A característica da multiplicação de matrizes

Novos Talentos em Matemática

Eduardo Dias

10 de Setembro de 2006

O problema

Quantas multiplicações escalares são necessárias para efectuar a multiplicação de duas matrizes reais do tipo 2×2?

O problema

Quantas multiplicações escalares são necessárias para efectuar a multiplicação de duas matrizes reais do tipo 2×2?

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} xa + yc & xb + yd \\ za + wc & zb + wd \end{bmatrix}$$

8 multiplicações.

O problema

Quantas multiplicações escalares são necessárias para efectuar a multiplicação de duas matrizes reais do tipo 2×2?

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} xa + yc & xb + yd \\ za + wc & zb + wd \end{bmatrix}$$

8 multiplicações.

Problema

Conseguiremos efectuar este produto fazendo menos que ***oito*** *multiplicações? E qual é o mínimo de multiplicações necessárias?*

Exemplos

- Diferença de quadrados:

$$x^2 - y^2 = x \cdot x - y \cdot y = (x - y)(x + y).$$

À partida, efectuamos duas multiplicações, mas usando o caso notável, **apenas uma** multiplicação.

Exemplos

- Diferença de quadrados:

$$x^2 - y^2 = x \cdot x - y \cdot y = (x - y)(x + y).$$

À partida, efectuamos duas multiplicações, mas usando o caso notável, **apenas uma** multiplicação.

- Multiplicação de complexos:

Por definição: $(a + b\,i) \times (c + d\,i) = (a \cdot c - b \cdot d) + (a \cdot d + b \cdot c)i$.

À partida, 4 multiplicações. No entanto, calculando: $\mathbf{x} = a(c + d)$, $\mathbf{y} = (a + b)d$, $\mathbf{z} = b(d - c)$ obtém-se:

$$\boxed{(a + bi) \times (c + di) = (\mathbf{x} - \mathbf{y}) + (\mathbf{y} - \mathbf{z})i}$$

ou seja, **apenas três** multiplicações.

O algoritmo de Strassen

No caso da multiplicação de matrizes 2×2,calculando as expressões:

$$\begin{aligned} \mathrm{I} &= (x+w)(a+d) & \mathrm{V} &= (x+y)d \\ \mathrm{II} &= (z+w)a & \mathrm{VI} &= (-x+z)(a+b) \\ \mathrm{III} &= x(b-d) & \mathrm{VII} &= (y-w)(c+d) \\ \mathrm{IV} &= w(-a+c) \end{aligned}$$

O algoritmo de Strassen

No caso da multiplicação de matrizes 2×2,calculando as expressões:

$$\begin{array}{ll} \mathrm{I} = (x+w)(a+d) & \mathrm{V} = (x+y)d \\ \mathrm{II} = (z+w)a & \mathrm{VI} = (-x+z)(a+b) \\ \mathrm{III} = x(b-d) & \mathrm{VII} = (y-w)(c+d) \\ \mathrm{IV} = w(-a+c) & \end{array}$$

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \mathrm{I}+\mathrm{IV}-\mathrm{V}+\mathrm{VII} & \mathrm{III}+\mathrm{V} \\ \mathrm{II}+\mathrm{IV} & \mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{VI} \end{bmatrix}$$

ou seja, com este algoritmo, são necessárias **apenas sete** multiplicações. (V. Strassen 1969.)

Mínimo de multiplicações

Questões

- *É possível efectuar a multiplicação de duas matrizes* 2×2 *com menos de* **sete** *multiplicações?*

Mínimo de multiplicações

Questões

- *É possível efectuar a multiplicação de duas matrizes* 2×2 *com menos de* ***sete*** *multiplicações?*
- *Existe uma fórmula para a multiplicação de dois números complexos com menos de* ***três*** *multiplicações?*

O que é uma multiplicação?

Definição (Forma bilinear de característica 1)

Uma forma bilinear de característica 1 é uma a aplicação $\mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ dada por: $(\mathbf{x}, \mathbf{y}) \rightarrow \mathbf{x}(\mathbf{u}^T\mathbf{v})\mathbf{y}^T$ onde $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$ e $\mathbf{v} = (v_1, \ldots, v_m) \in \mathbb{R}^m$ são dois vectores. Denotemos esta aplicação por $\mathbf{u} \otimes \mathbf{v}$.

O que é uma multiplicação?

Definição (Forma bilinear de característica 1)

Uma forma bilinear de característica 1 é uma a aplicação $\mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ dada por: $(\mathbf{x}, \mathbf{y}) \to \mathbf{x}(\mathbf{u}^T\mathbf{v})\mathbf{y}^T$ onde $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$ e $\mathbf{v} = (v_1, \ldots, v_m) \in \mathbb{R}^m$ são dois vectores. Denotemos esta aplicação por $\mathbf{u} \otimes \mathbf{v}$.

Temos as seguintes relações:

$$(\mathbf{v}_1 + \mathbf{v}_2) \otimes \mathbf{u} = \mathbf{v}_1 \otimes \mathbf{u} + \mathbf{v}_2 \otimes \mathbf{u}$$
$$\mathbf{v} \otimes (\mathbf{u}_1 + \mathbf{u}_2) = \mathbf{v} \otimes \mathbf{u}_1 + \mathbf{v}_2 \otimes \mathbf{u}_2$$
$$(\lambda\mathbf{v}) \otimes \mathbf{u} = \lambda(\mathbf{v} \otimes \mathbf{u}) = \mathbf{v} \otimes (\lambda\mathbf{u})$$

No algoritmo de Strassen $\mathrm{I}, \ldots, \mathrm{VII}$ são formas bilineares de característica 1:

$$\begin{aligned}\mathrm{I} &= (x+w)\cdot(a+d) = (x+0y+0z+w)\cdot(a+0b+0c+d) = \\ &= (x,y,z,w)(1,0,0,1)^T(1,0,0,1)(a,b,c,d)^T,\end{aligned}$$

$$\begin{aligned}\mathrm{II} &= (z+w)\cdot(a) = (0x+0y+z+w)\cdot(a+0b+0c+0d) = \\ &= (x,y,z,w)(0,0,1,1)^T(1,0,0,0)(a,b,c,d)^T,\end{aligned}$$

$\mathrm{III} =$ etc.

No algoritmo de Strassen $\mathrm{I}, \ldots, \mathrm{VII}$ são formas bilineares de característica 1:

$$\begin{aligned}\mathrm{I} &= (x+w)\cdot(a+d) = (x+0y+0z+w)\cdot(a+0b+0c+d) = \\ &= (x,y,z,w)(1,0,0,1)^T(1,0,0,1)(a,b,c,d)^T,\end{aligned}$$

$$\begin{aligned}\mathrm{II} &= (z+w)\cdot(a) = (0x+0y+z+w)\cdot(a+0b+0c+0d) = \\ &= (x,y,z,w)(0,0,1,1)^T(1,0,0,0)(a,b,c,d)^T,\end{aligned}$$

$$\mathrm{III} = \text{etc.}$$

E dizer que o produto de matrizes pode ser feito com apenas sete multiplicações é equivalente a exibir a fórmula:

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix}\cdot\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \mathrm{I}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \mathrm{II}\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} + \mathrm{III}\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} +$$

$$+\mathrm{IV}\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + \mathrm{V}\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} + \mathrm{VI}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \mathrm{VII}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Problema da multiplicação de complexos

Podemos fazer o mesmo para os números complexos chegando a:

$$x = a(c+d) = (a,b)(1,0)^T(1,1)(c,d)^T$$
$$y = (a+b)d = (a,b)(1,1)^T(0,1)(c,d)^T$$
$$z = b(d-c) = (a,b)(0,1)^T(-1,1)(c,d)^T$$

Problema da multiplicação de complexos

Podemos fazer o mesmo para os números complexos chegando a:

$$x = a(c+d) = (a,b)(1,0)^T(1,1)(c,d)^T$$
$$y = (a+b)d = (a,b)(1,1)^T(0,1)(c,d)^T$$
$$z = b(d-c) = (a,b)(0,1)^T(-1,1)(c,d)^T$$

Dizer que o produto de dois números complexos pode ser realizado com apenas 3 multiplicações, equivale à igualdade:

$$(a+bi)(c+di) = x \cdot 1 + y \cdot (-1+i) + z \cdot (-i)$$

Problema da multiplicação de complexos

Suponhamos que existem $\mathbf{u}, \mathbf{v}, \mathbf{s}, \mathbf{t} \in \mathbb{R}^2$ e $W_1, W_2 \in \mathbb{C}$, tais que:

$$(a+bi)(c+di) = (a,b)(\mathbf{u}^T\mathbf{v})(c,d)^T \cdot W_1 + (a,b)(\mathbf{s}^T\mathbf{t})(c,d)^T \cdot W_2$$

Problema da multiplicação de complexos

Suponhamos que existem $\mathbf{u}, \mathbf{v}, \mathbf{s}, \mathbf{t} \in \mathbb{R}^2$ e $W_1, W_2 \in \mathbb{C}$, tais que:

$$(a+bi)(c+di) = (a,b)(\mathbf{u}^T\mathbf{v})(c,d)^T \cdot W_1 + (a,b)(\mathbf{s}^T\mathbf{t})(c,d)^T \cdot W_2$$

Mas com isto chegamos a uma contradição pois, tomando $0 \neq (a,b) \in \mathcal{L}\{\mathbf{u}\}^\perp$ e $0 \neq (c,d) \in \mathcal{L}\{\mathbf{t}\}^\perp$, obtém-se:

$$\begin{aligned}(a+bi)(c+di) &= (a,b)(\mathbf{u}^T\mathbf{v})(c,d)^T W_1 + (a,b)(\mathbf{s}^T\mathbf{t})(c,d)^T W_2 \\ &= \left((a,b)\mathbf{u}^T\right)\left(\mathbf{v}(c,d)^T\right) W_1 + \left((a,b)\mathbf{s}^T\right)\left(\mathbf{t}(c,d)^T\right) W_2 \\ &= \left[0 \cdot \mathbf{v}(c,d)^T\right] W_1 + \left[(a,b)\mathbf{s}^T \cdot 0\right] W_2 = 0\end{aligned}$$

Problema da multiplicação de complexos

Conclusão

Não existe uma fórmula para a multiplicação de dois números complexos a partir das suas partes real e imaginária que envolva apenas ***duas*** *multiplicações.*

Aplicações bilineares

Definição (Aplicação Bilinear)

Sejam E, V, W espaços vectoriais sobre um corpo. Seja $\varphi\colon E \times V \to W$ uma aplicação que satisfaz as seguintes propriedades:

- $\varphi(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{u}) = \varphi(\mathbf{v}_1, \mathbf{u}) + \varphi(\mathbf{v}_2, \mathbf{u})$
- $\varphi(\mathbf{v}, \mathbf{u}_1 + \mathbf{u}_2) = \varphi(\mathbf{v}, \mathbf{u}_1) + \varphi(\mathbf{v}, \mathbf{u}_2)$
- $\varphi(\lambda\mathbf{v}, \mathbf{u}) = \lambda\varphi(\mathbf{v}, \mathbf{u}) = \varphi(\mathbf{v}, \lambda\mathbf{u})$

então φ diz-se uma aplicação bilinear.

Notação

- Denotemos o espaço das matrizes reais, quadradas de ordem n por $\mathcal{M}_n$ e denotemos por μ_n a aplicação $\mathcal{M}_n \times \mathcal{M}_n \to \mathcal{M}_n$ dada pela multiplicação de matrizes, $(A, B) \mapsto AB$. (Abreviamos μ_2 para μ.)

- Sejam $\mathbf{u}, \mathbf{v}$ vectores de $\mathbb{R}^n$ e $\mathbb{R}^m$ respectivamente e $\mathbf{w}$ um vector de W então $(\mathbf{u} \otimes \mathbf{v})\mathbf{w}$ é a aplicação bilinear de $\mathbb{R}^n \times \mathbb{R}^m \to W$ dada por $(\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x}(\mathbf{u}^T\mathbf{v})\mathbf{y}^T)\mathbf{w}$.

Exemplos

- A aplicação $\mu_n \colon \mathcal{M}_n \times \mathcal{M}_n \longrightarrow \mathcal{M}_n$ é bilinear pois:

$$A(B+C) = AB + AC$$
$$(A+B)C = AC + BC$$
$$(\lambda A)B = \lambda(AB) = A(\lambda B)$$

Exemplos

- A aplicação $\mu_n \colon \mathcal{M}_n \times \mathcal{M}_n \longrightarrow \mathcal{M}_n$ é bilinear pois:

$$\begin{aligned} A(B+C) &= AB + AC \\ (A+B)C &= AC + BC \\ (\lambda A)B &= \lambda(AB) = A(\lambda B) \end{aligned}$$

- Encarando $\mathbb{C}$ como espaço real (de dimensão 2), a aplicação $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$ dada pelo produto de dois números complexos é bilinear.

Exemplos

- A aplicação $\mu_n \colon \mathcal{M}_n \times \mathcal{M}_n \longrightarrow \mathcal{M}_n$ é bilinear pois:

$$\begin{aligned} A(B+C) &= AB + AC \\ (A+B)C &= AC + BC \\ (\lambda A)B &= \lambda(AB) = A(\lambda B) \end{aligned}$$

- Encarando $\mathbb{C}$ como espaço real (de dimensão 2), a aplicação $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$ dada pelo produto de dois números complexos é bilinear.
- A aplicação $\varphi \colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ dada por $\varphi = \sum(\mathbf{u}_i \otimes \mathbf{v}_i)\mathbf{z}_i$ onde $\mathbf{u}_i \in \mathbb{R}^n$, $\mathbf{v}_i \in \mathbb{R}^m$ e $\mathbf{z}_i \in \mathbb{R}^p$ é bilinear.

Teorema

Seja $\varphi\colon \mathbb{R}^n \times \mathbb{R}^m \to W$ *uma aplicação bilinear. Então existem* $\mathbf{u}_i \in \mathbb{R}^n$, $\mathbf{v}_j \in \mathbb{R}^m$, $\mathbf{w}_{ij} \in W$ *tais que* $\varphi = \sum_{i,j}(\mathbf{u}_i \otimes \mathbf{v}_j)\mathbf{w}_{ij}$.

Teorema

Seja $\varphi\colon \mathbb{R}^n \times \mathbb{R}^m \to W$ uma aplicação bilinear. Então existem $\mathbf{u}_i \in \mathbb{R}^n$, $\mathbf{v}_j \in \mathbb{R}^m$, $\mathbf{w}_{ij} \in W$ tais que $\varphi = \sum_{i,j}(\mathbf{u}_i \otimes \mathbf{v}_j)\mathbf{w}_{ij}$.

Sejam (u_i) e (v_j) bases ortonormadas de $\mathbb{R}^n$ e $\mathbb{R}^m$ respectivamente.

- $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{y} \in \mathbb{R}^m \implies \mathbf{x} = \sum \alpha_i \mathbf{u}_i$ e $\mathbf{y} = \sum \beta_j \mathbf{v}_j$.

Teorema

Seja $\varphi\colon \mathbb{R}^n \times \mathbb{R}^m \to W$ uma aplicação bilinear. Então existem $\mathbf{u}_i \in \mathbb{R}^n$, $\mathbf{v}_j \in \mathbb{R}^m$, $\mathbf{w}_{ij} \in W$ tais que $\varphi = \sum_{i,j}(\mathbf{u}_i \otimes \mathbf{v}_j)\mathbf{w}_{ij}$.

Sejam (u_i) e (v_j) bases ortonormadas de $\mathbb{R}^n$ e $\mathbb{R}^m$ respectivamente.

- $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{y} \in \mathbb{R}^m \implies \mathbf{x} = \sum \alpha_i \mathbf{u}_i$ e $\mathbf{y} = \sum \beta_j \mathbf{v}_j$.
- (u_i) e (v_j) bases ortonormadas $\implies \mathbf{x}\mathbf{u}_i^T = \alpha_i$ e $\mathbf{v}_j\mathbf{y}^T = \beta_j$.

Teorema

Seja $\varphi\colon \mathbb{R}^n \times \mathbb{R}^m \to W$ uma aplicação bilinear. Então existem $\mathbf{u}_i \in \mathbb{R}^n$, $\mathbf{v}_j \in \mathbb{R}^m$, $\mathbf{w}_{ij} \in W$ tais que $\varphi = \sum_{i,j}(\mathbf{u}_i \otimes \mathbf{v}_j)\mathbf{w}_{ij}$.

Sejam (u_i) e (v_j) bases ortonormadas de $\mathbb{R}^n$ e $\mathbb{R}^m$ respectivamente.

- $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{y} \in \mathbb{R}^m \implies \mathbf{x} = \sum \alpha_i \mathbf{u}_i$ e $\mathbf{y} = \sum \beta_j \mathbf{v}_j$.
- (u_i) e (v_j) bases ortonormadas $\implies \mathbf{x}\mathbf{u}_i^T = \alpha_i$ e $\mathbf{v}_j\mathbf{y}^T = \beta_j$.
- φ bilinear $\implies \varphi(\mathbf{x}, \mathbf{y}) = \sum_{i,j} \alpha_i \beta_j \varphi(\mathbf{u}_i, \mathbf{v}_j)$.

Teorema

Seja $\varphi\colon \mathbb{R}^n \times \mathbb{R}^m \to W$ uma aplicação bilinear. Então existem $\mathbf{u}_i \in \mathbb{R}^n$, $\mathbf{v}_j \in \mathbb{R}^m$, $\mathbf{w}_{ij} \in W$ tais que $\varphi = \sum_{i,j}(\mathbf{u}_i \otimes \mathbf{v}_j)\mathbf{w}_{ij}$.

Sejam (u_i) e (v_j) bases ortonormadas de $\mathbb{R}^n$ e $\mathbb{R}^m$ respectivamente.

- $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{y} \in \mathbb{R}^m \implies \mathbf{x} = \sum \alpha_i \mathbf{u}_i$ e $\mathbf{y} = \sum \beta_j \mathbf{v}_j$.
- (u_i) e (v_j) bases ortonormadas $\implies \mathbf{x}\mathbf{u}_i^T = \alpha_i$ e $\mathbf{v}_j\mathbf{y}^T = \beta_j$.
- φ bilinear $\implies \varphi(\mathbf{x}, \mathbf{y}) = \sum_{i,j} \alpha_i\beta_j\varphi(\mathbf{u}_i, \mathbf{v}_j)$.
- Assim, $\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i,j}(\mathbf{x}\mathbf{u}_i^T)(\mathbf{v}_j^T\mathbf{y})\varphi(\mathbf{u}_i, \mathbf{v}_j)$.

Teorema

Seja $\varphi\colon \mathbb{R}^n \times \mathbb{R}^m \to W$ uma aplicação bilinear. Então existem $\mathbf{u}_i \in \mathbb{R}^n$, $\mathbf{v}_j \in \mathbb{R}^m$, $\mathbf{w}_{ij} \in W$ tais que $\varphi = \sum_{i,j}(\mathbf{u}_i \otimes \mathbf{v}_j)\mathbf{w}_{ij}$.

Sejam (u_i) e (v_j) bases ortonormadas de $\mathbb{R}^n$ e $\mathbb{R}^m$ respectivamente.

- $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{y} \in \mathbb{R}^m \implies \mathbf{x} = \sum \alpha_i \mathbf{u}_i$ e $\mathbf{y} = \sum \beta_j \mathbf{v}_j$.
- (u_i) e (v_j) bases ortonormadas $\implies \mathbf{x}\mathbf{u}_i^T = \alpha_i$ e $\mathbf{v}_j\mathbf{y}^T = \beta_j$.
- φ bilinear $\implies \varphi(\mathbf{x}, \mathbf{y}) = \sum_{i,j} \alpha_i \beta_j \varphi(\mathbf{u}_i, \mathbf{v}_j)$.
- Assim, $\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i,j}(\mathbf{x}\mathbf{u}_i^T)(\mathbf{v}_j^T\mathbf{y})\varphi(\mathbf{u}_i, \mathbf{v}_j)$.
- Conclui-se que $\varphi = \sum_{i,j}(\mathbf{u}_i \otimes \mathbf{v}_j)\mathbf{w}_{ij}$ com $\mathbf{w}_{ij} = \varphi(\mathbf{u}_i, \mathbf{v}_j) \in W$.

Característica de uma aplicação bilinear

Definição (Característica)

Seja $\varphi\colon \mathbb{R}^n \times \mathbb{R}^m \to W$ uma aplicação bilinear. A característica de φ é o menor natural k para o qual existem $\mathbf{u}_1, \ldots, \mathbf{u}_k \in \mathbb{R}^n$, $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^m$ e $\mathbf{w}_1, \ldots, \mathbf{w}_k \in W$ tais que $\varphi = \sum_{i=1}^{k} (\mathbf{u}_i \otimes \mathbf{v}_i)\mathbf{w}_i$

(No caso da forma bilinear a característica de φ coincide com a característica da matriz que a representa.)

Exemplos

- $(\mathbf{u} \otimes \mathbf{v})\mathbf{w}$ tem característica ≤ 1.
- A multiplicação de complexos tem característica 3.
- O produto interno de $\mathbb{R}^n$ tem característica n.
- μ tem característica ≤ 7.

Exemplos

- $(\mathbf{u} \otimes \mathbf{v})\mathbf{w}$ tem característica ≤ 1.
- A multiplicação de complexos tem característica 3.
- O produto interno de $\mathbb{R}^n$ tem característica n.
- μ tem característica ≤ 7.

$$\left[\text{Identificando } \mathcal{M}_2 \text{ com } \mathbb{R}^4 \text{ através de } \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \mapsto (a,b,c,d)\right]$$

$$\mu = \mathbf{u}_1 \otimes \mathbf{v}_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \mathbf{u}_2 \otimes \mathbf{v}_2 \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} + \mathbf{u}_3 \otimes \mathbf{v}_3 \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} +$$

$$+ \mathbf{u}_4 \otimes \mathbf{v}_4 \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + \mathbf{u}_5 \otimes \mathbf{v}_5 \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} + \mathbf{u}_6 \otimes \mathbf{v}_6 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \mathbf{u}_7 \otimes \mathbf{v}_7 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

$$\begin{array}{llll} \mathbf{u}_1 = (1,0,0,1) & & \mathbf{v}_1 = (1,0,0,1) & \\ \mathbf{u}_2 = (0,0,1,1) & \mathbf{u}_5 = (1,1,0,0) & \mathbf{v}_2 = (1,0,0,0) & \mathbf{v}_5 = (0,0,0,1) \\ \mathbf{u}_3 = (1,0,0,0) & \mathbf{u}_6 = (-1,0,1,0) & \mathbf{v}_3 = (0,1,0,-1) & \mathbf{v}_6 = (1,1,0,0) \\ \mathbf{u}_4 = (1,0,0,1) & \mathbf{u}_7 = (0,1,0,-1) & \mathbf{v}_4 = (-1,0,1,0) & \mathbf{v}_7 = (0,0,1,1) \end{array}$$

Problema

Qual a característica de μ?

Se característica de μ for 6, então existem seis pares de vectores $\mathbf{u}_i, \mathbf{v}_i \in \mathbb{R}^4$ e seis matrizes $A_i \in \mathcal{M}_2$ tais que

$$\mu = (\mathbf{u}_1 \otimes \mathbf{v}_1)A_1 + (\mathbf{u}_2 \otimes \mathbf{v}_2)A_2 + \cdots + (\mathbf{u}_6 \otimes \mathbf{v}_6)A_6$$

Problema

Qual a característica de μ?

Se característica de μ for 6, então existem seis pares de vectores $\mathbf{u}_i, \mathbf{v}_i \in \mathbb{R}^4$ e seis matrizes $A_i \in \mathcal{M}_2$ tais que

$$\mu = (\mathbf{u}_1 \otimes \mathbf{v}_1)A_1 + (\mathbf{u}_2 \otimes \mathbf{v}_2)A_2 + \cdots + (\mathbf{u}_6 \otimes \mathbf{v}_6)A_6$$

Aplicando a ideia usada no caso dos complexos, tomemos $0 \neq \mathbf{a} \in \mathcal{L}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}^\perp$ e $0 \neq \mathbf{b} \in \mathcal{L}\{\mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6\}^\perp$ e assim

$$\mu(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{3} \mathbf{a}(\mathbf{u}_i^T \mathbf{v}_i)\mathbf{b}^T A_i + \sum_{i=4}^{6} \mathbf{a}(\mathbf{u}_i^T \mathbf{v}_i)\mathbf{b}^T A_i = 0 + 0 = 0.$$

Problema

Qual a característica de μ?

Se característica de μ for 6, então existem seis pares de vectores $\mathbf{u}_i, \mathbf{v}_i \in \mathbb{R}^4$ e seis matrizes $A_i \in \mathcal{M}_2$ tais que

$$\mu = (\mathbf{u}_1 \otimes \mathbf{v}_1)A_1 + (\mathbf{u}_2 \otimes \mathbf{v}_2)A_2 + \cdots + (\mathbf{u}_6 \otimes \mathbf{v}_6)A_6$$

Aplicando a ideia usada no caso dos complexos, tomemos $0 \neq \mathbf{a} \in \mathcal{L}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}^\perp$ e $0 \neq \mathbf{b} \in \mathcal{L}\{\mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6\}^\perp$ e assim

$$\mu(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{3} \mathbf{a}(\mathbf{u}_i^T \mathbf{v}_i)\mathbf{b}^T A_i + \sum_{i=4}^{6} \mathbf{a}(\mathbf{u}_i^T \mathbf{v}_i)\mathbf{b}^T A_i = 0 + 0 = 0.$$

Mas isto não nos permite concluir pois, por exemplo,

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

O ideal gerado por uma matriz

Definição

Seja $\mathbf{a} = (a_1, a_2, a_3, a_4)$. Define-se o conjunto $\mathbf{a}\mathcal{M}_2$ através de

$$\mathbf{a}\mathcal{M}_2 = \left\{ \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} B : B \in \mathcal{M}_2 \right\}.$$

- $\mathbf{a}\mathcal{M}_2$ é um subespaço de $\mathcal{M}_2$.

O ideal gerado por uma matriz

Definição

Seja $\mathbf{a} = (a_1, a_2, a_3, a_4)$. Define-se o conjunto $\mathbf{a}\mathcal{M}_2$ através de

$$\mathbf{a}\mathcal{M}_2 = \left\{ \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} B : B \in \mathcal{M}_2 \right\}.$$

- $\mathbf{a}\mathcal{M}_2$ é um subespaço de $\mathcal{M}_2$.
- $\mathbf{a}\mathcal{M}_2$ é um ideal direito de $\mathcal{M}_2$, pois

$$C \in \mathbf{a}\mathcal{M}_2 \text{ e } B \in \mathcal{M}_2 \implies CB \in \mathbf{a}\mathcal{M}_2$$

O ideal gerado por uma matriz

Definição

Seja $\mathbf{a} = (a_1, a_2, a_3, a_4)$. Define-se o conjunto $\mathbf{a}\mathcal{M}_2$ através de

$$\mathbf{a}\mathcal{M}_2 = \left\{ \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} B : B \in \mathcal{M}_2 \right\}.$$

- $\mathbf{a}\mathcal{M}_2$ é um subespaço de $\mathcal{M}_2$.
- $\mathbf{a}\mathcal{M}_2$ é um ideal direito de $\mathcal{M}_2$, pois

$$C \in \mathbf{a}\mathcal{M}_2 \text{ e } B \in \mathcal{M}_2 \implies CB \in \mathbf{a}\mathcal{M}_2$$

Dado $\mathbf{b} \in \mathbb{R}^4$, define-se o ideal (esquerdo), $\mathcal{M}_2\mathbf{b}$, de forma análoga.

Demonstração

Supomos que μ tem característica 6:

$$\mu = (\mathbf{u}_1 \otimes \mathbf{v}_1)A_1 + (\mathbf{u}_2 \otimes \mathbf{v}_2)A_2 + \cdots + (\mathbf{u}_6 \otimes \mathbf{v}_6)A_6$$

- $\mathcal{L}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_6\} = \mathbb{R}^4$ e assim podemos supor, sem perda de generalidade, que $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ formam uma base de $\mathbb{R}^4$.

Demonstração

Supomos que μ tem característica 6:

$$\mu = (\mathbf{u}_1 \otimes \mathbf{v}_1)A_1 + (\mathbf{u}_2 \otimes \mathbf{v}_2)A_2 + \cdots + (\mathbf{u}_6 \otimes \mathbf{v}_6)A_6$$

- $\mathcal{L}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_6\} = \mathbb{R}^4$ e assim podemos supor, sem perda de generalidade, que $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ formam uma base de $\mathbb{R}^4$.
- Seja $0 \neq \mathbf{b} \in \mathcal{L}\{\mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6\}^\perp$ então para todo $\mathbf{a} \in \mathbb{R}^4$,

$$\begin{aligned} \mu(\mathbf{a}, \mathbf{b}) &= \textstyle\sum_{i=1}^{3} \mathbf{a}(\mathbf{u}_i^T \mathbf{v}_i)\mathbf{b}^T A_i + \sum_{i=4}^{6} \mathbf{a}(\mathbf{u}_i^T \mathbf{v}_i)\mathbf{b}^T A_i = \\ &= \mathbf{a}(\mathbf{u}_1^T \mathbf{v}_1)\mathbf{b}^T A_1 + \mathbf{a}(\mathbf{u}_2^T \mathbf{v}_2)\mathbf{b}^T A_2 + \mathbf{a}(\mathbf{u}_3^T \mathbf{v}_3)\mathbf{b}^T A_3. \end{aligned}$$

Logo $\mathcal{M}_2\mathbf{b} \subset \mathcal{L}\{A_1, A_2, A_3\} \neq \mathcal{M}_2$.

Demonstração

Supomos que μ tem característica 6:

$$\mu = (\mathbf{u}_1 \otimes \mathbf{v}_1)A_1 + (\mathbf{u}_2 \otimes \mathbf{v}_2)A_2 + \cdots + (\mathbf{u}_6 \otimes \mathbf{v}_6)A_6$$

- $\mathcal{L}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_6\} = \mathbb{R}^4$ e assim podemos supor, sem perda de generalidade, que $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ formam uma base de $\mathbb{R}^4$.
- Seja $0 \neq \mathbf{b} \in \mathcal{L}\{\mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6\}^{\perp}$ então para todo $\mathbf{a} \in \mathbb{R}^4$,

$$\begin{aligned} \mu(\mathbf{a}, \mathbf{b}) &= \textstyle\sum_{i=1}^{3} \mathbf{a}(\mathbf{u}_i^T \mathbf{v}_i)\mathbf{b}^T A_i + \sum_{i=4}^{6} \mathbf{a}(\mathbf{u}_i^T \mathbf{v}_i)\mathbf{b}^T A_i = \\ &= \mathbf{a}(\mathbf{u}_1^T \mathbf{v}_1)\mathbf{b}^T A_1 + \mathbf{a}(\mathbf{u}_2^T \mathbf{v}_2)\mathbf{b}^T A_2 + \mathbf{a}(\mathbf{u}_3^T \mathbf{v}_3)\mathbf{b}^T A_3. \end{aligned}$$

  Logo $\mathcal{M}_2\mathbf{b} \subset \mathcal{L}\{A_1, A_2, A_3\} \neq \mathcal{M}_2$.
- Conclui-se que $\mathcal{M}_2\mathbf{b}$ tem dimensão 2.

Demonstração. Se $\mu = (\mathbf{u_1} \otimes \mathbf{v_1})A_1 + (\mathbf{u_2} \otimes \mathbf{v_2})A_2 + \cdots + (\mathbf{u_6} \otimes \mathbf{v_6})A_6$...

Teorema

Se $\mathbf{a}, \mathbf{b} \neq 0$ *então a dimensão de* $\mathbf{a}\mathcal{M}_2$ *e de* $\mathcal{M}_2\mathbf{b}$ *é* 2 *ou* 4.

Demonstração. Se $\mu = (\mathbf{u_1} \otimes \mathbf{v_1})A_1 + (\mathbf{u_2} \otimes \mathbf{v_2})A_2 + \cdots + (\mathbf{u_6} \otimes \mathbf{v_6})A_6$...

Teorema

Se $\mathbf{a}, \mathbf{b} \neq 0$ *então a dimensão de* $\mathbf{a}\mathcal{M}_2$ *e de* $\mathcal{M}_2\mathbf{b}$ *é* 2 *ou* 4.

- De entre $\mathbf{v}_1\mathbf{b}^T, \mathbf{v}_2\mathbf{b}^T, \mathbf{v}_3\mathbf{b}^T$ existem dois não nulos.

$$\left[\begin{array}{l} \text{Suponhamos que } \mathbf{v_1}\mathbf{b}^T = \mathbf{v_2}\mathbf{b}^T = 0. \text{ Então de } \mathcal{M}_2\mathbf{b} \subset \mathcal{L}\{A_3\} \\ \text{deduz-se que } \dim \mathcal{M}_2\mathbf{b} = 1, \text{ contradição.} \end{array}\right]$$

Suponhamos que $\mathbf{v}_1\mathbf{b}^T, \mathbf{v}_2\mathbf{b}^T \neq 0$.

Demonstração. Se $\mu = (\mathbf{u_1} \otimes \mathbf{v_1})A_1 + (\mathbf{u_2} \otimes \mathbf{v_2})A_2 + \cdots + (\mathbf{u_6} \otimes \mathbf{v_6})A_6$...

Teorema

Se $\mathbf{a}, \mathbf{b} \neq 0$ *então a dimensão de* $\mathbf{a}\mathcal{M}_2$ *e de* $\mathcal{M}_2\mathbf{b}$ *é* 2 *ou* 4.

- De entre $\mathbf{v}_1\mathbf{b}^T, \mathbf{v}_2\mathbf{b}^T, \mathbf{v}_3\mathbf{b}^T$ existem dois não nulos.

$$\left[\begin{array}{l} \text{Suponhamos que } \mathbf{v}_1\mathbf{b}^T = \mathbf{v}_2\mathbf{b}^T = 0. \text{ Então de } \mathcal{M}_2\mathbf{b} \subset \mathcal{L}\{A_3\} \\ \text{deduz-se que } \dim \mathcal{M}_2\mathbf{b} = 1, \text{ contradição.} \end{array} \right]$$

Suponhamos que $\mathbf{v}_1\mathbf{b}^T, \mathbf{v}_2\mathbf{b}^T \neq 0$.

- $\mathcal{L}\{A_1, A_2\} \subset \mathcal{M}_2\mathbf{b}$.

$$\left[\begin{array}{l} \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \text{ lin. indep.} \implies \mathcal{L}\{\mathbf{u}_2, \mathbf{u}_3\}^\perp \not\subset \mathcal{L}\{\mathbf{u}_1\}^\perp \text{ logo} \\ \exists\, \mathbf{c} \in \mathcal{L}\{\mathbf{u}_2, \mathbf{u}_3\}^\perp \setminus \mathcal{L}\{\mathbf{u}_1\}^\perp \text{ e assim } \mathcal{M}_2\mathbf{b} \ni (\mathbf{c}\mathbf{u}_1^T\mathbf{v}_1\mathbf{b}^T)A_1 = \mu(\mathbf{c}, \mathbf{b}), \text{ etc.} \end{array} \right]$$

Demonstração. Se $\mu = (\mathbf{u_1} \otimes \mathbf{v_1})A_1 + (\mathbf{u_2} \otimes \mathbf{v_2})A_2 + \cdots + (\mathbf{u_6} \otimes \mathbf{v_6})A_6$...

- $\mathcal{L}\{\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_6\} = \mathbb{R}^4$.

Demonstração. Se $\mu = (\mathbf{u_1} \otimes \mathbf{v_1})A_1 + (\mathbf{u_2} \otimes \mathbf{v_2})A_2 + \cdots + (\mathbf{u_6} \otimes \mathbf{v_6})A_6$...

- $\mathcal{L}\{\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_6\} = \mathbb{R}^4$.
- Se $\mathcal{M}_2\mathbf{b} \subset \mathcal{L}\{v_3, v_4, v_5, v_6\}^\perp \implies \mathcal{L}\{v_3, v_4, v_5, v_6\} \subset \mathcal{M}_2\mathbf{b}^\perp$.

Demonstração. Se $\mu = (\mathbf{u_1} \otimes \mathbf{v_1})A_1 + (\mathbf{u_2} \otimes \mathbf{v_2})A_2 + \cdots + (\mathbf{u_6} \otimes \mathbf{v_6})A_6$...

- $\mathcal{L}\{\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_6\} = \mathbb{R}^4$.
- Se $\mathcal{M}_2\mathbf{b} \subset \mathcal{L}\{v_3, v_4, v_5, v_6\}^\perp \implies \mathcal{L}\{v_3, v_4, v_5, v_6\} \subset \mathcal{M}_2\mathbf{b}^\perp$.
- $\exists\, v_j \in \{v_3, v_4, v_5, v_6\}$ tal que $\mathcal{M}_2\mathbf{b} \not\subset \mathcal{L}\{v_j\}^\perp$.

Suponhamos que $\mathcal{M}_2\mathbf{b} \not\subset \mathcal{L}\{v_3\}^\perp$. Seja $0 \neq \mathbf{a} \in \mathcal{L}\{\mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_6\}^\perp$. Então $\mathbf{a}\mathcal{M}_2 \subset \mathcal{L}\{A_1, A_2, A_3\}$.

Demonstração. Se $\mu = (\mathbf{u_1} \otimes \mathbf{v_1})A_1 + (\mathbf{u_2} \otimes \mathbf{v_2})A_2 + \cdots + (\mathbf{u_6} \otimes \mathbf{v_6})A_6$...

- $\mathcal{L}\{\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_6\} = \mathbb{R}^4$.
- Se $\mathcal{M}_2\mathbf{b} \subset \mathcal{L}\{v_3, v_4, v_5, v_6\}^\perp \implies \mathcal{L}\{v_3, v_4, v_5, v_6\} \subset \mathcal{M}_2\mathbf{b}^\perp$.
- $\exists\, v_j \in \{v_3, v_4, v_5, v_6\}$ tal que $\mathcal{M}_2\mathbf{b} \not\subset \mathcal{L}\{v_j\}^\perp$.

Suponhamos que $\mathcal{M}_2\mathbf{b} \not\subset \mathcal{L}\{v_3\}^\perp$. Seja $0 \neq \mathbf{a} \in \mathcal{L}\{\mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_6\}^\perp$. Então $\mathbf{a}\mathcal{M}_2 \subset \mathcal{L}\{A_1, A_2, A_3\}$.

Teorema

Se $\mathbf{a} \neq 0$ *e* $\mathbf{a}\mathcal{M}_2 \subset \mathcal{M}_2\mathbf{b}$ *então* $\mathcal{M}_2\mathbf{b} = \mathcal{M}_2$.

Demonstração. Se $\mu = (\mathbf{u_1} \otimes \mathbf{v_1})A_1 + (\mathbf{u_2} \otimes \mathbf{v_2})A_2 + \cdots + (\mathbf{u_6} \otimes \mathbf{v_6})A_6$...

- $\mathcal{L}\{\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_6\} = \mathbb{R}^4$.
- Se $\mathcal{M}_2\mathbf{b} \subset \mathcal{L}\{v_3, v_4, v_5, v_6\}^\perp \implies \mathcal{L}\{v_3, v_4, v_5, v_6\} \subset \mathcal{M}_2\mathbf{b}^\perp$.
- $\exists\, v_j \in \{v_3, v_4, v_5, v_6\}$ tal que $\mathcal{M}_2\mathbf{b} \not\subset \mathcal{L}\{v_j\}^\perp$.

Suponhamos que $\mathcal{M}_2\mathbf{b} \not\subset \mathcal{L}\{v_3\}^\perp$. Seja $0 \neq \mathbf{a} \in \mathcal{L}\{\mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_6\}^\perp$. Então $\mathbf{a}\mathcal{M}_2 \subset \mathcal{L}\{A_1, A_2, A_3\}$.

Teorema

Se $\mathbf{a} \neq 0$ *e* $\mathbf{a}\mathcal{M}_2 \subset \mathcal{M}_2\mathbf{b}$ *então* $\mathcal{M}_2\mathbf{b} = \mathcal{M}_2$.

- Para cada $\mathbf{x} \in \mathbb{R}^4$ existe $\mathbf{y} \in \mathcal{M}_2\mathbf{b}$ tal que $\mathbf{v}_3\mathbf{x}^T = \mathbf{v}_3\mathbf{y}^T$.
- $\mu(\mathbf{a}, \mathbf{x}) - \mu(\mathbf{a}, \mathbf{y}) \in \mathcal{L}\{A_1, A_2\} \subset \mathcal{M}_2\mathbf{b}$
- Logo $\mathbf{a}\mathcal{M}_2 \subset \mathcal{M}_2\mathbf{b} + \mathcal{L}\{A_1, A_2\} = \mathcal{M}_2\mathbf{b}$.

Demonstração. Se $\mu = (\mathbf{u_1} \otimes \mathbf{v_1})A_1 + (\mathbf{u_2} \otimes \mathbf{v_2})A_2 + \cdots + (\mathbf{u_6} \otimes \mathbf{v_6})A_6$...

- $\mathcal{L}\{\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_6\} = \mathbb{R}^4$.
- Se $\mathcal{M}_2\mathbf{b} \subset \mathcal{L}\{v_3, v_4, v_5, v_6\}^\perp \implies \mathcal{L}\{v_3, v_4, v_5, v_6\} \subset \mathcal{M}_2\mathbf{b}^\perp$.
- $\exists\, v_j \in \{v_3, v_4, v_5, v_6\}$ tal que $\mathcal{M}_2\mathbf{b} \not\subset \mathcal{L}\{v_j\}^\perp$.

Suponhamos que $\mathcal{M}_2\mathbf{b} \not\subset \mathcal{L}\{v_3\}^\perp$. Seja $0 \neq \mathbf{a} \in \mathcal{L}\{\mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_6\}^\perp$. Então $\mathbf{a}\mathcal{M}_2 \subset \mathcal{L}\{A_1, A_2, A_3\}$.

Teorema

Se $\mathbf{a} \neq 0$ *e* $\mathbf{a}\mathcal{M}_2 \subset \mathcal{M}_2\mathbf{b}$ *então* $\mathcal{M}_2\mathbf{b} = \mathcal{M}_2$.

- Para cada $\mathbf{x} \in \mathbb{R}^4$ existe $\mathbf{y} \in \mathcal{M}_2\mathbf{b}$ tal que $\mathbf{v}_3\mathbf{x}^T = \mathbf{v}_3\mathbf{y}^T$.
- $\mu(\mathbf{a}, \mathbf{x}) - \mu(\mathbf{a}, \mathbf{y}) \in \mathcal{L}\{A_1, A_2\} \subset \mathcal{M}_2\mathbf{b}$
- Logo $\mathbf{a}\mathcal{M}_2 \subset \mathcal{M}_2\mathbf{b} + \mathcal{L}\{A_1, A_2\} = \mathcal{M}_2\mathbf{b}$. **Contradição.**

Outros Resultados

- *A característica de μ_n é maior ou igual a $3n^2 - 3n + 1$.*
- *Em particular, a característica de μ_3 é maior ou igual a 19.*
- *Existe uma fórmula para μ_3 que usa 23 multiplicações.*