Métodos Computacionales

Clase VII: Ortogonalidad y Mínimos Cuadrados

Producto Interno

Producto Interno - Definición

- Si $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{R}^n$ se los puede pensar como matrices de n x 1. El vector transpuesto $\mathbf{u}^\mathsf{T}$ es por lo tanto una matriz 1 x n
- Si se multiplica $\mathbf{u}^\mathsf{T}\mathbf{v}$ se obtiene un escalar y se lo define como el **producto interno** entre $\mathbf{u}$ y $\mathbf{v}$, comúnmente escrito como $\mathbf{u} \cdot \mathbf{v}$

$$\mathbf{u} \cdot \mathbf{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

Producto Interno

- Calcular $\mathbf{u} \cdot \mathbf{v}$, $\mathbf{v} \cdot \mathbf{u}$, $\mathbf{u} \cdot \mathbf{u}$, $\mathbf{v} \cdot \mathbf{v}$ siendo:

$$\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} \mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$$

Producto Interno - Propiedades

- Sean $\mathbf{u}, \mathbf{v}$ y $\mathbf{w}$ vectores en $\mathbb{R}^n$ y c un escalar:

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
3. $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$
4. $\mathbf{u} \cdot \mathbf{u} \geq 0$
5. $\mathbf{u} \cdot \mathbf{u} = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$

Producto Interno - Propiedades

- Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores en $\mathbb{R}^n$ y c un escalar:

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
3. $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$
4. $\mathbf{u} \cdot \mathbf{u} \geq 0$
5. $\mathbf{u} \cdot \mathbf{u} = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$

- Con las propiedades 1 y 2 se puede deducir que:

$$(c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p) \cdot \mathbf{w} = c_1(\mathbf{u}_1 \cdot \mathbf{w}) + \cdots + c_p(\mathbf{u}_p \cdot \mathbf{w})$$

Espacio Fila de una Matriz

Espacio Columna (Recap)

- Sea A una matriz de $\mathbb{R}^{mxn}$: $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{c_1} & \mathbf{c_2} & \dots & \mathbf{c_n} \end{bmatrix}$

$$\text{col } A = \text{Gen}\{\mathbf{c_1}, \mathbf{c_2}, \dots, \mathbf{c_n}\}$$

$$\text{col } A = \text{Gen}\left\{ \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix} \right\}$$

col A es el **espacio generado** por los n **vectores columna** de A y pertenece a $\mathbb{R}^m$

Espacio Fila

- Sea A una matriz de $\mathbb{R}^{mxn}$: $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \dots & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1^T \\ \mathbf{f}_2^T \\ \dots \\ \mathbf{f}_m^T \end{bmatrix}$

$$\text{fil}\, A = \text{Gen}\left\{\mathbf{f}_1^T, \mathbf{f}_2^T, \dots, \mathbf{f}_m^T, \right\}$$

fil A es el **espacio generado** por los m **vectores fila** de A y pertenece a $\mathbb{R}^n$

$$\text{fil}\, A = \text{Gen}\left\{(a_{11}, a_{12}, \dots, a_{1n})^T, (a_{21}, a_{22}, \dots, a_{2n})^T, \dots, (a_{m1}, a_{m2}, \dots, a_{mn})^T\right\}$$

Producto Matriz-Vector

$$A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \dots & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = x_1\mathbf{c_1} + x_2\mathbf{c_2} + \cdots + x_n\mathbf{c_n}$$

$$A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \dots & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} \mathbf{f_1^T x} \\ \mathbf{f_2^T x} \\ \dots \\ \mathbf{f_m^T x} \end{pmatrix}$$

El vector resultante de $A\mathbf{x}$ en $\mathbb{R}^m$ se puede interpretar como:

1. Una combinación lineal de los vectores columnas de A con las componentes de $\mathbf{x}$ como pesos.
2. Un vector cuyas coordenadas son el resultado del producto interno de los vectores filas de A y el vector **x.**

Ortogonalidad

Norma (longitud) de un vector

- Supongamos que $\mathbf{v}$ es un vector en $\mathbb{R}^2$:

$$\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$$

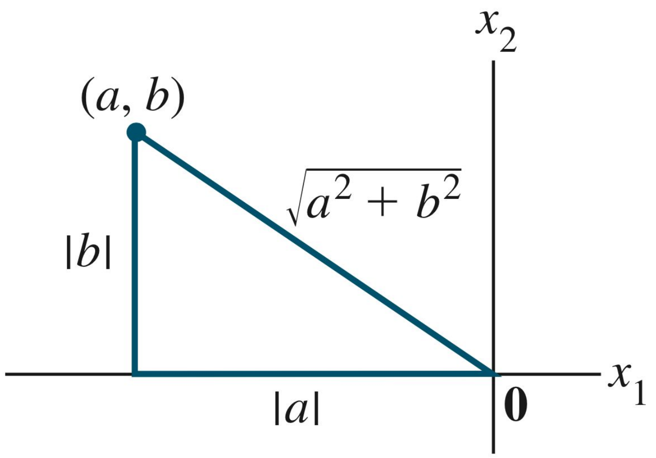
$$||\mathbf{v}|| = \sqrt{a^2 + b^2}$$

Norma (longitud) de un vector

- Supongamos que **v** es un vector en $\mathbb{R}^2$:

$$\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$$

$$\|\mathbf{v}\| = \sqrt{a^2 + b^2}$$



Norma (longitud) de un vector

- En general para $\mathbb{R}^n$:

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

$$\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$$

Norma (longitud) de un vector

- En general para $\mathbb{R}^n$:

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

$$||\mathbf{v}||^2 = \mathbf{v} \cdot \mathbf{v}$$

- Notar que:

$$||c\mathbf{v}||^2 = (c\mathbf{v}) \cdot (c\mathbf{v}) = c^2 \mathbf{v} \cdot \mathbf{v} = c^2 ||\mathbf{v}||^2$$

$$||c\mathbf{v}|| = |c|\,||\mathbf{v}||$$

Normalización - Vector Unitario

- Un vector cuya norma es 1 es llamado **vector unitario**:

$$\mathbf{u} \text{ tal que } ||\mathbf{u}|| = 1 \rightarrow \quad \text{vector unitario}$$

- Para cualquier vector $\mathbf{v}$ distinto de cero podemos encontrar un vector unitario con la misma dirección que $\mathbf{v}$:

$$\mathbf{u} = \frac{\mathbf{v}}{||\mathbf{v}||}$$

A este proceso de crear el vector unitario $\mathbf{u}$ a partir de $\mathbf{v}$ se lo llama **normalización.**

Normalización

- Encontrar el vector unitario $\mathbf{u}$ en la misma dirección que $\mathbf{v}$:

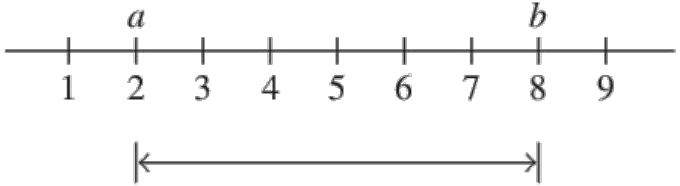
$$\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix}$$

Normalización

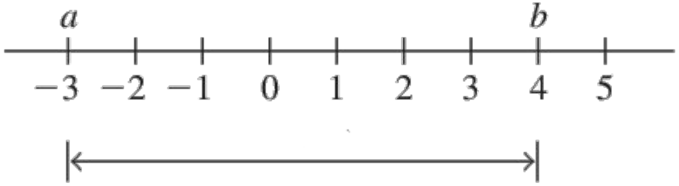
- Sea W el subespacio en $\mathbb{R}^2$ generado por $\mathbf{x} = \begin{bmatrix} 2/3 \\ 1 \end{bmatrix}$

1. Encontrar el vector unitario $\mathbf{z}$ que sea una base de W
2. ¿Hay más de una solución?

Distancia en $\mathbb{R}$



$|2-8| = |-6| = 6$ or $|8-2| = |6| = 6$



$|(-3)-4| = |-7| = 7$ or $|4-(-3)| = |7| = 7$

$$|a - b|$$

Distancia en $\mathbb{R}^n$

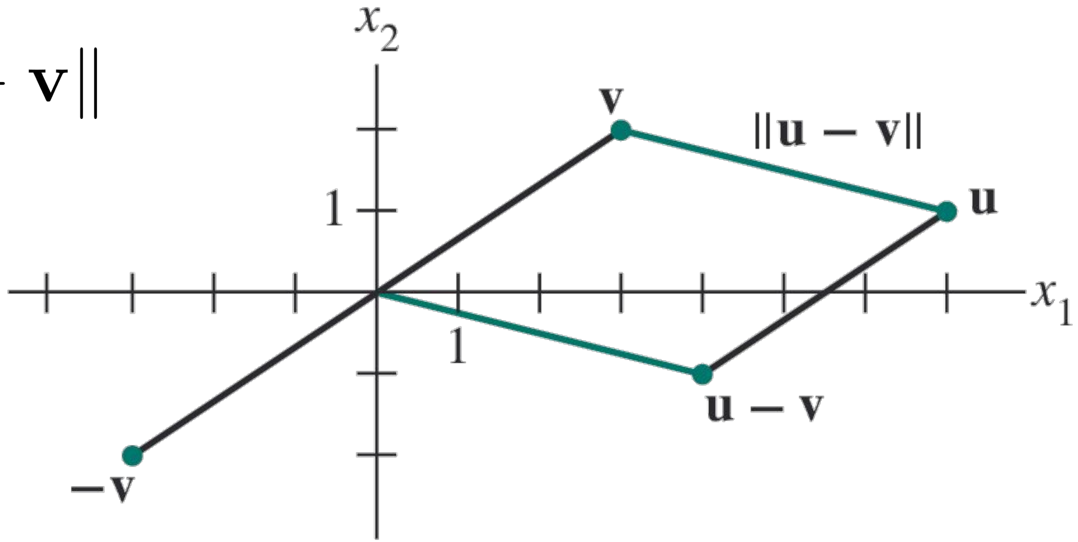
- Dados $\mathbf{u}, \mathbf{v}$ vectores en $\mathbb{R}^n$, la distancia entre $\mathbf{u}$ y $\mathbf{v}$ se define como la norma del vector $\mathbf{u}$ - $\mathbf{v}$:

$$\operatorname{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

Distancia en $\mathbb{R}^n$

- Dados $\mathbf{u}, \mathbf{v}$ vectores en $\mathbb{R}^n$, la distancia entre $\mathbf{u}$ y $\mathbf{v}$ se define como la norma del vector $\mathbf{u}$ - $\mathbf{v}$:

$$\operatorname{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$



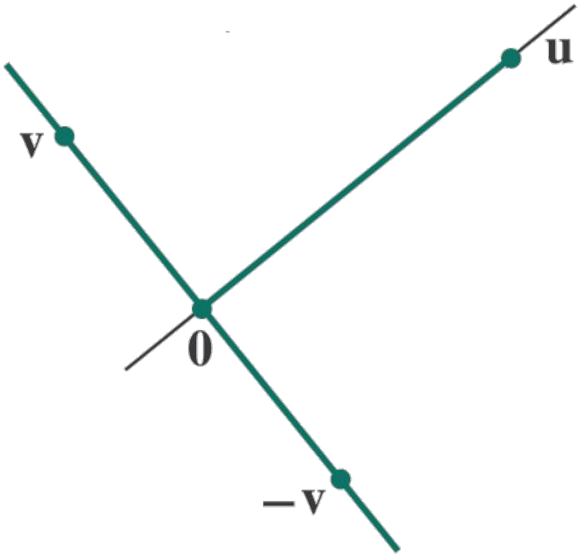
Distancia en $\mathbb{R}^n$

- Encontrar la distancia entre $\mathbf{u}$ y $\mathbf{v}$:

$$\mathbf{u} = \begin{bmatrix} 7 \\ 1 \end{bmatrix} \quad \mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

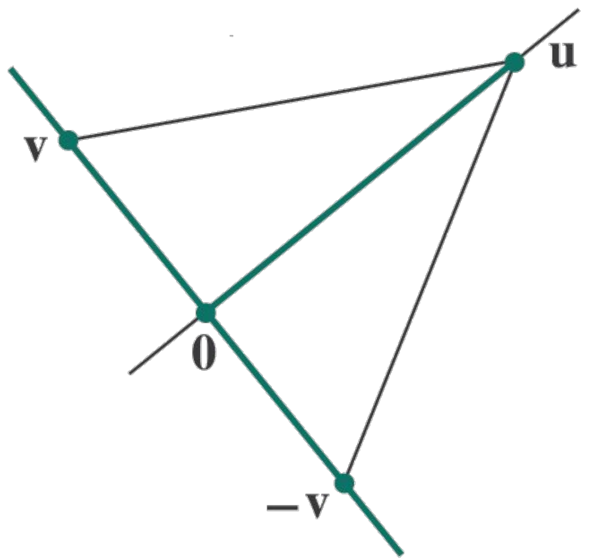
Ortogonalidad

- El concepto de **perpendicularidad** en $\mathbb{R}^2$ y $\mathbb{R}^3$ se puede extender a $\mathbb{R}^n$ y se generaliza bajo el nombre de **ortogonalidad.**



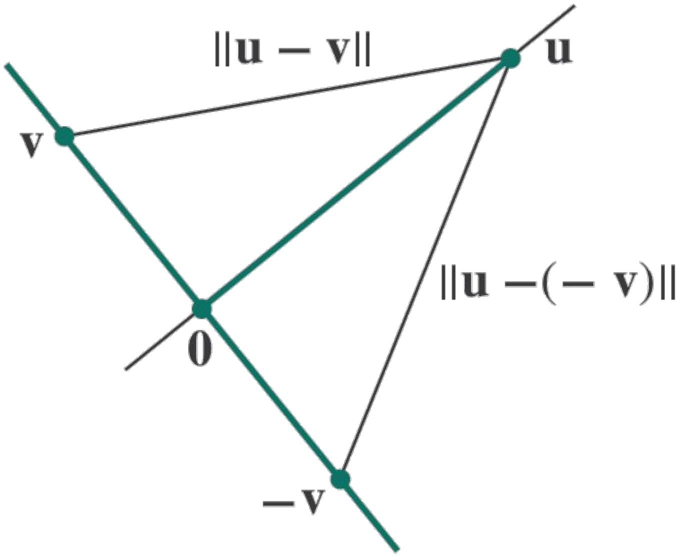
Ortogonalidad

- El concepto de **perpendicularidad** en $\mathbb{R}^2$ y $\mathbb{R}^3$ se puede extender a $\mathbb{R}^n$ y se generaliza bajo el nombre de **ortogonalidad.**
- Dos rectas son geométricamente perpendiculares si y sólo si la distancia de $\mathbf{u}$ a $\mathbf{v}$ es la misma que la distancia de $\mathbf{u}$ a $\mathbf{-v}$.



Ortogonalidad

- El concepto de **perpendicularidad** en $\mathbb{R}^2$ y $\mathbb{R}^3$ se puede extender a $\mathbb{R}^n$ y se generaliza bajo el nombre de **ortogonalidad.**
- Dos rectas son geométricamente perpendiculares si y sólo si la distancia de $\mathbf{u}$ a $\mathbf{v}$ es la misma que la distancia de $\mathbf{u}$ a $-\mathbf{v}$.

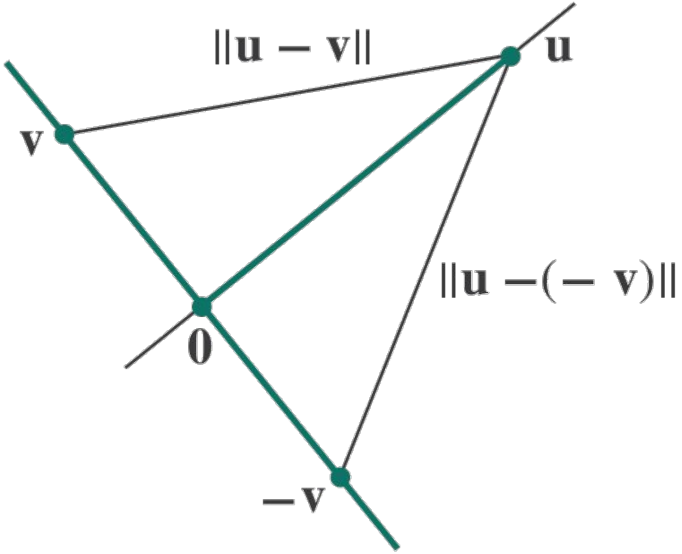


Ortogonalidad

- Demostración gráfica en $\mathbb{R}^2$:

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \text{dist}(\mathbf{u}, -\mathbf{v})$$

$$\|\mathbf{u} - \mathbf{v}\| = \|\mathbf{u} - (-\mathbf{v})\|$$



Ortogonalidad

- Demostración gráfica en $\mathbb{R}^2$

$$
\begin{aligned}
[\operatorname{dist}(\mathbf{u}, -\mathbf{v})]^2 &= \|\mathbf{u} - (-\mathbf{v})\|^2 = \|\mathbf{u} + \mathbf{v}\|^2 \\
&= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) \\
&= \mathbf{u} \cdot (\mathbf{u} + \mathbf{v}) + \mathbf{v} \cdot (\mathbf{u} + \mathbf{v}) \\
&= \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} \\
&= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v}
\end{aligned}
$$

Ortogonalidad

- Demostración gráfica en $\mathbb{R}^2$

$$
\begin{aligned}
[\mathrm{dist}(\mathbf{u}, -\mathbf{v})]^2 &= \|\mathbf{u} - (-\mathbf{v})\|^2 = \|\mathbf{u} + \mathbf{v}\|^2 \\
&= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) \\
&= \mathbf{u} \cdot (\mathbf{u} + \mathbf{v}) + \mathbf{v} \cdot (\mathbf{u} + \mathbf{v}) \\
&= \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} \\
&= \underbrace{\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v}}
\end{aligned}
$$

$$
\begin{aligned}
[\mathrm{dist}(\mathbf{u}, \mathbf{v})]^2 &= \|\mathbf{u} - \mathbf{v}\|^2 \\
&= (\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v}) \\
&= \mathbf{u} \cdot (\mathbf{u} - \mathbf{v}) - \mathbf{v} \cdot (\mathbf{u} - \mathbf{v}) \\
&= \mathbf{u} \cdot \mathbf{u} - \mathbf{u} \cdot \mathbf{v} - \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} \\
&= \underbrace{\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u} \cdot \mathbf{v}}
\end{aligned}
$$

Ortogonalidad

- Demostración gráfica en $\mathbb{R}^2$

$$\begin{aligned}[\operatorname{dist}(\mathbf{u},-\mathbf{v})]^2 &= \|\mathbf{u}-(-\mathbf{v})\|^2 = \|\mathbf{u}+\mathbf{v}\|^2 \\ &= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) \\ &= \mathbf{u}\cdot(\mathbf{u}+\mathbf{v})+\mathbf{v}\cdot(\mathbf{u}+\mathbf{v}) \\ &= \mathbf{u}\cdot\mathbf{u}+\mathbf{u}\cdot\mathbf{v}+\mathbf{v}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v} \\ &= \|\mathbf{u}\|^2+\|\mathbf{v}\|^2+2\mathbf{u}\cdot\mathbf{v}\end{aligned}$$

$$\begin{aligned}[\operatorname{dist}(\mathbf{u},\mathbf{v})]^2 &= \|\mathbf{u}-\mathbf{v}\|^2 \\ &= (\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v}) \\ &= \mathbf{u}\cdot(\mathbf{u}-\mathbf{v})-\mathbf{v}\cdot(\mathbf{u}-\mathbf{v}) \\ &= \mathbf{u}\cdot\mathbf{u}-\mathbf{u}\cdot\mathbf{v}-\mathbf{v}\cdot\mathbf{u}-\mathbf{v}\cdot\mathbf{v} \\ &= \|\mathbf{u}\|^2+\|\mathbf{v}\|^2-2\mathbf{u}\cdot\mathbf{v}\end{aligned}$$

$$\operatorname{dist}(\mathbf{u},-\mathbf{v}) = \operatorname{dist}(\mathbf{u},\mathbf{v}) \leftrightarrow -2\mathbf{u}\cdot\mathbf{v} = 2\mathbf{u}\cdot\mathbf{v}$$

Ortogonalidad

- Demostración gráfica en $\mathbb{R}^2$

$$
\begin{aligned}
[\mathrm{dist}(\mathbf{u}, -\mathbf{v})]^2 &= \|\mathbf{u} - (-\mathbf{v})\|^2 = \|\mathbf{u} + \mathbf{v}\|^2 \\
&= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) \\
&= \mathbf{u} \cdot (\mathbf{u} + \mathbf{v}) + \mathbf{v} \cdot (\mathbf{u} + \mathbf{v}) \\
&= \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} \\
&= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v}
\end{aligned}
$$

$$
\begin{aligned}
[\mathrm{dist}(\mathbf{u}, \mathbf{v})]^2 &= \|\mathbf{u} - \mathbf{v}\|^2 \\
&= (\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v}) \\
&= \mathbf{u} \cdot (\mathbf{u} - \mathbf{v}) - \mathbf{v} \cdot (\mathbf{u} - \mathbf{v}) \\
&= \mathbf{u} \cdot \mathbf{u} - \mathbf{u} \cdot \mathbf{v} - \mathbf{v} \cdot \mathbf{u} - \mathbf{v} \cdot \mathbf{v} \\
&= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u} \cdot \mathbf{v}
\end{aligned}
$$

Dos vectores **u** y **v** en $\mathbb{R}^n$ son ortogonales entre sí $\leftrightarrow$ **u** • **v** = 0

Ortogonalidad (resumen)

- Entonces, $\mathbf{u}$ y $\mathbf{v}$ son perpendiculares si:

$$\mathrm{dist}(\mathbf{u}, \mathbf{v}) = \mathrm{dist}(\mathbf{u}, -\mathbf{v})$$

- Dichas distancias son iguales si:

$$\mathrm{dist}(\mathbf{u}, -\mathbf{v}) = \mathrm{dist}(\mathbf{u}, \mathbf{v}) \leftrightarrow -2\mathbf{u} \cdot \mathbf{v} = 2\mathbf{u} \cdot \mathbf{v}$$

- Y se cumple solo si:

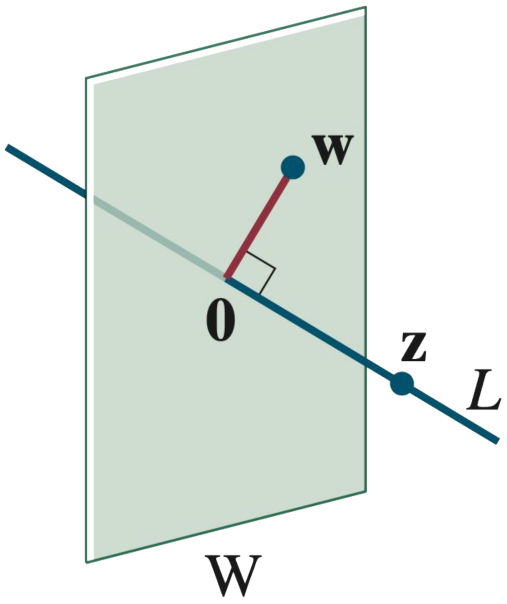
$$\mathbf{u} \cdot \mathbf{v} = 0$$

Complementos Ortogonales

Complementos Ortogonales

- Si un vector $\mathbf{z} \in \mathbb{R}^n$ es ortogonal a cualquier vector en un subespacio $W \in \mathbb{R}^n$, se dice entonces que $\mathbf{z}$ es ortogonal a W.

- El conjunto de todos los vectores $\mathbf{z} \in \mathbb{R}^n$ que son ortogonales a W, se llama el **complemento ortogonal** de W y se lo denota $W^{\perp}$



$L = W^{\perp} \quad W = L^{\perp}$

Complementos Ortogonales - Definición

- Un vector $\mathbf{z}$ está en $W^{\perp}$ si y sólo si es ortogonal a todos los vectores de un conjunto generador de $W \in \mathbb{R}^n$

$$W = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$$

$$\rightarrow \mathbf{z} \in W^{\perp} \qquad \mathbf{z} \perp \mathbf{v}_1, \mathbf{z} \perp \mathbf{v}_2, \ldots, \mathbf{z} \perp \mathbf{v}_p$$

- $W^{\perp}$ es un subespacio de $\mathbb{R}^n$

Complementos Ortogonales

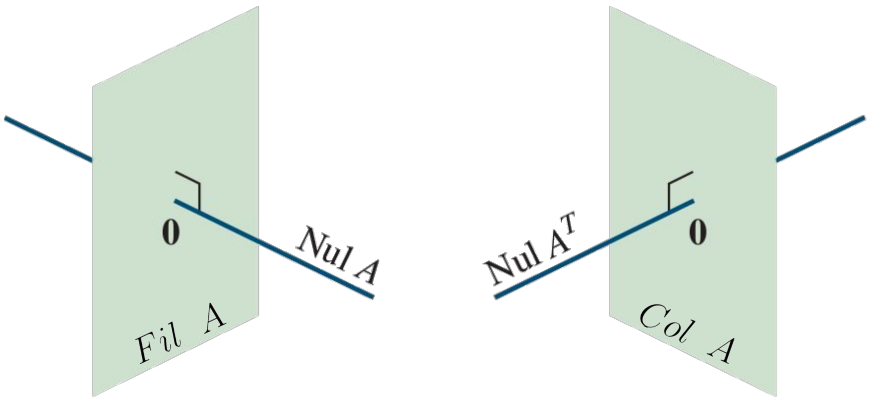
Dada una matriz A de $\mathbb{R}^{\mathrm{m\,x\,n}}$:

- El complemento ortogonal del espacio filas de A es el espacio nulo de A:

$$(\mathrm{Fila}\,A)^{\perp} = \mathrm{Nul}\,A$$

- El complemento ortogonal del espacio columnas de A es el espacio nulo de A^{T}:

$$(\mathrm{Col}\,A)^{\perp} = \mathrm{Nul}\,A^{T}$$



Complementos Ortogonales - Demostración

Dada una matriz A de $\mathbb{R}^{\mathrm{m \times n}}$:

- El complemento ortogonal del espacio filas de A es el espacio nulo de A: $(\mathrm{Fila}\, A)^{\perp} = \mathrm{Nul}\, A$

$$A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \dots & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} \mathbf{f}_1^T\mathbf{x} \\ \mathbf{f}_2^T\mathbf{x} \\ \dots \\ \mathbf{f}_m^T\mathbf{x} \end{pmatrix}$$

$$\mathrm{nul}\, A \rightarrow \mathbf{x} \mid A\mathbf{x} = 0 \rightarrow \mathbf{f}_1^T\mathbf{x} = 0, \mathbf{f}_2^T\mathbf{x} = 0, \cdots, \mathbf{f}_m^T\mathbf{x} = 0$$

Complementos Ortogonales - Demostración

Dada una matriz A de $\mathbb{R}^{\mathrm{m \times n}}$:

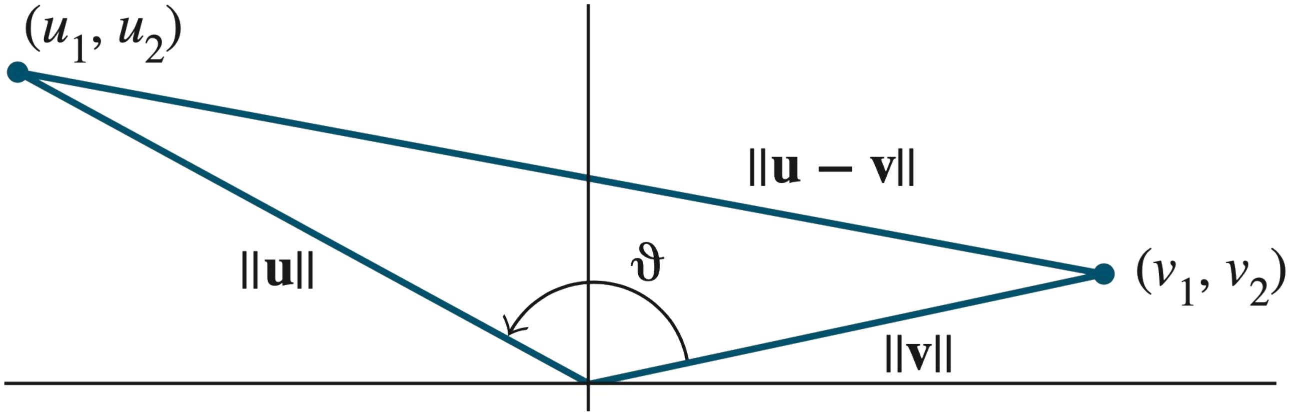
- El complemento ortogonal del espacio columnas de A es el espacio nulo de A^{T}: $(\mathrm{Col}\, A)^{\perp} = \mathrm{Nul}\, A^{T}$

  (Basta hacer el mismo ejercicio anterior con A^{T})

Ángulos en $\mathbb{R}^2$ y $\mathbb{R}^3$

- Interpretación geométrica del producto interno en $\mathbb{R}^2$ y $\mathbb{R}^3$:

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \vartheta$$



Conjuntos Ortogonales

Conjuntos Ortogonales

- Un conjunto de vectores $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p\}$ en $\mathbb{R}^n$ es un **conjunto ortogonal** si cada par de vectores distintos en el conjunto es ortogonal, es decir:

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0 \text{ para todo } i \neq j$$

Conjuntos Ortogonales

- Mostrar que $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto ortogonal:

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

Conjuntos Ortogonales

- Mostrar que $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto ortogonal:

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 3(-1) + 1(2) + 1(1) = 0$$

$$\mathbf{u}_1 \cdot \mathbf{u}_3 = 3\left(-\frac{1}{2}\right) + 1(-2) + 1\left(\frac{7}{2}\right) = 0$$

$$\mathbf{u}_2 \cdot \mathbf{u}_3 = -1\left(-\frac{1}{2}\right) + 2(-2) + 1\left(\frac{7}{2}\right) = 0$$

Conjuntos Ortogonales

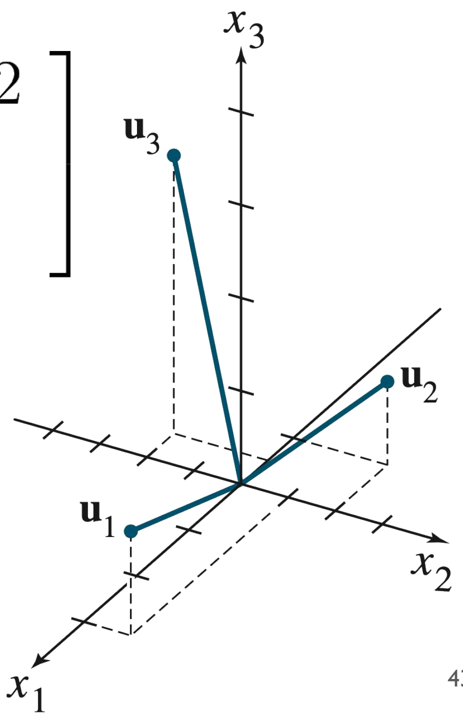
- Mostrar que $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto ortogonal:

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 3(-1) + 1(2) + 1(1) = 0$$

$$\mathbf{u}_1 \cdot \mathbf{u}_3 = 3\left(-\frac{1}{2}\right) + 1(-2) + 1\left(\frac{7}{2}\right) = 0$$

$$\mathbf{u}_2 \cdot \mathbf{u}_3 = -1\left(-\frac{1}{2}\right) + 2(-2) + 1\left(\frac{7}{2}\right) = 0$$



Base Ortogonal

- Si $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ en $\mathbb{R}^n$ es un conjunto ortogonal de vectores distintos de cero, entonces S es **linealmente independiente** y por lo tanto es una **base del subespacio** generado por S.
- A esta base se la llama **base ortogonal**.

Base Ortogonal - Demostración

- Recordar que para que $\{\mathbf{u}_1, ..., \mathbf{u}_p\}$ sea linealmente independiente, se tiene que cumplir que:

$$\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

sólo tenga la solución trivial.

Base Ortogonal - Demostración

$$\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

- Si esto se cumple para algunos escalares c_1, ..., c_p, entonces:

$$\begin{aligned}
0 = \mathbf{0}\cdot\mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\
&= (c_1\mathbf{u}_1)\cdot\mathbf{u}_1 + (c_2\mathbf{u}_2)\cdot\mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\
&= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) + c_2(\mathbf{u}_2\cdot\mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{u}_1) \\
&= c_1(\mathbf{u}_1\cdot\mathbf{u}_1)
\end{aligned}$$

Base Ortogonal - Demostración

$$\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

- Si esto se cumple para algunos escalares c_1, ..., c_p, entonces:

$$\begin{aligned} 0 = \mathbf{0} \cdot \mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p) \cdot \mathbf{u}_1 \\ &= (c_1\mathbf{u}_1) \cdot \mathbf{u}_1 + (c_2\mathbf{u}_2) \cdot \mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p) \cdot \mathbf{u}_1 \\ &= c_1 (\mathbf{u}_1 \cdot \mathbf{u}_1) + c_2 (\mathbf{u}_2 \cdot \mathbf{u}_1) + \cdots + c_p (\mathbf{u}_p \cdot \mathbf{u}_1) \\ &= c_1 (\mathbf{u}_1 \cdot \mathbf{u}_1) \longrightarrow c_1 = 0 \end{aligned}$$

Base Ortogonal

- Sea $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ una base ortogonal para un subespacio W en $\mathbb{R}^n$, para cada $\mathbf{y}$ en W, los pesos de la combinación lineal:

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

están dados por:

$$c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j} \quad (j = 1, \ldots, p)$$

Base Ortogonal

- A partir del conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ que es una base ortogonal para $\mathbb{R}^3$, expresar el vector $\mathbf{y}$ como una combinación lineal de los vectores en S:

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$$

Base Ortogonal

- A partir del conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ que es una base ortogonal para $\mathbb{R}^3$, expresar el vector $\mathbf{y}$ como una combinación lineal de los vectores en S:

$$\mathbf{y} \cdot \mathbf{u}_1 = 11, \quad \mathbf{y} \cdot \mathbf{u}_2 = -12, \quad \mathbf{y} \cdot \mathbf{u}_3 = -33$$
$$\mathbf{u}_1 \cdot \mathbf{u}_1 = 11, \quad \mathbf{u}_2 \cdot \mathbf{u}_2 = 6, \quad \mathbf{u}_3 \cdot \mathbf{u}_3 = 33/2$$

Base Ortogonal

- A partir del conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ que es una base ortogonal para $\mathbb{R}^3$, expresar el vector $\mathbf{y}$ como una combinación lineal de los vectores en S:

$$\begin{aligned}\mathbf{y} &= \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\mathbf{u}_3 \\ &= \frac{11}{11}\mathbf{u}_1 + \frac{-12}{6}\mathbf{u}_2 + \frac{-33}{33/2}\mathbf{u}_3 \\ &= \mathbf{u}_1 - 2\mathbf{u}_2 - 2\mathbf{u}_3\end{aligned}$$

Conjunto Ortonormal

- Un conjunto de vectores $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p\}$ es un **conjunto ortonormal** si es un conjunto ortogonal de vectores unitarios.
- Es decir:

$$||\mathbf{u}_i|| = 1 \text{ para todo } i$$

Conjunto Ortonormal

- Un conjunto de vectores $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p\}$ es un **conjunto ortonormal** si es un conjunto ortogonal de vectores unitarios.
- Es decir:

$$||\mathbf{u}_i|| = 1 \text{ para todo } i$$

- Como todo conjunto ortogonal es linealmente independiente, si W es el espacio generado por $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p\}$, entonces S es una **base ortonormal** para W.

Conjunto Ortonormal - Ejemplos

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Conjunto Ortonormal - Ejemplos

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

Conjunto Ortonormal - Ejemplos

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

Hay que probar:

1. $\mathbf{v}_i \cdot \mathbf{v}_j = 0 \text{ si } i \neq j$
2. $\mathbf{v}_i \cdot \mathbf{v}_j = 1 \text{ si } i = j$

Conjunto Ortonormal - Ejemplos

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

1. $\mathbf{v}_i \cdot \mathbf{v}_j = 0 \text{ si } i \neq j$

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0$$

$$\mathbf{v}_1 \cdot \mathbf{v}_3 = -3/\sqrt{726} - 4/\sqrt{726} + 7/\sqrt{726} = 0$$

$$\mathbf{v}_2 \cdot \mathbf{v}_3 = 1/\sqrt{396} - 8/\sqrt{396} + 7/\sqrt{396} = 0$$

Conjunto Ortonormal - Ejemplos

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

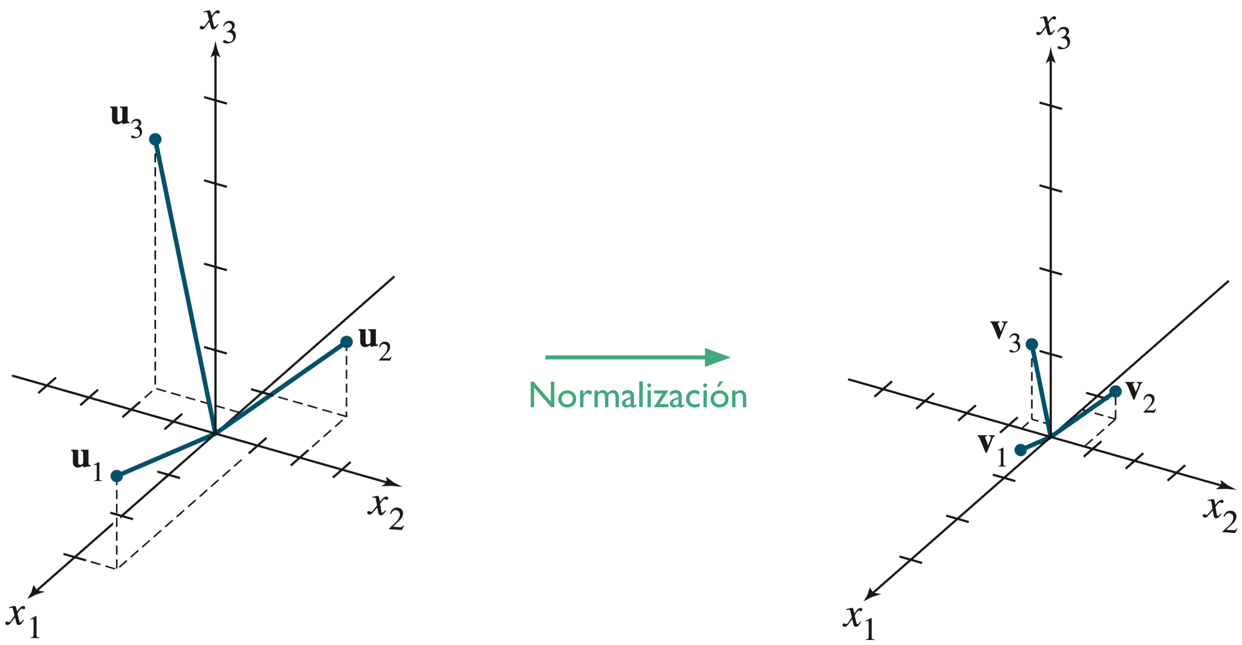
2. $\mathbf{v}_i \cdot \mathbf{v}_j = 1 \text{ si } i = j$

$$\mathbf{v}_1 \cdot \mathbf{v}_1 = 9/11 + 1/11 + 1/11 = 1$$

$$\mathbf{v}_2 \cdot \mathbf{v}_2 = 1/6 + 4/6 + 1/6 = 1$$

$$\mathbf{v}_3 \cdot \mathbf{v}_3 = 1/66 + 16/66 + 49/66 = 1$$

Conjunto Ortonormal - Ejemplos



$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

Conjunto Ortonormal

- Si partimos de un conjunto ortogonal $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p\}$ y lo normalizamos, los vectores resultantes seguirán siendo también ortogonales. Por lo tanto, formarán un conjunto ortonormal.

Matrices con Columnas Ortonormales

- Una matriz U de m x n tiene columnas ortonormales sí y sólo sí $U^{\top}U = I$

$$U^T U = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \mathbf{u}_1^T\mathbf{u}_3 \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \mathbf{u}_2^T\mathbf{u}_3 \\ \mathbf{u}_3^T\mathbf{u}_1 & \mathbf{u}_3^T\mathbf{u}_2 & \mathbf{u}_3^T\mathbf{u}_3 \end{bmatrix}$$

$$\mathbf{u}_i^\top \mathbf{u}_j = 1 \text{ si } i = j$$

$$\mathbf{u}_i^\top \mathbf{u}_j = 0 \text{ s } i \neq j$$

Matrices con Columnas Ortonormales - Propiedades

- Sea U una matriz de m x n con columnas ortonormales, $\mathbf{x}$ e $\mathbf{y}$ vectores en $\mathbb{R}^n$, entonces:

1. $||U\mathbf{x}|| = ||\mathbf{x}||$
2. $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$
    a. $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ sí y sólo sí $\mathbf{x} \cdot \mathbf{y} = 0$

Definición: Matrices Ortogonales - Definición

- Sea U una matriz cuadrada invertible tal que $U^{-1} = U^{T}$, entonces U es una matriz ortogonal.

- En particular: toda matriz cuadrada con columnas ortonormales, es una matriz ortogonal.

Observación: Se cumple también que dicha matriz tiene filas ortonormales también

Ejemplo

- Verificar que U es una matriz ortogonal y que sus filas son ortonormales:

$$U = \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

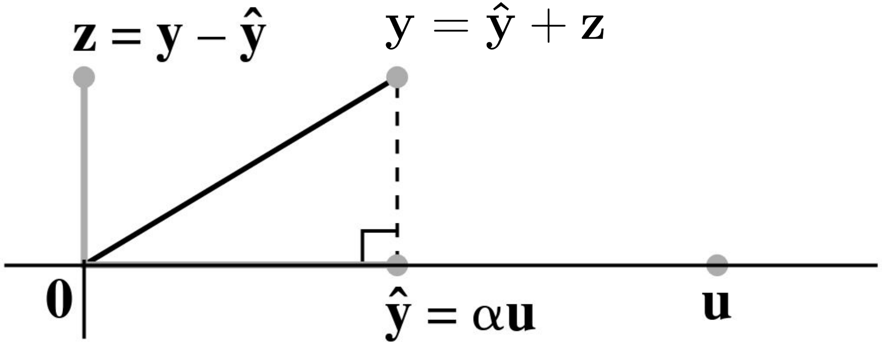
Proyecciones Ortogonales

Proyecciones Ortogonales

- Dado un vector $\mathbf{u}$ diferente de $\mathbf{0}$ en $\mathbb{R}^n$. Consideremos el problema de descomponer otro vector $\mathbf{y}$ en $\mathbb{R}^n$ en la suma de dos vectores, uno múltiplo de $\mathbf{u}$ y otro ortogonal a $\mathbf{u}$:

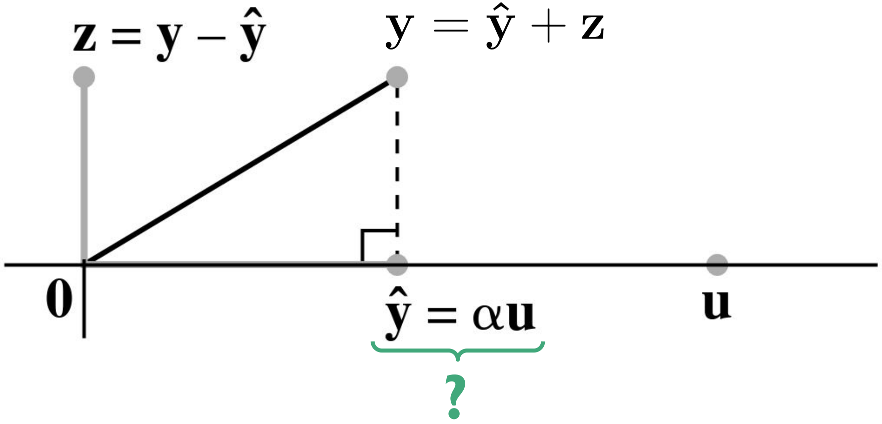
Proyecciones Ortogonales

- Dado un vector $\mathbf{u}$ diferente de $\mathbf{0}$ en $\mathbb{R}^n$. Consideremos el problema de descomponer otro vector $\mathbf{y}$ en $\mathbb{R}^n$ en la suma de dos vectores, uno múltiplo de $\mathbf{u}$ y otro ortogonal a $\mathbf{u}$:



Proyecciones Ortogonales

- Dado un vector $\mathbf{u}$ diferente de $\mathbf{0}$ en $\mathbb{R}^n$. Consideremos el problema de descomponer otro vector $\mathbf{y}$ en $\mathbb{R}^n$ en la suma de dos vectores, uno múltiplo de $\mathbf{u}$ y otro ortogonal a $\mathbf{u}$:



Proyecciones Ortogonales

- Dado un vector $\mathbf{u}$ diferente de $\mathbf{0}$ en $\mathbb{R}^n$. Consideremos el problema de descomponer otro vector $\mathbf{y}$ en $\mathbb{R}^n$ en la suma de dos vectores, uno múltiplo de $\mathbf{u}$ y otro ortogonal a $\mathbf{u}$:

$$\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$$

$$\mathbf{z} = \mathbf{y} - \alpha\mathbf{u}$$

Proyecciones Ortogonales

- Dado un vector $\mathbf{u}$ diferente de $\mathbf{0}$ en $\mathbb{R}^n$. Consideremos el problema de descomponer otro vector $\mathbf{y}$ en $\mathbb{R}^n$ en la suma de dos vectores, uno múltiplo de $\mathbf{u}$ y otro ortogonal a $\mathbf{u}$:

$$\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$$

$$\mathbf{z} = \mathbf{y} - \alpha\mathbf{u}$$

$$\mathbf{z} \perp \mathbf{u} \longrightarrow \mathbf{z} \cdot \mathbf{u} = 0$$

$$(\mathbf{y} - \alpha\mathbf{u}) \cdot \mathbf{u} = 0$$

Proyecciones Ortogonales

- Dado un vector $\mathbf{u}$ diferente de $\mathbf{0}$ en $\mathbb{R}^n$. Consideremos el problema de descomponer otro vector $\mathbf{y}$ en $\mathbb{R}^n$ en la suma de dos vectores, uno múltiplo de $\mathbf{u}$ y otro ortogonal a $\mathbf{u}$:

$$0 = (\mathbf{y} - \alpha\mathbf{u}) \cdot \mathbf{u} = \mathbf{y} \cdot \mathbf{u} - (\alpha\mathbf{u}) \cdot \mathbf{u} = \mathbf{y} \cdot \mathbf{u} - \alpha(\mathbf{u} \cdot \mathbf{u})$$

Proyecciones Ortogonales

- Dado un vector $\mathbf{u}$ diferente de $\mathbf{0}$ en $\mathbb{R}^n$. Consideremos el problema de descomponer otro vector $\mathbf{y}$ en $\mathbb{R}^n$ en la suma de dos vectores, uno múltiplo de $\mathbf{u}$ y otro ortogonal a $\mathbf{u}$:

$$0 = (\mathbf{y} - \alpha\mathbf{u}) \cdot \mathbf{u} = \mathbf{y} \cdot \mathbf{u} - (\alpha\mathbf{u}) \cdot \mathbf{u} = \mathbf{y} \cdot \mathbf{u} - \alpha(\mathbf{u} \cdot \mathbf{u})$$

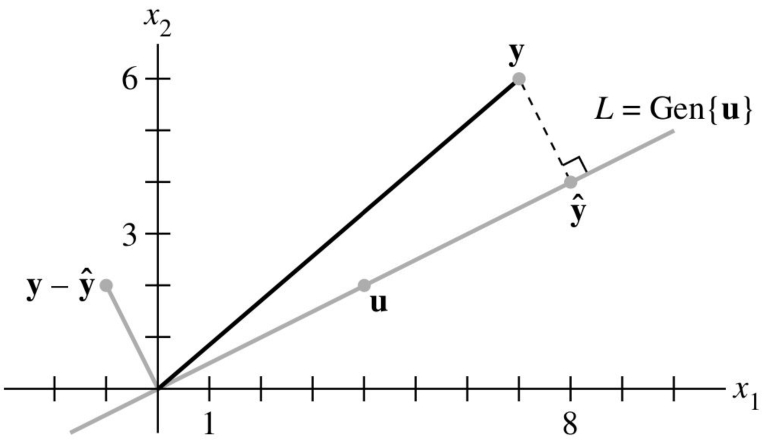
$$\alpha = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \rightarrow \hat{\mathbf{y}} = \underbrace{\frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u}}_{\text{Proyección ortogonal de } \mathbf{y} \text{ sobre } \mathbf{u}}$$

Proyecciones Ortogonales

- Más generalmente, cualquier vector en $\text{Gen}\{\mathbf{u}\}$ es de la forma $c\,\mathbf{u}$. Podemos ver que la proyección de $\mathbf{y}$ sobre $c\,\mathbf{u}$ es la misma que sobre $\mathbf{u}$

$$L = \text{Gen}\;\{\mathbf{u}\}$$

$$\hat{\mathbf{y}} = \text{proj}_L\,\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$$



Proyecciones Ortogonales: ejercicio

- Encontrar la proyección de $\mathbf{y}$ sobre $\mathbf{u}$. Escribir $\mathbf{y}$ como la suma de dos vectores ortogonales, uno en $\mathrm{Gen}\{\mathbf{u}\}$ y otro ortogonal a $\mathbf{u}$:

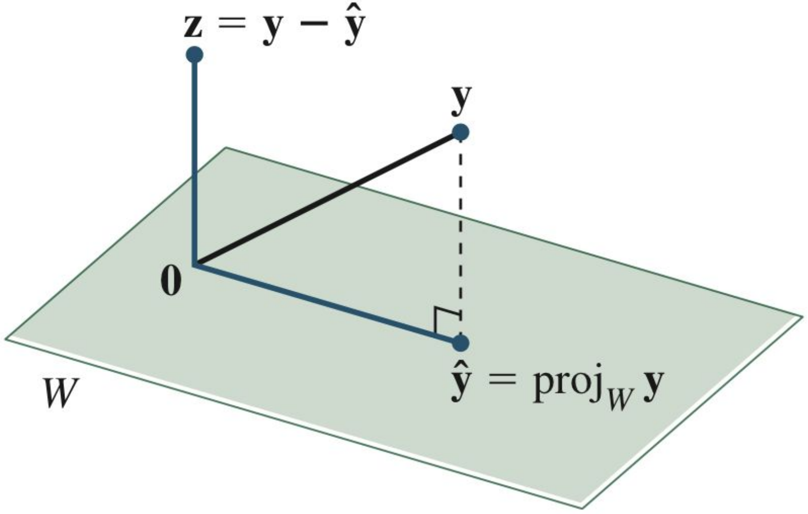
$$\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} \qquad \mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

Proyecciones Ortogonales

- Sea W un subespacio de $\mathbb{R}^n$. Entonces cada $\mathbf{y}$ en $\mathbb{R}^n$ se puede escribir de manera única en la forma:

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

$$\begin{cases} \hat{\mathbf{y}} \text{ está en } W \\ \mathbf{z} \text{ está en } W^\perp \end{cases}$$



Proyecciones Ortogonales

- Sea W un subespacio de $\mathbb{R}^n$. Entonces cada $\mathbf{y}$ en $\mathbb{R}^n$ se puede escribir de manera única en la forma:

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

- Si $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ es una base ortogonal de W:

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p$$

Proyecciones Ortogonales: ejercicio

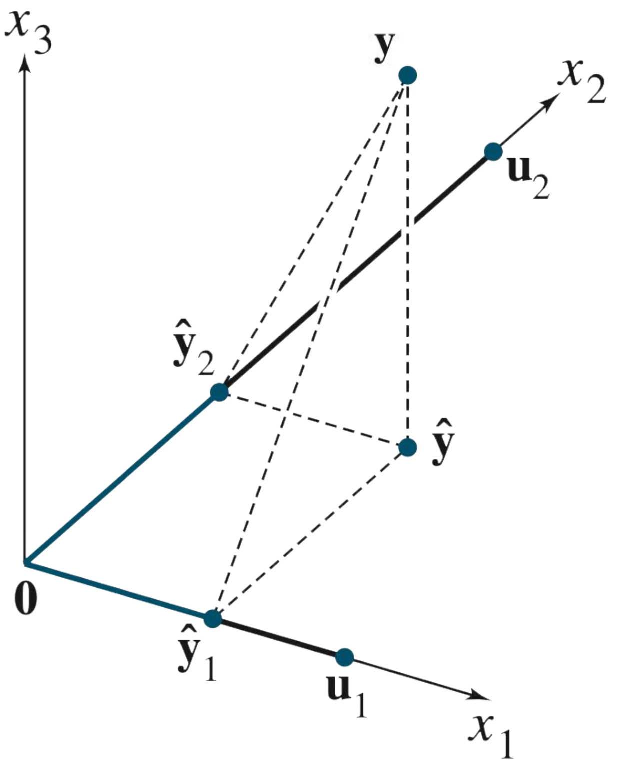
- Escribir el vector $\mathbf{y}$ como la suma de un vector en W y un vector ortogonal a W, donde $\mathbf{u}_1$, $\mathbf{u}_2$ es una base ortogonal para $W = \text{Gen}\{\mathbf{u_1}, \mathbf{u_2}\}$:

$$\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Proyecciones Ortogonales: Interpretación Geométrica

- Cada término es la proyección ortogonal de $\mathbf{y}$ en los subespacios unidimensionales generados por cada $\mathbf{u}_i$ en W

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 = \hat{\mathbf{y}}_1 + \hat{\mathbf{y}}_2$$

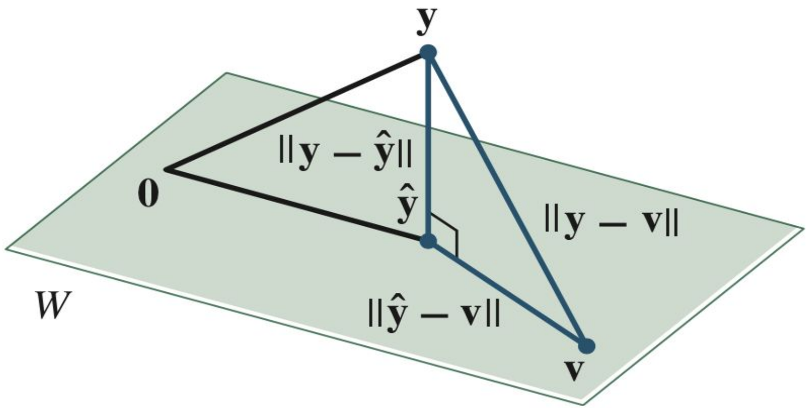


Proyecciones Ortogonales como mejor aproximación

- Sea W un subespacio de $\mathbb{R}^n$. Sea $\mathbf{y}$ un vector en $\mathbb{R}^n$, y sea $\hat{\mathbf{y}}$ la proyección ortogonal de $\mathbf{y}$ en W, entonces $\hat{\mathbf{y}}$ es **el punto más cercano** a $\mathbf{y}$ en W.

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|$$

$$\text{para todo } \mathbf{v} \text{ diferente de } \hat{\mathbf{y}}$$



Proceso de Gram-Schmidt y Factorización QR

El proceso de Gram-Schmidt

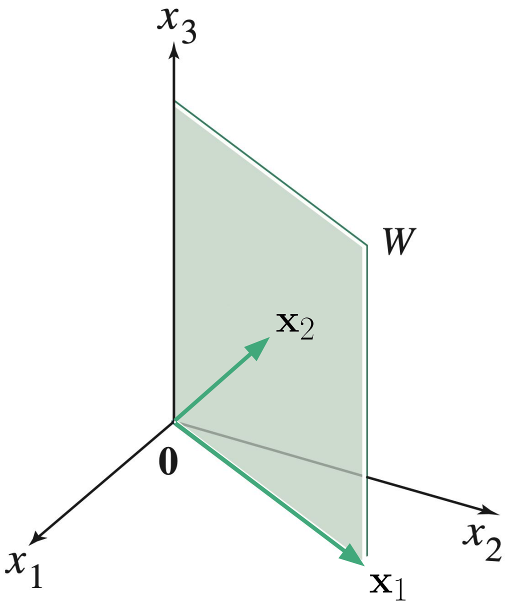
- Sea $W = \text{Gen}\,\{\mathbf{x}_1, \mathbf{x}_2\}$, encontrar una base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2\}$:

$$\mathbf{x}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

El proceso de Gram-Schmidt

- Sea $W = \text{Gen}\,\{\mathbf{x}_1, \mathbf{x}_2\}$, encontrar una base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2\}$:

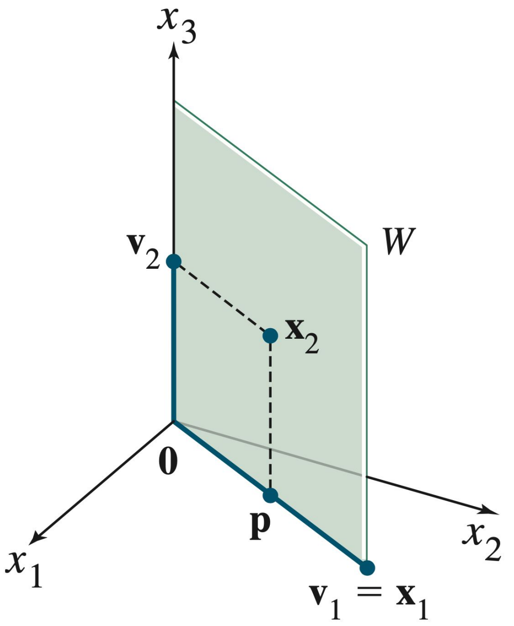
$$\mathbf{x}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$



El proceso de Gram-Schmidt

- Sea $W = \text{Gen}\,\{\mathbf{x}_1, \mathbf{x}_2\}$, encontrar una base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2\}$:

$$\mathbf{x}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$



El proceso de Gram-Schmidt

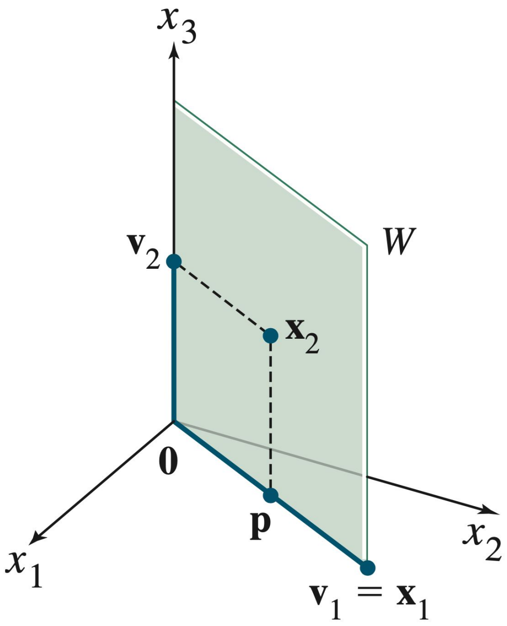
- Sea $W = \text{Gen}\,\{\mathbf{x}_1, \mathbf{x}_2\}$, encontrar una base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2\}$:

$$\mathbf{v_1} = \mathbf{x_1}$$

$$\mathbf{p} = \text{proj}_{\text{Gen}\{\text{x}_1\}}\,\mathbf{x_2} = \frac{\mathbf{x_2} \cdot \mathbf{x_1}}{\mathbf{x_1} \cdot \mathbf{x_1}}\mathbf{x_1}$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{p} = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{x}_1}{\mathbf{x}_1 \cdot \mathbf{x}_1}\mathbf{x}_1$$

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45}\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$



El proceso de Gram-Schmidt

- Dada una base $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p\}$, para un espacio W en $\mathbb{R}^n$ si definimos:

$$
\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}
\end{aligned}
$$

- Entonces $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ es una base ortogonal para W.

El proceso de Gram-Schmidt

- Dada una base $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p\}$, para un espacio W en $\mathbb{R}^n$ si definimos:

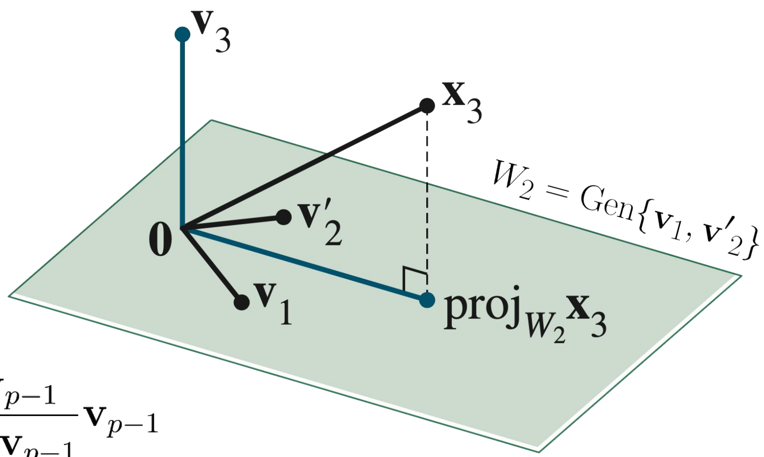
$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2$$

$$\vdots$$

$$\mathbf{v}_p = \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}$$



- Entonces $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$ es una base ortogonal para W.

Factorización QR

- Sea A una matriz de $m \times n$ con columnas linealmente independientes, entonces A se puede factorizar como $A = QR$, donde:
  - Q es una matriz de $m \times n$ cuyas columnas forman una base ortonormal para columnas de A
  - R es una matriz triangular superior invertible de $n \times n$ con entradas positivas en su diagonal

Factorización QR: procedimiento

1. Construir una base ortogonal para $W = \mathrm{Col}\, A$ con el procedimiento de Gram-Schmidt.
2. Normalizar los vectores del punto 1 para obtener una base ortonormal.
3. Utilizar los vectores ortonormales como columnas de Q.
4. Para encontrar R:
   a. $A = QR \rightarrow Q^T A = Q^T QR$
   b. como Q es una matriz ortogonal $\longrightarrow$ $Q^T Q = I$
   c. entonces $R = Q^T A$

Factorización QR: ejercicio

- Hallar la factorización QR de A:

$$A = \begin{bmatrix} 3 & 1 \\ 6 & 2 \\ 0 & 2 \end{bmatrix}$$

Factorización QR: ejercicio

- Hallar la factorización QR de A:

$$A = \begin{bmatrix} 3 & 1 \\ 6 & 2 \\ 0 & 2 \end{bmatrix} \quad \underbrace{\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}}_{\text{1. Base ortogonal para Col A}}$$

Factorización QR: ejercicio

- Hallar la factorización QR de A:

$$A = \begin{bmatrix} 3 & 1 \\ 6 & 2 \\ 0 & 2 \end{bmatrix} \quad \underbrace{\mathbf{u}_1 = \begin{bmatrix} \frac{3}{\sqrt{45}} \\ \frac{6}{\sqrt{45}} \\ 0 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{}$$

2. Base ortonormal para Col A (normalizamos)

$$\|\mathbf{v}_1\| = \sqrt{45} \qquad \|\mathbf{v}_2\| = 2$$

Factorización QR: ejercicio

- Hallar la factorización QR de A:

$$Q = [\mathbf{u_1} \;\; \mathbf{u_2}] = \begin{bmatrix} \frac{3}{\sqrt{45}} & 0 \\ \frac{6}{\sqrt{45}} & 0 \\ 0 & 1 \end{bmatrix}$$

3. Construimos matriz Q

Factorización QR: ejercicio

- Hallar la factorización QR de A:

$$R = Q^T A = \begin{bmatrix} \frac{3}{\sqrt{45}} & \frac{6}{\sqrt{45}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 6 & 2 \\ 0 & 2 \end{bmatrix} = \underbrace{\begin{bmatrix} \sqrt{45} & \sqrt{5} \\ 0 & 2 \end{bmatrix}}_{\text{4. Construimos matriz } R}$$

Factorización QR: ejercicio

- Hallar la factorización QR de A:

$$A = QR$$

$$\begin{bmatrix} 3 & 1 \\ 6 & 2 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{45}} & 0 \\ \frac{6}{\sqrt{45}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{45} & \sqrt{5} \\ 0 & 2 \end{bmatrix}$$

Problemas de Mínimos Cuadrados

Mínimos cuadrados

- Muchas veces, un sistema lineal $A\mathbf{x} = \mathbf{b}$ no tiene solución (es inconsistente).

- Podemos pensar en $A\mathbf{x}$ como **una aproximación** a $\mathbf{b}$, y entonces, un sustituto aceptable de solución es un vector $\mathbf{x}$ que reduce la distancia entre $A\mathbf{x}$ y $\mathbf{b}$ lo más posible.

Mínimos cuadrados

- Si A es de m x n, y $\mathbf{b}$ está en $\mathbb{R}^{\mathrm{m}}$, una solución por mínimos cuadrados de $A\,\mathbf{x} = \mathbf{b}$ es:

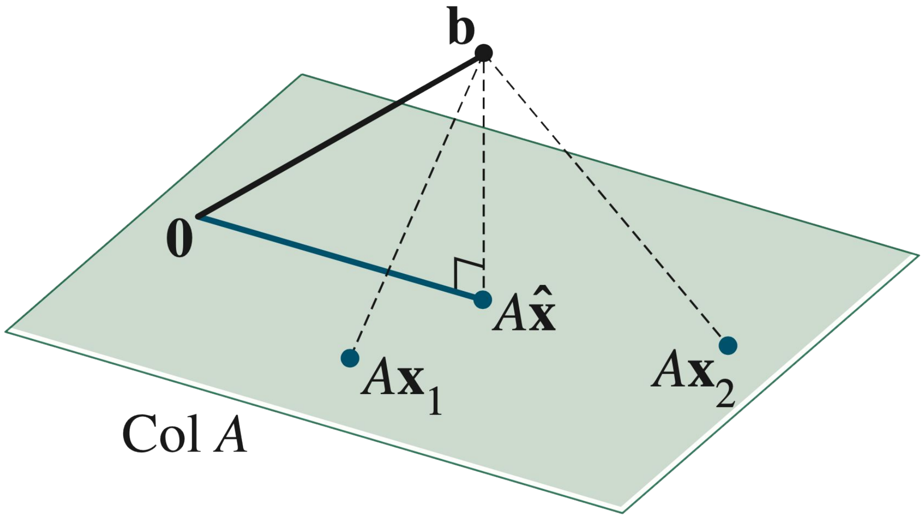
$$\hat{\mathbf{x}} \text{ en } \mathbb{R}^n \text{ tal que:}$$

$$||\mathbf{b} - A\hat{\mathbf{x}}|| \leq ||\mathbf{b} - A\mathbf{x}|| \text{ para todo } \mathbf{x} \text{ en } \mathbb{R}^n$$

Mínimos cuadrados

- Sin importar que $\mathbf{x}$ se elija, el vector $A\,\mathbf{x}$ que mejor aproxima a $\mathbf{b}$, necesariamente estará en el espacio columnas, $\mathrm{Col}\,A$

- Buscamos un $\mathbf{x}$ que convierta $A\,\mathbf{x}$ en el punto de $\mathrm{Col}\,A$ más cercano a $\mathbf{b}$.



$$\|\mathbf{b}-A\hat{\mathbf{x}}\| \leq \|\mathbf{b}-A\mathbf{x}\| \text{ para todo } \mathbf{x} \text{ en } \mathbb{R}^n$$

Solución al problema de mínimos cuadrados

- Definimos $\hat{\mathbf{b}}$:

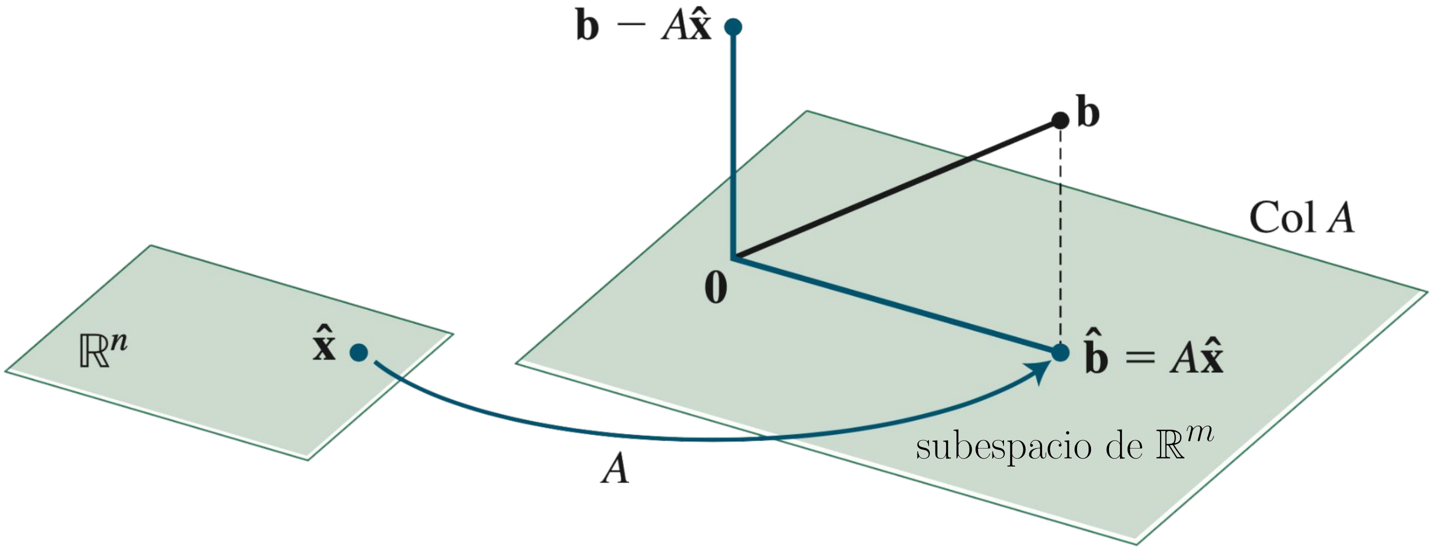
$$\hat{\mathbf{b}} = \mathrm{proj}_{\mathrm{Col}\,A}\,\mathbf{b}$$

- Como $\hat{\mathbf{b}}$ está en el espacio columna de A (Col A), sabemos que existe una $\hat{\mathbf{x}}$ en $\mathbb{R}^{\mathrm{n}}$ tal que:

$$A\hat{\mathbf{x}} = \hat{\mathbf{b}}$$

Solución al problema de mínimos cuadrados

- Dado que $\hat{\mathbf{b}}$ es el punto de $\mathrm{Col}\,A$ más cercano a $\mathbf{b}$, el vector $\hat{\mathbf{x}}$ es una solución de mínimos cuadrados de $A\,\mathbf{x} = \mathbf{b}$.



Mínimos cuadrados: Solución

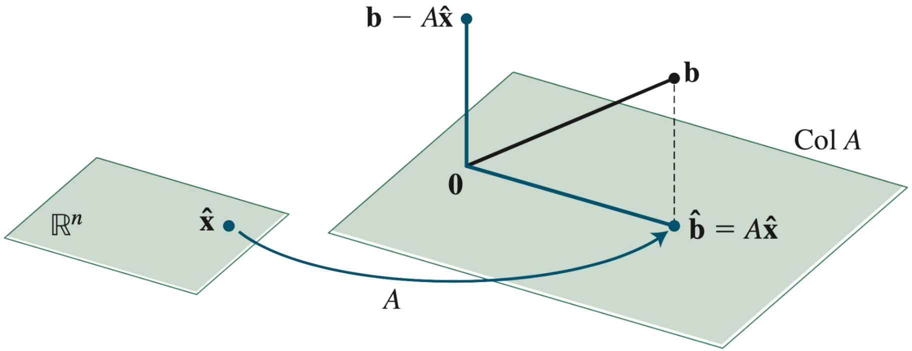
- Según la descomposición ortogonal vista antes:

$$\mathbf{b} - A\hat{\mathbf{x}} \perp \operatorname{Col} A$$

$$A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$$

$$A^T\mathbf{b} - A^T A\hat{\mathbf{x}} = \mathbf{0}$$

$$\boxed{A^T A\hat{\mathbf{x}} = A^T\mathbf{b}}$$



→ Ecuaciones Normales
(única o infinitas soluciones)

Mínimos cuadrados: ejercicio

- Encontrar la solución de mínimos cuadrados del siguiente sistema inconsistente:

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Mínimos cuadrados: ejercicio

- Encontrar la solución de mínimos cuadrados del siguiente sistema:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$

Mínimos cuadrados: ejercicio

- Encontrar la solución de mínimos cuadrados del siguiente sistema:

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$$

Mínimos cuadrados: ejercicio

- Encontrar la solución de mínimos cuadrados del siguiente sistema:

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}$$

Mínimos cuadrados: ejercicio

- Encontrar la solución de mínimos cuadrados del siguiente sistema:

$$A^T A\mathbf{x} = A^T\mathbf{b}$$

$$\begin{bmatrix} 6 & 2 & 2 & 2 & 4 \\ 2 & 2 & 0 & 0 & -4 \\ 2 & 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 2 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Mínimos cuadrados: ejercicio

- Encontrar la solución de mínimos cuadrados del siguiente sistema:

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \longrightarrow \hat{\mathbf{x}} = \begin{bmatrix} 3 \\ -5 \\ -2 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Infinitas soluciones

Mínimos cuadrados con solución única

- Sea A una matriz de m x n, los siguientes enunciados son equivalentes:
  - La ecuación $A\,\mathbf{x} = \mathbf{b}$ tiene una única solución por mínimos cuadrados para todo $\mathbf{b}$ en $\mathbb{R}^{\mathrm{m}}$
  - Las columnas de A son linealmente independientes
  - La matriz $A^{\mathrm{T}}A$ es invertible
- Cuando estos enunciados son verdaderos, la solución $\hat{\mathbf{x}}$ de mínimos cuadrados está dada por:

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$$

Matriz de Proyección

- Dado que:

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$$

entonces siendo $\hat{\mathbf{b}} = A\,\hat{\mathbf{x}}$:

$$\hat{\mathbf{b}} = A(A^T A)^{-1} A^T \mathbf{b}$$

- A la matriz $A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ de m x m se la llama **Matriz de proyección.**