Teorema da Triangularização de Schur e Diagonalização de Matrizes Normais

Reginaldo J. Santos
Departamento de Matemática-ICEx
Universidade Federal de Minas Gerais
http://www.mat.ufmg.br/~regi

16 de novembro de 2011

1 Resultados Preliminares

Vamos chamar o conjunto das matrizes $m \times n$ cujas entradas são números complexos de $\mathcal{M}_{mn}(\mathbb{C})$. Para uma matriz $A = (a_{ij}) \in \mathcal{M}_{mn}(\mathbb{C})$, definimos o **conjugado da matriz** A, denotado por $\overline{A}$ como sendo a matriz $B = (b_{ij}) \in \mathcal{M}_{mn}(\mathbb{C})$ dada por $b_{ij} = \bar{a}_{ij}$, em que, se $a_{ij} = \alpha_{ij} + i\beta_{ij}$, então $\bar{a}_{ij} = \alpha_{ij} - i\beta_{ij}$.

Para as matrizes de $\mathcal{M}_{mn}(\mathbb{C})$ além das propriedades que são válidas para matrizes com entradas que são números reais, são válidas as seguintes propriedades, cuja demonstração deixamos a cargo do leitor:

(p) Se $A \in \mathcal{M}_{mp}(\mathbb{C})$ e $B \in \mathcal{M}_{pn}(\mathbb{C})$, então

$$\overline{AB} = \overline{A}\,\overline{B}.$$

(q) Se $A \in \mathcal{M}_{mn}(\mathbb{C})$ e $\alpha \in \mathbb{C}$, então

$$\overline{\alpha A} = \bar{\alpha}\overline{B}.$$

Definição 1. Para cada inteiro positivo n, o **espaço (vetorial)** $\mathbb{C}^n$ é definido pelo conjunto de todas as n-uplas ordenadas $X = (x_1, \ldots, x_n)$ de números reais.

Definição 2. (a) Definimos o **produto escalar ou interno** de dois vetores $X = (x_1, \ldots, x_n) \in \mathbb{C}^n$ e $Y = (y_1, \ldots, y_n) \in \mathbb{C}^n$ por

$$X \cdot Y = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n = \sum_{i=1}^{n} x_i\bar{y}_i\,.$$

(b) Definimos a **norma** de um vetor $X = (x_1, \ldots, x_n) \in \mathbb{C}^n$ por

$$||X|| = \sqrt{X \cdot X} = \sqrt{|x_1|^2 + \ldots + |x_n|^2} = \sqrt{\sum_{i=1}^{n} |x_i|^2}\,.$$

Escrevendo os vetores como matrizes colunas, o produto interno de dois vetores

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

pertencentes a $\mathbb{C}^n$ pode ser escrito em termos do produto de matrizes como

$$X \cdot Y = X^t \overline{Y}.$$

São válidas as seguintes propriedades para o produto escalar e a norma de vetores de $\mathbb{C}^n$ que são análogas àquelas que são válidas para vetores de $\mathbb{R}^n$.

Proposição 1. *Se X, Y e Z são vetores de $\mathbb{C}^n$ e $\alpha \in \mathbb{C}$, então*

(a) $X \cdot Y = \overline{Y \cdot X}$;

(b) $(X + Y) \cdot Z = X \cdot Z + Y \cdot Z$ *e* $X \cdot (Y + Z) = X \cdot Y + X \cdot Z$ *(distributividade em relação à soma);*

(c) $(\alpha X) \cdot Y = \alpha(X \cdot Y)$ *e* $X \cdot (\alpha Y) = \overline{\alpha}(X \cdot Y)$;

(d) $X \cdot X = ||X||^2 \geq 0$ *e* $||X|| = 0$ *se, e somente se,* $X = \bar{0}$;

(e) $||\alpha X|| = |\alpha| \, ||X||$;

(f) $|X \cdot Y| \leq ||X|| ||Y||$ *(desigualdade de Cauchy-Schwarz);*

(g) $||X + Y|| \leq ||X|| + ||Y||$ *(desigualdade triangular).*

Demonstração. Sejam $X, Y, Z \in \mathbb{C}^n$ e $\alpha \in \mathbb{C}$. Usando o fato de que se os vetores são escritos como matrizes colunas, então o produto escalar pode ser escrito como o produto de matrizes, $X \cdot Y = X^t \overline{Y}$, e as propriedades da álgebra matricial, temos que

(a) $X \cdot Y = x_1 \overline{y}_1 + x_2 \overline{y}_2 + \ldots + x_n \overline{y}_n = \overline{y_1 \overline{x}_1 + \cdots + y_n \overline{x}_n} = \overline{Y \cdot X}$.

(b) $(X + Y) \cdot Z = (X + Y)^t \overline{Z} = X^t \overline{Z} + Y^t \overline{Z} = X \cdot Z + Y \cdot Z$ e
$X \cdot (Y + Z) = X^t (\overline{Y + Z}) = X^t \overline{Y} + X^t \overline{Z} = X \cdot Y + X \cdot Z$.

(c) $\alpha(X \cdot Y) = \alpha(X^t\overline{Y}) = (\alpha X^t)\overline{Y} = (\alpha X)^t\overline{Y} = (\alpha X) \cdot Y$ e
$\alpha(X \cdot Y) = \alpha(X^t\overline{Y}) = X^t(\alpha\overline{Y}) = X^t(\overline{\overline{\alpha}Y}) = X \cdot (\overline{\alpha}Y)$

(d) $X \cdot X$ é uma soma de quadrados, por isso é sempre maior ou igual a zero e é zero se, e somente se, todas as parcelas são iguais a zero.

(e) $||\alpha X||^2 = |\alpha x_1|^2 + \cdots + |\alpha x_n|^2 = |\alpha|^2(|x_1|^2 + \cdots + |x_n|^2) = |\alpha|^2||X||^2$. Tomando a raiz quadrada, segue-se o resultado.

(f) A norma de $X - \lambda Y$ é maior ou igual a zero, para qualquer λ complexo. Assim,

$$0 \leq ||X - \lambda Y||^2 = (X - \lambda Y) \cdot (X - \lambda Y) = ||X||^2 - \lambda Y \cdot X - \overline{\lambda} X \cdot Y + |\lambda|^2 ||Y||^2,$$

para qualquer λ complexo. Tomando $\lambda = \dfrac{X \cdot Y}{||Y||^2}$ obtemos

$$0 \leq ||X||^2 - \frac{X \cdot Y}{||Y||^2} Y \cdot X - \frac{Y \cdot X}{||Y||^2} X \cdot Y + \frac{|X \cdot Y|^2}{||Y||^4} ||Y||^2 = ||X||^2 - \frac{|X \cdot Y|^2}{||Y||^2}$$

Logo,

$$|X \cdot Y| \leq ||X||\,||Y||.$$

(g) Pelo item anterior temos que

$$\begin{aligned} ||X + Y||^2 &= (X + Y) \cdot (X + Y) = ||X||^2 + X \cdot Y + Y \cdot X + ||Y||^2 \\ &\leq ||X||^2 + 2\Re(X \cdot Y) + ||Y||^2 \\ &\leq ||X||^2 + 2|X \cdot Y| + ||Y||^2 \\ &\leq ||X||^2 + 2||X||||Y|| + ||Y||^2 = (||X|| + ||Y||)^2. \end{aligned}$$

Tomando a raiz quadrada, segue-se o resultado. □

Com o resultado anterior, podemos estender a definição de bases ortogonais e ortonormais para o espaço $\mathbb{C}^n$ e provar, exatamente da mesma forma que é provado o resultado análogo para o $\mathbb{R}^n$ (ver por exemplo [2]), o Teorema que garante que todo subespaço $\mathbb{C}^n$ tem uma base ortonormal.

Teorema 2 (Gram-Schmidt). *Seja $\{V_1, \ldots, V_k\}$ uma base de um subespaço $\mathbb{W}$ de $\mathbb{C}^n$. Então, existe uma base $\{U_1, \ldots, U_k\}$ de $\mathbb{W}$ que é ortonormal e tal que o subespaço gerado por $U_1, \ldots, U_j$ é igual ao subespaço gerado por $V_1, \ldots, V_j$ para $j = 1, \ldots, k$.*

---

2 Resultados Principais

Teorema 3 (Schur)**.** *Se A é uma matriz $n \times n$ com entradas que são números complexos, então existe uma matriz unitária P (isto é, $\overline{P}^t = P^{-1}$) e uma matriz triangular superior T tal que*

$$A = PT\overline{P}^t.$$

Demonstração. O resultado é obvio se $n = 1$. Vamos supor que o resultado seja verdadeiro para matrizes $(n-1) \times (n-1)$ e vamos provar que ele é verdadeiro para matrizes $n \times n$. Como todo polinômio com coeficientes complexos tem uma raiz, a matriz A tem um autovalor $\lambda_1 \in \mathbb{C}$. Isto significa que existem autovetores associados a λ_1. Seja V_1 um autovetor de norma igual a 1 associado a λ_1. Sejam $V_2, \ldots, V_n$ vetores tais que $\{V_1, \ldots, V_n\}$ é uma base ortonormal de $\mathbb{C}^n$ (isto pode ser conseguido aplicando-se o processo de ortogonalização de Gram-Schmidt a uma base de $\mathbb{C}^n$ que contenha V_1.) Seja $P_1 = [\, V_1 \ \ldots \ V_n \,]$. Como $AV_1 = \lambda_1 V_1$ e $AV_2, \ldots, AV_n$ são combinações lineares de $V_1, \ldots, V_n$, temos que

$$AP_1 = [\, AV_1 \ \ldots \ AV_n \,] = [\, V_1 \ \ldots \ V_n \,]M = P_1 M, \tag{1}$$

em que $M = \left[\begin{array}{c|c} \lambda_1 & M_2 \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right]$. Como estamos supondo o resultado verdadeiro para matrizes $(n-1) \times (n-1)$, então existe uma matriz unitária $\tilde{P}_2$ e uma matriz triangular superior T_2, ambas $(n-1) \times (n-1)$, tais que $B = \tilde{P}_2 T_2 \overline{\tilde{P}}_2^t$. Seja $P_2 = \left[\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{P}_2 & \\ 0 & & & \end{array}\right]$. Seja $P = P_1 P_2$. P é unitária (verifique!) e pela equação (1)

$$AP = (AP_1)P_2 = P_1 M P_2 = P_1 \left[\begin{array}{c|c} \lambda_1 & M_2\tilde{P}_2 \\ \hline 0 & \\ \vdots & B\tilde{P}_2 \\ 0 & \end{array}\right]$$

Mas, $B\tilde{P}_2 = \tilde{P}_2 T_2$ e assim,

$$AP = P_1 \left[\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{P}_2 & \\ 0 & & & \end{array}\right] \left[\begin{array}{c|c} \lambda_1 & M_2\tilde{P}_2 \\ \hline 0 & \\ \vdots & T_2 \\ 0 & \end{array}\right] = P_1 P_2 \left[\begin{array}{c|c} \lambda_1 & M_2\tilde{P}_2 \\ \hline 0 & \\ \vdots & T_2 \\ 0 & \end{array}\right] = PT,$$

em que $T = \left[\begin{array}{c|c} \lambda_1 & M_2\tilde{P}_2 \\ \hline 0 & \\ \vdots & T_2 \\ 0 & \end{array}\right]$. Multiplicando-se à direita por $\overline{P}^t$ obtemos o resultado.

$\square$

Teorema 4. *Se A é uma matriz normal (isto é, $A\overline{A}^t = \overline{A}^t A$) $n \times n$ com entradas que são números complexos, então existe uma matriz unitária P (isto é, $\overline{P}^t = P^{-1}$) e uma matriz diagonal D tal que*

$$A = PD\overline{P}^t.$$

Demonstração. O resultado é obvio se $n = 1$. Vamos supor que o resultado seja verdadeiro para matrizes $(n-1) \times (n-1)$ e vamos provar que ele é verdadeiro para matrizes $n \times n$. Como todo polinômio com coeficientes complexos tem uma raiz, então A matriz A tem um autovalor $\lambda_1 \in \mathbb{C}$. Isto significa que existem autovetores associados a λ_1. Seja V_1 um autovetor de norma igual a 1 associado a λ_1. Sejam $V_2, \dots, V_n$ vetores tais que $\{V_1, \dots, V_n\}$ é uma base ortonormal de $\mathbb{C}^n$ (isto pode ser conseguido aplicando-se o processo de ortogonalização de Gram-Schmidt a uma base de $\mathbb{C}^n$ que contenha V_1.) Seja $P_1 = [\, V_1 \ \dots \ V_n \,]$. Como $AV_1 = \lambda_1 V_1$ e $AV_2, \dots, AV_n$ são combinações lineares de $V_1, \dots, V_n$, temos que

$$AP_1 = [\, AV_1 \ \dots \ AV_n \,] = [\, V_1 \ \dots \ V_n \,]M = P_1 M, \tag{2}$$

em que $M = \left[\begin{array}{c|ccc} \lambda_1 & m_{12} & \dots & m_{1n} \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right]$. Multiplicando-se à esquerda (2) por $\overline{P}_1^t$ obtemos $M = \overline{P}_1^t A P_1$. Mas,

$$\overline{M}^t M = (P_1^t \overline{AP}_1)^t (\overline{P}_1^t A P_1) = \overline{P}_1^t \overline{A}^t A^t P_1 = \overline{P}_1^t A^t \overline{A}^t P_1 = (\overline{P}_1^t A P_1)(\overline{P}_1^t \overline{A} P_1) = M\overline{M}^t,$$

ou seja, a matriz M é normal.

$$(M\overline{M}^t)_{11} = |\lambda_1|^2 + |m_{12}|^2 + \cdots + |m_{1n}|^2,$$

enquanto

$$(\overline{M}^t M)_{11} = |\lambda_1|^2.$$

Portanto,

$$M = \left[\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right]$$

com B uma matriz normal $(n-1) \times (n-1)$. Como estamos supondo o resultado verdadeiro para matrizes $(n-1) \times (n-1)$, então existe uma matriz unitária $\tilde{P}_2$, $(n-1) \times (n-1)$, tal que $D_2 = \overline{\tilde{P}}_2^t B \tilde{P}_2$ é diagonal. Seja $P_2 = \left[\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{P}_2 & \\ 0 & & & \end{array}\right]$. Seja $P = P_1 P_2$. P é unitária (verifique!) e pela equação (2)

$$AP = (AP_1)P_2 = P_1 M P_2 = P_1 \left[\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & B\tilde{P}_2 & \\ 0 & & & \end{array}\right]$$

Mas, $B\tilde{P}_2 = \tilde{P}_2 D_2$ e assim,

$$AP = P_1 P_2 \left[\begin{array}{c|ccc} \lambda_1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & D_2 & \\ 0 & & & \end{array}\right] = PD,$$

em que $D = \left[\begin{array}{c|ccc} \lambda_1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & D_2 & \\ 0 & & & \end{array}\right]$. Multiplicando-se à direita por $\overline{P}^t$ obtemos o resultado. $\square$

Referências

[1] Michael Artin. *Algebra*. Prentice Hall, New Jersey, 1991.

[2] Reginaldo J. Santos. *Um Curso de Geometria Analítica e Álgebra Linear*. Imprensa Universitária da UFMG, Belo Horizonte, 2010.