- PROBLEMA: calcular autovalores é caro, principalmente para sistemas grandes. Teremos que ser espertos!