MS211 - Cálculo Numérico

Aula 5 – Normas de Vetores e Matrizes. Condicionamento de uma Matriz.

Marcos Eduardo Valle

Na aula anterior, falamos sobre os erros de arredondamento na representação de ponto flutuante e suas operações aritméticas.

Em termos gerais, erros sempre existirão quando resolvemos um problema contínuo (e.g. em $\mathbb{R}$) num computador.

Consequentemente, um método numérico raramente produz a solução exata de um problema matemático contínuo.

Na aula de hoje, veremos quando o arredondamento em ponto flutuante e suas operações aritméticas influenciam na credibilidade do resultado produzido por um método numérico.

Iniciaremos apresentando a definição de norma, conceito matemático utilizado para medir tamanho ou distância.

Definição 1 (Norma)

Uma norma é uma função $\|\cdot\|$, de um espaço vetorial $\mathcal{V}$ no conjunto dos reais $\mathbb{R}$, que satisfaz as seguintes propriedades para quaisquer $u, v \in \mathcal{V}$ e $\alpha \in \mathbb{R}$:

1. $\|v\| \geqslant 0$ com $\|v\| = 0$ se e somente se $v = 0$.
2. $\|\alpha v\| = |\alpha|\|v\|$.
3. $\|u + v\| \leqslant \|u\| + \|v\|$. (desigualdade triangular)

Não entraremos nos detalhes do que é um espaço vetorial. Por ora, basta saber que os conjuntos $\mathbb{R}^n$ e $\mathbb{R}^{n\times n}$ dos vetores com n componentes e as matrizes $n \times n$ são ambos espaços vetoriais com a soma e multiplicação por escalar.

Uma norma $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$, que associa à um vetor de $\mathbb{R}^n$ um número real, será chamada **norma vetorial**. Similarmente, uma função $\|\cdot\| : \mathbb{R}^{n\times n} \to \mathbb{R}$ é uma **norma matricial**.

Norma Vetorial

Muitas normas vetoriais são dadas pela equação

$$\|\mathbf{x}\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p} = \sqrt[p]{|x_1|^p + |x_2|^p + \ldots + |x_n|^p}, \quad 1 \leqslant p < +\infty.$$

Exemplos de normas para $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ incluem:

- $\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \ldots + |x_n|$.
- $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$. (norma Euclidiana)
- $\|\mathbf{x}\|_\infty = \max_{i=1:n} |x_i| = \max\{|x_1|, |x_2|, |x_n|\}$.

Exemplo 2

Para o vetor $\mathbf{x} = [3 \; -4]^T$, temos:

$$\|\mathbf{x}\|_1 = |3| + |-4| = 7,$$
$$\|\mathbf{x}\|_2 = \sqrt{3^2 + (-4)^2} = 5,$$
$$\|\mathbf{x}\|_\infty = \max\{|3|, |-4|\} = 4.$$

No `GNU Octave`, podemos calcular as normas $\|\cdot\|_1$, $\|\cdot\|_2$ e $\|\cdot\|_\infty$ de um vetor $\mathbf{x}$ usando respectivamente os comandos:

» `norm(x,1)`

» `norm(x)` ou » `norm(x,2)`

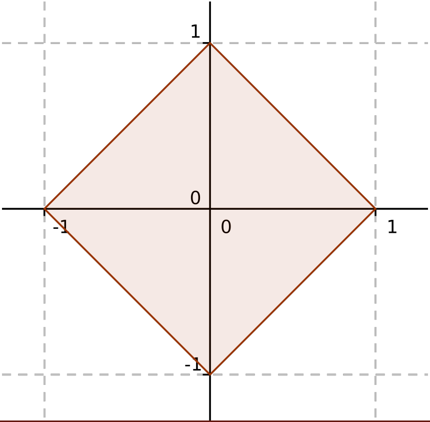
» `norm(x,inf)`

Interpretação Geométrica

O conjunto

$$\mathcal{B}_1 = \{\mathbf{x} = (x, y) \in \mathbb{R}^2 : \|\mathbf{x}\|_1 \leqslant 1\},$$

corresponde à figura geométrica

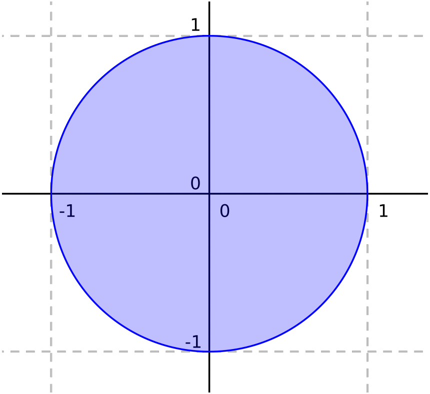


Interpretação Geométrica

O conjunto

$$\mathcal{B}_2 = \{\mathbf{x} = (x, y) \in \mathbb{R}^2 : \|\mathbf{x}\|_2 \leqslant 1\},$$

corresponde à figura geométrica

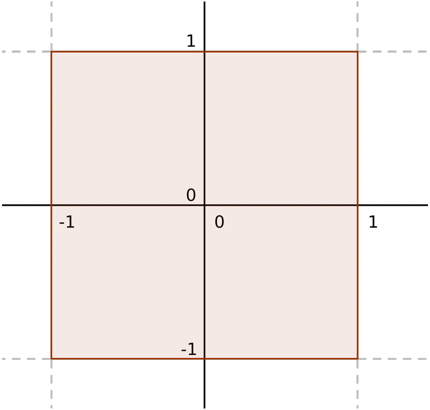


Interpretação Geométrica

O conjunto

$$\mathcal{B}_\infty = \{\mathbf{x} = (x, y) \in \mathbb{R}^2 : \|\mathbf{x}\|_\infty \leqslant 1\},$$

corresponde à figura geométrica



Normas Matriciais Subordinadas

Podemos identificar uma matriz **A** com uma transformação linear

$$\mathbf{x} \mapsto \mathbf{A}\mathbf{x}.$$

Uma norma matricial subordinada $\|\mathbf{A}\|$ mede **a maior distorção** efetuada pela transformação linear $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$. Formalmente, temos

$$\|\mathbf{A}\| = \max_{\mathbf{x} \in \mathbb{R}^n} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}.$$

Equivalentemente, escrevendo $\mathbf{v} = \mathbf{x}/\|\mathbf{x}\|$, concluímos que

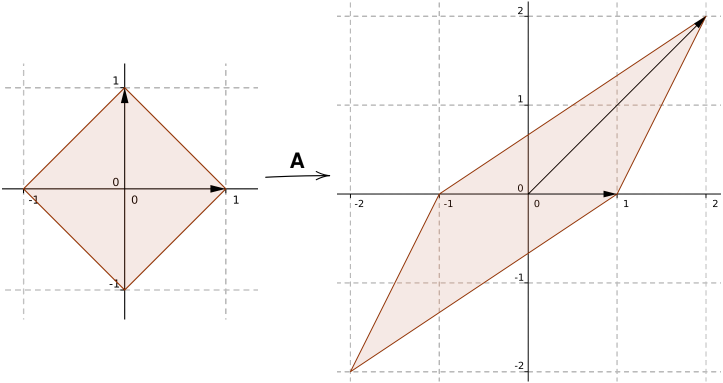
$$\|\mathbf{A}\| = \max\{\|\mathbf{A}\mathbf{v}\| : \|\mathbf{v}\| = 1\},$$

ou seja, $\|\mathbf{A}\|$ é a maior distorção que **A** faz em $\{\mathbf{v} : \|\mathbf{v}\| = 1\}$.

Observe que a norma matricial acima está subordinada ou é induzida pela norma vetorial!

Exemplo da $\| \cdot \|_1$

Considere a matriz $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$. Geometricamente, temos



$$\|A\|_1 = \max\{\|\mathbf{Av}\|_1 : \|\mathbf{v}\|_1 = 1\} = 4.$$

A Norma Subordinada $\|\cdot\|_1$

Pode-se mostrar que $\{\mathbf{Av} : \|\mathbf{v}\|_1 = 1\}$ é um politopo (generalização de um polígono) e o valor máximo de $\|\mathbf{Av}\|_1$, para $\|\mathbf{v}\|_1 = 1$, é obtido num dos vértices.

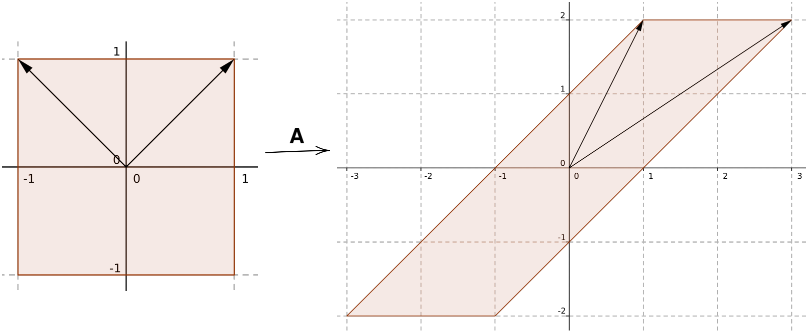
Os vértices do politopo $\{\|\mathbf{Av} : \|\mathbf{v}\|_1 = 1\}$ são $\pm\mathbf{a}_1, \pm\mathbf{a}_2, \ldots, \pm\mathbf{a}_n$, em que $\mathbf{a}_1, \ldots, \mathbf{a}_n$ são as colunas de $\mathbf{A}$, i.e., $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_n]$.

Portanto, $\|\mathbf{A}\|_1$ é o valor máximo da soma dos valores absolutos das colunas de $\mathbf{A}$, ou seja,

$$\|\mathbf{A}\|_1 = \max\{\|\mathbf{a}_1\|_1, \ldots, \|\mathbf{a}_n\|_1\} = \max_{j=1:n} \left\{ \sum_{i=1}^{n} |a_{ij}| \right\}.$$

Exemplo da $\|\cdot\|_\infty$

Considere a matriz $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$. Geometricamente, temos



$$\|A\|_\infty = \max\{\|\mathbf{Av}\|_\infty : \|\mathbf{v}\|_\infty = 1\} = 3.$$

A Norma Subordinada $\|\cdot\|_\infty$

De um modo similar, pode-se mostrar que $\{\mathbf{Av} : \|\mathbf{v}\|_\infty = 1\}$ é um politopo e o valor máximo de $\|\mathbf{Av}\|_\infty$ é obtido num dos vértices.

Contudo, os vértices do politopo $\{\|\mathbf{Av} : \|\mathbf{v}\|_\infty = 1\}$ são obtidos pelo produto $\mathbf{Av}$, em que $\mathbf{v} = [\pm 1, \pm 1, \ldots, \pm 1]^T$.

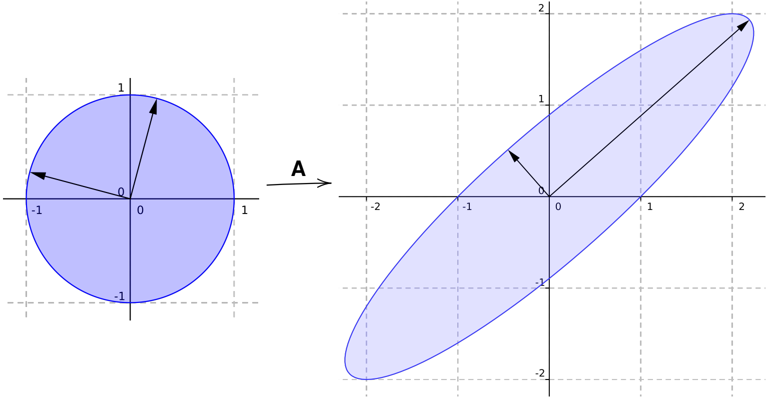
Usando esse fato, podemos concluir que $\|\mathbf{A}\|_\infty$ é o valor máximo da soma dos valores absolutos das linhas de $\mathbf{A}$, ou seja,

$$\|\mathbf{A}\|_\infty = \max\{\|\mathbf{a}_1^T\|_1, \ldots, \|\mathbf{a}_n^T\|_1\} = \max_{i=1:n} \left\{ \sum_{j=1}^{n} |a_{ij}| \right\},$$

em que $\mathbf{a}_i^T$ denota a i-ésima linha de $\mathbf{A}$.

Exemplo da $\| \cdot \|_2$

Considere a matriz $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$. Geometricamente, temos



$$\|A\|_2 = \max\{\|\mathbf{Av}\|_2 : \|\mathbf{v}\|_2 = 1\} = 2.92081.$$

A Norma Subordinada $\|\cdot\|_2$

Pode-se mostrar que $\{\mathbf{Av} : \|\mathbf{v}\|_2 = 1\}$ é um hiper-elipse (generalização de uma elipse) e o valor máximo de $\|\mathbf{Av}\|_2$, para $\|\mathbf{v}\|_1 = 2$, é a metade do maior eixo.

Formalmente, $\|\mathbf{A}\|_2$ é o maior valor singular de matriz $\mathbf{A}$.

No `GNU Octave`, podemos calcular as normas $\|\cdot\|_1$, $\|\cdot\|_2$ e $\|\cdot\|_\infty$ de uma matriz $\mathbf{A}$ usando respectivamente os comandos:

```
» norm(A,1)
» norm(A) ou » norm(A,2)
» norm(A,inf)
```

Norma Consistente

Dizemos que uma norma matricial $\|\cdot\|$ é consistente se

$$\|\mathbf{AB}\| \leqslant \|\mathbf{A}\|\|\mathbf{B}\|,$$

para quaisquer matrizes **A** e **B**.

Todas as normas subordinadas são consistentes!

As normas subordinadas também satisfazem a desigualdade

$$\|\mathbf{Ax}\| \leqslant \|\mathbf{A}\|\|\mathbf{x}\|,$$

para quaisquer matriz **A** e vetor **x**.

Erro Absoluto e Erro Relativo

Suponha que resolvemos um sistema linear $\mathbf{Ax} = \mathbf{b}$, em que $\mathbf{A} \in \mathbb{R}^{n \times n}$ é uma matriz não-singular e $\mathbf{b} \in \mathbb{R}^n$, usando um método numérico como, por exemplo, o método da eliminação de Gauss.

Vamos denotar por $\tilde{\mathbf{x}}$ a solução encontrada pelo método numérico e $\mathbf{x}^* = \mathbf{A}^{-1}\mathbf{b}$ a solução exata.

Para avaliar a qualidade da solução produzida pelo método numérico, comparamos a solução numérica $\tilde{\mathbf{x}}$ com a solução exata $\mathbf{x}^*$ usando uma norma vetorial. Especificamente, calculamos o erro absoluto E_a ou o erro relativo E_r dados por:

$$E_a = \|\mathbf{x}^* - \tilde{\mathbf{x}}\| \quad \text{e} \quad E_r = \frac{\|\mathbf{x}^* - \tilde{\mathbf{x}}\|}{\|\mathbf{x}^*\|}.$$

Resíduo

Uma desvantagem da análise que descrevemos anteriormente é que, na prática, não conhecemos a solução exata $\mathbf{x}^*$ do sistema linear $\mathbf{A}\mathbf{x} = \mathbf{b}$. Portanto, não podemos calcular o erro.

Como alternativa, calculamos o chamado **resíduo absoluto** R_a ou o **resíduo relativo** R_r definidos por

$$R_a = \|\mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}\| \quad \text{e} \quad R_r = \frac{\|\mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}\|}{\|\mathbf{b}\|}.$$

Em algumas situações, é conveniente definir os vetores

$$\mathbf{e} = \mathbf{x}^* - \tilde{\mathbf{x}} \quad \text{e} \quad \mathbf{r} = \mathbf{b} - \mathbf{A}\tilde{\mathbf{x}},$$

chamados respectivamente **erro** e **resíduo**.

Matriz Gerada Aleatoriamente

Para valiar a qualidade da solução de um sistema linear $\mathbf{Ax} = \mathbf{b}$ produzida pelo método da eliminação de Gauss, geramos uma matriz $\mathbf{A} \in \mathbb{R}^{100\times 100}$ cujos elementos a_{ij} possuem distribuição normal padrão.

Além disso, definimos $\mathbf{x}^* = [1, \dots, 1]^T$ e $\mathbf{b} = \mathbf{Ax}^*$.

Determinamos a solução numérica $\tilde{\mathbf{x}}$ usando o método da eliminação de Gauss e calculamos o erro e o resíduo relativo.

Repetimos o processo 1000 vezes. A média dos erros e resíduos relativos foram 1.3×10^{-12} e 6.0×10^{-14}, respectivamente.

Note que o resíduo forneceu uma boa estimativa para o erro!

Comandos do `GNU Octave`

```
» for i=1:1000
    A = randn(100,100); xs=ones(100,1);
    b=A*xs; xt=A\b;
    E_r(i)=norm(xs-xt,inf);
    R_r(i)=norm(b-A*xt,inf)/norm(b,inf);
end
» [mean(E_r),mean(R_r)];
ans =
    1.6535e-13    2.2155e-15
```

Matriz de Hilbert

Considere a matriz $\mathbf{A} \in \mathbb{R}^{100 \times 100}$ cujos elementos são

$$a_{ij} = \frac{1}{i + j - 1}.$$

Definimos $\mathbf{x}^* = [1, \ldots, 1]^T$, $\mathbf{b} = \mathbf{A}\mathbf{x}^*$ e determinamos a solução numérica $\tilde{\mathbf{x}}$ usando o método da eliminação de Gauss.

O erro relativo e o resíduo relativo foram

$$E_r = \frac{\|\mathbf{x}^* - \tilde{\mathbf{x}}\|_\infty}{\|\mathbf{x}^*\|_\infty} = 1.38 \quad \text{e} \quad R_r = \frac{\|\mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}\|_\infty}{\|\mathbf{b}\|_\infty} = 1.46 \times 10^{-15}.$$

Ao contrário do exemplo anterior, temos resíduo relativo muito pequeno mas um erro relativo grande.

Comandos do GNU Octave

```
» A = hilb(100); xs=ones(100,1);
» b=A*xs; xt=A\b;
warning:  matrix singular to machine precision,
rcond = 1.14558e-21
» R_r=norm(b-A*xt,inf)/norm(b,inf)
R_r = 1.4554e-15
» E_r=norm(xs-xt,inf)
E_r = 1.3774
```

Sistemas Lineares Mal-Condicionados

Em termos gerais, um sistema linear $\mathbf{Ax} = \mathbf{b}$ é mal-condicionado se pequenas perturbações na matriz $\mathbf{A}$ ou no vetor $\mathbf{b}$ causam grandes variações na solução.

Nesse caso, devido aos erros na representação e operações de pontos flutuantes, não devemos esperar uma solução precisa de um método numérico!

Além disso, num sistema linear mal-condicionado, o resíduo pode não revelar a natureza do erro!

O condicionamento de uma matriz $\mathbf{A}$ é definido em termos de sua norma e a norma de sua inversa.

Sabemos que os vetores **e** (erro) e **r** (resíduo) satisfazem:

$$\mathbf{r} = \mathbf{A}\mathbf{e} \quad \text{e} \quad \mathbf{e} = \mathbf{A}^{-1}\mathbf{r}.$$

Sendo $\|\cdot\|$ uma norma subordinada, tem-se:

$$\|\mathbf{r}\| \leqslant \|\mathbf{A}\|\|\mathbf{e}\| \quad \text{e} \quad \|\mathbf{e}\| \leqslant \|\mathbf{A}^{-1}\|\|\mathbf{r}\| \quad \Longrightarrow \quad \frac{\|\mathbf{r}\|}{\|\mathbf{A}\|} \leqslant \|\mathbf{e}\| \leqslant \|\mathbf{A}^{-1}\|\|\mathbf{r}\|.$$

Analogamente, das equações $\mathbf{b} = \mathbf{A}\mathbf{x}^*$ e $\mathbf{x}^* = \mathbf{A}^{-1}\mathbf{b}$, obtemos

$$\|\mathbf{b}\| \leqslant \|\mathbf{A}\|\|\mathbf{x}^*\| \text{ e } \|\mathbf{x}^*\| \leqslant \|\mathbf{A}^{-1}\|\|\mathbf{b}\| \quad \Longrightarrow \quad \frac{1}{\|\mathbf{A}^{-1}\|\|\mathbf{b}\|} \leqslant \frac{1}{\|\mathbf{x}^*\|} \leqslant \frac{\|\mathbf{A}\|}{\|\mathbf{b}\|}.$$

Combinamos as inequações, obtemos:

$$\frac{1}{\|\mathbf{A}\|\|\mathbf{A}^{-1}\|}\frac{\|\mathbf{r}\|}{\|\mathbf{b}\|} \leqslant \frac{\|\mathbf{e}\|}{\|\mathbf{x}^*\|} \leqslant \|\mathbf{A}\|\|\mathbf{A}^{-1}\|\frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}.$$

Número de Condição de uma Matriz

O número de condição de uma matriz $\mathbf{A}$, também chamado condicionamento de $\mathbf{A}$, é definido por

$$cond(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|.$$

Vale a seguinte relação entre o erro relativo e o resíduo relativo:

$$\frac{R_r}{cond(\mathbf{A})} \leqslant E_r \leqslant cond(\mathbf{A})R_r.$$

- Se $cond(\mathbf{A})$ é próximo de 1, o erro relativo E_r e o resíduo relativo R_r são próximos.
- Se $cond(\mathbf{A})$ é grande, o erro relativo pode ser muito maior que o resíduo relativo.

O número de condição de uma matriz **A** pode ser calculada no `GNU Octave` usando o comando:

```
» cond(A)
```

O comando

```
» rcond(A)
```

fornece o recíproco $1/\mathit{cond}(\mathbf{A})$, obtido usando a norma-1.
Note que **A** é mal-condicionada se `rcond(A)` é próximo de zero.

Calculamos o condicionamento da matriz de Hilbert como segue:

```
» cond(hilb(100))
ans = 5.9832e+19
```

Com base nos números do exemplo anterior, temos

$$\underbrace{1.38}_{E_r} \leqslant \underbrace{5.98 \times 10^{19}}_{\mathit{cond}(\mathbf{A})} \underbrace{1.46 \times 10^{-15}}_{R_r} = 8.71 \times 10^{4}.$$

Considerações Finais

Na aula de hoje apresentamos os conceitos de norma vetorial e norma matricial subordinada.

Vimos como esses conceitos podem ser usados para determinar o erro relativo de um sistema linear.

Vimos também o conceito de resíduo relativo e mostramos que ele representa o erro relativo se o condicionamento de **A** for próximo de 1.

Se $cond(\mathbf{A})$ é grande, dizemos que a matriz é mal-condicionada. Nesse caso, não podemos confiar na solução fornecida por um método numérico.

Muito grato pela atenção!