AULA 2

Sumário. Produto interno canónico no espaço das matrizes complexas. Processo de ortonormalização de Gram-Schmidt.

▷ Para quaisquer matrizes $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m\times n}$ de coeficientes complexos, definimos

$$\langle \mathbf{A}|\mathbf{B}\rangle = \operatorname{tr}(\mathbf{A}^*\mathbf{B})$$

onde $\mathbf{A}^* = \overline{\mathbf{A}}^{\mathsf{T}} \in \mathbb{C}^{n\times m}$ denota a MATRIZ-TRANSCONJUGADA de $\mathbf{A}$ (isto é, $\mathbf{A}^* = \overline{\mathbf{A}^{\mathsf{T}}}$) e onde, para qualquer matrix quadrada $\mathbf{X}$ de tipo $m \times m$, $\operatorname{tr}(\mathbf{X})$ denota o TRAÇO de $\mathbf{X}$ (isto é, a soma de todos os coeficientes da diagonal de $\mathbf{X}$); sendo assim, temos

$$\langle \mathbf{A}|\mathbf{B}\rangle = \sum_{1\leqslant i\leqslant m} \sum_{1\leqslant j\leqslant n} \overline{a_{i,j}}\, b_{i,j}$$

onde $\mathbf{A} = [a_{i,j}]$ e $\mathbf{B} = [b_{i,j}]$.

A correspondência $(\mathbf{A}, \mathbf{B}) \mapsto \langle \mathbf{A}|\mathbf{B}\rangle$ define um PRODUTO INTERNO no espaço vectorial complexo $\mathbb{C}^{m\times n}$, isto é, para quaisquer $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{C}^{m\times n}$ e qualquer $\alpha \in \mathbb{C}$, tem-se:

(a) $\langle \mathbf{A}|\mathbf{A}\rangle \in \mathbb{R}_0^+$ e $\langle \mathbf{A}|\mathbf{A}\rangle = 0 \iff \mathbf{A} = \mathbf{0}$;

(b) $\langle \mathbf{A}|\mathbf{B}\rangle = \overline{\langle \mathbf{B}|\mathbf{A}\rangle}$;

(c) $\langle \mathbf{A} + \mathbf{B}|\mathbf{C}\rangle = \langle \mathbf{A}|\mathbf{C}\rangle + \langle \mathbf{B}|\mathbf{C}\rangle$;

(d) $\langle \mathbf{A}|\alpha\mathbf{B}\rangle = \alpha\langle \mathbf{A}|\mathbf{B}\rangle$[†].

Referimo-nos a este produto interno como o PRODUTO INTERNO CANÓNICO em $\mathbb{C}^{m\times n}$.

De modo análogo, a correspondência $(\mathbf{A}, \mathbf{B}) \mapsto \langle \mathbf{A}|\mathbf{B}\rangle$ define um produto interno no espaço vectorial real $\mathbb{R}^{m\times n}$; nesta situação, a transconjugada $(\star)^*$ é simplesmente a transposta $(\star)^{\mathsf{T}}$.

▷ No caso particular em que $n = 1$, o produto interno canónico em $\mathbb{C}^{m\times 1}$ é dado por

$$\langle \mathbf{u}|\mathbf{v}\rangle = \operatorname{tr}(\mathbf{u}^*\mathbf{v}) = \mathbf{u}^*\mathbf{v}, \qquad \mathbf{u}, \mathbf{v} \in \mathbb{C}^{m\times 1};$$

[†] NOTA IMPORTANTE. Este axioma difere do que é usualmente exigido

$$\langle \alpha\mathbf{A}|\mathbf{B}\rangle = \alpha\langle \mathbf{A}|\mathbf{B}\rangle$$

na definição de um produto interno. Esta alteração não afecta os resultados que iremos usar e é feita só por razões de conveniência.

sendo assim, para quaisquer $\mathbf{u}, \mathbf{v} \in \mathbb{C}^{m\times 1}$, temos

$$\langle \mathbf{u}|\mathbf{v}\rangle = \overline{u_1}\, v_1 + \cdots + \overline{u_m}\, v_m$$

onde $u_1, \ldots, u_m, v_1, \ldots, v_m \in \mathbb{C}$ são as componentes de $\mathbf{u}$ e $\mathbf{v}$, respectivamente.

Dizemos que dois vectores $\mathbf{u}, \mathbf{v} \in \mathbb{C}^{m\times 1}$ são ORTOGONAIS, e escrevemos $\mathbf{u} \perp \mathbf{v}$, se $\langle \mathbf{u}|\mathbf{v}\rangle = 0$; sendo assim,

$$\mathbf{u} \perp \mathbf{v} \qquad \Longleftrightarrow \qquad \mathbf{v}^*\mathbf{u} = 0.$$

Um subconjunto $\mathcal{S} = \{\mathbf{v}_1, \ldots, \mathbf{v}_s\} \subseteq \mathbb{C}^{m\times 1}$ diz-se um SUBCONJUNTO ORTONORMADO se

$$\langle \mathbf{v}_i|\mathbf{v}_j\rangle = \begin{cases} 1, & \text{se } i = j, \\ 0, & \text{se } i \neq j, \end{cases} \qquad 1 \leqslant i, j \leqslant s.$$

Sendo assim, os vectores de um subconjunto ortonormado são ortogonais dois-a-dois e normados; para qualquer $\mathbf{v} \in \mathbb{C}^{m\times 1}$, definimos a NORMA de $\mathbf{v}$ por

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}|\mathbf{v}\rangle}$$

e dizemos que $\mathbf{v}$ é um VECTOR NORMADO se $\|\mathbf{v}\| = 1$.

É óbvio que qualquer subconjunto ortonormado é constituído por vectores não-nulos e ortogonais dois-a-dois. Como exemplo, a base canónica de $\mathbb{C}^{m\times n}$ é um subconjunto ortonormado (isto é, a base canónica é uma BASE ORTONORMADA).

LEMA 2.1. *Qualquer subconjunto ortonormado de $\mathbb{C}^{m\times 1}$ é linearmente independente.*

DEMONSTRAÇÃO. Seja $\mathcal{S} = \{\mathbf{v}_1, \ldots, \mathbf{v}_s\} \subseteq \mathbb{C}^{m\times 1}$ um subconjunto ortogonal e sejam $\alpha_1, \ldots, \alpha_s \in \mathbb{C}$ tais que

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_s\mathbf{v}_s = \mathbf{0}.$$

Pondo $\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_s\mathbf{v}_s$, temos

$$0 = \langle \mathbf{v}_i|\mathbf{0}\rangle = \langle \mathbf{v}_i|\alpha_1\mathbf{v}_1 + \cdots + \alpha_s\mathbf{v}_s\rangle = \alpha_1\langle \mathbf{v}_i|\mathbf{v}_1\rangle + \cdots + \alpha_n\langle \mathbf{v}_i|\mathbf{v}_n\rangle = \alpha_i$$

para qualquer $1 \leqslant i \leqslant s$. □

LEMA 2.2. *Se $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ for uma base ortonormada de $\mathbb{C}^{m\times 1}$, então qualquer vector $\mathbf{v} \in \mathbb{C}^{m\times 1}$ pode ser escrito (de maneira única) na forma*

$$\mathbf{v} = \langle \mathbf{v}_1|\mathbf{v}\rangle\mathbf{v}_1 + \langle \mathbf{v}_2|\mathbf{v}\rangle\mathbf{v}_2 + \cdots + \langle \mathbf{v}_m|\mathbf{v}\rangle\mathbf{v}_m.$$

[*A esta expressão chamamos a* EXPANSÃO DE FOURIER *de* $\mathbf{v}$ *e* $\langle \mathbf{v}_1|\mathbf{v}\rangle, \cdots, \langle \mathbf{v}_m|\mathbf{v}\rangle$ *dizem-se os* COEFICIENTES DE FOURIER *de* $\mathbf{v}$.]

DEMONSTRAÇÃO. Exercício. □

Proposição 2.3 (Processo de ortonormalização de Gram-Schmidt). *Seja $\mathcal{S}$ um subespaço vectorial de $\mathbb{C}^{m\times 1}$ e seja $\{\mathbf{u}_1,\dots,\mathbf{u}_s\}$ uma base de $\mathcal{S}$. Definamos os vectores $\mathbf{v}_1,\dots,\mathbf{v}_s\in\mathbb{C}^{m\times 1}$ recursivamente por*

$$\begin{aligned}\mathbf{v}_1 &= \frac{1}{\|\mathbf{u}_1\|}\,\mathbf{u}_1,\\ \mathbf{v}_k &= \frac{1}{\|\mathbf{u}_k-\sum_{1\leqslant i\leqslant k-1}\langle\mathbf{v}_i|\mathbf{u}_k\rangle\mathbf{v}_i\|}\left(\mathbf{u}_k-\sum_{1\leqslant i\leqslant k-1}\langle\mathbf{v}_i|\mathbf{u}_k\rangle\mathbf{v}_i\right),\qquad 2\leqslant k\leqslant s.\end{aligned}$$

Então, $\{\mathbf{v}_1,\dots,\mathbf{v}_m\}$ é uma base ortonormada de $\mathcal{S}$. Os vectores $\mathbf{v}_1,\dots,\mathbf{v}_m$ podem ser expressos (matricialmente) na forma

$$\mathbf{v}_k=\frac{1}{\|(\mathbf{I}_m-\mathbf{V}_k\mathbf{V}_k^*)\mathbf{u}_k\|}\,(\mathbf{I}_m-\mathbf{V}_k\mathbf{V}_k^*)\mathbf{u}_k,\qquad 1\leqslant k\leqslant m,$$

onde

$$\mathbf{V}_1=\mathbf{0}\in\mathbb{C}^{m\times 1}\qquad e\qquad \mathbf{V}_k=\begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_{k-1}\end{bmatrix}\in\mathbb{C}^{m\times(k-1)},\quad 2\leqslant k\leqslant s.$$

Demonstração. Exercício. Para exprimir os vectores $\mathbf{v}_1,\dots,\mathbf{v}_s$ na forma matricial, notamos que

$$\mathbf{V}_k^*\mathbf{u}_k=\begin{bmatrix}\mathbf{v}_1^*\mathbf{u}_k\\ \vdots\\ \mathbf{v}_{k-1}^*\mathbf{u}_k\end{bmatrix}=\begin{bmatrix}\langle\mathbf{v}_1|\mathbf{u}_k\rangle\\ \vdots\\ \langle\mathbf{v}_{k-1}|\mathbf{u}_k\rangle\end{bmatrix},$$

logo

$$\mathbf{V}_k\mathbf{V}_k^*\mathbf{u}_k=\sum_{1\leqslant i\leqslant k-1}\mathbf{v}_i(\mathbf{v}_i^*\mathbf{u}_k)=\sum_{1\leqslant i\leqslant k-1}\mathbf{v}_i\langle\mathbf{v}_i|\mathbf{u}_k\rangle=\sum_{1\leqslant i\leqslant k-1}\langle\mathbf{v}_i|\mathbf{u}_k\rangle\mathbf{v}_i$$

e, portanto,

$$(\mathbf{I}_m-\mathbf{V}_k\mathbf{V}_k^*)\mathbf{u}_k=\mathbf{u}_k-\sum_{1\leqslant i\leqslant k-1}\langle\mathbf{v}_i|\mathbf{u}_k\rangle\mathbf{v}_i,\qquad 1\leqslant k\leqslant s.$$ □

▷ Dizemos que uma matriz complexa $\mathbf{U}\in\mathbb{C}^{n\times n}$ é uma matriz unitária se as colunas de $\mathbf{U}$ formarem uma base ortonormada de $\mathbb{C}^{n\times 1}$.

Por outro lado, dizemos que uma matriz real $\mathbf{P}\in\mathbb{R}^{n\times n}$ é uma matriz ortogonal se as colunas de $\mathbf{P}$ formarem uma base ortonormada de $\mathbb{R}^{n\times 1}$.

Teorema 2.4. (a) *Uma matriz complexa $\mathbf{U}\in\mathbb{C}^{n\times n}$ será unitária se e só se $\mathbf{U}$ for invertível e $\mathbf{U}^{-1}=\mathbf{U}^*$. Em particular, se $\mathbf{U}\in\mathbb{C}^{n\times n}$ for unitária, então $\mathbf{U}^*$ também será unitária.*

(b) *Uma matriz real $\mathbf{P}\in\mathbb{R}^{n\times n}$ será ortogonal se e só se $\mathbf{P}$ for invertível e $\mathbf{P}^{-1}=\mathbf{P}^{\mathsf{T}}$. Em particular, se $\mathbf{P}\in\mathbb{R}^{n\times n}$ for ortogonal, então $\mathbf{P}^{\mathsf{T}}$ também será unitária.*

DEMONSTRAÇÃO. (a) Suponhamos que $\mathbf{U} \in \mathbb{C}^{n\times n}$ é unitária e sejam $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{C}^{n\times 1}$ as colunas de $\mathbf{U}$. Como $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ é uma base de $\mathbb{C}^{n\times 1}$, temos $r(\mathbf{U}) = n$ e, portanto, $\mathbf{U}$ é invertível. Por outro lado, temos

$$\mathbf{U}^* = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix}^* = \begin{bmatrix} \mathbf{u}_1^* \\ \mathbf{u}_2^* \\ \vdots \\ \mathbf{u}_n^* \end{bmatrix}$$

e, portanto,

$$\mathbf{U}^*\mathbf{U} = \begin{bmatrix} \mathbf{u}_1^* \\ \mathbf{u}_2^* \\ \vdots \\ \mathbf{u}_n^* \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^*\mathbf{u}_1 & \mathbf{u}_1^*\mathbf{u}_2 & \cdots & \mathbf{u}_1^*\mathbf{u}_n \\ \mathbf{u}_2^*\mathbf{u}_1 & \mathbf{u}_2^*\mathbf{u}_2 & \cdots & \mathbf{u}_2^*\mathbf{u}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{u}_n^*\mathbf{u}_1 & \mathbf{u}_n^*\mathbf{u}_2 & \cdots & \mathbf{u}_n^*\mathbf{u}_n \end{bmatrix}.$$

Como $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ é base ortonormada, temos

$$\mathbf{u}_i^*\mathbf{u}_j = \langle \mathbf{u}_i | \mathbf{u}_j \rangle = \begin{cases} 1, & \text{se } i = j, \\ 0 & \text{se } i \neq j, \end{cases} \qquad 1 \leqslant i, j \leqslant n,$$

logo $\mathbf{U}^*\mathbf{U} = \mathbf{I}_n$ e, portanto,

$$\mathbf{U}^* = \mathbf{U}^*\mathbf{I}_n = \mathbf{U}^*(\mathbf{U}\mathbf{U}^{-1}) = (\mathbf{U}^*\mathbf{U})\mathbf{U}^{-1} = \mathbf{I}_n\mathbf{U}^{-1} = \mathbf{U}^{-1}.$$

Reciprocamente, suponhamos que $\mathbf{U}$ é invertível com $\mathbf{U}^{-1} = \mathbf{U}^*$ e, como antes, sejam $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{C}^{n\times 1}$ as colunas de $\mathbf{U}$. Como $\mathbf{U}$ é invertível, temos $r(\mathbf{U}) = n$, logo $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ são uma base de $\mathbb{C}^{n\times 1}$ (porque são n vectores linearmente independentes). Por outro lado, como $\mathbf{U}^{-1} = \mathbf{U}^*$, temos $\mathbf{U}^*\mathbf{U} = \mathbf{I}_n$ e, portanto, como acima, concluímos que

$$\mathbf{u}_i^*\mathbf{u}_j = \langle \mathbf{u}_i | \mathbf{u}_j \rangle = \begin{cases} 1, & \text{se } i = j, \\ 0 & \text{se } i \neq j, \end{cases} \qquad 1 \leqslant i, j \leqslant n,$$

isto é, $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ é uma base ortonormada de $\mathbb{C}^{n\times 1}$.

Para a última afirmação, basta observar que $\mathbf{U}^{-1}$ é invertível e que

$$(\mathbf{U}^{-1})^* = (\mathbf{U}^*)^* = \mathbf{U} = (\mathbf{U}^{-1})^{-1}.$$

(b) É como (a) substituindo $(\star)^*$ por $(\star)^{\mathsf{T}}$. $\square$

TEOREMA 2.5. (a) *Uma matriz complexa* $\mathbf{U} \in \mathbb{C}^{n\times n}$ *será unitária se e só se* $\|\mathbf{U}\mathbf{v}\| = \|\mathbf{v}\|$ *para qualquer* $\mathbf{v} \in \mathbb{C}^{n\times 1}$.

(b) *Uma matriz real* $\mathbf{P} \in \mathbb{R}^{n\times n}$ *será ortogonal se e só se* $\|\mathbf{P}\mathbf{v}\| = \|\mathbf{v}\|$ *para qualquer* $\mathbf{v} \in \mathbb{R}^{n\times 1}$.

Demonstração. (a) Seja $\mathbf{U} \in \mathbb{C}^{n\times n}$ uma matriz unitária e seja $\mathbf{v} \in \mathbb{C}^{n\times 1}$. Como $\mathbf{U}^*\mathbf{U} = \mathbf{I}_n$, obtemos

$$\|\mathbf{U}\mathbf{v}\|^2 = \langle \mathbf{U}\mathbf{v}|\mathbf{U}\mathbf{v}\rangle = \mathbf{v}^*\mathbf{U}^*\mathbf{U}\mathbf{v} = \mathbf{v}^*\mathbf{v} = \langle \mathbf{v}|\mathbf{v}\rangle = \|\mathbf{v}\|^2$$

e, portanto, $\|\mathbf{U}\mathbf{v}\| = \|\mathbf{v}\|$ (porque a norma de um vector é um número real positivo).

Reciprocamente, suponhamos que $\mathbf{U} \in \mathbb{C}^{n\times n}$ é tal que $\|\mathbf{U}\mathbf{v}\| = \|\mathbf{v}\|$ para qualquer $\mathbf{v} \in \mathbb{C}^{n\times 1}$. Sejam $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{C}^{n\times 1}$ as colunas de $\mathbf{U}$, de modo que

$$\mathbf{u}_1 = \mathbf{U}\mathbf{e}_1,\ \mathbf{u}_2 = \mathbf{U}\mathbf{e}_2,\ \ldots,\ \mathbf{u}_n = \mathbf{U}\mathbf{e}_n$$

onde $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ a base canónica de $\mathbb{C}^{n\times 1}$. Provamos que $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ é uma base ortonormada de $\mathbb{C}^{n\times 1}$. Por um lado, temos

$$\|\mathbf{u}_i\| = \|\mathbf{U}\mathbf{e}_i\| = \|\mathbf{e}_i\| = 1, \qquad 1 \leqslant i \leqslant n.$$

Por outro lado, sejam $1 \leqslant i \neq j \leqslant n$. Para provar que $\langle \mathbf{u}_i|\mathbf{u}_j\rangle = 0$, verificamos que

$$\|\alpha\mathbf{u}_i + \beta\mathbf{u}_j\|^2 = \|\alpha\mathbf{u}_i\|^2 + \|\beta\mathbf{u}_j\|^2$$

para quaisquer $\alpha, \beta \in \mathbb{C}$. De facto, temos

$$\|\alpha\mathbf{u}_i + \beta\mathbf{u}_j\|^2 = \|\mathbf{U}(\alpha\mathbf{e}_i + \beta\mathbf{e}_j)\|^2 = \|\alpha\mathbf{e}_i + \beta\mathbf{e}_j\|^2 = |\alpha|^2 + |\beta|^2,$$

enquanto que

$$\|\alpha\mathbf{u}_i\|^2 + \|\beta\mathbf{u}_j\|^2 = \|\mathbf{U}(\alpha\mathbf{e}_i)\|^2 + \|\mathbf{U}(\beta\mathbf{e}_j)\|^2 = \|\alpha\mathbf{e}_i\|^2 + \|\beta\mathbf{e}_j\|^2 = |\alpha|^2 + |\beta|^2.$$

Agora, temos

$$\begin{aligned}\|\alpha\mathbf{u}_i + \beta\mathbf{u}_j\|^2 &= \langle \alpha\mathbf{u}_i + \beta\mathbf{u}_j|\alpha\mathbf{u}_i + \beta\mathbf{u}_j\rangle = \|\alpha\mathbf{u}_i\|^2 + \|\beta\mathbf{u}_j\|^2 + \langle \alpha\mathbf{u}_i|\beta\mathbf{u}_j\rangle + \langle \beta\mathbf{u}_j|\alpha\mathbf{u}_i\rangle \\ &= \|\alpha\mathbf{u}_i\|^2 + \|\beta\mathbf{u}_j\|^2 + \langle \alpha\mathbf{u}_i|\beta\mathbf{u}_j\rangle + \overline{\langle \alpha\mathbf{u}_i|\beta\mathbf{u}_j\rangle}\end{aligned}$$

e, portanto,

$$\langle \alpha\mathbf{u}_i|\beta\mathbf{u}_j\rangle + \overline{\langle \alpha\mathbf{u}_i|\beta\mathbf{u}_j\rangle} = 0.$$

Em particular, obtemos

$$\begin{cases} \langle \mathbf{u}_i|\mathbf{u}_j\rangle + \overline{\langle \mathbf{u}_i|\mathbf{u}_j\rangle} = 0 & (\text{com } \alpha = \beta = 1), \\ \langle \mathbf{u}_i|\mathbf{u}_j\rangle - \overline{\langle \mathbf{u}_i|\mathbf{u}_j\rangle} = 0 & (\text{com } \alpha = 1 \text{ e } \beta = -1). \end{cases}$$

Daqui resulta que $\langle \mathbf{u}_i|\mathbf{u}_j\rangle = 0$, como se queria.

(b) É como (a) substituindo $(\star)^*$ por $(\star)^{\mathsf{T}}$. □