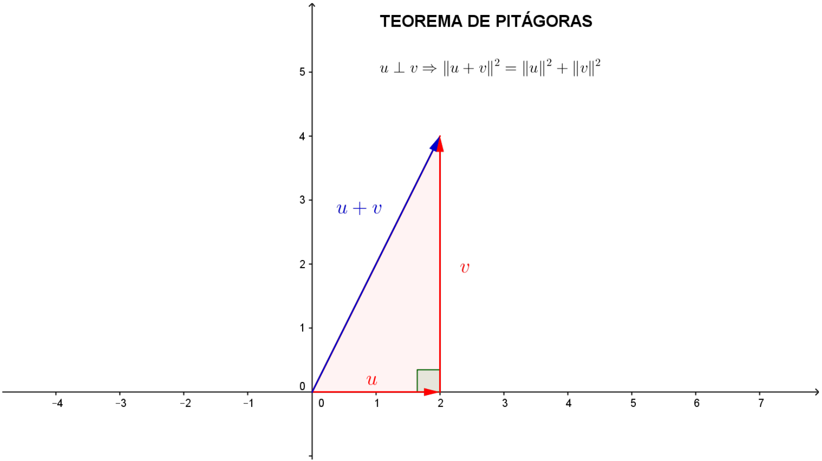
*"...No veo televisión ni las revistas*
*No veo ya nada que no pueda ser*
*Por eso yo no voy en tren, voy en avión...."*
Charly García

# Producto interno
# Complemento ortogonal-Proyección ortogonal

Curso 1

**TEOREMA DE PITÁGORAS**

$$u \perp v \Rightarrow \|u+v\|^2 = \|u\|^2 + \|v\|^2$$



# Teorema de Pitágoras

Si $\mathbb{V}$ es un $\mathbb{K}$-espacio vectorial con Producto Interno, se cumple:

$$\boxed{u \perp v \Rightarrow \|u+v\|^2 = \|u\|^2 + \|v\|^2}$$

Demostración:

$u \perp v \iff \langle u, v\rangle = 0$

$$\begin{aligned} \|u+v\|^2 &= \langle u+v, u+v\rangle = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\ &= \|u\|^2 + \underbrace{\langle u, v\rangle}_{=0} + \underbrace{\langle v, u\rangle}_{\overline{\langle u,v\rangle}=0} + \|v\|^2 \end{aligned}$$

$$\therefore\ \|u+v\|^2 = \|u\|^2 + \|v\|^2 \ \checkmark$$

# Observaciones:

- El recíproco del teorema de Pitágoras sólo es valido si $\mathbb{V}$ es un $\mathbb{R}$- espacio vectorial.
  Si para un par de vectores se cumple:

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2 \iff$$

$$\begin{aligned} \iff \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle &= \langle u,u\rangle + \langle v,v\rangle \iff \\ \iff \langle u,v\rangle + \underbrace{\langle v,u\rangle}_{=\overline{\langle u,v\rangle}} &= 0 \\ \langle u,v\rangle + \overline{\langle u,v\rangle} &= 0 \\ 2\mathbb{R}e(\langle u,v\rangle) &= 0 \end{aligned}$$

  Si $\mathbb{V}$ es un $\mathbb{R}$-espacio
  vectorial, $\langle u,v\rangle \in \mathbb{R} \Rightarrow \langle u,v\rangle = 0 \Rightarrow u \perp v$

Si $\mathbb{V}$ es un $\mathbb{C}$-espacio vectorial, la única implicación es que $\mathbb{R}e(\langle u, v\rangle) = 0$ y esto no significa que los vectores sean ortogonales, sólo implica que el producto interno entre los dos vectores es un complejo de la forma $z = bi$.

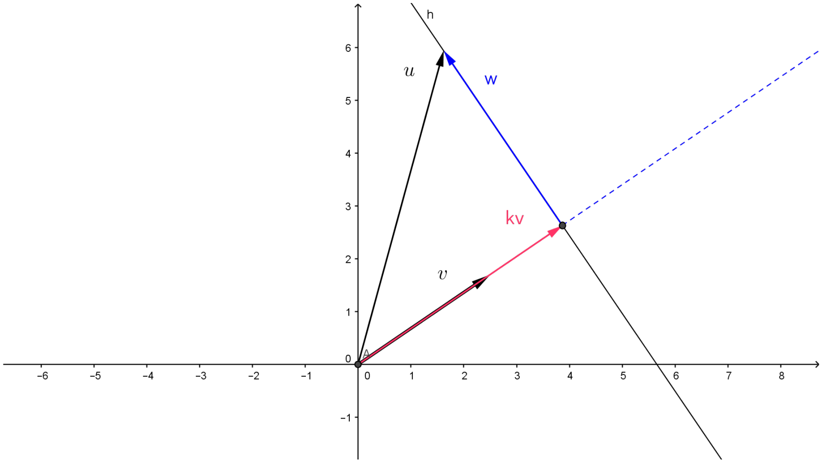
Como contraejemplo inmediato: Tomemos en $\mathbb{C}^n$ el P.I. canónico, $\langle x, y\rangle = y^* x = \bar{y}^T x$. Si $x = (1\ \ 0)^T$ e $y = (3i\ \ 1)^T$, entonces:

$$\left\langle (1\ \ 0)^T (3i\ \ 1)^T \right\rangle = (\bar{3}i\ \ \bar{1}) \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$(-3i\ \ 1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -3i = -3i \neq 0$$

$\therefore x \not\perp y$, pero:
$\|x + y\|^2 = \|(1 + 3i\ \ 1)^T\|^2 = |1 + 3i|^2 + 1^2 = 1 + 9 + 1 = 11$
$\|x\|^2 = 1,$ y $\|y\|^2 = 9 + 1 = 10.$
Entonces $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ y $x \not\perp y$.

# Descomposición Ortogonal



$\forall u, v \in \mathbb{V}, v \neq 0_{\mathbb{V}}$, si tomamos $c = \frac{\langle u,v\rangle}{\|v\|^2}$, y $w = u - (\frac{\langle u,v\rangle}{\|v\|^2})v$, se cumple:

$$u = cv + w \text{ con } w \perp u$$

En todo todo lo que sigue, siempre que hablamos de $\mathbb{V}$, estamos hablando de $\mathbb{V} - \mathbb{K}$ espacio vectorial con P.I.

Definición: Se dice que $\{v_1, v_2, \ldots, v_k\} \subset \mathbb{V}$ es un **conjunto ortogonal**, si $\langle v_i, v_j \rangle = 0 \; \forall \; i \neq j$.
Aclaración: Todo conjunto con un sólo elemento, $\{v_1\}$ se considera ortogonal.

Definición: Se dice que $\{v_1, v_2, \ldots, v_k\} \subset \mathbb{V}$ es un **conjunto ortonormal**, si $\langle v_i, v_j \rangle = 0 \; \forall \; i \neq j$ y $\|v_i\|^2 = 1$, $\forall i = 1, \ldots, k$.

# Observación

- Si $\{v_1, v_2, \ldots, v_k\}$ es un **conjunto ortogonal que no contiene al vector nulo** $\Rightarrow \{v_1, v_2, \ldots, v_k\}$ es l.i.

Definición:
Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial con producto interno. Se dice que $B = \{v_1, v_2, \ldots, v_n\}$ es una **base ortogonal de** $\mathbb{V}$ si es una base de $\mathbb{V}$ y es un conjunto ortogonal.
$(\langle v_i, v_j\rangle = 0 \ \forall i \neq j.)$

Definición:
Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial con producto interno. Se dice que $B = \{v_1, v_2, \ldots, v_n\}$ es una **base ortonormal de** $\mathbb{V}$ si es una base de $\mathbb{V}$ y es un conjunto ortonormal.
O sea, $\langle v_i, v_j\rangle = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{cases}$ .

## Descomposición con respecto a una base ortonormal

Sea $B = \{v_1, v_2, \dots v_n\}$ es una base ortogonal de $\mathbb{V}$, *si* $u \in \mathbb{V}$ $u = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$, aplicando el teorema de Pitágoras sucesivamente en los $n$ términos y usando el hecho de que la base es un conjunto ortonormal obtenemos:

$\|u\|^2 = |\lambda_1|^2 + |\lambda_2|^2 + \cdots + |\lambda_n|^2$
Pero el resultado es más "sabroso"todavía, porque los escalares, o sea las coordenadas de cualquier vector con respecto a una base ortonormal quedan explicitamente determinados por el vector $u$ y su p.i. con respecto a los vectores de la base ortonormal.

$$\boxed{\lambda_i = \langle u, v_i \rangle}$$

Si $B = \{v_1, v_2, \dots, v_n\}$ es una base ortonormal : $u = \langle u, v_1 \rangle v_1 + \langle u, v_2 \rangle v_2 + \cdots + \langle u, v_n \rangle v_n$

$\|u\|^2 = |\langle u, v_1 \rangle|^2 + |\langle u, v_2 \rangle|^2 + \cdots + |\langle u, v_n \rangle|^2$

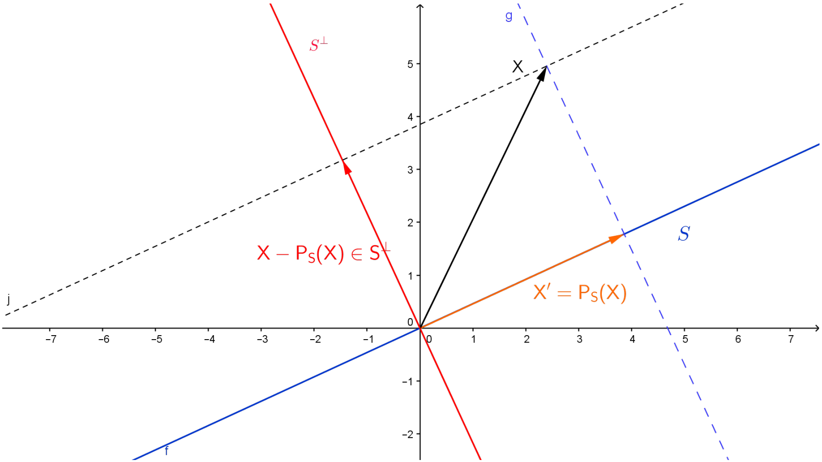
# Complemento ortogonal

Definición:
Si $A \subseteq \mathbb{V}, A \neq \emptyset$, se llama **complemento ortogonal de A** al conjunto $A^{\perp} = \{w \in \mathbb{V} : \langle w, v \rangle = 0 \; \forall v \in A\}$, el conjunto formado por todos los vectores de $\mathbb{V}$ que son ortogonales a cada elemento de $A$.

- $A^\perp$ es un subespacio de $\mathbb{V}$, $\forall A \subseteq \mathbb{V}$.
  Es inmediato, pues:
  a $0_\mathbb{V} \in A^\perp$ pues $\langle 0_\mathbb{V}, v\rangle \;\; v \in \mathbb{V}$, en particular entonces $\langle 0_\mathbb{V}, v\rangle \;\; v \in A$
  b si $w_1$ y $w_2$ son elementos de $A^\perp$
  $\langle w_1 + w_2, v\rangle = \langle w_1, v\rangle + \langle w_2, v\rangle = 0 + 0 = 0$
  c Tarea para el hogar
- $\{0_\mathbb{V}\}^\perp = \mathbb{V}$. Sabemos que $\forall\ \langle ., .\rangle$ se cumple que $\langle 0_\mathbb{V}, v\rangle = 0\ \forall v \in \mathbb{V} \Rightarrow v \in \{0_\mathbb{V}\}^\perp\ \forall v \in \mathbb{V} \Rightarrow \mathbb{V} = \{0_\mathbb{V}\}^\perp$.
- $\mathbb{V}^\perp = \{0_\mathbb{V}\}$, pues si $v \in \mathbb{V}^\perp \Rightarrow \ \langle v, u\rangle = 0\ \forall u \in \mathbb{V}$ en particular $\langle v, v\rangle = 0 \Rightarrow v = 0_\mathbb{V} \Rightarrow \mathbb{V}^\perp = \{0_\mathbb{V}\}$.
- Si $S$ y $T$ son **subconjuntos** de $\mathbb{V}$, $S \subseteq T \Rightarrow T^\perp \subseteq S^\perp$. Pues si $v \in T^\perp \Rightarrow \langle v, v_t\rangle = 0\ \forall\ v_t \in T$ como $S \subseteq T$ en particular, $\langle v, v_S\rangle = 0\ \forall\ v_S \in S\ \Rightarrow v \in S^\perp \Rightarrow T^\perp \subseteq S^\perp$
- Si $S \subseteq \mathbb{V}$ es un subespacio $\Rightarrow S \cap S^\perp = \{0_\mathbb{V}\}$. Pues si $v \in S \cap S^\perp, \langle v, v\rangle = 0 \Rightarrow v = 0_\mathbb{V} \Rightarrow S \cap S^\perp = \{0_\mathbb{V}\}$

# Complemento ortogonal de un subespacio de dimensión finita.

# Proyección Ortogonal



# Proyección Ortogonal

Sea $S \subseteq \mathbb{V}$ un subespacio de $\mathbb{V}$ y $v \in \mathbb{V}$, se dice que $v'$ es la **proyección ortogonal** de $v$ sobre $S$ si:

1. $v' \in S$.
2. $v - v' \in S^{\perp}$.

Notación: Se escribe $P_S(v) = v'$

# Dos observaciones importantes

a $\boxed{\text{Si existe, } P_S(v) \text{ es única}}$

b Para todo $v \in \mathbb{V}$, $P_S(v)$ es el punto de $S$ más cercano a $v$:
$\boxed{\mathrm{d}(v, P_S(v)) \leq \mathrm{d}(v, v_S)\ \forall\ v_S \in S.}$

# Más observaciones

Si existe, $P_S(v)\ \forall\ v \in \mathbb{V}$ :

- $v - P_S(v) = P_{S^\perp}(v),\ \forall v \in \mathbb{V}$.
  Basta con probar que cumple las condiciones de la definición de proyección ortogonal:
  1. $(v - P_S(v)) \in S^\perp$, por la definición de $P_S(v)$
  2. $v - ((v - P_S(v)) = P_S(v) \in S$, por lo tanto $v - ((v - P_S(v)) \in (S^\perp)^\perp$ pues es ortogonal a todos los elementos de $S^\perp$

  Entonces $v - P_S(v)$ cumple con la definición de proyección ortogonal sobre $S^\perp$. Luego $P_{S^\perp}(v) = v - P_S(v)$
- $v = P_S(v) + P_{S^\perp}(v)\ \forall v \in \mathbb{V}$.
- $\mathbb{V} = S \oplus S^\perp$. Además, si $\mathbb{V}$ es de dimensión finita $(S^\perp)^\perp = S$
- Si $v \in \mathbb{V}$ y $v = v_S + v_{S^\perp}$ con $v_S \in S$ y $v_{S^\perp} \in S^\perp$.
  $v_S = P_S(v)$ y $v_{S^\perp} = P_{S^\perp}(v)$ (Tarea: ver que $v_S$ y $v_{S^\perp}$ cumplen con la definición de proy. ortogonal sobre $S$ y $S^\perp$ respectivamente.)

# Otras Propiedades Importantes

a $P_S(v) = v \iff v \in S$.
Si $P_S(v) = v \Rightarrow v \in S$. Si $v \in S$, cumple la definición de proyección ortogonal:$v \in S$ y $v - v = 0_{\mathbb{V}} \in S^{\perp} \Rightarrow P_S(v) = v$

b $P_S(v) = 0_{\mathbb{V}} \iff v \in S^{\perp}$. (Tarea)

c $P_S(\lambda v + w) = \lambda P_S(v) + P_S(w)$, $\forall\ v, w \in \mathbb{V} y \forall\ \lambda \in K$.
1) $\lambda P_S(v) + P_S(w) \in S.\checkmark$
2)$\lambda v + w - (\lambda P_S(v) + P_S(w)) =$
$$= \lambda(v - P_S(v)) + (w - P_S(w)) \in S^{\perp}.\checkmark$$
Por lo tanto, $P_S(\lambda v + w) = \lambda P_S(v) + P_S(w)$

d Lo anterior demuestra que, $P_S : \mathbb{V} \to \mathbb{V}$ es t.l. y además por a. y b. $\boxed{\mathrm{Im}(P_S) = S,\ \mathrm{Nu}(P_S) = S^{\perp}.}$

e $P_S(P_S(v)) = P_S(v)\ \forall\ v \in \mathbb{V}$

# Fórmula de la proyección ortogonal.

Sea $S$ un subespacio en $\mathbb{V}$, $v \in \mathbb{V}$, y $B_S = \{v_1, v_2, \ldots, v_k\}$ una base ortogonal de $S$:

$$P_S(v) = \frac{\langle v, v_1\rangle}{\|v_1\|^2} v_1 + \frac{\langle v, v_2\rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle v, v_k\rangle}{\|v_k\|^2} v_k$$

$$P_S(v) \in S \Rightarrow P_S(v) = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k, \lambda_i \in \mathbb{K},\ i = 1, \ldots, k.(1)$$

$$v - P_S(v) \in S^{\perp} \Longleftrightarrow v - P_S(v) \perp v_i \ \forall\ i = 1, \ldots, k.(2)$$

Reemplazando (1) en (2), obtenemos que:

$$\langle v - (\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k), v_i \rangle = 0 \ \forall \ i = 1, \ldots, k.$$

$$\langle v, v_i \rangle - \langle \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k, v_i \rangle = 0 \ \forall \ i = 1, \ldots, k.$$

$$\langle v, v_i \rangle = \langle \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k, v_i \rangle \ \forall \ i = 1, \ldots, k.$$

Como $\{v_1, v_2, \ldots, v_k\}$ es un conjunto ortogonal:
$\langle v, v_i \rangle = \lambda_i \langle v_i, v_i \rangle \ \forall \ i = 1, \ldots, k.$
Además $\langle v_i, v_i \rangle = \|v_i\|^2 \neq 0$ pues $v_i \neq 0_{\mathbb{V}} \ \forall \ i = 1, \ldots, k$ porque forman parte de una base .

$$\lambda_i = \frac{\langle v, v_i \rangle}{\|v_i\|^2} \text{ y reemplazando en (1):}$$

$$P_S(v) = \frac{\langle v, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle v, v_2 \rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle v, v_k \rangle}{\|v_k\|^2} v_k \checkmark$$

1. Ya tenemos una fórmula para encontrar la proyección ortogonal si tenemos una base ortogonal del subespacio. Queda la pregunta ¿ siempre podré encontar una base ortogonal ? El procedimiento de Gram Scmidt, que vamos a presentar en la próxima reunión nos asegura que 'para todo espacio de dimensión finita puedo encontrar una base ortogonal. Entonces, cuando el subespacio sobre el que proyectamos es de dimensión finita, siempre vas a poder calcular la proyección ortogonal.
2. Para un subespacio de dimensión 2 ya conocemos una manera de construir una base ortogonal, a través de la fórmula de descomposición ortogonal que probamos al inicio de este episodio.

# Ejemplos.

- Sea $S \in \mathbb{R}^3$ con el P.I. canónico,
  $S = \{(x_1\ x_2\ x_3)^T \in \mathbb{R}^3 / x_1 - x_2 + 2x_3 = 0\}$, se pide:

  a Hallar $S^\perp$.
  b Hallar una base ortogonal de $S$.
  c Encuentre $P_S((1\ 2\ 3)^T)$
  d Encuentre $P_S((x_1 x_2 x_3)^T)\ \forall\ (x_1 x_2 x_3)^T \in \mathbb{R}^3$.

  Resolución:

  a Para encontrar $S^\perp$, miremos la condición de $S$:
  $x_1 - x_2 + x_3 = 0 \Leftrightarrow \left\langle (1 - 1\ 2)^T, (x_1\ x_2\ x_3)^T \right\rangle = 0$. Entonces $S = (\text{gen}\{(1 - 1\ 2)^T\})^\perp \Rightarrow S^\perp = \text{gen}\{(1 - 1\ 2)^T\}$ Por lo tanto proponemos $B_{S^\perp} = \{(1 - 1\ 2)^T\}$
  b Para hallar una base **ortogonal** de $S$, busquemos los generadores de $S$.
  $(x_1\ x_2\ x_3)^T \in S \Leftrightarrow x_1 - x_2 + 2x_3 = 0 \Leftrightarrow x_1 = x_2 - 2x_3$.
  $(x_1\ x_2\ x_3)^T = (x_2 - 2x_3\ x_2\ x_3)^T = x_2(1\ 1\ 0)^T + x_3(-2\ 0\ 1)^T$

Entonces: $S = \text{gen}\{(1\ 1\ 0)^T, (-2\ 0\ 1)^T\}$
Estos vectores no son ortogonales, pero justamente en el comienzo de nuestra episodio hemos demostrado que si $u$ y $v$ eran dos vectores cualesquiera ($v \neq 0_{\mathbb{V}}$), tomando $c = \frac{\langle u,v\rangle}{\|v\|^2}$, obteníamos $w = u - cv = u - \frac{\langle u,v\rangle}{\|v\|^2} v$, de manera tal que $w \perp v$ y $cv$, ahora que hemos definido proyección ortogonal, sabemos que es la proyección ortogonal de $u$ sobre el subespacio $\text{gen}\{v\}$ .
Entonces, llamando $u = (1\ 1\ 0)^T$ y $v = (-2\ 0\ 1)^T$, obtenemos:

$$w = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{\left\langle (1\ 1\ 0)^T, (-2\ 0\ 1)^T \right\rangle}{\|[(-2\ 0\ 1)^T\|^2} \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{(-2)}{5} \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

$$w = \begin{pmatrix} \frac{1}{5} \\ 1 \\ \frac{2}{5} \end{pmatrix} \Rightarrow \text{una base ortogonal de } S \text{ es } B_s = \left\{ \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix} \right\}$$

c Nos piden encontrar $P_S((1\ 2\ 3)^T)$.Como $\dim(S^{\perp})=1$ y $P_S(v) = v - P_{S^{\perp}(v)}$, calculamos primero $P_{S^{\perp}}((1\ 2\ 3)^T)$ con la fórmula de proyección ortogonal:

$$P_{S^{\perp}}((1\ 2\ 3)^T) = \frac{\langle (1\ 2\ 3)^T, (1\ \ -1\ 2)^T \rangle}{\|(\ 1 - 1\ \ 2)^T\|^2} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

$$P_{S^{\perp}}((1\ 2\ 3)^T) = \frac{5}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} =$$

$$P_S((1\ 2\ 3)^T) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \frac{5}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{6} \\ \\ \frac{17}{6} \\ \\ \frac{4}{3} \end{pmatrix}.$$

d Para encontrar $P_S\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right)$ podemos repetir la estrategia ya utilizada. Calculamos en primer lugar la proyección sobre $S^{\perp}$:

$$P_{S^\perp}((x_1\ x_2\ x_3)^T) = \frac{\langle (x_1\ x_2\ x_3)^T, (1\ -1\ 2)^T \rangle}{\|(\ 1 - 1\ \ 2)^T\|^2} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

$$P_{S^\perp}((x_1\ x_2\ x_3)^T) = \frac{(x_1 - x_2 + 2x_3)}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

$$P_S((x_1\ x_2\ x_3)^T) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} - \frac{(x_1 - x_2 + 2x_3)}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

Obtenemos:

$$P_S((x_1\ x_2\ x_3)^T) = \begin{pmatrix} \frac{5x_1+x_2-2x_3}{6} \\ \\ \frac{x_1+5x_2+2x_3}{6} \\ \\ \frac{-2x_1+2x_2+2x_3}{6} \end{pmatrix} = \begin{pmatrix} \frac{5}{6} & \frac{1}{6} & \frac{-1}{3} \\ \\ \frac{1}{6} & \frac{5}{6} & \frac{1}{3} \\ \\ \frac{-1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} x_1 \\ \\ x_2 \\ \\ x_3 \end{pmatrix}$$

Donde

$$[P_S]_E^E = \begin{pmatrix} \frac{5}{6} & \frac{1}{6} & \frac{-1}{3} \\ \\ \frac{1}{6} & \frac{5}{6} & \frac{1}{3} \\ \\ \frac{-1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

# MATRIZ DE PROYECCIÓN ORTOGONAL EN $\mathbb{C}^n$

Supongamos que buscamos la proyección ortogonal sobre un subespacio $S$ en $\mathbb{C}^n$ o $\mathbb{R}^n$ con respecto al producto interno canónico. Podemos trabajar directamente en $\mathbb{C}^n$ y mirar el caso de $\mathbb{R}^n$ como un caso particular.
Recordemos que el producto interno canónico en $\mathbb{C}^n$ está definido por : $\langle x, y \rangle = \overline{y}^T x$.
En $\mathbb{R}^n$: $\langle x, y \rangle = y^T x$.
Si $B_S = \{v_1 \ldots, , v_k\}$ es una base ortogonal de $S \subset \mathbb{C}^n$, aplicando la fórmula tenemos:

$$P_S(v) = \frac{\langle v, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle v, v_2 \rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle v, v_k \rangle}{\|v_k\|^2} v_k$$

Esto es una combinación lineal de vectores columna, puedo escribirlo como un producto matricial:

$$P_S(v) = \left[ \frac{v_1}{||v_1||} \middle| \frac{v_2}{||v_2||} \middle| \dots \middle| \frac{v_k}{||v_k||} \right] \begin{pmatrix} \frac{\overline{v_1}^T}{||v_1||} v \\ \\ \frac{\overline{v_2}^T}{||v_2||} v \\ \\ \vdots \\ \frac{\overline{v_k}^T}{||v_k||} v \end{pmatrix}$$

$$P_S(v) = \left[ \frac{v_1}{||v_1||} \middle| \frac{v_2}{||v_2||} \middle| \dots \middle| \frac{v_k}{||v_k||} \right] \begin{pmatrix} \frac{\overline{v_1}^T}{||v_1||} \\ \\ \frac{\overline{v_2}^T}{||v_2||} \\ \\ \vdots \\ \frac{\overline{v_k}^T}{||v_k||} \end{pmatrix} v$$

Entonces, encontramos una matriz $P_S = Q\overline{Q}^T$ que cumple $P_S v = proy_S(v)$. Donde Q es una matriz que tiene como columnas una BON de $S$. Esta matriz $P_S$ es la matriz de la proyección ortogonal sobre $S$ con respecto a la base canónica, pues cumple $P_S[x]^E = [proy_S(x)]^E$

Si buscáramos, dado $S$ un subespacio en $\mathbb{C}^n$, una matriz que cumpla $Px = proy_S(x)$ esa matriz debe cumplir:

- $Px = proy_S(x) \in S \; \forall x \in \mathbb{C}^n \Rightarrow \boxed{\mathrm{Col(P)} \subset S}$ y además se cumple que si $v \in S \Rightarrow Pv = v \Rightarrow v \in \mathrm{Col}(P) \Rightarrow \boxed{\mathrm{S} \subset \mathrm{Col}(P)}$.

  Por lo tanto, $\boxed{\mathrm{Col}(P) = S}$ (a)

- Además se debe cumplir que $x - Px \in S^{\perp}$.
  Luego, $\forall x, y \in \mathbb{C}^n$ se cumple: $x - Px \perp Py$:
  $\overline{(Py)}^T(x - Px) = 0 \Leftrightarrow \overline{y}^T\overline{P}^T x - \overline{y}^T\overline{P}^T Px = 0 \; \forall x, y \in \mathbb{C}^n$.
  $\overline{y}^T(\overline{P}^T - \overline{P}^T P)x = 0 \Rightarrow \overline{P}^T - \overline{P}^T P = 0_{\mathbb{C}^{n\times n}}$

$\overline{P}^T = \overline{P}^T P$ (1). si conjugo y transpongo m. a m:

$\overline{\overline{P}^T}^T = \overline{\overline{P}^T P}^T$

$P = \overline{P}^T P$ (2)

Pero entonces, por (1) y (2):

$$\boxed{P = \overline{P}^T.}$$

- Además como
  $Px = v \in S\ \forall x \in \mathbb{C}^n \Rightarrow P^2 x = Pv = v \Rightarrow \boxed{P^2 = P}$
  Entonces, toda matriz que cumpla $Px = proy_S(x)$ debe cumplir:

$$\boxed{P = \overline{P}^T.}$$

$$\boxed{P^2 = P}$$

Por esto se define:

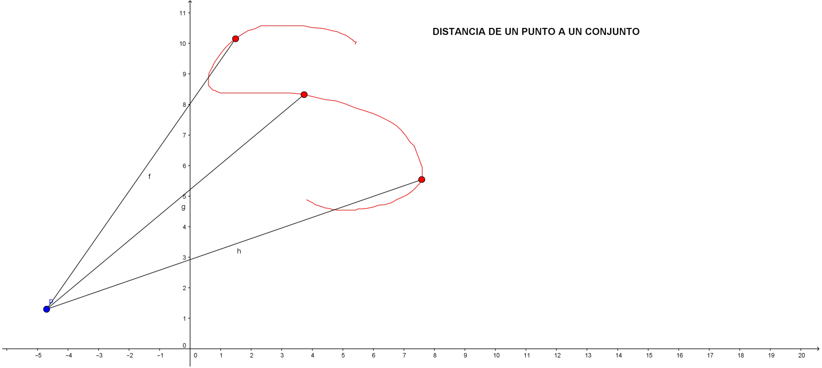
Definición: Se dice que una matriz $P \in \mathbb{C}^{n\times n}$ o $P \in \mathbb{R}^{n\times n}$es una matriz de proyección ortogonal, si cumple:

- $P = \overline{P}^T$ (Osea, P hermítica en $\mathbb{C}^{n\times n}$ o simétrica en $\mathbb{R}^{n\times n}$).
- $P^2 = P$ (P idempotente).

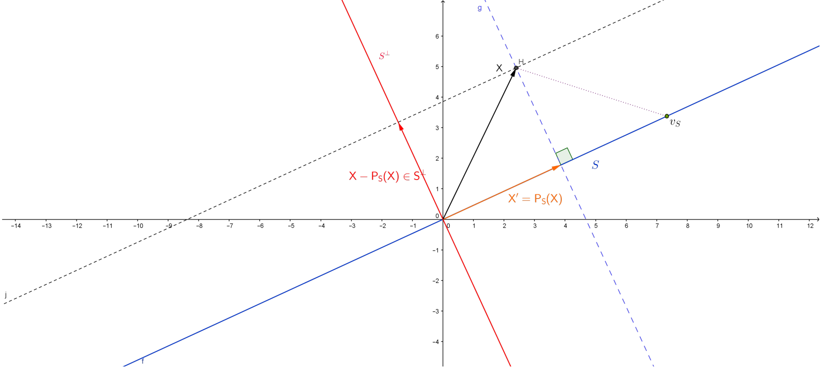
Observaciones:

a Si $P$ es una matriz de proyección ortogonal, proyecta sobre el subespacio $S = \text{Col}(P)$.

b La matriz de proyección ortogonal $P_S$ es única pero la matriz $Q$ no, pues no hay una única base ortogonal de $S$.




Definición: Si $P \in \mathbb{V}$ y $A \subseteq \mathbb{V}$, $A \neq \emptyset$ la distancia de $P$ al conjunto $A$ es
$\text{dist}(P, A) = \text{mín}\{dist(P, v_A), v_A \in A\}$.




- Si $X \in \mathbb{V}$ y $S \subseteq \mathbb{V}$ es un subespacio de $\mathbb{V}$, entonces por lo ya estudiado en proyección ortogonal, $\text{dist}(X, S) = \text{mín}\{dist(X, v_S), v_S \in S\} = dist(X, P_S(X)) = \|X - P_S(X)\| = \|P_{S^{\perp}}(X)\|$.

Empecemos por plantear el problema en $\mathbb{K}^n$ y después generalizamos. Recordemos que en general en $\mathbb{K}^n$, el P.I. canónico es $\langle x, y\rangle = y^*x = \bar{y}^T x$
Sabemos que si $\boxed{A \in \mathbb{K}^{m\times n}}$ y $b \in \mathbb{K}^m$ el sistema $AX = b$ es compatible sí y sólo sí
$b \in \mathrm{Col}(A)$.
Entonces, si $b \notin \mathrm{Col}(A)$, podemos intentar encontrar $\hat{x} \in \mathbb{K}^n$ tal que $\mathrm{dist}(A\hat{x}, b)$ sea "la menor posible".
Por lo visto entonces, la menor distancia la alcanzaremos cuando $A\hat{x} = P_{Col(A)}(b)$
Vamos a buscar un vector $\hat{x}$ tal que $A\hat{x} = P_{Col(A)}(b)$, entonces $A\hat{x}$ tienen que cumplir las dos condiciones de proyección ortogonal:

- $A\hat{x} \ \in Col(A)$. Esto se cumple $\forall \hat{x} \in \mathbb{K}^n$.
- $b - A\hat{x} \ \in Col(A)^{\perp}$

Analicemos la segunda condición, que será la que determine a $\hat{x}$:

$b - A\hat{x} \in Col(A)^{\perp} \iff \langle b - A\hat{x}, Col_k(A)\rangle = 0 \; \forall k = 1, \ldots, n.$
Entonces obtenemos en principio $n$ ecuaciones:

$$\begin{cases} \overline{Col_1(A)}^T (b - A\hat{x}) & = 0 \\ \overline{Col_2(A)}^T (b - A\hat{x}) & = 0 \\ \vdots & = \vdots \\ \overline{Col_n(A)}^T (b - A\hat{x}) & = 0 \end{cases}$$

Estas $n$ ecuaciones se pueden expresar como un producto matricial:

$$\overline{A}^{\,T}(b - A\hat{x}) = 0_{\mathbb{K}^n} \iff \boxed{\overline{A}^{T} A\hat{x} = \overline{A}^{\,T} b}$$ **ECUACIONES NORMALES.**

Si estamos trabajando en $\mathbb{R}^n$ los signos de conjugación no son necesarios.
Formalicemos:

Definición : Dada $A \in \mathbb{K}^{m \times n}$ y $b \in \mathbb{K}^m$, se dice que $\hat{x}$ es la **solución por mínimos cuadrados** de $Ax = b$ si $A\hat{x}$ **realiza** la mínima distancia a $b$, o sea
$\text{dist}(A\hat{x}, b) \leq \text{dist}(Ax, b) \Longleftrightarrow A\hat{x} = P_{Col(A)}(b)$.

Se llama error cuadrático medio a la diferencia:
ECM$=\|b - A\hat{x}\|^2$

Observaciones:

- El problema de mínimos cuadrados siempre tiene solución.
- $Nul(\bar{A}^T A) = Nul(A)$
- $rg(\bar{A}^T A) = rg(A)$
- Por lo anterior el problema de mínimos cuadrados tiene solución única $\Longleftrightarrow rg(\bar{A}^T A) = rg(A) = n$, se dice que $A$ tiene rango columna máximo.
- Si $A$ es de rango columna máximo, $\bar{A}^T A\hat{x} = \bar{A}^T b \Longleftrightarrow \hat{x} = (\bar{A}^T A)^{-1}\bar{A}^T b$

Definición: Si $A \in \mathbb{K}^{m\times n}$, $rg(A) = n$, se llama seudoinversa de $A$ a la matriz
$A^{\#} = (\bar{A}^T A)^{-1}\bar{A}^T$

Propiedades:

- $A^{\#}A = (\bar{A}^TA)^{-1}\bar{A}^TA = (\bar{A}^TA)^{-1}(\bar{A}^TA) = I_n$
- $AA^{\#}$ es la matriz de la proyección ortogonal sobre $Col(A) \subset \mathbb{K}^m$.

Ahora bien, cuando trabajamos buscando la pre-imagen de una transformación lineal ya vimos que las situaciones con las que nos encontramos, son las mismas con las que nos encontramos al resolver un sistema lineal. Sabemos que la ecuación del tipo $T(v) = w$ tiene solución $\iff w \in Im(T)$, entonces podemos hacer el mismo razonamiento que hicimos antes para el sistema lineal.

Definición: Se dice que $\bar{v}$ es solución por mínimos cuadrados de la ecuación $T(v) = w$ si $\bar{v}$ **realiza** la mínima distancia en $Im(T)$, o sea $\text{dist}(T(\hat{v}), w) \leq \text{dist}(T(v), w) \iff T(\hat{v}) = P_{Im(T)}(w)$

## Curva que mejor ajusta-Regresión lineal

Tenemos un conjunto de datos $(x_1, y_1),\ (x_2, y_2)\dots(x_n, y_n)$. Si suponemos una relacion entre las variables y buscamos una curva que ajuste estos datos:
Podemos intentar buscar una función polinómica que pase por estos puntos. Si suponemos una relación lineal por ejemplo, estaremos buscando una curva de la forma $y = mx + b$.
Tendría que cumplirse:

$$\begin{aligned} mx_1 + b &= y_1 \\ mx_2 + b &= y_2 \\ \vdots \quad + \vdots \quad &= \vdots \\ mx_n + b &= y_n \end{aligned}$$

Este sistema de ecuaciones lineales, puede plantearse en forma matricial:

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Si el sistema es incompatible, siempre podemos resolver por mínimos cuadrados. Entonces, tendremos las ecuaciones normales correspondientes.

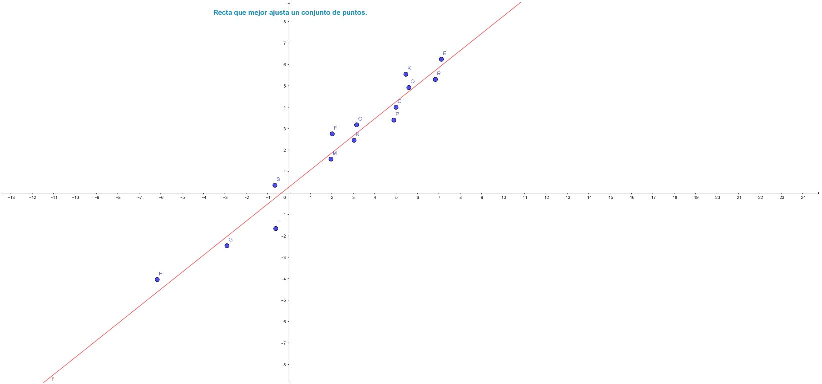
Si $A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}$ planteamos las ecuaciones normales y queda:

$$A^T A \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & \vdots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} =$$

$$\begin{bmatrix} \sum_1^n x_i^2 & \sum_1^n x_i \\ \sum_1^n x_i & n \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = A^T Y$$

Las ecuaciones normales quedan:

$$\begin{bmatrix} \sum_1^n x_i^2 & \sum_1^n x_i \\ \sum_1^n x_i & n \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \sum_1^n x_i y_i \\ \sum_1^n y_i \end{bmatrix}$$


Recta que mejor ajusta un conjunto de puntos.
E
K
R
Q
C
P
O
F
N
M
S
T
G
H
f

Si en lugar de suponer una relación lineal, suponemos una relación cuadrática, tendremos que encontrar tres coeficientes para encontrar la función cuadrática:
$y = ax^2 + bx + c$
Imponemos: $y_i = ax_i^2 + bx_i + c$, nuestro objetivo es encontrar los coeficientes $a, b, c$ que mejor *ajustan* a estos datos. Buscamos $a, b, c \in \mathbb{R}$ tal que:

$$\begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

, si este sistema no tiene solución, sabemos que la solución por mínimos cuadrados siempre existe.

Para encontrar, los coeficientes $a, b, c \in \mathbb{R}$ que mejor aproximan, buscamos $a, b, c \in \mathbb{R}$, tal que hagan mínima la distancia :

$$dist\left(\begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}\right)$$

Eso equivale a resolver el problema por mínimos cuadrados y por lo tanto deberemos plantear las ecuaciones normales. Lo dejamos aquí para que cada **une** continúe el razonamiento