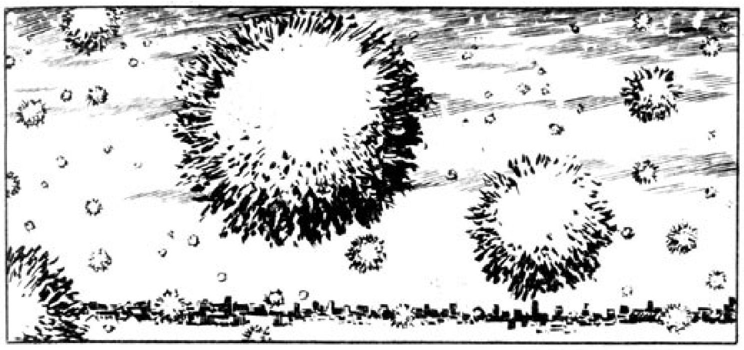
Álgebra II (Curso 23)
Primer cuatrimestre, 2021

Notas en la emergencia sanitaria:

borradores para la clase del 7 de julio

Sebastian Grynberg



El único héroe válido es el héroe "en grupo",
nunca el héroe individual, el héroe solo.

H. G. Oesterheld

Índice

1. Propiedades de las matrices simétricas

1.1. Subespacios fundamentales.

Lema 1.1. *Si $A \in \mathbb{R}^{n\times n}$ es una matriz simétrica, el subespacio nulo de A es ortogonal al subespacio columna de A. En otras palabras,*

$$\mathrm{nul}(A) \perp \mathrm{col}(A).$$

Demostración. Inmediata: $\mathrm{nul}(A) \perp \mathrm{col}(A^T) = \mathrm{col}(A)$. □

1.2. Espectro.

Lema 1.2. *Todas las raíces del polinomio característico de una matriz simétrica $A \in \mathbb{R}^{n\times n}$ son números reales.*

Demostración. Si $\lambda \in \sigma_{\mathbb{C}}(A)$ y $z \in \mathrm{nul}(A - \lambda I) \setminus \{0\}$, entonces

$$\begin{aligned}\lambda\|z\|^2 &= \lambda \langle z, z\rangle = \langle \lambda z, z\rangle = \langle Az, z\rangle = \langle z, A^* z\rangle \\ &= \langle z, Az\rangle = \langle z, \lambda z\rangle = \overline{\lambda} \langle z, z\rangle = \overline{\lambda}\|z\|^2.\end{aligned}$$

Como $\|z\|^2 \neq 0$, resulta que $\lambda = \overline{\lambda}$. Por lo tanto, $\lambda \in \mathbb{R}$. □

1.3. Autoespacios.

Lema 1.3. *Si $A \in \mathbb{R}^{n\times n}$ es una matriz simétrica, autovectores correspondientes a distintos autovalores son ortogonales entre sí. En otras palabras,*

$$\mathrm{nul}(A - \lambda_i I) \perp \mathrm{nul}(A - \lambda_j I) \quad \textit{para} \quad \lambda_i \neq \lambda_j.$$

Demostración. Sean $x \in \mathrm{nul}(A - \lambda_i I)$ e $y \in \mathrm{nul}(A - \lambda_j I)$, vale que

$$\lambda_i \langle x, y\rangle = \langle \lambda_i x, y\rangle = \langle Ax, y\rangle = \langle x, Ay\rangle = \langle x, \lambda_j y\rangle = \lambda_j \langle x, y\rangle.$$

Como $\lambda_i \neq \lambda_j$, resulta que $\langle x, y\rangle = 0$. Por lo tanto, $x \perp y$. □

1.4. Teorema espectral.

Teorema 1.4. *Una matriz $A \in \mathbb{R}^{n\times n}$ con espectro $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$ es simétrica si y sólo si existen matrices $\{P_1, P_2, \dots, P_m\}$ tales que*

$$A = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_m P_m, \tag{1}$$

donde las P_i tienen las siguientes propiedades

a) *P_i es la proyección ortogonal sobre $\mathrm{nul}(A - \lambda_i I)$.*
b) *$P_i P_j = 0$ para $i \neq j$.*
c) *$P_1 + P_2 + \cdots + P_m = I$.*

Nota Bene. Nótese que si $\{u_1, \dots, u_r\}$ es un sistema ortonormal de vectores de $\mathbb{R}^n$ y $U_r = \begin{bmatrix} u_1 & \cdots & u_r \end{bmatrix} \in \mathbb{R}^{n \times r}$, entonces

$$P = U_r U_r^T$$

es la matriz de la proyección ortogonal sobre el subespacio $\text{gen}\{u_1, \dots, u_r\}$. En efecto,

$$Px = U_r U_r^T x = \left(\sum_{j=1}^{r} u_j u_j^T \right) x = \sum_{j=1}^{r} \left(u_j^T x \right) u_j = \sum_{j=1}^{r} \langle x, u_j \rangle \, u_j.$$

1.5. Courant-Fisher.

Como los autovalores de una matriz simétrica $A \in \mathbb{R}^{n \times n}$ son números reales, se los puede organizar en forma decreciente: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Además, existe una matriz ortogonal $U \in \mathbb{R}^{n \times n}$ tal que $U^T A U = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, o, lo que es lo mismo $A = U \Lambda U^T$. Definiendo $y = U^T x$ podemos escribir

$$x^T A x = x^T U \Lambda U^T x = y^T \Lambda y = \sum_{j=1}^{n} \lambda_j y_j^2. \tag{2}$$

Como $\|y\| = \|U^T x\| = \|x\|$, las relaciones

$$\sum_{j=1}^{n} \lambda_j y_j^2 \leq \lambda_1 \sum_{j=1}^{n} y_j^2 = \lambda_1 \|y\|^2,$$
$$\sum_{j=1}^{n} \lambda_j y_j^2 = \lambda_1 \text{ para } y = e_1,$$

junto a las identidades (2), implican que

$$\lambda_1 = \max_{\|x\|=1} x^T A x.$$

Del mismo modo, las relaciones

$$\sum_{j=1}^{n} \lambda_j y_j^2 \geq \lambda_n \sum_{j=1}^{n} y_j^2 = \lambda_n \|y\|^2,$$
$$\sum_{j=1}^{n} \lambda_j y_j^2 = \lambda_n \text{ para } y = e_n,$$

implican que

$$\lambda_n = \min_{\|x\|=1} x^T A x.$$

Con un poco más de trabajo se puede demostrar el siguiente resultado

Teorema 1.5 (Courant-Fischer)**.** *Los autovalores $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ de una matriz simétrica $A \in \mathbb{R}^{n \times n}$ son*

$$\lambda_i = \min_{\dim(\mathbb{S}) = n-i+1} \; \max_{x \in \mathbb{S} : \|x\|=1} x^T A x$$

Demostración. Usando una diagonalización ortogonal de A el problema se reduce a demostrar que

$$\lambda_i = \min_{\dim(\mathbb{S})=n-i+1} \max_{y\in\mathbb{S}:\|y\|=1} y^T\Lambda y$$

Sea $\mathbb{T} = \text{gen}\{e_1, \dots, e_i\} \neq \{0\}$. Si $\mathbb{S}$ tiene dimensión $n+i-1$, entonces $\mathbb{S}\cap\mathbb{T} \neq \{0\}$. Si se supone lo contrario se obtiene una contradicción. Consideramos

$$\Sigma_{\mathbb{S}} = \{y \in \mathbb{S} : \|y\| = 1\}.$$

Como $\mathbb{S}\cap\mathbb{T} \neq \{0\}$, $\Sigma_{\mathbb{S}} \cap \mathbb{T} \neq \emptyset$. Para cada $y \in \Sigma_{\mathbb{S}} \cap \mathbb{T}$ vale que

$$y^T\Lambda y = \sum_{j=1}^{i} \lambda_j y_j^2 \geq \lambda_i \sum_{j=1}^{i} y_j^2 = \lambda_i$$

Como $\Sigma_{\mathbb{S}} \cap \mathbb{T} \subseteq \Sigma_{\mathbb{S}}$, tenemos

$$\max_{y\in\Sigma_{\mathbb{S}}} y^T\Lambda y \geq \max_{y\in\Sigma_{\mathbb{S}}\cap\mathbb{T}} y^T\Lambda y = \lambda_i,$$

y en consecuencia,

$$\min_{\dim(\mathbb{S})=n-i+1} \max_{y\in\Sigma_{\mathbb{S}}} y^T\Lambda y \geq \lambda_i.$$

Si $\mathbb{S}^* = \{e_1, \dots, e_{i-1}\}$

$$y^T\Lambda y = \sum_{j=1}^{n} \lambda_j y_j^2 \leq \lambda_i \sum_{j=i}^{n} y_j^2 = \lambda_i \text{ para todo } y \in \Sigma_{\mathbb{S}^*}.$$

En consecuencia,

$$\min_{\dim(\mathbb{S})=n-i+1} \max_{y\in\Sigma_{\mathbb{S}}} y^Y\Lambda y \leq \max_{y\in\Sigma_{\mathbb{S}^*}} y^Y\Lambda y \leq \lambda_i,$$

y por lo tanto,

$$\lambda_i = \min_{\dim(\mathbb{S})=n-i+1} \max_{y\in\mathbb{S}:\|y\|=1} y^T\Lambda y.$$

□

2. Para un estudio de A^TA

Nota Bene. Nótese que si $A \in \mathbb{R}^{m\times n}$, entonces $A^TA \in \mathbb{R}^{n\times n}$ es una matriz simétrica.

2.1. Subespacios fundamentales.

Lema 2.1. *Sea $A \in \mathbb{R}^{m\times n}$. Vale que*

$$\text{nul}(A^TA) = \text{nul}(A)$$
$$\text{col}(A^TA) = \text{col}(A^T)$$

Demostración. Ejercicio. □

2.2. Espectro.

Lema 2.2. *Sea $A \in \mathbb{R}^{m\times n}$. Vale que $\sigma\left(A^T A\right) \subset \mathbb{R}^+$.*

Demostración. Si $\lambda \in \sigma\left(A^T A\right)$ y $x \in \text{nul}(A^T A - \lambda I) \setminus \{0\}$, entonces

$$\lambda\|x\|^2 = \lambda \langle x, x\rangle = \langle \lambda x, x\rangle = \left\langle A^T A x, x\right\rangle = \langle Ax, Ax\rangle = \|Ax\|^2.$$

Como $\|x\|^2 > 0$ y $\|Ax\|^2 \geq 0$, resulta que $\lambda \geq 0$. Por lo tanto, $\sigma\left(A^T A\right) \subset \mathbb{R}^+$. $\square$

2.3. Diagonalización ortogonal.

Sea $A \in \mathbb{R}^{m\times n}$ una matriz de rango r. Como $\text{rango}(A^T) = \text{rango}(A)$ y $\text{col}(A^T A) = \text{col}(A^T)$ tenemos que $\text{rango}(A^T A) = r$. Ordenando los autovalores no nulos de $A^T A$ de manera decreciente podemos escribir $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$. Como $A^T A$ es simétrica, existe una matriz ortogonal $V \in \mathbb{R}^{n\times n}$ tal que

$$A^T A = V\Lambda V^T,$$

donde $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0)$.

Nota Bene. Nótese que si $V = \begin{bmatrix} v_1 & \cdots & v_n\end{bmatrix}$, entonces

$$A^T A v_i = \lambda_i v_i.$$

2.4. Estructura geométrica de A.

Lema 2.3. *Sea $A \in \mathbb{R}^{m\times n}$ una matriz de rango r. Si $\{v_1, v_2, \ldots, v_r\}$ es una base ortonormal de $\text{col}(A^T)$ compuesta por autovectores de $A^T A$ correspondientes a los autovalores $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$, entonces los vectores de $\mathbb{R}^m$ definidos por*

$$u_i := \frac{1}{\sqrt{\lambda_i}} A v_i, \quad i \in \mathbb{I}_r, \tag{3}$$

constituyen una base ortonormal del espacio columna de A y además vale que

$$A v_i = \sqrt{\lambda_i} u_i. \tag{4}$$

Demostración. Una cuenta:

$$\begin{aligned}\langle u_i, u_j\rangle &= \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} \langle A v_i, A v_j\rangle = \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} \left\langle v_i, A^T A v_j\right\rangle \\ &= \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} \langle v_i, \lambda_j v_j\rangle = \frac{\sqrt{\lambda_j}}{\sqrt{\lambda_i}} \langle v_i, v_j\rangle = 0.\end{aligned}$$

$\square$

Nota Bene. Nótese que las igualdades (4) significan que

$$A \begin{bmatrix} v_1 & v_2 & \ldots & v_r\end{bmatrix} = \begin{bmatrix} u_1 & u_2 & \ldots & u_r\end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_r}\end{bmatrix}. \tag{5}$$

Definiendo

- $\sigma_i := \sqrt{\lambda_i}$ para cada $i \in \mathbb{I}_r$,

- $\Sigma_r := \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix}$,
- $V_r := \begin{bmatrix} v_1 & v_2 & \cdots & v_r \end{bmatrix}$,
- $U_r := \begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix}$,

tenemos que

$$AV_r = U_r\Sigma_r. \tag{6}$$

Podemos completar la base ortonormal de $\mathrm{col}(A^T)$, $\{v_1, \ldots, v_r\}$, a una base ortonormal de $\mathbb{R}^n$, $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ y la base ortonormal de $\mathrm{col}(A)$, a una base ortonormal de $\mathbb{R}^m$, $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_m\}$. Definiendo

$$V_{n-r} = \begin{bmatrix} v_{r+1} & \cdots & v_n \end{bmatrix} \text{ y } U_{m-r} = \begin{bmatrix} u_{r+1} & \cdots & u_m \end{bmatrix},$$

tenemos que

$$A\underbrace{\begin{bmatrix} V_r & V_{n-r} \end{bmatrix}}_{V\in\mathbb{R}^{n\times n}} = \begin{bmatrix} AV_r & AV_{n-r} \end{bmatrix} = \begin{bmatrix} U_r\Sigma_r & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} U_r & U_{m-r} \end{bmatrix}}_{U\in\mathbb{R}^{m\times m}}\underbrace{\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}}_{\Sigma\in\mathbb{R}^{m\times n}}. \tag{7}$$

Teorema 2.4 (Descomposición en valores singulares)**.** *Sea $A \in \mathbb{R}^{m\times n}$ una matriz de rango r. Existe una factorización de la forma*

$$A = U\Sigma V^T,$$

donde $U \in \mathbb{R}^{m\times m}$ y $V \in \mathbb{R}^{n\times n}$ son matrices ortogonales y $\Sigma \in \mathbb{R}^{m\times n}$ es una matriz de la forma

$$\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix},$$

con $\Sigma_r = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r) \in \mathbb{R}^{r\times r}$, donde $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.