Semana 11

- Se sugiere antes de resolver los ejercicios ver los videos de YouTube de los temas correspondientes así como también leer la bibliografía recomendada y el material teórico subido en el campus del curso.
- A continuación se presentan algunos ejercicios resueltos y algunas observaciones para resolver los ejercicios 17 a 24 de la Guía 3. Los ejercicios propuestos que no están en la guía (pero que se relacionan con los mismos) no tienen numeración.

Proyecciones ortogonales

Antes de comenzar con los ejercicios de la semana 11, recordemos la definición de proyección ortogonal.

Sean $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo, $\mathcal{S}$ un subespacio de $\mathbb{V}$ y $v \in \mathbb{V}$, decimos que $\hat{v}$ es la *proyección ortogonal* del vector v sobre $\mathcal{S}$ si verifica:

$$\hat{v} \in \mathcal{S} \text{ y } v - \hat{v} \in \mathcal{S}^{\perp}.$$

En el **Ejercicio 16** de la semana 10, probamos que si $\mathcal{S}$ es de dimensión finita, la proyección ortogonal del vector v sobre $\mathcal{S}$ existe y es única. Lo anterior es un resultado muy importante de espacios euclídeos así que lo vamos a resaltar:

Sean $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo, $\mathcal{S}$ un subespacio de $\mathbb{V}$ (de dimensión finita) y $v \in \mathbb{V}$, entonces $\hat{v}$ (la proyección ortogonal del vector v sobre $\mathcal{S}$) existe y es única.

Recordemos que en el **Ejercicio 16** de la semana 10, probamos que si $\{u_1, u_2, \cdots, u_r\}$ es una base ortonormal de $\mathcal{S}$, entonces

$$\hat{v} = \langle v, u_1 \rangle u_1 + \langle v, u_2 \rangle u_1 + \cdots + \langle v, u_r \rangle u_r.$$

Más aún, si $\{v_1, v_2, \cdots, v_r\}$ es una base ortogonal de $\mathcal{S}$, entonces, tomando $u_1 := \frac{v_1}{\|v_1\|}$, $u_2 := \frac{v_2}{\|v_2\|}, \cdots, u_r := \frac{v_r}{\|v_r\|}$, tenemos que $\{u_1, u_2, \cdots, u_r\}$ es una base ortonormal de $\mathcal{S}$. Entonces,

$$\begin{aligned} \hat{v} &= \langle v, u_1 \rangle + \langle v, u_2 \rangle + \cdots + \langle v, u_r \rangle = \frac{\langle v, v_1 \rangle}{\|v_1\|} \frac{v_1}{\|v_1\|} + \frac{\langle v, v_2 \rangle}{\|v_2\|} \frac{v_2}{\|v_2\|} + \cdots + \frac{\langle v, v_r \rangle}{\|v_r\|} \frac{v_r}{\|v_r\|} \\ &= \frac{\langle v, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle v, v_2 \rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle v, v_r \rangle}{\|v_r\|^2} v_r. \end{aligned}$$

Sean $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo, $\mathcal{S} \subseteq \mathbb{V}$ un subespacio de dimensión finita y $\mathcal{B}_{\mathcal{S}} = \{v_1, v_2, \cdots, v_r\}$ una **base ortogonal** de $\mathcal{S}$. Definimos la función $P_{\mathcal{S}} : \mathbb{V} \to \mathbb{V}$ como la función que a cada elemento $v \in \mathbb{V}$ le asigna su proyección ortogonal sobre $\mathcal{S}$, es decir

$$P_{\mathcal{S}}(v) := \hat{v} = \frac{\langle v, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle v, v_2 \rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle v, v_r \rangle}{\|v_r\|^2} v_r.$$

Como vimos en la Semana 10, la función $P_{\mathcal{S}}$ está bien definida porque a cada elemento $v \in \mathbb{V}$ le corresponde un único elemento de $\mathbb{V}$ que es $\hat{v}$. Por otra parte $P_{\mathcal{S}}$ es una transformación lineal, pues si $v, w \in \mathbb{V}$ y $\alpha, \beta \in \mathbb{K}$ entonces

$$\begin{aligned} P_{\mathcal{S}}(\alpha v + \beta w) &= \sum_{i=1}^{r} \frac{\langle\, \alpha v + \beta w, v_i \,\rangle}{\|v_i\|^2} v_i = \alpha \sum_{i=1}^{r} \frac{\langle\, v, v_i \,\rangle}{\|v_i\|^2} v_i + \beta \sum_{i=1}^{r} \frac{\langle\, w, v_i \,\rangle}{\|v_i\|^2} v_i \\ &= \alpha P_{\mathcal{S}}(v) + \beta P_{\mathcal{S}}(w). \end{aligned}$$

Finalmente, como vimos en la Semana 10, observar que como $P_{\mathcal{S}}(v) = \sum_{i=1}^{r} \frac{\langle v, v_i \rangle}{\|v_i\|^2} v_i \in \mathcal{S}$, entonces

$$\begin{aligned} P_{\mathcal{S}}(P_{\mathcal{S}}(v)) &= P_{\mathcal{S}}(\sum_{i=1}^{r} \frac{\langle\, v, v_i \,\rangle}{\|v_i\|^2} v_i) = \sum_{j=1}^{r} \sum_{i=1}^{r} \left\langle \frac{\langle\, v, v_i \,\rangle v_i}{\|v_i\|^2}, v_j \right\rangle \frac{v_j}{\|v_j\|^2} \\ &= \sum_{j=1}^{r} \sum_{i=1}^{r} \frac{\langle\, v, v_i \,\rangle}{\|v_i\|^2} \langle\, v_i, v_j \,\rangle \frac{v_j}{\|v_j\|^2} = \sum_{i=1}^{r} \frac{\langle\, v, v_i \,\rangle}{\|v_i\|^2} v_i = P_{\mathcal{S}}(v), \end{aligned}$$

donde usamos que $\langle\, v_i, v_j \,\rangle = 0$ si $i \neq j$. Por lo tanto $P_{\mathcal{S}}^2 = P_{\mathcal{S}} \circ P_{\mathcal{S}} = P_{\mathcal{S}}$, entonces $P_{\mathcal{S}}$ es un proyector.

Finalmente, observar que

$$Im(P_{\mathcal{S}}) = \mathcal{S} \text{ y } Nu(P_{\mathcal{S}}) = \mathcal{S}^{\perp}.$$

De hecho, como $P_{\mathcal{S}}$ es un proyector, en la Semana 7 (pág. 12) vimos que $v \in Im(P_{\mathcal{S}})$ si y sólo si $v = P_{\mathcal{S}}(v) = \hat{v} \in \mathcal{S}$, donde para la última igualdad usamos la definición de proyección ortogonal del vector v sobre el subespacio $\mathcal{S}$. Meditar por qué esto demuestra que $Im(P_{\mathcal{S}}) = \mathcal{S}$.

Por último, es claro que si $v \in \mathcal{S}^{\perp}$, entonces $P_{\mathcal{S}}(v) = 0_{\mathbb{V}}$, porque $\langle\, v, v_i \,\rangle = 0$ para $i = 1, 2, \cdots, r$. Por otra parte, si $v \in Nu(P_{\mathcal{S}})$ entonces $P_{\mathcal{S}}(v) = \hat{v} = 0_{\mathbb{V}}$, entonces, por definición de proyección ortogonal del vector v sobre el subespacio $\mathcal{S}$, tenemos que $v = v - 0_{\mathbb{V}} = v - \hat{v} \in \mathcal{S}^{\perp}$. Concluimos entonces que $Nu(P_{\mathcal{S}}) = \mathcal{S}^{\perp}$.

Vamos a remarcar todas estas propiedades que probamos para tenerlas presentes:

Sean $(\mathbb{V}, \langle\, \cdot, \cdot \,\rangle)$ un $\mathbb{K}$-espacio euclídeo, $\mathcal{S} \subseteq \mathbb{V}$ un subespacio de dimensión finita y $\mathcal{B}_{\mathcal{S}} = \{v_1, v_2, \cdots, v_r\}$ una **base ortogonal** de $\mathcal{S}$. La función $P_{\mathcal{S}} : \mathbb{V} \to \mathbb{V}$ definida como

$$P_{\mathcal{S}}(v) := \hat{v} = \frac{\langle\, v, v_1 \,\rangle}{\|v_1\|^2} v_1 + \frac{\langle\, v, v_2 \,\rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle\, v, v_r \,\rangle}{\|v_r\|^2} v_r, \tag{1}$$

es una tranformación lineal que cumple:

- $P_{\mathcal{S}}^2 = P_{\mathcal{S}}$, es decir $P_{\mathcal{S}}$ es un proyector.
- $Im(P_{\mathcal{S}}) = \mathcal{S}$ y $Nu(P_{\mathcal{S}}) = \mathcal{S}^{\perp}$. Entonces, claramente $\mathbb{V} = Im(T) \oplus Nu(T)$.
- Como $P_{\mathcal{S}}$ es un proyector y tenemos que $Im(P_{\mathcal{S}}) = \mathcal{S}$ y $Nu(P_{\mathcal{S}}) = \mathcal{S}^{\perp}$, entonces (ver Semana 7 **Ejercicio 24**):
$$P_{\mathcal{S}^{\perp}} = I_{\mathbb{V}} - P_{\mathcal{S}}. \tag{2}$$

Para que no queden dudas: observar que si T es un proyector, tenemos que $Im(T)$ y $Nu(T)$ están en suma directa (no necesariamente son ortogonales). Por otra parte, si T es un proyector ortogonal entonces T es un proyector y además $Im(T) = Nu(T)^{\perp}$, es decir $Im(T)$ y $Nu(T)$ son ortogonales entre sí (en particular están en suma directa). Todo proyector ortogonal es un proyector, pero obviamente no todo proyector es un proyector ortogonal. Por ejemplo, en $\mathbb{R}^2$ con el producto interno canónico, definimos $T : \mathbb{R}^2 \to \mathbb{R}^2$, como

$$T(\begin{bmatrix} 1 \\ 1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad T(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Entonces T es un proyector (verificarlo). Pero como $Nu(T)$ NO es ortogonal a $Im(T)$, concluimos que T NO es un proyector ortogonal.

A partir de la existencia de la proyección ortogonal podemos enunciar las siguientes propiedades muy importantes que usaremos a lo largo de la materia.

Proposición 1. *Sea* $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ *un* $\mathbb{K}$*-espacio euclídeo de dimensión finita y* $\mathcal{S} \subseteq \mathbb{V}$ *un subespacio. Entonces:*

1. $\mathbb{V} = \mathcal{S} \oplus \mathcal{S}^{\perp}$.
2. $\dim(\mathcal{S}^{\perp}) + \dim(\mathcal{S}) = \dim(\mathbb{V})$.
3. $(\mathcal{S}^{\perp})^{\perp} = \mathcal{S}$.

Notar que la Proposición 1 sólo vale cuando $\mathcal{S}$ es un subespacio y $\mathbb{V}$ es de dimensión finita.

Dem. 1. : Es claro que $\mathcal{S} \cap \mathcal{S}^{\perp} = \{0_{\mathbb{V}}\}$. Por otra parte, si $v \in \mathbb{V}$ tenemos que

$$v = P_{\mathcal{S}}(v) + (v - P_{\mathcal{S}}(v)).$$

Por definición de proyección ortogonal, $P_{\mathcal{S}}(v) \in \mathcal{S}$ y $v - P_{\mathcal{S}}(v) \in \mathcal{S}^{\perp}$. Por lo tanto $v \in \mathcal{S} \oplus \mathcal{S}^{\perp}$ y se sigue que $\mathbb{V} \subseteq \mathcal{S} \oplus \mathcal{S}^{\perp}$. Como la otra inclusión siempre vale, tenemos que $\mathbb{V} = \mathcal{S} \oplus \mathcal{S}^{\perp}$.

2. : Usando el item 1. y el Teorema de la dimensión para suma de subespacios, tenemos que

$$\dim(\mathbb{V}) = \dim(\mathcal{S} \oplus \mathcal{S}^{\perp}) = \dim(\mathcal{S}) + \dim(\mathcal{S}^{\perp}).$$

3. : Es claro que $\mathcal{S} \subseteq (\mathcal{S}^{\perp})^{\perp}$, pues si $s \in \mathcal{S}$ entonces $\langle s, t \rangle = 0$ para todo $t \in \mathcal{S}^{\perp}$ entonces $s \in (\mathcal{S}^{\perp})^{\perp}$. Por otra parte, usando el item 2. dos veces, tenemos que

$$\dim((\mathcal{S}^{\perp})^{\perp}) = \dim(\mathbb{V}) - \dim(\mathcal{S}^{\perp}) = \dim(\mathcal{S}).$$

Entonces, como $\mathcal{S} \subseteq (\mathcal{S}^{\perp})^{\perp}$ y $\dim(\mathcal{S}) = \dim((\mathcal{S}^{\perp})^{\perp})$ se sigue que $(\mathcal{S}^{\perp})^{\perp} = \mathcal{S}$. $\square$

El **Ejercicio 18** es un ejemplo de cálculo de la proyección ortogonal $P_{\mathcal{S}}$. Antes de resolverlo, recordemos la definición de distancia de un punto a un subespacio: sean $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo, $\mathcal{S}$ un subespacio de $\mathbb{V}$ y $v \in \mathbb{V}$. Se define la *distancia del punto* v *al subespacio* $\mathcal{S}$ como

$$d(v, \mathcal{S}) := \inf\{\|v - s\| : s \in \mathcal{S}\}.$$

Donde ínf denota el ínfimo del conjunto en cuestión. En el **Ejercicio 16** probamos que cuando $\mathcal{S}$ es de dimensión finita, ese ínfimo de hecho es un mínimo y el vector que realiza el mínimo es justamente $\hat{v} = P_{\mathcal{S}}(v)$. Por lo tanto

$$d(v, \mathcal{S}) = \|v - P_{\mathcal{S}}(v)\| = \|P_{\mathcal{S}^\perp}(v)\|,$$

donde para la última igualdad usamos que por (2), tenemos que $P_{\mathcal{S}^\perp} = I_{\mathbb{V}} - P_{\mathcal{S}}$. Entonces

$$P_{\mathcal{S}^\perp}(v) = I_{\mathbb{V}}(v) - P_{\mathcal{S}}(v) = v - P_{\mathcal{S}}(v).$$

Ejercicio 18: En $\mathbb{R}^{2\times 2}$ con el producto interno $\langle A, B \rangle = \operatorname{tr}(B^T A)$, consideramos el subespacio $\mathcal{S}$ de todas las matrices simétricas.

a) Dada $X \in \mathbb{R}^{2\times 2}$, hallar la expresión de $P_{\mathcal{S}}(X)$.

b) Hallar la proyección ortogonal de $B = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ sobre $\mathcal{S}$. Calcular $d(B, \mathcal{S})$.

Dem. *a)* : El subespacio $\mathcal{S}$ en cuestión es $\mathcal{S} = \{A \in \mathbb{R}^{2\times 2} : A^T = A\}$. Si $A \in \mathcal{S}$, entonces $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ con $a, b, c, d \in \mathbb{R}$ y $A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} = A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Entonces, $b = c$. Por lo tanto $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ y tenemos que

$$\mathcal{S} = gen\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \}.$$

Observar que

$$\left\langle \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\rangle = \operatorname{tr}(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}) = \operatorname{tr}(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}) = 0,$$

$$\left\langle \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle = \operatorname{tr}(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}) = \operatorname{tr}(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}) = 0,$$

$$\left\langle \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle = \operatorname{tr}(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}) = \operatorname{tr}(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}) = 0.$$

Entonces, $B_{\mathcal{S}} = \{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \}$ es una base ortogonal de $\mathcal{S}$. Y podemos usar la ecuación (1). Entonces,

$$P_{\mathcal{S}}(X) = \frac{\left\langle X, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\rangle}{\| \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \|^2} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{\left\langle X, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\rangle}{\| \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \|^2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \frac{\left\langle X, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle}{\| \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \|^2} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Haciendo cuentas, tenemos que $\| \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \|^2 = 1$, $\| \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \|^2 = 1$ y $\| \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \|^2 = 2$. Por otra parte, si $X = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} \in \mathbb{R}^{2\times 2}$, se sigue que

$$\left\langle \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\rangle = \operatorname{tr}(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}) = \operatorname{tr}(\begin{bmatrix} x_1 & x_2 \\ 0 & 0 \end{bmatrix}) = x_1.$$

Operando de manera similar, tenemos que $\left\langle \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\rangle = x_2 + x_3$ y $\left\langle \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle = x_4$. Entonces, reemplazando, nos queda:

$$P_{\mathcal{S}}(X) = x_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + (\frac{x_2 + x_3}{2}) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} x_1 & \frac{x_2+x_3}{2} \\ \frac{x_2+x_3}{2} & x_4 \end{bmatrix}.$$

Observar que (obviamente) $P_{\mathcal{S}}(X)$ siempre es simétrica pues por definición $P_{\mathcal{S}}(X) \in \mathcal{S}$.

Otra manera de resolver este ejercicio es observar que por un lado $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathcal{S}^\perp$ (hacer la cuenta) y como, por la Proposición 2, $\dim(\mathcal{S}^\perp) = \dim(\mathbb{R}^{2\times 2}) - \dim(\mathcal{S}) = 4 - 3 = 1$, tenemos que $\mathcal{S}^\perp = gen\{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \}$. Una base de $\mathcal{S}^\perp$ (ortogonal) es $B_{\mathcal{S}^\perp} = \{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \}$. Entonces (verificarlo haciendo las cuentas),

$$P_{\mathcal{S}^\perp}(X) = \frac{\left\langle X, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\rangle}{\| \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \|^2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{x_2-x_3}{2} \\ \frac{x_3-x_2}{2} & 0 \end{bmatrix}.$$

Entonces, por la ecuación (2),

$$P_{\mathcal{S}}(X) = X - P_{\mathcal{S}^\perp}(X) = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} - \begin{bmatrix} 0 & \frac{x_2-x_3}{2} \\ \frac{x_3-x_2}{2} & 0 \end{bmatrix} = \begin{bmatrix} x_1 & \frac{x_2+x_3}{2} \\ \frac{x_2+x_3}{2} & x_4 \end{bmatrix}.$$

Y (por suerte) obtuvimos el mismo resultado.

b) : Usando la fórmula que acabamos de probar, tenemos que

$$P_{\mathcal{S}}(B) = P_{\mathcal{S}}(\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}) = \begin{bmatrix} 1 & \frac{-1+1}{2} \\ \frac{-1+1}{2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Finalmente,

$$d(B, \mathcal{S}) = \|B - P_{\mathcal{S}}(B)\| = \| \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \| = \| \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \| =$$

$$[\operatorname{tr}(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix})]^{1/2} = [\operatorname{tr}(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix})]^{1/2} = \sqrt{2}.$$

$\square$

Ejercicio 20 : Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo y sea $T \in \mathcal{L}(\mathbb{V})$ un proyector, es decir $T^2 = T$. Verificar que las siguientes afirmaciones son equivalentes:

a) T es una proyección ortogonal,

b) para todo $x, y \in \mathbb{V}$ vale $\langle T(x), y \rangle = \langle x, T(y) \rangle$,

c) para todo $x \in \mathbb{V}$ vale $\|T(x)\| \leq \|x\|$.

Dem. *a)* $\Rightarrow$ *b)* : Supongamos que T es una proyección ortogonal y veamos que para todo $x, y \in \mathbb{V}$ vale $\langle T(x), y \rangle = \langle x, T(y) \rangle$.

Sean $x, y \in \mathbb{V}$, entonces como T es un proyector, vimos que vale que $\mathbb{V} = Im(T) \oplus Nu(T)$ (lo vimos en la Semana 7). Entonces $x = x_1 + x_2$ e $y = y_1 + y_2$, con (únicos) $x_1, y_1 \in Im(T)$ y $x_2, y_2 \in Nu(T)$. Entonces,

$$T(x) = T(x_1 + x_2) = T(x_1) + T(x_2) = T(x_1) + 0_{\mathbb{V}} = T(x_1) = x_1 \text{ y}$$

$$T(y) = T(y_1 + y_2) = T(y_1) + T(y_2) = T(y_1) + 0_{\mathbb{V}} = T(y_1) = y_1,$$

donde usamos que como T es un proyector $T(z) = z$, para todo $z \in Im(T)$, entonces como $x_1, y_1 \in Im(T)$, tenemos que $T(x_1) = x_1$ y $T(y_1) = y_1$.

Como T es una proyección ortogonal tenemos también que $Im(T) \perp Nu(T)$, entonces

$$\langle x_1, y_2 \rangle = \langle x_2, y_1 \rangle = 0.$$

Por lo tanto

$$\begin{aligned} \langle T(x), y \rangle &= \langle x_1, y_1 + y_2 \rangle = \langle x_1, y_1 \rangle + \langle x_1, y_2 \rangle \\ &= \langle x_1, y_1 \rangle + 0 = \langle x_1, y_1 \rangle + \langle x_2, y_1 \rangle = \langle x_1 + x_2, y_1 \rangle = \langle x, T(y) \rangle . \end{aligned}$$

b) $\Rightarrow$ *c)* : Supongamos que para todo $x, y \in \mathbb{V}$ vale $\langle T(x), y \rangle = \langle x, T(y) \rangle$ y veamos que para todo $x \in \mathbb{V}$ vale $\|T(x)\| \leq \|x\|$.

Sea $x \in \mathbb{V}$, usando la hipótesis (es decir que vale a)) y que T es proyector, tenemos que

$$\langle T(x), x - T(x) \rangle = \langle x, T(x - T(x)) \rangle = \left\langle x, T(x) - T^2(x) \right\rangle = \langle x, T(x) - T(x) \rangle = \langle x, 0_{\mathbb{V}} \rangle = 0.$$

Por lo tanto, por el Teorema de Pitágoras (o haciendo la cuenta),

$$\|x\|^2 = \|T(x) + [x - T(x)]\|^2 = \|T(x)\|^2 + \|x - T(x)\|^2 \geq \|T(x)\|^2,$$

donde usamos que $\|x - T(x)\|^2 \geq 0$. Finalmente, tomando raíz (que es una función creciente), tenemos que para todo $x \in \mathbb{V}$ vale $\|T(x)\| \leq \|x\|$.

c) $\Rightarrow$ *d)* : Supongamos que para todo $x \in \mathbb{V}$, $\|T(x)\| \leq \|x\|$ y veamos que T es una proyección ortogonal.

Como T es un proyector (por hipótesis), para probar que T es un proyector ortogonal sólo nos resta probar que $Nu(T)^{\perp} = Im(T)$.

Veamos primero que $Nu(T)^{\perp} \subseteq Im(T)$. Sea $x \in Nu(T)^{\perp}$ queremos ver que $x \in Im(T)$ es decir $T(x) = x$ (recordemos que como T es proyector, $x \in Im(T)$ si y sólo si $T(x) = x$). Como $x - T(x) \in Nu(T)$ pues $T(x - T(x)) = T(x) - T^2(x) = T(x) - T(x) = 0_{\mathbb{V}}$, y $x \in Nu(T)^{\perp}$ se sigue que

$$0 = \langle x, x - T(x) \rangle = \langle x, x \rangle - \langle x, T(x) \rangle = \|x\|^2 - \langle x, T(x) \rangle .$$

Entonces, nos queda que

$$\|x\|^2 = \langle x, T(x) \rangle .$$

Entonces

$$\|x - T(x)\|^2 = \|x\|^2 - 2\ Re \langle x, T(x) \rangle + \|T(x)\|^2 = \|x\|^2 - 2\|x\|^2 + \|T(x)\|^2 = \|T(x)\|^2 - \|x\|^2.$$

Usando la hipótesis, tenemos que

$$\|x - T(x)\|^2 = \|T(x)\|^2 - \|x\|^2 \leq 0.$$

Entonces, como siempre tenemos que $\|x - T(x)\|^2 \geq 0$, se sigue que $\|x - T(x)\|^2 = 0$, entonces $x - T(x) = 0_{\mathbb{V}}$, por lo tanto $T(x) = x$ y $x \in Im(T)$.

Recíprocamente, supongamos que $x \in Im(T)$. Por la Proposición 2, $Nu(T) \oplus Nu(T)^{\perp} = \mathbb{V}$ y tenemos que $x = x_1 + x_2$, con (únicos) $x_1 \in Nu(T)$ y $x_2 \in Nu(T)^{\perp}$. Entonces, por un lado $x = T(x)$ pues $x \in Im(T)$. Por otro lado, acabamos de probar que $Nu(T)^{\perp} \subseteq Im(T)$, entonces, como $x_2 \in Nu(T)^{\perp}$ tenemos que $x_2 \in Im(T)$ y (nuevamente) como T es proyector, se sigue que $T(x_2) = x_2$. Por lo tanto

$$x = T(x) = T(x_1 + x_2) = T(x_1) + T(x_2) = 0_{\mathbb{V}} + T(x_2) = T(x_2) = x_2.$$

En conclusión,

$$x = x_2 \in Nu(T)^{\perp},$$

entonces $x \in Nu(T)^{\perp}$ y tenemos que $Im(T) \subseteq Nu(T)^{\perp}$. Por lo tanto, como probamos la doble inclusión, se sigue que $Nu(T)^{\perp} = Im(T)$ y T es un proyector ortogonal.

Como probamos $a) \Rightarrow b) \Rightarrow c) \Rightarrow a)$, se sigue que $a), b)$ y $c)$ son equivalentes. $\square$

Ejercicio 19 : En $\mathbb{R}_2[x]$ se considera el producto interno definido por

$$\langle p, q \rangle := \int_0^{\infty} p(x)q(x)e^{-x}dx.$$

Calcular

$$\min_{a_1, a_2 \in \mathbb{R}} \int_0^{\infty} (1 + a_1 x + a_2 x^2)^2 e^{-x} dx.$$

Sugerencia para la resolución:

Observar que
$$\int_0^\infty (1 + a_1x + a_2x^2)^2 e^{-x} dx = \|1 + a_1x + a_2x^2\|^2.$$
Donde (abusando de la notación), con $\|1+a_1x+a_2x^2\|^2$ (por ejemplo) queremos decir $\|p\|^2$ con $p(x) = 1 + a_1x + a_2x^2$.

Entonces
$$\begin{aligned}\min_{a_1,a_2\in\mathbb{R}} \int_0^\infty (1 + a_1x + a_2x^2)^2 e^{-x} dx &= \min\{\|1 + (a_1x + a_2x^2)\|^2 : a_1, a_2 \in \mathbb{R}\}\\ &= \min\{\|1 + p\|^2 : p \in gen\{x, x^2\}\}\\ &= \min\{\|1 - p\|^2 : p \in gen\{x, x^2\}\}\\ &= d(1, gen\{x, x^2\})^2\\ &= \|1 - P_{gen\{x,x^2\}}(1)\|^2.\end{aligned}$$
Si no están convencidos, verificar que vale la igualdad de los siguientes conjuntos (probando la doble inclusión):
$$\{\|1 + (a_1x + a_2x^2)\|^2 : a_1, a_2 \in \mathbb{R}\} = \{\|1 + p\|^2 : p \in gen\{x, x^2\}\} = \{\|1 - p\|^2 : p \in gen\{x, x^2\}\}.$$
Entonces, basta obtener una base ortogonal del subespacio $gen\{x, x^2\}$, aplicar la ecuación (1) y calcular $\|1 - P_{gen\{x,x^2\}}(1)\|^2$. Recordar que la ecuación (1) sólo se puede aplicar cuando tenemos una base ortogonal del subespacio en cuestión.

Ejercicio de examen : En $\mathbb{R}_2[x]$ se considera el producto interno definido por
$$\langle p, q \rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1).$$
Sea $\mathcal{U} := \{p \in \mathbb{R}_2[x] : p'(0) = 0\}$ y sea $q(x) := -x^2 + x$. Hallar todos los $p \in \mathcal{U}$ tales que $p \perp q$ y $\|p\| = 20$.

Dem. Si $p \in \mathcal{U}$, entonces $p \in \mathbb{R}_2[x]$, es decir $p(x) = ax^2 + bx + c$ con $a, b, c \in \mathbb{R}$ y $p'(0) = 0$. Entoces, como $p'(x) = 2ax + b$, tenemos que $p'(0) = b = 0$ y, volviendo a la expresión de p, nos queda que $p(x) = ax^2 + c$, con $a, c \in \mathbb{R}$.

Por otra parte, $p \perp q$, entonces $0 = \langle p, q \rangle = \left\langle ax^2 + c, -x^2 + x \right\rangle$. Entonces, como $p(-1) = a + c,\ p(0) = c,\ p(1) = a + c,\ q(-1) = -2,\ q(0) = 0,\ q(1) = 0$. Tenemos que
$$0 = \left\langle ax^2 + b, -x^2 + x \right\rangle = (a + c)(-2) + c \cdot 0 + (a + c) \cdot 0 = -2(a + c).$$
Entonces $a = -c$ y, volviendo a la expresión de p nos queda que $p(x) = ax^2 - a = a(x^2 - 1)$, con $a \in \mathbb{R}$. Finalmente
$$20 = \|p\| = \|a(x^2 - 1)\| = |a|\|x^2 - 1\|.$$
Entonces, como $\|x^2 - 1\|^2 = 0 \cdot 0 + -1 \cdot -1 + 0 \cdot 0 = 1$. Entonces $\|x^2 - 1\| = 1$ y
$$20 = |a|\|x^2 - 1\| = |a|,$$
por lo tanto $a = 20$ ó $a = -20$.

Conclusión, los polinomios $p \in \mathcal{U}$ tales que $p \perp q$ y $\|p\| = 20$, son $p(x) = 20(x^2 - 1)$ y $p(x) = -20(x^2 - 1)$.

$\square$

Mínimos cuadrados

Sean $\mathbb{V}, \mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales y sea $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Dado $w_0 \in \mathbb{W}$, recordemos que en la guía de transformaciones lineales vimos que:

$$\text{existe } v_0 \in \mathbb{V} \text{ tal que } T(v_0) = w_0 \text{ si y sólo si } w_0 \in Im(T).$$

Cuando $w_0 \notin Im(T)$, sabemos que la ecuación $T(v) = w_0$ NO tiene solución. En ese caso, nos interesa obtener una solución aproximada (en algún sentido) de dicha ecuación.

Si $\mathbb{W}, \langle \cdot, \cdot \rangle$ es un $\mathbb{K}$-espacio euclídeo, el problema de *mínimos cuadrados* consiste en encontrar (si existen) todos los vectores $\hat{v} \in \mathbb{V}$ tales que la distancia de $T(\hat{v})$ a w_0 sea mínima, es decir, buscamos $\hat{v} \in \mathbb{V}$ tal que

$$\|w_0 - T(\hat{v})\| = \text{mín}\{\|w_0 - T(v)\| : v \in \mathbb{V}\}.$$

Notar que

$$\{\|w_0 - T(v)\| : v \in \mathbb{V}\} = \{\|w_0 - w\| : w \in Im(T)\}$$

(verificarlo probando la doble inclusión).

Entonces,

$$\|w_0 - T(\hat{v})\| = \text{mín}\{\|w_0 - w\| : w \in Im(T)\}.$$

Ya vimos en el **Ejercicio 16** que si $Im(T)$ es de dimensión finita, entonces el vector $P_{Im(T)}(w_0)$ es el (único) vector de $Im(T)$ que minimiza la distancia a w_0. Por lo tanto los vectores $\hat{v}$ que buscamos son las soluciones del sistema

$$T(v) = P_{Im(T)}(w_0),$$

como $P_{Im(T)}(w_0) \in Im(T)$, la ecuación $T(v) = P_{Im(T)}(w_0)$ siempre tiene solución y será única si y sólo si T es un monomorfismo. Veamos un ejemplo para fijar ideas.

Ejercicio de examen : Sea la transformación lineal $T : \mathbb{R}_3[x] \to \mathbb{R}^4$ definida por

$$T(p) = \begin{bmatrix} p(-1) \\ p(0) \\ p(1) \\ \int_{-1}^{1} p(x)dx \end{bmatrix}.$$

Consideremos en $\mathbb{R}^3$ el producto interno canónico $\langle x, y \rangle = x^T y$ y la ecuación

$$T(p) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}.$$

a) Decidir si existe $p \in \mathbb{R}_3[x]$ tal que $T(p) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}$.

b) Si la respuesta de a) es negativa, hallar todas las soluciones aproximadas (en el sentido de mínimos cuadrados) de dicha ecuación y estimar el error cometido.

Dem. $a)$: Primero vamos a obtener una base de $Im(T)$. Recordemos que como $\{1, x, x^2, x^3\}$ es un sistema de generadores de $\mathbb{R}_3[x]$, tenemos que $Im(T) = gen\{T(1), T(x), T(x^2), T(x^3)\}$.

Observar que $\int_{-1}^{1} 1dx = 2$, $\int_{-1}^{1} xdx = 0$, $\int_{-1}^{1} x^2dx = \frac{2}{3}$ $\int_{-1}^{1} x^3dx = 0$. Entonces

$$T(1) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix},\ T(x) = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix},\ T(x^2) = \begin{bmatrix} 1 \\ 0 \\ 1 \\ \frac{2}{3} \end{bmatrix} \text{ y } T(x^3) = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Entonces,

$$Im(T) = gen\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ \frac{2}{3} \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \}.$$

Y una base de $Im(T)$ puede ser (verificarlo)

$$\mathcal{B}_{Im(T)} = \{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ \frac{2}{3} \end{bmatrix} \}.$$

Como $\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix} \notin Im(T)$ (verificarlo). No existe p tal que $T(p) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}$.

$b)$: Como vimos arriba, todas las soluciones aproximadas (en el sentido de mínimos cuadrados) de dicha ecuación son las soluciones del sistema (ahora compatible) $T(p) = P_{Im(T)}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix})$.

Observar que (calcularlo) $Im(T)^{\perp} = gen\{ \begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix} \}$. Por la ecuación (2), tenemos que

$$P_{Im(T)}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix} - P_{Im(T)^{\perp}}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}).$$

Como $\{ \begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix} \}$ es una base ortogonal de $Im(T)^{\perp}$ (tiene un sólo elemento), podemos aplicar la ecuación (1), para obtener que

$$P_{Im(T)^{\perp}}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}) = \frac{\left\langle (\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix} , \begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix} \right\rangle}{\| \begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix} \|^2} \begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix} = \frac{-27}{27} \begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 1 \\ -3 \end{bmatrix}.$$

Entonces,

$$P_{Im(T)}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix} - P_{Im(T)^{\perp}}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 \\ 4 \\ 1 \\ -3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}.$$

Por lo tanto, buscamos $p \in \mathbb{R}_3[x]$ tales que $T(p) = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}$.

Si $p \in \mathbb{R}_3[x]$ es tal que $T(p) = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}$ entonces, $p(x) = a + bx + cx^2 + dx^3$ con $a, b, c, d \in \mathbb{R}$, tal que $T(p) = aT(1) + bT(x) + cT(x^2) + dT(x^3) = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}$. Entonces,

$$a \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix} + b \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 1 \\ 0 \\ 1 \\ \frac{2}{3} \end{bmatrix} + d \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}.$$

Resolviendo el sistema, nos queda que $a = 1, b = 1 - d, c = 0$. Volviendo a la expresión de p nos queda que $p(x) = 1 + x + d(x^3 - x)$, con $d \in \mathbb{R}$. Por lo tanto, todas las soluciones de mínimos cuadrados de la ecuación $T(p) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ 1 \end{bmatrix}$ son

$$\hat{p}(x) = 1 + x + d(x^3 - x), \text{ con } d \in \mathbb{R}.$$

Finalmente, el error cometido, se obtiene como

$$e = \|T(\hat{p}) - \begin{bmatrix} 1 \\ 5 \\ 3 \\ 1 \end{bmatrix} \| = \| \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 5 \\ 3 \\ 1 \end{bmatrix} \| = \| \begin{bmatrix} -1 \\ -4 \\ -1 \\ 1 \end{bmatrix} \| = \sqrt{19}.$$

$\square$

En el siguiente ejercicio vamos a estudiar un caso particular del problema de mínimos cuadrados cuando consideramos matrices $A \in \mathbb{R}^{m\times n}$ y el espacio $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ euclídeo canónico.

Antes de resolver el **Ejercicio 21** veamos las siguiente propiedades que usaremos ampliamente.

Proposición 2. *Sea $A \in \mathbb{R}^{n\times m}$ y consideremos el espacio euclídeo canónico correspondiente. Entonces:*

1. $nul(A^T A) = nul(A)$.
2. $col(A^T A) = col(A^T)$.
3. $nul(A) = col(A^T)^{\perp}$.
4. $nul(A^T) = col(A)^{\perp}$.

Dem. 1. : Vamos a probar $nul(A^T A) = nul(A)$ viendo la doble inclusión. Ya sabemos que siempre vale $nul(A) \subseteq nul(A^T A)$ (si no se acuerdan, verificarlo). Por otra parte, si $x \in nul(A^T A)$ entonces $A^T Ax = 0$. Entonces,

$$\|Ax\|^2 = (Ax)^T(Ax) = x^T A^T(Ax) = x^T(A^T Ax) = x^T 0 = 0.$$

Entonces $Ax = 0$, $x \in nul(A)$ y tenemos que $nul(A^T A) \subseteq nul(A)$. Como probamos la doble inclusión, se sigue que $nul(A^T A) = nul(A)$.

2. : Recordemos que en el **Ejercicio 11**, probamos que si $A \in \mathbb{R}^{n\times m}$ y $(\mathbb{R}^m, \langle \cdot, \cdot \rangle)$ es el espacio euclídeo canónico entonces

$$nul(A)^{\perp} = col(A^T).$$

Por lo tanto, usando esa propiedad y que (acabamos de ver que) $nul(A^T A) = nul(A)$, tenemos que

$$col(A^T A) = col((A^T A)^T) = nul(A^T A)^{\perp} = nul(A)^{\perp} = col(A^T).$$

3. : Como $nul(A)$ es un subespacio, por la Proposición 1, tenemos que

$$nul(A) = (nul(A)^{\perp})^{\perp}$$

y usando que $nul(A)^{\perp} = col(A^T)$ se sigue que

$$nul(A) = (nul(A)^{\perp})^{\perp} = col(A^T)^{\perp}.$$

4. : Finalmente, por 3.,

$$nul(A^T) = col((A^T)^T)^{\perp} = col(A)^{\perp}.$$

$\square$

Ejercicio 21: Sea $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ el espacio euclídeo canónico. Se considera un subespacio $\mathcal{S}$ de $\mathbb{R}^n$ y una matriz $A \in \mathbb{R}^{n\times m}$ cuyas columnas son una base de $\mathcal{S}$.

a) Demostrar que $P_{\mathcal{S}}(v)$ satisface el siguiente sistema de ecuaciones:

$$\begin{cases} P_{\mathcal{S}}(v) = A\hat{x} & \text{para algún } \hat{x} \in \mathbb{R}^m, \\ A^T(v - P_{\mathcal{S}}(v)) = 0. & \end{cases}$$

Comprobar que $\hat{x} = A^{\#}v$, donde $A^{\#} := (A^TA)^{-1}A^T$.

b) Deducir que $AA^{\#} = [P_{\mathcal{S}}]_E^E$ donde E es la base canónica de $\mathbb{R}^n$ y que $A^{\#}A = I_{\mathbb{R}^m}$. **Observar que hay dos errores en la guía**.

c) Concluir que $d(v, \mathcal{S})^2 = \|v - A\hat{x}\|^2$. Motivo por el cual $\hat{x}$ se denomina (la) *solución por mínimos cuadrados de la ecuación Ax=v.*

Dem. $a)$: Notar que $\mathcal{S} = col(A)$. Más aún, como las columnas de A forman una base de $\mathcal{S}$ tenemos que $\dim(\mathcal{S}) = \dim(col(A)) = rg(A) = m$. Recordemos que la ecuación $Ax = y$ tiene solución si y sólo si $y \in col(A)$, como en este caso $P_{\mathcal{S}}(v) = P_{col(A)}(v) \in \mathcal{S} = col(A)$ (por definición de proyección ortogonal) entonces existe algún vector (llamémoslo) $\hat{x} \in \mathbb{R}^m$ tal que

$$A\hat{x} = P_{\mathcal{S}}(v).$$

Además, como $rg(A) = m$, por el teorema de la dimensión tenemos que

$$\dim(nul(A)) = m - rg(A) = m - m = 0.$$

Entonces $nul(A) = \{0\}$ y entonces, la ecuación $P_{\mathcal{S}}(v) = Ax$, tiene solución única y es $\hat{x}$.

Por otra parte, por definición de proyección ortogonal, tenemos que

$$v - P_{\mathcal{S}}(v) \in \mathcal{S}^{\perp} = col(A)^{\perp} = nul(A^T),$$

donde usamos la Proposición 2. Entonces

$$A^T(v - P_{\mathcal{S}}(v)) = 0.$$

Y tenemos que $P_{\mathcal{S}}(v)$ satisface el sistema de ecuaciones.

Finalmente, observar que por la Proposición 2, $col(A^TA) = col(A^T)$. Como $A^TA \in \mathbb{R}^{m\times m}$ y $\dim(col(A^TA)) = \dim(col(A^T)) = rg(A^T) = rg(A) = m$, se sigue que A^TA es inversible y entonces tiene sentido la matriz $A^{\#}$ definida por $A^{\#} := (A^TA)^{-1}A^T$.

Sea $\hat{x} := A^{\#}v = (A^TA)^{-1}A^Tv$, veamos que $\hat{x}$ (así definido) cumple que $A\hat{x} = P_{\mathcal{S}}(v)$.

De hecho, por un lado

$$A\hat{x} = A[(A^TA)^{-1}A^Tv] \in col(A) = \mathcal{S}$$

y por otro lado $v - A\hat{x} \in nul(A^T) = col(A)^\perp$. De hecho,

$$A^T(v - A\hat{x}) = A^T(v - A(A^TA)^{-1}A^Tv) = A^Tv - A^TA(A^TA)^{-1}A^Tv =$$

$$= A^Tv - IA^Tv = A^Tv - A^Tv = 0.$$

Entonces, probamos que $A\hat{x} \in \mathcal{S}$ y $v - A\hat{x} \in \mathcal{S}^\perp$. Como $P_{\mathcal{S}}(v)$ es el único vector de $\mathbb{R}^n$ tal que $P_{\mathcal{S}}(v) \in \mathcal{S}$ y $v - P_{\mathcal{S}}(v) \in \mathcal{S}^\perp$ (recordar el **Ejercicio 16**) entonces

$$A\hat{x} = P_{\mathcal{S}}(v)$$

y probamos lo que queríamos.

b) : Veamos que $AA^\# = [P_{\mathcal{S}}]_E^E$ y que $A^\#A = I_{\mathbb{R}^m}$.

Queremos probar una igualdad de matrices, recordar que dos matrices $A, B \in \mathbb{R}^{n\times m}$ son iguales si y sólo si $Az = Bz$ para todo $z \in \mathbb{R}^m$ (si no se acuerda de esto, traten de demostrarlo pensando qué pasa si tomamos como z a los vectores de la base canónica de $\mathbb{R}^m$).

Sea $v \in \mathbb{R}^n$, usando lo que probamos en el item *a*) y llamando E a la base canónica de $\mathbb{R}^n$, tenemos que

$$AA^\#v = A\hat{x} = P_{\mathcal{S}}(v) = [P_{\mathcal{S}}(v)]^E = [P_{\mathcal{S}}]_E^E\ [v]^E = [P_{\mathcal{S}}]_E^E\ v.$$

Como la igualdad anterior vale para todo $v \in \mathbb{R}^n$, concluimos que $AA^\# = [P_{\mathcal{S}}]_E^E$.

Finalmente, sea $z \in \mathbb{R}^m$ entonces,

$$A^\#Az = (A^TA)^{-1}A^T(Az) = (A^TA)^{-1}(A^TA)z = I_{\mathbb{R}^m}z.$$

Entonces, como la igualdad anterior vale para todo $z \in \mathbb{R}^m$, concluimos que $A^\#A = I_{\mathbb{R}^m}$.

c) : Usando lo probado en *a*) tenemos que

$$d(v, \mathcal{S})^2 = \|v - P_{\mathcal{S}}(v)\|^2 = \|v - A\hat{x}\|^2.$$

Una aclaración respecto de por qué está entre paréntesis el artículo “la”. En este caso, $\hat{x}$ es “la” solución por mínimos cuadrados de la ecuación $Ax = v$ porque como $nul(A) = \{0\}$ la solución de mínimos cuadrados es única. Como vimos en el ejemplo previo a este ejercicio si $\dim(nul(A)) > 0$ hay infinitas soluciones por mínimos cuadrados. $\square$

Vamos a remarcar las conclusiones del ejercicio anterior para tenerlas presentes:

Sea $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ el espacio euclídeo canónico, $A \in \mathbb{R}^{n\times m}$ e $y \in \mathbb{R}^n$. Entonces, son equivalentes:

a) $\hat{x} \in \mathbb{R}^m$ es una solución de mínimos cuadrados de la ecuación $Ax = y$.

b) $\|A\hat{x} - y\| = \text{mín}\{\|Ax - y\| : x \in \mathbb{R}^m\}$ ó, equivalentemente, $\|A\hat{x} - y\| \leq \|Ax - y\|$ para todo $x \in \mathbb{R}^m$.

c) $\hat{x} \in \mathbb{R}^m$ es una solución de la ecuación $Ax = P_{col(A)}(y)$, es decir $A\hat{x} = P_{col(A)}(y)$.

d) $\hat{x} \in \mathbb{R}^m$ es una solución de la *ecuación normal* $A^T Ax = A^T y$, es decir $A^T A\hat{x} = A^T y$.

En este caso, si $\hat{x}_p$ es una solución de mínimos cuadrados de $Ax = y$ entonces todas las soluciones de mínimos cuadrados de $Ax = y$ son

$$\hat{x} = \hat{x}_p + z,$$

con $z \in nul(A)$. En particular, existe una única solución de mínimos cuadrados de $Ax = y$ si y sólo si $nul(A) = \{0\}$.

Para hacer alguna cuenta, consideremos el siguiente ejercicio.

Ejercicio: La siguiente tabla muestra, para algunos instantes t (en segundos), la posición de un móvil $p(t)$ (en metros).

t	10	20	30	40
p(t)	2,5	5,2	6,6	9,4

a) Suponiendo velocidad constante. Hallar la recta $p(t)$ que mejor ajusta los datos (en el sentido de mínimos cuadrados).

b) Determinar en qué posición se estima que estará el móvil en $t = 50$ seg.

c) Calcular el error de estimación.

Dem. $a)$: Buscamos una recta que mejor ajuste los datos. La recta será de la forma

$$p(t) = p_0 + vt,$$

con p_0 y v a determinar. Además, la recta que buscamos, queremos que cumpla:

$$2,5 = p(10) = p_0 + v \cdot 10.$$

$$5,2 = p(20) = p_0 + v \cdot 20.$$

$$6,6 = p(30) = p_0 + v \cdot 30.$$

$$9,4 = p(40) = p_0 + v \cdot 40.$$

Si llamamos $b := \begin{bmatrix} 2,5 \\ 5,2 \\ 6,6 \\ 9,4 \end{bmatrix}$ y $A := \begin{bmatrix} 1 & 10 \\ 1 & 20 \\ 1 & 30 \\ 1 & 40 \end{bmatrix}$. Buscamos la solución del sistema $Ax = b$, con $x := \begin{bmatrix} p_0 \\ v \end{bmatrix}$. Como claramente ese sistema no tiene solución, vamos a buscar una solución aproximada aplicando mínimos cuadrados.

Por lo que vimos arriba, $\hat{x}$ es una solución de mínimos cuadrados de $Ax = b$ si y sólo si

$$A^T A\hat{x} = A^T b.$$

Entonces, como $A^T A = \begin{bmatrix} 3000 & 100 \\ 100 & 4 \end{bmatrix}$ y $A^T b = \begin{bmatrix} 703 \\ 23,7 \end{bmatrix}$.

Buscamos las soluciones del sistema

$$\begin{bmatrix} 3000 & 100 \\ 100 & 4 \end{bmatrix} x = \begin{bmatrix} 703 \\ 23,7 \end{bmatrix}.$$

Como $rg(A) = rg(A^T A) = 2$, el sistema tiene única solución y es $\hat{x} = \begin{bmatrix} 0,4 \\ 0,221 \end{bmatrix}$. Entonces, la recta que mejor ajusta los datos (en el sentido de mínimos cuadrados) es

$$p(t) = 0,4 + 0,221 \cdot t.$$

$b)$: Usando la recta que mejor ajusta los datos (en el sentido de mínimos cuadrados) tenemos que $p(50) = 11,45$ m.

$c)$: El error cometido se obtiene calculando

$$e = \|A\hat{x} - b\| = \| \begin{bmatrix} 2,61 \\ 4,82 \\ 7,03 \\ 9,24 \end{bmatrix} - \begin{bmatrix} 2,5 \\ 5,2 \\ 6,6 \\ 9,4 \end{bmatrix} \| = [(0,11)^2 + (-0,38)^2 + (0,43)^2 + (0,16)^2]^{1/2}$$

$$= 0,613.$$

$\square$

Sugerencia para resolver el Ejercicio 23 :

Observar que hay algunos errores en la guía.

Se consideran n datos experimentales $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \cdots, \begin{bmatrix} x_n \\ y_n \end{bmatrix}$, donde $x_i \neq x_j$ para todo $i \neq j$. Se postula la existencia de un polinomio de $\mathbb{R}_m[x]$ (**importante:** $m < n-1$) tal que se minimice el *error cuadrático medio*. Es decir, se busca (si existe)

$$\text{mín}\{ \frac{1}{n} \sum_{i=1}^{n} (p_m(x_i) - y_i)^2 : p_m \in \mathbb{R}_m[x]\}.$$

$a)$: Primero, notar que (si existe) $\text{mín}\{\sum_{i=1}^{n}\ (p_m(x_i) - y_i)^2 : p_m \in \mathbb{R}_m[x]\}$ entonces existe el mínimo que nos interesa y además se cumple que

$$\text{mín}\{\frac{1}{n}\ \sum_{i=1}^{n}\ (p_m(x_i) - y_i)^2 : p_m \in \mathbb{R}_m[x]\} = \frac{1}{n}\ \text{mín}\{\sum_{i=1}^{n}\ (p_m(x_i) - y_i)^2 : p_m \in \mathbb{R}_m[x]\}.$$

Si no se entiende la igualdad anterior, demostrarla recordando qué significa que m sea el mínimo de un conjunto. Entonces, nos podemos concentrar en estudiar directamente (si existe)

$$\text{mín}\{\sum_{i=1}^{n}\ (p_m(x_i) - y_i)^2 : p_m \in \mathbb{R}_m[x]\}.$$

Desarrollar todas las sumatorias si no se entiende y recordar que como estamos trabajando con el producto interno canónico de $\mathbb{R}^n$ vale que $\|[x_1\ x_2\ \cdots\ x_n]^T\|^2 = x_1^2 + x_2^2 + \cdots + x_n^2$. Haciendo eso, veremos que se siguen las siguientes igualdades de conjuntos:

$$\begin{aligned}
&\{\sum_{i=1}^{n}\ (p_m(x_i) - y_i)^2 : p_m \in \mathbb{R}_m[x]\} = \{\sum_{i=1}^{n}\ [(\sum_{j=0}^{m}\ a_j x_i^j) - y_i]^2 : a_0, a_1, \cdots, a_m \in \mathbb{R}\} \\
&= \{(a_0 + a_1 x_1 + \cdots + a_m x_1^m - y_1)^2 + \cdots + (a_0 + a_1 x_n + \cdots + a_m x_n^m - y_n)^2 : a_0, a_1, \cdots, a_m \in \mathbb{R}\} \\
&= \{\| \begin{bmatrix} a_0 + a_1 x_1 + \cdots + a_m x_1^m - y_1 \\ a_0 + a_1 x_2 + \cdots + a_m x_2^m - y_2 \\ \vdots \\ a_0 + a_1 x_n + \cdots + a_m x_n^m - y_n \end{bmatrix} \|^2 : a_0, a_1, \cdots, a_m \in \mathbb{R}\} \\
&= \{\|a_0 \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + a_1 \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \cdots + a_m \begin{bmatrix} x_1^m \\ x_2^m \\ \vdots \\ x_n^m \end{bmatrix} - y\|^2 : a_0, a_1, \cdots, a_m \in \mathbb{R}\} \\
&= \{\|z - y\|^2 : z \in col(V_m(x_1, x_2, \cdots, x_m))\}.
\end{aligned}$$

Por lo tanto, $\text{mín}\{\sum_{i=1}^{n}\ (p_m(x_i) - y_i)^2 : p_m \in \mathbb{R}_m[x]\}$ existe y además (recordar el **Ejercicio 16** o la definición de distancia de un punto a un subespacio):

$$\begin{aligned}
\text{mín}\{\sum_{i=1}^{n}\ (p_m(x_i) - y_i)^2 : p_m \in \mathbb{R}_m[x]\} &= \text{mín}\ \{\|z - y\|^2 : z \in col(V_m(x_1, x_2, \cdots, x_m))\} \\
&= \|P_{col(V_m(x_1, x_2, \cdots, x_n))}(y) - y\|^2.
\end{aligned}$$

$b)$: Primero observar que en el ejercicio debería decir $V_{m-1}(x_1, x_2, \cdots, x_m)$. En ese caso, la matriz $V_{m-1}(x_1, x_2, \cdots, x_m) \in \mathbb{R}^{m\times m}$, nos queda:

$$V_{m-1}(x_1, x_2, \cdots, x_m) = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{m-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{m-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^{m-1} \end{bmatrix}.$$

La base $\mathcal{L}$ (de $\mathbb{R}_{m-1}[x]$) de los polinomios interpoladores de Lagrange correspondientes al conjunto de absisas $\{x_1, \cdots, x_m\}$ nos queda

$$\mathcal{L} = \{\frac{(x-x_1)(x-x_2)\cdots(x-x_m)}{(x_1-x_2)(x_1-x_3)\cdots(x_1-x_m)}, \cdots, \frac{(x-x_1)(x-x_2)\cdots(x-x_m)}{(x_m-x_1)(x_m-x_2)\cdots(x_m-x_{m-1})}\}.$$

En el **Ejercicio 1.30**, vimos que para $0 \leq k \leq m-1$, vale que

$$[x^k]^{\mathcal{L}} = [x_1^k \; x_2^k \; \cdots \; x_m^k]^T.$$

Entonces

$$[1]^{\mathcal{L}} = [1 \; 1 \cdots \; 1]^T, \; [x]^{\mathcal{L}} = [x_1 \; x_2 \; \cdots x_m]^T, \; \cdots, \; [x^{m-1}]^{\mathcal{L}} = [x_1^{m-1} \; x_2^{m-1} \; \cdots x_m^{m-1}]^T.$$

Por lo tanto si $\mathcal{E} = \{1, x, x^2, \cdots, x^{m-1}\}$ es la base canónica de $\mathbb{R}_{m-1}[x]$, escribiendo cómo quedan las coordenadas en base $\mathcal{L}$ de cada vector de la base canónica de $\mathbb{R}_{m-1}[x]$ (usando lo que vimos del **Ejercicio 1.30**), tenemos que :

$$[M]_{\mathcal{E}}^{\mathcal{L}} = [\; [1]^{\mathcal{L}} \; [x]^{\mathcal{L}} \; [x^2]^{\mathcal{L}} \; \cdots \; [x^{m-1}]^{\mathcal{L}}] = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{m-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{m-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^{m-1} \end{bmatrix} = V_{m-1}(x_1, x_2, \cdots, x_m).$$

Es claro entonces que como $[M]_{\mathcal{E}}^{\mathcal{L}}$ es inversible (pues es una matriz de cambio de base), se sigue que

$$rg(V_{m-1}(x_1, x_2, \cdots, x_m)) = rg([M]_{\mathcal{E}}^{\mathcal{L}}) = m.$$

c) : Como $m < n-1$ o lo que es lo mismo que $m+1 < n$, tiene sentido la matriz $V_m(x_1, x_2, \cdots, x_m, x_{m+1})$ (meditar por qué).

Por el item b), sabemos que $rg(V_{m-1}(x_1, x_2, \cdots, x_m)) = m$, entonces es inmediato ver que

$$rg(V_m(x_1, x_2, \cdots, x_m, x_{m+1})) = m+1$$

(meditar por qué vale eso).

Observar que

$$V_m(x_1, x_2, \cdots, x_n) \in \mathbb{R}^{n \times (m+1)}.$$

Además recordando que $m+1 < n$, tenemos que

$$V_m(x_1, x_2, \cdots, x_n) = V_m(x_1, x_2, \cdots, x_m, x_{m+1}, \cdots, x_n)$$
$$= \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^m \\ 1 & x_{m+1} & x_{m+1}^2 & \cdots & x_{m+1}^m \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{bmatrix} = \begin{bmatrix} V_m(x_1, x_2, \cdots, x_{m+1}) \\ \vdots \\ 1 \; x_n \; x_n^2 \cdots x_n^{m-1} \; x_n^m \end{bmatrix}.$$

Por lo tanto

$$rg(V_m(x_1, x_2, \cdots, x_n)) = rg(V_m(x_1, x_2, \cdots, x_m, x_{m+1})) = m + 1.$$

$d)$: Si ahora $\mathcal{E}$ es la base canónica de $\mathbb{R}_m[x]$, tenemos que

$$[p_m]^{\mathcal{E}} = [a_0 \ a_1 \ \cdots \ a_m]^T.$$

Por otra parte, como $rg(V_m(x_1, x_2, \cdots, x_n) = m + 1$ (es decir $V_m(x_1, x_2, \cdots, x_n)$ tiene rango máximo), por el **Ejercicio 21**, tenemos que el vector $\hat{a} := [a_0 \ a_1 \ \cdots \ a_m]^T$ es el (único) vector que cumple

$$P_{col(V_m(x_1,x_2,\cdots,x_n))}(y) = V_m(x_1, x_2, \cdots, x_n)\hat{a}.$$

Nuevamente, por el **Ejercicio 21**, $\hat{a} = V_m(x_1, x_2, \cdots, x_n)^{\#}y$. Entonces

$$[p_m]^{\mathcal{E}} = \hat{a} = V_m(x_1, x_2, \cdots, x_n)^{\#}y.$$