ÁLGEBRA LINEAL I
NOTAS DE CLASE
UNIDAD IV

Nota

Las siguientes páginas contienen una presentación del concepto de proyección ortogonal. El objetivo de esta presentación es el de complementar la exposición de clase.

1. Definiciones

Definición 1.: sea $V \leq \mathbb{R}^n$. Definimos el *complemento ortogonal de* V *en* $\mathbb{R}^n$ como el conjunto de todos los vectores $w \in \mathbb{R}^n$, tales que $w^t v = 0$ para cualquier $v \in V$, a tal conjunto lo denotamos por $V^{\perp}$. Entonces, el complemento ortogonal de un subespacio vectorial de $\mathbb{R}^n$ es el conjunto de todos los vectores que son perpendiculares a cada vector en V,

$$V^{\perp} = \{w \in \mathbb{R}^n : w \cdot v = 0 \; \forall \; v \in V\}.$$

2. Proyección ortogonal

Lema 1.: sea $V \leq \mathbb{R}^n$. Entonces $V^{\perp} \leq \mathbb{R}^n$, decir, el complemento ortogonal de un subespacio vectorial V, también es un subespacio vectorial.

<u>Prueba</u>: (ejercicio).

Lema 2.: sea $V \leq \mathbb{R}^n$, tal que $\dim(V) = k$ y considere una base ortonormal de V, $\beta = \{q_1, \ldots, q_k\}$. Defínase $\mathbf{P} \in \mathbb{R}^{n\times n}$ de la forma siguiente:

$$\mathbf{P} := q_1 q_1^t + q_2 q_2^t + \ldots + q_k q_k^t.$$

Entonces, para cualquier $v \in V$, $\mathbf{P}v = v$.

<u>Prueba</u>: dado que $v \in V$ y β es una base de V, entonces existen $\alpha_1, \alpha_2, \ldots, \alpha_k$ en $\mathbb{R}$, definidos de manera única, tales que $v = \alpha_1 q_1 + \alpha_2 q_2 + \cdots + \alpha_k q_k$; por lo tanto,

$$\begin{aligned}
\mathbf{P}v &= (q_1 q_1^t + q_2 q_2^t + \cdots + q_k q_k^t)(\alpha_1 q_1 + \alpha_2 q_2 + \cdots + \alpha_k q_k) \\
&= \sum_{\ell=1}^{k} q_\ell q_\ell^t(\alpha_\ell q_\ell) + \sum_{\substack{j,k \\ j\neq k}} q_k q_k^t(\alpha_j q_j) \\
&= \sum_{\ell=1}^{k} \alpha_\ell q_\ell \underbrace{(q_\ell^t q_\ell)}_{\overset{(*)}{=}1} + \sum_{\substack{j,k \\ j\neq k}} \alpha_j q_k \underbrace{(q_k^t q_j)}_{\overset{(*)}{=}0} \\
&= \sum_{\ell=1}^{k} \alpha_\ell q_\ell = v,
\end{aligned}$$

$(*)$ indica que se ha empleado la propiedad de ortonormalidad de la base β. `Q.E.D.`

Lema 3.: sean V y β como en el lema 2 y $b \in \mathbb{R}^n$ fijo. Entonces, para cualquier $v \in V$,

$$(b - \mathbf{P}b)^t v = 0,$$

Date: April 26, 2010.

es decir $(b - \mathbf{P}b) \in V^{\perp}$.

Prueba: tenemos que demostrar que, para cualquier $v \in V$, $(b - \mathbf{P}b)^t v = 0$ (equivalentemente, $b - \mathbf{P}b \in V^{\perp}$).

$$\begin{aligned}(b - \mathbf{P}b)^t v &= b^t v - (\mathbf{P}b)^t v \\ &= b^t v - b^t(\mathbf{P}^t v) \\ &\overset{(1)}{=} b^t v - b^t(\mathbf{P}v) \\ &\overset{(2)}{=} b^t v - b^t v \,=\, 0\,;\end{aligned}$$

(1) $\mathbf{P^t} = \mathbf{P}$ se sigue de la definición de $\mathbf{P}$, (2) $\mathbf{P}v = v$ por el lema 2. Entonces, para $b \in \mathbb{R}^n$ fijo pero arbitrario, se tiene que $(b - \mathbf{P}b) \cdot v = 0$ para cualquier $v \in V$; es decir, $(b - \mathbf{P}b) \in V^{\perp}$. `Q.E.D.`

Lema 4.: sean V y β y $\mathbf{P}$ como en el lema 2 y $b \in \mathbb{R}^n$. Entonces,

$$|b - \mathbf{P}b| = \min\{|b - v| : v \in V\}\,.$$

Prueba: sea $v \in V$ arbitrario, entonces,

$$|b - v| \,=\, |(b - \mathbf{P}b) - (v - \mathbf{P}b)|\,.$$

Llamemos $y = b - \mathbf{P}b$, entonces,

$$\begin{aligned}|b - v|^2 &= |y - (v - \mathbf{P}b)|^2 \\ &= (y - (v - \mathbf{P}b)) \cdot (y - (v - \mathbf{P}b)) \\ &= |y|^2 - 2y \cdot (v - \mathbf{P}b) + |v - \mathbf{P}b|^2\,.\end{aligned}$$

Observe que, por definición de $\mathbf{P}$, $\mathbf{P}b \in V$, luego, dado que V es un subespacio vectorial y $v \in V$, $v - \mathbf{P}b \in V$. Ahora, por el lema 3, $y = b - \mathbf{P}b \in V^{\perp}$, en particular, $y \cdot (v - \mathbf{P}b) = 0$, substituyendo esto en la última igualdad de la cadena de igualdades arriba, concluimos que

$$\begin{aligned}|b - v|^2 &= |y|^2 + |v - \mathbf{P}b|^2 \\ &\geq |y|^2 = |b - \mathbf{P}b|^2\,,\end{aligned}$$

tomando raíces de ambos lados de la desigualdad,

$$|b - v| \,\geq\, |b - \mathbf{P}b|\,,$$

luego, como ya hicimos notar, $\mathbf{P}b \in V$, de donde se sigue el enunciado del lema. `Q.E.D.`

Lema 5.: sean V y β como en el lema 2 y sea $\tilde{\beta} = \{\tilde{q}_1, \ldots \tilde{q}_k\}$ otra base ortonormal de V. De manera similar a como se definió $\mathbf{P}$, defínase ahora

$$\tilde{\mathbf{P}} \,:=\, \tilde{q}_1\tilde{q}_1^t + \ldots + \tilde{q}_k\tilde{q}_k^t\,.$$

Entonces, para cualquier $b \in \mathbb{R}^n$,

$$|b - \mathbf{P}b| \,=\, |b - \tilde{\mathbf{P}}b|\,.$$

Prueba: definamos $b_P = \mathbf{P}b$ y $b_{\tilde{P}} = \tilde{\mathbf{P}}b$; observe que b_P y $b_{\tilde{P}}$ ambos pertencen a V y satisfacen el lema 4. Entonces,

$$|b - b_P| \overset{(*)}{\leq} |b - b_{\tilde{P}}| \qquad \text{y} \qquad |b - b_{\tilde{P}}| \overset{(*)}{\leq} |b - b_P|\,,$$

en donde (*) se sigue del lema 4. Ambas desigualdades nos llevan a que

$$|b - b_P| \,=\, |b - b_{\tilde{P}}|\,,$$

como se quería demostrar. `Q.E.D.`

Lema 6.: sean b_P y $b_{\tilde{P}}$ como en el lema anterior. Entonces

$$b_P \,=\, b_{\tilde{P}}\,.$$

Prueba: defina $v_t = tb_P + (1-t)b_{\tilde{P}}$, tal que $t \in \mathbb{R}$. Observe que $v_t \in V$ para toda $t \in \mathbb{R}$ y que $v_0 = b_{\tilde{P}}$ y $v_1 = b_P$. Considere la siguiente función:

$$g(t) \;:=\; |b - v_t|^2 \,,$$

es claro que $g(0) = |b - b_{\tilde{P}}|^2$ y que $g(1) = |b - b_P|^2$. Por el lema 5, $g(0) = g(1)$ son los mínimos de $g(t)$. Por otra parte,

$$\begin{aligned} g(t) &= |b - b_{\tilde{P}} + t(b_{\tilde{P}} - b_P)|^2 \\ &= |b - b_{\tilde{P}}|^2 + 2t(b - b_{\tilde{P}}) \cdot (b_{\tilde{P}} - b_P) + t^2 |b_{\tilde{P}} - b_P|^2 \,. \end{aligned}$$

Ahora $b_P - b_{\tilde{P}} \in V$, entonces, por el lema 3 (reemplazando b_P por $b_{\tilde{P}}$), $(b - b_{\tilde{P}}) \cdot (b_{\tilde{P}} - b_P) = 0$, por lo tanto,

$$g(t) \;=\; |b - b_{\tilde{P}}|^2 + t^2 |b_{\tilde{P}} - b_P|^2 \,.$$

De la ecuación anterior se sigue que $g(0)$ es el mínimo absoluto de g, luego $g(1) = g(0) + |b_{\tilde{P}} - b_P|^2$, entonces $g(1) = g(0)$ si y sólo si $|b_{\tilde{P}} - b_P| = 0$, es decir, si y sólo si $b_{\tilde{P}} = b_P$, como se quería demostrar.
`Q.E.D.`

De todo lo anterior se concluye que, dadas cualesquier bases ortonormales β y $\tilde{\beta}$ de $V \leq \mathbb{R}^n$, y $b \in \mathbb{R}^n$ arbitrario, las proyecciones ortogonales, $\mathbf{P}$ y $\tilde{\mathbf{P}}$, como se definieron arriba, dan como resultado el mismo vector, es decir, $\mathbf{P}b = \tilde{\mathbf{P}}b$, esto motiva la siguiente definición:

Definición 2: sea $V \leq \mathbb{R}^n$ y β una base ortonormal de V. Dado cualquier $b \in \mathbb{R}^n$, definimos la *proyección ortogonal de b en V* como

$$b_{_V} \;:=\; \mathbf{P}b \,,$$

en donde $\mathbf{P} \in \mathbb{R}^{n \times n}$ está definida en términos de los elementos de la base ortonormal β, como se hizo anteriormente.

Corolario: sean $\mathbf{P}$ y $\tilde{\mathbf{P}}$ somo se definieron en el lema 5, entonces

$$\mathbf{P} \;=\; \tilde{\mathbf{P}} \,.$$

Prueba: sea e_j el vector canónico tal que todas sus componentes son cero, excepto la j-ésima, entonces, por el lema 6, se tiene que

$$\mathbf{P}e_j \;=\; \tilde{\mathbf{P}}e_j \,,$$

es decir,

$$0 \;=\; (\mathbf{P} - \tilde{\mathbf{P}})e_j \;=\; C_j(\mathbf{P} - \tilde{\mathbf{P}}) \,,$$

en donde $C_j(\mathbf{P} - \tilde{\mathbf{P}})$ denota la columna j-ésima de la diferencia entre $\mathbf{P}$ y $\tilde{\mathbf{P}}$. Ahora,

$$C_j(\mathbf{P} - \tilde{\mathbf{P}}) \;=\; C_j(\mathbf{P}) - C_j(\tilde{\mathbf{P}}) \,,$$

entonces

$$C_j(\mathbf{P}) = C_j(\tilde{\mathbf{P}}) \,,$$

dado que $j \in \{1, \ldots, n\}$ es arbitraria, se sigue que

$$\mathbf{P} \;=\; \tilde{\mathbf{P}} \,.$$

`Q.E.D.`

El anterior corolario justifica entonces la siguiente definición:

Definición 3: sea $V \leq \mathbb{R}^n$ y β una base ortonormal de V. Entonces la matriz $\mathbf{P}$ como se definió en el lema 2, es la *matriz de proyección ortogonal sobre V*.

Jorge Viveros. Centro de Investigación en Matemáticas, Universidad Autónoma del Estado de Hidalgo. Hgo, México.

E-mail address: `jviveros@uaeh.edu.mx`