

Atividade para casa #3

Soluções de Sistemas lineares *Método de Gauss-Jordan e decomposição L.U*

May 9, 2025

1 Contextualização

Até esse momento temos chamado a técnica de eliminação Gaussiana apresentada de *ingênua*. Essa palavra tem sido usada propositalmente para antecipar a ideia de que existem alguns problemas potenciais no algoritmo de eliminação Gaussiana apresentado até o momento. Um resumo geral desses problemas encontra-se listado abaixo.

1. Erros de arredondamento que vão crescendo conforme aumenta-se a ordem do sistema;
2. Divisão por zero ou por pivôs muito pequenos durante a etapa de eliminação;
3. Sistemas mal-condicionados;

O problema 1 é de mais fácil solução. Para evitarmos a propagação numérica de erros de arredondamento na etapa de eliminação basta aumentarmos o número de algarismos significativos das variáveis usadas no algoritmo. Essa solução entretanto vem acompanhada de um aumento no consumo de memória por parte do hardware que irá processar o programa. Em linguagens compiláveis, como C++ e FORTRAN podemos especificar na definição das variáveis reais utilizadas o número de casas decimais que serão consideradas para cada variável. Em C++ essa é a diferença entre variáveis do tipo float (7 dígitos decimais de precisão) e double (15 dígitos decimais), já em FORTRAN essa diferença se traduz nas variáveis do tipo real (8 dígitos decimais de precisão) e real*16 (16 dígitos decimais de precisão). Esse tipo de problema aumenta com a ordem do sistema, uma vez que o número de cálculos para o algoritmo de eliminação Gaussiana é $\mathcal{O}(n^3)$.

O problema 2 é um pouco mais complexo de ser abordado e para isso envolve uma estratégia algorítmica mais complexa. Em termos gerais a solução para esse problema consiste em utilizarmos sempre o maior pivô disponível em cada linha para realizarmos a etapa de eliminação. Essa estratégia se chama *pivoteamento parcial*. No fundo, o pivoteamento parcial consiste em dar para o programador que implementa o algoritmo de eliminação Gaussiana a liberdade para triangularizar a matriz dos coeficientes sem necessariamente seguir uma ordem crescente de linhas ou colunas.

Para entendermos melhor esse problema e como o pivoteamento parcial atua dentro do algoritmo de eliminação Gaussiana, vamos usar um exemplo de um sistema 2×2 como referência. Considere então o seguinte sistema linear:

$$0.0003x_1 + 3.0000x_2 = 2.0001 \quad (1)$$

$$1.0000x_1 + 1.0000x_2 = 1.0000 \quad (2)$$

A aplicação da técnica de eliminação Gaussiana em sua forma ingênua, como nos foi apresentada até o momento nos diria para começar pelo cálculo do fator $f_{21} = a_{21}/a_{11}$, em que a_{11} é o elemento pivô. E em seguida, multiplicar esse fator pela linha (1) para finalmente subtraí-lo da linha (2),

eliminando assim o termo x_1 da equação (2). Esse procedimento leva a uma nova equação (2) dada por

$$-9999x_2 = -6666. \quad (3)$$

O resultado para x_2 a partir de (3) é $x_2 = 2/3$. A substituição desse valor na equação (1) nos leva à

$$x_1 = \frac{2.0001 - 3.0000 \times (2/3)}{0.0003}. \quad (4)$$

Note porém que o valor $x_2 = 2/3$ consiste numa dízima infinita, sujeita a arredondamento na última casa decimal considerada pelo computador durante o processamento do programa. Veja que na equação (4) temos uma subtração de números quase iguais no numerador e uma divisão por um número muito pequeno no denominador. Esse procedimento gera uma dependência muito grande do valor de x_1 em função do número de algarismos significativos utilizados para calcular x_2 . Veja a tabela (1).

x_2	x_1
0.667	-3.333
0.6667	0.0000
0.66667	0.30000
0.666667	0.330000
0.6666667	0.3330000

Table 1: Sensibilidade de x_1 em função do número de algarismos significativos utilizados para o cálculo x_2 na versão ingênua do algoritmo de eliminação Gaussiana

Vamos considerar agora a aplicação do mesmo procedimento, porém invertendo a linha (1) pela linha (2) do sistema, ou seja, agora iremos resolver o seguinte sistema de equações:

$$1.0000x_1 + 1.0000x_2 = 1.0000 \quad (5)$$

$$0.0003x_1 + 3.0000x_2 = 2.0001. \quad (6)$$

A aplicação do mesmo procedimento de eliminação leva agora a uma nova versão da equação (6), dada por:

$$2.9997x_2 = 1.9998 \rightarrow x_1 = \frac{1 - 1 \times (2/3)}{1}. \quad (7)$$

Nesse novo cenário, a sensibilidade de x_1 com relação ao número de algarismos significativos utilizados para x_2 diminui consideravelmente. Veja a tabela (2).

x_2	x_1
0.667	0.333
0.6667	0.3333
0.66667	0.33333
0.666667	0.333333
0.6666667	0.3333333

Table 2: Sensibilidade de x_1 em função do número de algarismos significativos utilizados para o cálculo x_2 no algoritmo de eliminação Gaussiana com pivoteamento parcial

Essa lógica de buscar o maior elemento nas linhas disponíveis para realizar o pivoteamento é equivalente a ter a liberdade de escolher a ordem das linhas nas quais iremos realizar a etapa de

eliminação. Quando estendemos essa lógica não só para as linhas, mas também para as colunas, o processo é chamado de pivoteamento completo. Esse procedimento é raramente usado, uma vez que complica muito o algoritmo e geralmente o pivoteamento parcial já resolve o problema relacionado à divisão por pivôs nulos ou muito pequenos.

Finalmente, o problema (3), relacionado a questão do condicionamento, apesar de já ter sido abordado aqui, pode ser ainda discutido em maiores detalhes, uma vez que sistemas lineares são blocos construtores de diversos tipos de algoritmos utilizados em diferentes problemas de modelagem, sendo assim fundamental que compreendamos com bastante clareza esse problema associado. Até aqui vimos a ideia de condicionamento como uma métrica do comportamento do sistema, relacionada ao determinante da matriz dos coeficientes. Uma outra interpretação do problema de mau condicionamento se deu pelo método gráfico aplicado a um sistema 2×2 . Nesse caso, um sistema mal condicionado é aquele para o qual as inclinações das retas que descrevem $x_2(x_1)$ são muito próximas, de tal sorte que ficaria muito difícil identificar o ponto de intersecção dessas retas (solução do sistema) no método gráfico. Esse caso foi apresentado como equivalente a um determinante muito pequeno na matriz dos coeficientes. O caso extremo de um sistema com determinante nulo na matriz dos coeficientes é o que nós chamamos de sistemas singulares, para os quais não há nenhuma ou infinitas soluções.

Apesar do método gráfico ser útil para entendermos a ideia do que significa o condicionamento de um sistema, ele não nos atende para sistemas de alta ordem. Ao mesmo tempo, o valor absoluto do determinante de um sistema nem sempre é pequeno para um sistema mal condicionado. Para avançarmos um pouco mais nessa discussão, vamos definir aqui uma outra ideia do que realmente significa o condicionamento de um sistema.

Minha proposta é a de que pensemos em condicionamento como o nível de correlação de uma relação causal. Em outras palavras, se um sistema é bem condicionado isso significa que uma pequena mudança em um dos coeficientes do sistema vai gerar uma pequena mudança na solução do mesmo. Ao mesmo tempo, um sistema mal condicionado é aquele para o qual uma pequena mudança em um dos coeficientes do sistema gera grandes mudanças na solução final. Essa ideia guarda semelhança ao que chamamos de caos em sistemas dinâmicos. Dentre outras características, um sistema dinâmico é caótico quando possui grande sensibilidade às condições iniciais. Nesse tipo de sistema, basta mudar um único parâmetro em uma casa decimal e a solução ao longo do tempo evolui para estados completamente diferentes. Um recurso gráfico útil para entendermos as

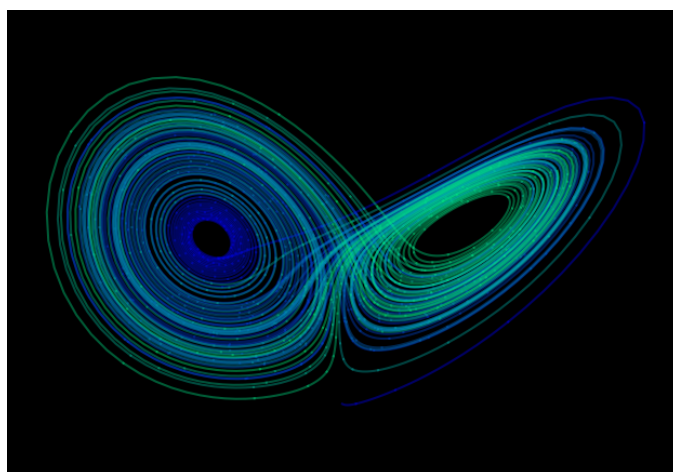


Figure 1: Exemplo de caminhos percorridos por um sistema dinâmico no espaço de fases quando o comportamento deste sistema é caótico.

características desse tipo de sistema é a identificação dos caminhos que o mesmo percorre em seu espaço de fase. Esse espaço por sua vez é uma representação dos pares de posição e velocidade pelos

quais o sistema passa. Para um sistema que possui uma resposta harmônica do tipo $x(t) = \cos(\omega t)$, esse espaço de fase $x \times \dot{x}$, em que $\dot{x} = dx/dt$ é um círculo por exemplo. Imagens de atratores como o representado na figura (1) representam comportamentos típicos de sistemas caóticos.

Para entendermos como essa ideia de sensibilidade às condições iniciais se relaciona ao condicionamento de um sistema linear, considere novamente um sistema linear 2×2 como exemplo:

$$x_1 + 2x_2 = 10 \quad (8)$$

$$1.1x_1 + 2x_2 = 10.4. \quad (9)$$

A solução desse sistema pela regra de Cramer nos leva a $x_1 = 4$ e $x_2 = 3$ e o determinante do mesmo é $D = -0.2$. Considere agora alterar o valor de a_{21} de 1.1 para 1.05. A aplicação da regra de Cramer sobre esse novo sistema leva ao novo par de soluções $x_1 = 8$ e $x_2 = 1$. Veja que a alteração em apenas 1 dos 4 coeficientes da matriz $[A]$ levou a primeira raiz a dobrar de valor e a segunda a diminuir em $1/3$. Esse é um sistema mal condicionado e nesse caso seu determinante é de fato um número que poderia ser considerado pequeno. Entretanto, essa relação entre um determinante pequeno e um sistema mal condicionado não é sempre obedecida. Para entendermos melhor essa questão vamos a um outro exemplo. Considere agora o seguinte sistema:

$$3x_1 + 2x_2 = 18 \quad (10)$$

$$-x_1 + 2x_2 = 2. \quad (11)$$

Faça um teste em casa com esse novo sistema e brinque de mudar ligeiramente um dos coeficientes de $[A]$. Veja o impacto que essa mudança possui na solução final. Você verá que esse é um sistema bem condicionado. O determinante D para esse novo sistema é $D = 8$. Considere agora multiplicar por 10 as equações (8) e (9) do nosso sistema mal condicionado. Ao fazermos isso, alteramos o determinante D para $D = 20$, porém não alteramos as equações originais e nem as relações entre os coeficientes do sistema. Se pensamos em termos do método gráfico, a multiplicação por 10 das equações desse novo sistema não altera a relação entre essas retas que se cruzam numa região de difícil detecção do ponto de intersecção entre ambas. O sistema continua sendo mal condicionado, porém seu determinante agora é mais que o dobro do determinante $D = 8$ do sistema bem condicionado dado pelas equações (10) e (11).

No fundo, o condicionamento de um sistema linear tem muito mais a ver com a relação entre os valores numéricos dos coeficientes da matriz $[A]$ do que com o valor absoluto do determinante. Uma forma de não nos deixarmos enganar pelo valor do determinante e ainda assim utilizarmos este como métrica para avaliar o nível de condicionamento do sistema é dividirmos cada linha do sistema pelo valor do maior coeficiente, garantindo assim que o maior coeficiente de cada linha seja unitário e em seguida calculando o determinante do sistema normalizado. Esse procedimento anula efeitos de escala e dá confiabilidade para o uso do determinante como uma métrica de avaliação do seu condicionamento. A figura (2) ilustra como esse procedimento atua nos sistemas 1 (10-11) e 2 (8-9) agora normalizados

	Sistema 1	Sistema 2
Equações	$x_1 + 0,667x_2 = 6$ $-0,5x_1 + x_2 = 1$	$0,5x_1 + x_2 = 5$ $0,55x_1 + x_2 = 5,2$
Determinante	1,333	-0,05
Condicionamento	Bom	Ruim

Figure 2: Exemplo da anulação do efeito de escala pela normalização do sistema

O que percebemos é que ao normalizarmos o sistema, podemos voltar a usar o valor absoluto do determinante do novo sistema como uma métrica confiável do seu nível de condicionamento. O problema aqui é justamente o alto custo do cálculo do determinante. Entretanto, até nesse ponto a técnica de eliminação Gaussiana é vantajosa, uma vez que para matrizes triangulares o determinante é calculado simplesmente pelo produto dos termos diagonais, ou seja:

$$D = \prod_{i=1}^n a_{ii}. \quad (12)$$

Faça o teste em casa para sistemas 3×3 e 4×4 e cheque a relação dada na equação (12).

2 Decomposição L.U

Uma outra estratégia possível de solução de um sistema linear, baseada no método da eliminação Gaussiana consiste em triangularizar o sistema linear tanto por cima quanto por baixo. Em outras palavras, decompor a matriz dos coeficientes em duas matrizes triangulares, uma superior $[U]$ e outra inferior $[L]$, sem precisar manipular o vetor $\{b\}$ durante o processo de eliminação e solução. Essa técnica é conhecida como decomposição L.U e será apresentada a seguir.

A ideia geral da decomposição L.U se baseia no seguinte racional:

- Nosso sistema linear original é dado por

$$[A]\{x\} = \{b\} \rightarrow [A]\{x\} - \{b\} = \{0\} \quad (13)$$

- Suponha agora que a equação (13) possa ser escrita em termos de um sistema triangular superior, como se estivéssemos implementado a etapa de eliminação do algoritmo de eliminação Gaussiana descrito aqui, de tal sorte que

$$[U]\{x\} = \{d\} \rightarrow [U]\{x\} - \{d\} = \{0\} \quad (14)$$

- Agora, suponha a existência de uma matriz triangular inferior com 1's na diagonal, que chamaremos aqui de $[L]$. Se multiplicarmos $[L]$ pela equação (14) teremos

$$[L]([U]\{x\}) = [L]\{d\} \rightarrow [L]([U]\{x\} - \{d\}) = \{0\} \quad (15)$$

- Como o lado direito da equação (15) é igual o lado direito da equação (13) podemos fazer

$$[L]([U]\{x\} - \{d\}) = [A]\{x\} - \{b\} \quad (16)$$

- Comparando os termos da equação (16) concluímos que:

$$[A] = [L][U] \quad (17)$$

e

$$[L]\{d\} = \{b\} \quad (18)$$

Baseado nesses procedimentos, a ideia do algoritmo de decomposição L.U consistem em

1. Decompor $[A]$ em $[L]$ e $[U]$;
2. Calcular $\{d\}$ por substituição progressiva, usando $[L]\{d\} = \{b\}$. Note que isso só é possível pelo fato de $[L]$ ser uma matriz triangular inferior;

3. Com o valor de $\{d\}$ calculamos $\{x\}$ por substituição regressiva utilizando $[U]\{x\} = \{d\}$, o que só é possível pelo fato de $[U]$ ser uma matriz triangular superior.

A grande pergunta aqui é: como decompor $[A]$ em $[L]$ e $[U]$? E a resposta para essa pergunta está no próprio algoritmo de eliminação Gaussiana. Quando aplicamos o algoritmo de eliminação Gaussiana para triangularizar o sistema, acabamos jogando fora os fatores utilizados na etapa de eliminação. Esses fatores, se armazenados nas posições adequadas em uma outra matriz acabam dando origem à matriz $[L]$ quando usados para triangularizar a matriz $[A]$ na matriz $[U]$.

Para exemplificar o procedimento vamos aplicar a estratégia de decomposição L.U para um sistema 3×3 genérico, dado por:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}. \quad (19)$$

A ideia aqui é aplicar a eliminação Gaussiana apenas sobre a matriz $[A]$ sem manipular $\{b\}$ para obter $[U]$ e armazenar os fatores utilizados no processo sobre uma matriz identidade $[I]$ para formar a matriz $[L]$. Do ponto de vista de sequência de procedimentos fazemos:

1. Calcula-se o fator $f_{21} = a_{21}/a_{11}$;
2. Multiplica-se a linha 1 por f_{21} e subtrai-se o resultado da linha 2:

$$[U] = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a'_{22} & a'_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad (20)$$

3. Armazenamos o fator f_{21} na posição ℓ_{21} de uma matriz $[I]$ para começarmos a construir $[L]$:

$$[L] = \begin{pmatrix} 1 & 0 & 0 \\ f_{21} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (21)$$

4. Calcula-se agora o fator $f_{31} = a_{31}/a_{11}$;
5. Multiplica-se a linha 1 por f_{31} e subtrai-se o resultado da linha 3:

$$[U] = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a'_{22} & a'_{23} \\ 0 & a'_{32} & a'_{33} \end{pmatrix} \quad (22)$$

6. Armazenamos o fator f_{31} na posição ℓ_{31} de $[L]$:

$$[L] = \begin{pmatrix} 1 & 0 & 0 \\ f_{21} & 1 & 0 \\ f_{31} & 0 & 1 \end{pmatrix} \quad (23)$$

7. Calcula-se agora o fator $f_{32} = a'_{32}/a'_{22}$;
8. Multiplica-se a linha 2 por f_{32} e subtrai-se o resultado da linha 3:

$$[U] = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a'_{22} & a'_{23} \\ 0 & 0 & a''_{33} \end{pmatrix} \quad (24)$$

9. Armazenamos o fator f_{32} na posição ℓ_{32} de $[L]$:

$$[L] = \begin{pmatrix} 1 & 0 & 0 \\ f_{21} & 1 & 0 \\ f_{31} & f_{32} & 1 \end{pmatrix} \quad (25)$$

Esse procedimento decompõe $[A]$ em $[A] = [L][U]$ e não envolve operações em $\{b\}$, nesse sentido ele tem um custo computacional ligeiramente menor do que o custo da mesma etapa no processo original de eliminação Gaussiana, gastando apenas um pouco mais de memória para armazenar os fatores na matriz $[L]$.

Em termos de terminologia, existem dois tipos de métodos de decomposição L.U. Quando as matrizes $[L]$ e $[U]$ possuem as características descritas na equação (26) dizemos que o procedimento é denominado *decomposição de Doolittle ou fatoração L.U.*

$$[L] = \begin{pmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{pmatrix} \quad \text{e} \quad [U] = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix} \quad (26)$$

Quando as matrizes $[L]$ e $[U]$ possuem as características descritas na equação (27) dizemos que o procedimento é denominado *decomposição de Crout*.

$$[L] = \begin{pmatrix} \ell_{11} & 0 & 0 \\ \ell_{21} & \ell_{22} & 0 \\ \ell_{31} & \ell_{32} & \ell_{33} \end{pmatrix} \quad \text{e} \quad [U] = \begin{pmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{pmatrix} \quad (27)$$

2.1 Fórmulas fechadas para a decomposição de Crout

O algoritmo de decomposição de Crout é baseado nas fórmulas fechadas apresentadas na figura (3).

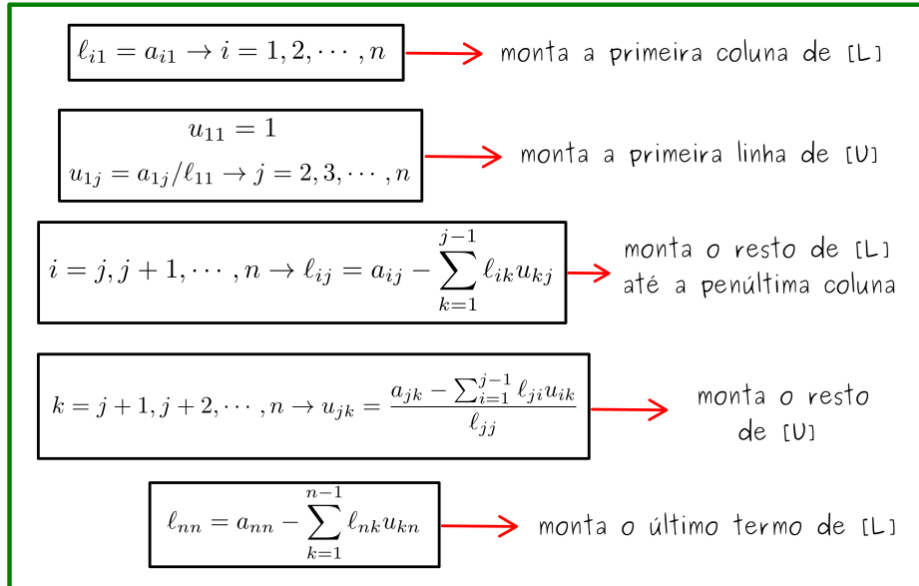


Figure 3: Fórmulas fechadas para a decomposição de Crout

3 Enunciado da tarefa

1. Prove que o número de operações de ponto flutuante relacionadas à divisões e/ou multiplicações envolvidos no método de Gauss-Jordan sem pivoteamento é dado por:

$$\text{FLOPS}_{\times/\div} = \frac{n^3}{2} + n^2 - \frac{n}{2} \rightarrow \frac{n^3}{2} + \mathcal{O}(n^2); \quad (28)$$

2. Considere o sistema linear dado pela equação (29)

$$\begin{pmatrix} 3 & -0.1 & -0.2 \\ 0.1 & 7 & -0.3 \\ 0.3 & -0.2 & 10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7.85 \\ -19.3 \\ 71.4 \end{pmatrix}, \quad (29)$$

usando a decomposição de Doolittle obtenha $[L]$ e $[U]$ e mostre que $[A] = [L][U]$. Em seguida, calcule o vetor $\{d\}$ e obtenha a solução $\{x\}$ do sistema. Tudo isso na mão, sem usar computador.

4 Referências bibliográficas

1. S. C. Chapra, R. P. Canale. “Métodos Numéricos para Engenharia.” McGrawHill, 5a edição (2008): 1-825.