

# Álgebra Lineal

## Lectura 21: Producto Interno

Ing. Aldo Jiménez Arteaga

Enero 2020

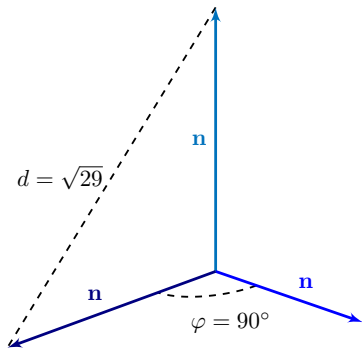


Figura 1. La métrica en el espacio vectorial incluye medir distancias entre puntos y medir ángulos entre vectores.

Los elementos de un espacio vectorial pueden ubicarse en una referencia; además, puede investigarse qué espacio vectorial generan dichos elementos al analizar su independencia lineal. En conceptos más avanzados un espacio vectorial puede transformarse en otro espacio, conservando todas sus propiedades que inicialmente posee. Al ubicar y transformar, se puede cuantificar la ubicación y la deformación sufrida por el vector.

Estos procesos requieren de una medición. Geométricamente, la métrica se divide en dos: medición de distancias y medición de direcciones. Cuando se habla de medir distancia, se hace mediante líneas rectas; en cambio cuando se habla de medir direcciones, se hace mediante arcos de circunferencia (véase la figura 1).

La medición en el espacio vectorial crea la necesidad de introducir

una herramienta que arroje métricas y conserve las propiedades de la linealidad. Éste es el concepto de producto interno.

Sea el espacio vectorial  $V$ , definido sobre el campo  $K$ . La función  $f : V \times V \rightarrow K$  es un producto interno (interior o escalar), si satisface las siguientes propiedades:

- Linealidad,  $f(\alpha \mathbf{u}, \mathbf{v} + \mathbf{w}) = \alpha f(\mathbf{u}, \mathbf{v}) + \alpha f(\mathbf{u}, \mathbf{w})$
- Simetría,  $f(\mathbf{u}, \mathbf{v}) = \alpha f(\mathbf{v}, \mathbf{u})$
- Positividad,  $f(\alpha \mathbf{u}, \mathbf{u}) > 0, \forall \mathbf{u} \neq \mathbf{0}$

El producto interno permite calcular magnitudes, distancias y ángulos en un espacio vectorial. Los espacios vectoriales con estas propiedades se llaman espacios prehilbertianos.

Un producto interno trascendente en Geometría Analítica y Cálculo Vectorial es el producto punto, el cual dota a los espacios  $\mathbb{R}^n$  con la métrica propuesta por la Geometría Euclídeana; es por ello que a estos espacios se les llama espacios euclídeanos.

Las propiedades adicionales del producto interno son:

- $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$ , si los vectores pertenecen a un espacio vectorial sobre los complejos.
- $\langle \mathbf{u}, \alpha \mathbf{v} \rangle = \bar{\alpha} \langle \mathbf{u}, \mathbf{v} \rangle$ , si  $\alpha \in \mathbb{C}$ .
- $\langle \mathbf{0}, \mathbf{0} \rangle = 0$ .
- $\langle \mathbf{u}, \mathbf{0} \rangle = 0$

donde  $\overline{\langle \mathbf{v}, \mathbf{u} \rangle}$  y  $\bar{\alpha}$  son los conjugados de  $\langle \mathbf{v}, \mathbf{u} \rangle$  y  $\alpha$ , respectivamente.

**Ejemplo.** En el espacio vectorial  $\mathbb{R}^2$  se define la función

$$f : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R} \Rightarrow f(\mathbf{u}, \mathbf{v}) = 2u_1v_1 + 3u_2v_2$$

$$\forall \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2. \text{ ¿}f \text{ es un producto interno?}$$

Definiendo a los vectores  $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ , cada propiedad se desarrolla como sigue.

Linealidad:

$$\begin{aligned} \langle \mathbf{u} + \mathbf{v}, \alpha \mathbf{w} \rangle &= \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \alpha \langle \mathbf{v}, \mathbf{w} \rangle \\ \left\langle \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}, \begin{bmatrix} \alpha w_1 \\ \alpha w_2 \end{bmatrix} \right\rangle &= \alpha (2u_1w_1 + 3u_2w_2) + \\ &\quad + \alpha (2v_1w_1 + 3v_2w_2) \\ 2\alpha (u_1 + v_1) w_1 + 3\alpha (u_2 + v_2) w_2 &= \alpha (2u_1w_1 + 2v_1w_1) + \\ &\quad + \alpha (3u_2w_2 + 3v_2w_2) \\ &= 2\alpha (u_1 + v_1) w_1 + 3\alpha (u_2 + v_2) w_2 \end{aligned}$$

Simetría:

$$\left\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right\rangle = \left\langle \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right\rangle$$

$$\begin{aligned} 2u_1v_1 + 3u_2v_2 &= 2v_1u_1 + 3v_2u_2 \\ &= 2u_1v_1 + 3u_2v_2 \end{aligned}$$

Positividad:

$$\begin{aligned} \left\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right\rangle &> 0 \\ 2u_1 \cdot u_1 + 3u_2 \cdot u_2 &> 0 \\ 2u_1^2 + 3u_2^2 &> 0 \end{aligned}$$

Ésta última propiedad se cumple pues 2 y 3 son positivos, los números reales  $u_1$  y  $u_2$  al elevarlos al cuadrado se vuelven positivos, y las sumas y multiplicaciones entre números positivos siempre devuelven un resultado positivo.

Las tres propiedades se cumplen, por lo tanto la función es un producto interno.

Para un espacio vectorial existe una infinidad de productos internos; cada uno de ellos arroja sus propios resultados. La necesidad de tener varios productos internos radica en la individualidad que tienen diversos fenómenos matemáticos, estadísticos o físicos, entre otros. En los conceptos de norma y ángulo se observará esa variedad de resultados para una misma herramienta definida de diferentes formas.

# Álgebra Lineal

## Lectura 22: Norma y Vector Unitario

Ing. Aldo Jiménez Arteaga

Enero 2020

### 1. Norma del Vector

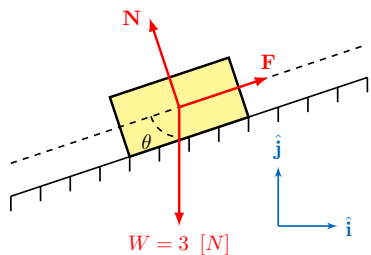


Figura 1. La definición física de vector se obtiene mediante una base y el producto interno.

Se ha comentado que el producto interno introduce la métrica en el espacio vectorial. Dicha métrica se basa en la propiedad de positividad, ya que al ser un escalar real mayor a cero puede indicar una magnitud.

En Física, una cantidad vectorial posee tres características en su definición: magnitud, dirección y sentido. El sentido se establece al momento de ubicar al vector en una referencia; la dirección se conoce a partir del ángulo con la referencia; la magnitud es la longitud del vector, llamada el módulo. Estas tres características se definen a partir de los espacios vectoriales mediante los conceptos de base, módulo y ángulo (figura 1). El primero se llama vector de coordenadas; los dos últimos se definen a partir del producto interno definido en el espacio vectorial.

El primer concepto de la medición es la norma (módulo) de un vector, que es la positividad enmarcada en el ámbito del teorema de Pitágoras.

En un espacio vectorial  $V$  con producto interno  $\langle \mathbf{u}, \mathbf{v} \rangle \quad \forall \mathbf{u}, \mathbf{v} \in V$  definido, la norma de un vector se establece como

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$$

Dentro de un espacio vectorial se pueden definir varios productos internos, lo cual implica que para un mismo vector se tendrá una norma diferente por cada producto interno definido. A pesar de esto, la norma siempre cumple las siguientes propiedades:

- ❑  $\|\mathbf{v}\| > 0$ . Solo el vector nulo posee norma nula.
- ❑  $\|\alpha \mathbf{v}\| = |\alpha| \|\mathbf{v}\|$ ,  $\forall \alpha \in K$ . El escalar  $|\alpha|$  denota el valor absoluto si  $\alpha \in \mathbb{R}$ , y denota el conjugado si  $\alpha \in \mathbb{C}$ .
- ❑  $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ . Esta propiedad es conocida como desigualdad del triángulo, la cual es la extensión de la misma propiedad del valor absoluto pero en espacios vectoriales diferentes a  $\mathbb{R}$ .

Se hace notar que la definición de norma, bajo el producto punto, es

idéntica a la definición de módulo en Geometría Analítica:

$$\begin{aligned}
 |\mathbf{v}| &= \|\mathbf{v}\| \\
 \sqrt{x^2 + y^2 + z^2} &= \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} \\
 &= \sqrt{\begin{bmatrix} x \\ y \\ z \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}} \\
 &= \sqrt{x^2 + y^2 + z^2}
 \end{aligned}$$

La norma basada en el producto punto no es la única forma de medir la longitud de un vector en  $\mathbb{R}^n$ . Otros productos internos en dichos espacios vectoriales se expresan como

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}$$

donde  $A$  es una matriz simétrica con valores propios positivos. Este tipo de productos se conocen como productos ponderados, los cuales permiten analizar las magnitudes de ciertos vectores de formas más eficientes que el producto punto. Por ejemplo, un ejercicio estadístico donde se desea calcular las magnitudes de las parejas de datos mostradas en la tabla 1.

	$P_1$	$P_2$	$P_3$	$P_4$
$x$	$1.0 \times 10^4$	$2.5 \times 10^4$	$3.1 \times 10^4$	$3.5 \times 10^4$
$y$	$3.2 \times 10^{-5}$	$4.1 \times 10^{-5}$	$4.5 \times 10^{-5}$	$5.0 \times 10^{-5}$

Tabla 1

Mediante la norma del producto punto, los datos en el renglón  $y$  caerían de significado al momento de establecer la métrica:

$$\begin{aligned}
 \|P_1\| &= \sqrt{(10000)^2 + (0.000032)^2} \\
 &= 10000
 \end{aligned}$$

En cambio, si se utiliza el producto interno ponderado

$$\langle \mathbf{u}, \mathbf{v} \rangle = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 10^{-8} & 0 \\ 0 & 10^9 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

la norma conserva toda la información de la pareja de datos:

$$\begin{aligned}
 \|P_1\| &= \sqrt{\begin{bmatrix} 10000 & 0.000032 \end{bmatrix} \begin{bmatrix} 10^{-8} & 0 \\ 0 & 10^9 \end{bmatrix} \begin{bmatrix} 10000 \\ 0.000032 \end{bmatrix}} \\
 &= \sqrt{\begin{bmatrix} 0.0001 & 32000 \end{bmatrix} \begin{bmatrix} 10000 \\ 0.000032 \end{bmatrix}} \\
 &= \sqrt{1 + 1.024} \\
 &= 1.4227
 \end{aligned}$$

Se observa que en este caso un dato no ha prevalecido sobre el otro.

La existencia de diversos productos internos permite escalar vectores al momento de medirlos. El escalamiento puede ser de ampliación, de reducción o mixto, según sea la necesidad del fenómeno o evento que se esté analizando.

**Ejemplo.** Calcule la norma del vector  $\mathbf{p} = -2x^2 + x - 1$  bajo tres productos internos diferentes en el espacio vectorial

$$P_2 = \{ax^2 + bx + c | a, b, c \in \mathbb{R}\}$$

Un primer producto interno es

$$\langle \mathbf{p}, \mathbf{q} \rangle = \sum_{i=0}^2 p(i) q(i)$$

bajo el cual se obtiene la siguiente norma:

$$\begin{aligned}\langle \mathbf{p}, \mathbf{p} \rangle &= \langle -2x^2 + x - 1, -2x^2 + x - 1 \rangle \\ &= (-1)(-1) + (-2)(-2) + (-8 + 2 - 1)(-8 + 2 - 1) \\ &= 1 + 4 + 49 \Rightarrow \|\mathbf{p}\| = \sqrt{54}\end{aligned}$$

Mientras tanto, con el producto interno

$$\langle \mathbf{p}, \mathbf{q} \rangle = \sum_{i=0}^2 a_i b_i$$

la norma es:

$$\begin{aligned}\langle \mathbf{p}, \mathbf{p} \rangle &= \langle -2x^2 + x - 1, -2x^2 + x - 1 \rangle \\ &= (-2)(-2) + (1)(1) + (-1)(-1) \\ &= 4 + 1 + 1 \Rightarrow \|\mathbf{p}\| = \sqrt{6}\end{aligned}$$

Y con el tercer producto interno

$$\langle \mathbf{p}, \mathbf{q} \rangle = \int_0^1 p(x) q(x) dx$$

la norma también cambia, y se calcula como:

$$\begin{aligned}\langle \mathbf{p}, \mathbf{p} \rangle &= \langle -2x^2 + x - 1, -2x^2 + x - 1 \rangle \\ &= \int_0^1 (-2x^2 + x - 1)(-2x^2 + x - 1) dx \\ &= \int_0^1 (4x^4 - 4x^3 + 5x^2 - 2x + 1) dx \\ &= \left[ \frac{4}{5}x^5 - x^4 + \frac{5}{3}x^3 - x^2 + x \right]_0^1 \\ &= \frac{4}{5} - 1 + \frac{5}{3} - 1 + 1 \Rightarrow \|\mathbf{p}\| = \sqrt{\frac{22}{15}}\end{aligned}$$

## 2. Vector Unitario

Dentro de los vectores, aquéllos con norma unitaria son especialmente importantes; permiten definir al operador lineal adjunto, definir la base canónica de los espacios  $\mathbb{R}^n$  o  $\mathbb{C}^n$ , y la llamada base ortonormal.

Sea un espacio vectorial  $V$  sobre el campo  $K$ , donde se define un producto interno. Un vector unitario tiene norma igual a uno:

$$\|\hat{\mathbf{v}}\| = 1$$

Para encontrar un vector unitario simplemente se divide cualquier vector entre su norma.

$$\hat{\mathbf{v}} = \frac{1}{\|\mathbf{v}\|} \mathbf{v}$$

Cuando se obtiene un vector unitario  $\|\hat{\mathbf{v}}\|$  a partir de un vector  $\mathbf{v}$  cualquiera, se dice que éste último fue normalizado.

La normalización de vectores permite definir los vectores principales en los sistema de coordenadas; por ejemplo, la base canónica  $B = \{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$  contiene los vectores fundamentales del espacio  $\mathbb{R}^3$ .

**Ejemplo.** Bajo el producto interno

$$\langle A, B \rangle \text{ tr } (A^T B)$$

Normalice al vector  $\mathbf{m} = \begin{bmatrix} -2 & 3 \\ 4 & 0 \end{bmatrix}$ .

Calculando la norma del vector:

$$\left\langle \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix} \right\rangle = \text{tr} \begin{bmatrix} 4+9 & -8 \\ -8 & 16 \end{bmatrix} \\ = 13 + 16$$

Por lo que el vector unitario es

$$\hat{\mathbf{m}} = \frac{1}{\sqrt{29}} \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix}$$

La normalización de vectores muestra una geometría interesante. Los vectores unitarios bajo el producto punto forman una circunferencia:

$$\begin{bmatrix} x \\ y \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 1 \quad \Rightarrow \quad x^2 + y^2 = 1 \quad (1)$$

En cambio, bajo un producto ponderado la circunferencia se vería deformada bajo la métrica usual:

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 1 \quad \Rightarrow \quad \frac{x^2}{2} + y^2 = 1 \quad (2)$$

La figura 2 muestra la geometría de ambos productos internos.

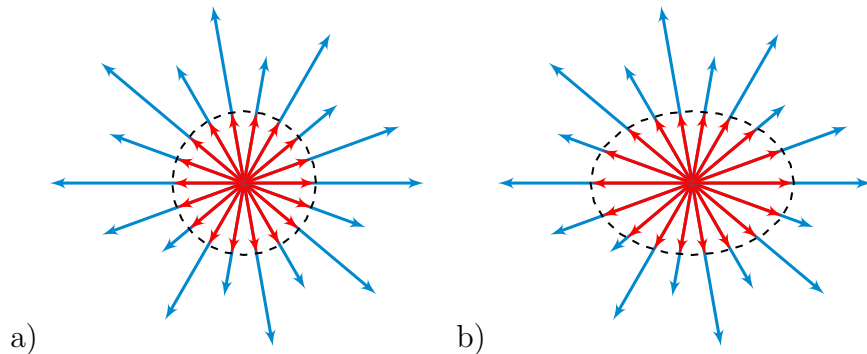


Figura 2. a) Circunferencia unitaria bajo el producto punto (1). b) Circunferencia unitaria bajo el producto interno ponderado (2).

# Álgebra Lineal

## Lectura 23: Distancia y Ángulo entre Vectores

Ing. Aldo Jiménez Arteaga

Enero 2020, Rev. Agosto 2021

### 1. Distancia entre Vectores

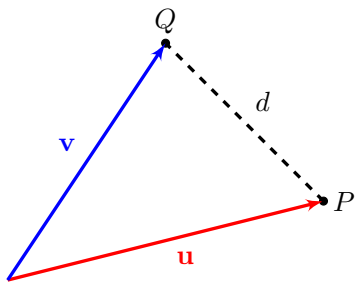


Figura 1. La distancia entre dos puntos implica geometría entre dos vectores.

Una vez que se estableció la métrica primordial (la norma) en el espacio vectorial, pueden calcularse distancias entre los elementos de todo el espacio. Mediante la interpretación geométrica de distancia (véase la figura 1) entre los puntos  $P(p_1, p_2)$  y  $Q(q_1, q_2)$ , se puede extender el concepto a los espacios vectoriales con producto interno:

$$d = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2}$$

que puede reescribirse como

$$d = \sqrt{(P - Q) \cdot (P - Q)}$$

y que incluye el módulo y producto escalar ordinarios.

Considerando que los puntos pueden sustituirse por sus vectores de

posición, se involucran los conceptos de resta y norma para calcular la distancia entre dos elementos de un espacio vectorial.

La distancia entre los vectores  $\mathbf{u}$  y  $\mathbf{v}$  cualesquiera se calcula como

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

La distancia siempre será la misma, sin importar el orden de la resta.

Al ser una norma, la distancia mantiene las mismas propiedades.

**Ejemplo.** En el espacio  $\mathbb{C}^2$  el producto punto complejo es

$$\left\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right\rangle = u_1 \bar{v}_1 + u_2 \bar{v}_2$$

donde  $\bar{v}_1$  y  $\bar{v}_2$  son los conjugados de  $v_1$  y  $v_2$ , respectivamente. Calcule la distancia entre los vectores  $\mathbf{x} = \begin{bmatrix} 3 + i \\ -2i \end{bmatrix}$  y  $\mathbf{y} = \begin{bmatrix} -2 + 5i \\ -6 \end{bmatrix}$ .

Al restar los vectores:

$$\begin{aligned} \mathbf{x} - \mathbf{y} &= \begin{bmatrix} 3 + i \\ -2i \end{bmatrix} - \begin{bmatrix} -2 + 5i \\ -6 \end{bmatrix} \\ &= \begin{bmatrix} 5 - 4i \\ 6 - 2i \end{bmatrix} \end{aligned}$$

Para calcular la norma:

$$\begin{aligned} \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle &= (5 - 4i)(5 + 4i) + (6 - 2i)(6 + 2i) \\ &= 25 + 16 + 36 + 4 \\ &= 81 \end{aligned}$$

Por lo que la distancia entre los vectores involucrados es  $d(\mathbf{x}, \mathbf{y}) = 9$ .

## 2. Ángulo entre Vectores

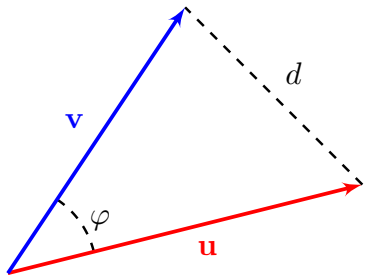


Figura 2. El ángulo  $\varphi$  está formado por lados conocidos, por lo que induce la ley del coseno.

Una vez que se ha establecido la magnitud, la dirección entre vectores considerará el menor ángulo que se forma con la referencia utilizada. Cuando son dos vectores cualesquiera, uno será la referencia mutua del otro. Si se considera la distancia entre los vectores involucrados, se obtiene la ley del coseno desde la cual se puede inducir la expresión para calcular el ángulo entre vectores, como se muestra en la figura 2. El cálculo del ángulo incluye la ley del

coseno en términos de normas y propiedades del producto interno:

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\|\cos\varphi \quad (1)$$

Al desarrollar la parte izquierda de (1), por la propiedad de linealidad, se obtiene

$$\begin{aligned} \|\mathbf{u} - \mathbf{v}\|^2 &= \langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} - \mathbf{v} \rangle - \langle \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} - \mathbf{v} \rangle - \langle \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle - \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \end{aligned} \quad (2)$$

Al aplicar la simetría y la positividad a (2), se obtiene la expresión

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 - 2\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 \quad (3)$$

Al sustituir (3) en (1) puede despejarse el coseno del ángulo:

$$\begin{aligned} \|\mathbf{u} - \mathbf{v}\|^2 &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\|\cos\varphi \\ \|\mathbf{u}\|^2 - 2\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\|\cos\varphi \\ -2\langle \mathbf{u}, \mathbf{v} \rangle &= -2\|\mathbf{u}\|\|\mathbf{v}\|\cos\varphi \\ \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|} &= \cos\varphi \end{aligned}$$

El ángulo entre dos vectores  $\mathbf{u}$  y  $\mathbf{v}$  cualesquiera se calcula como

$$\angle(\mathbf{u}, \mathbf{v}) = \frac{\text{Re}\{\langle \mathbf{u}, \mathbf{v} \rangle\}}{\|\mathbf{u}\|\|\mathbf{v}\|}$$

donde  $\text{Re}\{*\}$  denota la parte real del producto interno, en caso que sea un espacio vectorial complejo.

En espacios complejos el producto interno entre dos vectores puede resultar en un escalar complejo, el cual no se encuentra en el intervalo  $-1 \leq \cos\varphi \leq 1$  que son los valores que el coseno puede tomar. Por esta razón se especifica que para el ángulo en vectores complejos se descarta la parte imaginaria.



**Ejemplo.** Calcule el ángulo entre las funciones  $f(x) = \sin x + \cos x$  y  $g(x) = 2 \cos x$ , bajo el producto interno

$$\langle f(x), g(x) \rangle = \int_0^\pi f(x) g(x) dx$$

Se calculan las normas de manera independiente:

$$\begin{aligned} \|f(x)\|^2 &= \int_0^\pi (\sin x + \cos x)^2 dx \\ &= [x + \sin^2 x]_0^\pi \\ &= \pi + 0 - 0 - 0 \\ &= \pi \end{aligned}$$

$$\begin{aligned} \|g(x)\|^2 &= \int_0^\pi 4 \cos^2 x dx \\ &= [2x + \sin 2x]_0^\pi \\ &= 2\pi + 0 - 0 - 0 \\ &= 2\pi \end{aligned}$$

Ahora el producto interno entre las funciones:

$$\begin{aligned} \langle f(x), g(x) \rangle &= \int_0^\pi 2 \sin x \cos x + 2 \cos^2 x dx \\ &= \left[ \sin^2 x + x + \frac{1}{2} \sin 2x \right]_0^\pi \\ &= 0 + \pi + 0 - 0 - 0 - 0 \\ &= \pi \end{aligned}$$

Por lo que el ángulo entre las funciones es

$$\cos \varphi = \frac{\pi}{\sqrt{\pi} \sqrt{2\pi}} \Rightarrow \varphi = \arccos \frac{1}{\sqrt{2}}$$

Finalmente, el ángulo entre las funciones es:

$$\sphericalangle (f, g) = 45^\circ$$

# Álgebra Lineal

## Lectura 24: Ortogonalidad

Ing. Aldo Jiménez Arteaga

Enero 2020

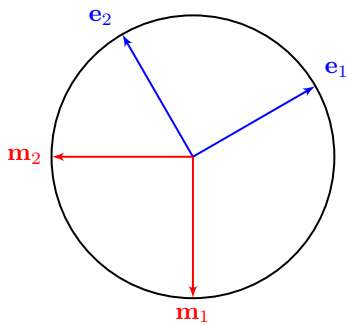


Figura 1. Los sistemas de coordenadas rectangulares son una referencia ortogonal, aunque no son los únicos con esta característica.

Los sistemas coordenados rectangulares (véase la figura 1) poseen dos características trascendentales que permiten manejarlos de una manera muy eficiente en términos del Álgebra Lineal: la ortogonalidad y la normalización. La segunda se logra al obtener vectores unitarios, en tanto que la primera se logra con la condición de ortogonalidad:

$$\cos 90^\circ = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|}$$
$$0 = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|}$$

En un espacio vectorial  $V$  con producto interno definido, la condición de ortogonalidad entre vectores es

$$\langle \mathbf{u}, \mathbf{v} \rangle = 0$$

Los sistemas cartesianos, esféricos y cilíndricos son ortogonales.

### 1. Conjunto Ortogonal

La ortogonalidad no es específica de un par de vectores; se pueden tener conjuntos completos de vectores ortogonales a un vector, o incluso conjuntos de vectores ortogonales entre sí. Por ejemplo la base canónica de  $\mathbb{R}^3$ , en la cual sus vectores son ortogonales entre sí.

Sea  $V$  un espacio vectorial de dimensión  $n$  con producto interno definido. El conjunto  $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$  es ortogonal si

$$\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0 \quad \forall \mathbf{b}_i \neq \mathbf{b}_j$$

Si además el conjunto ortogonal contiene solo vectores unitarios, entonces el conjunto es ortonormal.

**Ejemplo.** Obtenga un conjunto ortogonal del espacio vectorial real  $P = \{ax^2 + bx + c \mid a, b, c \in \mathbb{R}\}$ , bajo el producto interno

$$\langle p(x), q(x) \rangle = \sum_{i=-1}^1 p(i) q(i)$$

Para encontrar este conjunto se debe partir de un vector dado o propuesto. Se partirá de un vector arbitrario:  $p_1 = x^2$ . A partir de este vector se aplicará sucesivamente la condición de ortogonalidad con un vector genérico  $p = ax^2 + bx + c$ .

$$\begin{aligned} \langle x^2, ax^2 + bx + c \rangle &= 0 \\ 0 &= (1)(a - b + c) + (0)(c) + (1)(a + b + c) \\ 0 &= 2a + 2c \Rightarrow c = -a \end{aligned}$$

Por lo que el conjunto de todos los vectores ortogonales a  $p_1$  tiene la forma  $p = ax^2 + bx - a$ . Dando valores aleatorios a  $a$  y  $b$ , el segundo vector ortogonal es  $p_2 = x$ . Para obtener el siguiente vector, se repite el proceso utilizando conjuntamente a  $p_1$  y  $p_2$ .

$$\begin{aligned} \langle x^2, ax^2 + bx + c \rangle &= 0 \\ 0 &= 2a + 2c \Rightarrow a = -c \end{aligned} \tag{1}$$

$$\begin{aligned} \langle x, ax^2 + bx + c \rangle &= 0 \\ 0 &= 2b \Rightarrow b = 0 \end{aligned} \tag{2}$$

Por lo que el tercer vector ortogonal tiene la forma  $p = -cx^2 + c$ . Seleccionando un valor aleatorio para  $c$ , el vector es  $p_3 = -x^2 + 1$ . Finalmente, el conjunto ortogonal pedido es

$$B_{\perp} = \{x^2, x, -x^2 + 1\}$$

## 2. Base Ortogonal

La ortogonalidad es una característica adicional al sistema de referencia, y por lo tanto puede caracterizar a una base. Considerando el conjunto ortogonal  $B_{\perp} = \{e_1, e_2, \dots, e_i, \dots, e_n\}$ , se aplica el producto interno a la ecuación de dependencia lineal con cualquiera de los

vectores del conjunto:

$$\begin{aligned} \mathbf{0} &= \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_i e_i + \dots + \alpha_n e_n \\ \langle \mathbf{0}, e_i \rangle &= \langle \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_i e_i + \dots + \alpha_n e_n, e_i \rangle \end{aligned}$$

Por linealidad del producto interno:

$$\langle \mathbf{0}, e_i \rangle = \alpha_1 \langle e_1, e_i \rangle + \alpha_2 \langle e_2, e_i \rangle + \dots + \alpha_i \langle e_i, e_i \rangle + \dots + \alpha_n \langle e_n, e_i \rangle$$

Al ser un conjunto ortogonal, todo producto interno entre vectores diferentes es igual a cero en tanto que el producto interno con el mismo vector es positivo:

$$\begin{aligned} 0 &= \alpha_1 (0) + \alpha_2 (0) + \dots + \alpha_i \langle e_i, e_i \rangle + \dots + \alpha_n (0) \\ 0 &= \alpha_i \langle e_i, e_i \rangle \end{aligned}$$

La única solución es  $\alpha_i = 0$ . Si se realiza el mismo proceso para todos los vectores se obtendrá la solución trivial que indica un conjunto linealmente independiente; por lo tanto, la ortogonalidad define independencia lineal.

Si ahora se plantea una combinación lineal con un vector cualquiera  $\mathbf{v}$  y se aplica el producto interno, cada escalar  $\alpha_i$  puede calcularse como

$$\alpha_i = \frac{\langle \mathbf{v}, e_i \rangle}{\langle e_i, e_i \rangle}$$

que es la expresión conocida como la proyección de un vector sobre otro, o componente vectorial. Esto quiere decir que bajo cualquier producto interno, toda combinación lineal es la suma de componentes vectoriales sobre una base ortogonal, tal como se muestra en la figura 2.

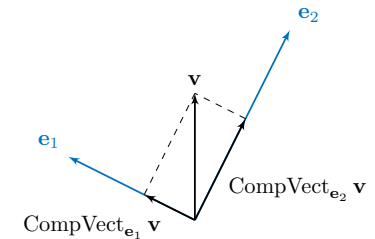


Figura 2. Una combinación lineal con componentes vectoriales solo es posible con bases ortogonales.

**Ejemplo.** Obtenga una base ortonormal  $B$  del espacio vectorial real  $M = \left\{ \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \mid a, b, c \in \mathbb{R} \right\}$ , bajo el producto interno

$$\langle A, B \rangle = \text{tr}(A^T B) \quad \forall A, B \in M$$

y calcule el vector de coordenadas de  $G = \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}$  en  $B$ .

Se partirá del vector  $E_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$  para encontrar el resto de vectores de la base. La normalización de vectores se realizará al final del proceso. El segundo vector de la base debe ser  $E_2 = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$ , por lo tanto

$$\begin{aligned} \langle E_1, E_2 \rangle &= \text{tr} \left( \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^T \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \right) \\ &= \text{tr} \left( \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \right) \\ &= \text{tr} \begin{bmatrix} a & b \\ a & b+c \end{bmatrix} \\ 0 &= a + b + c \end{aligned}$$

El vector buscado es  $E_2 = \begin{bmatrix} a & b \\ 0 & -a-b \end{bmatrix}$ . Al seleccionar los valores  $a = 1$  y  $b = -1$  el vector que formará parte de la base es  $E_2 = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$ . Ahora se buscará el tercer vector  $E_3 = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$ .

Podrá observarse que este proceso reutilizará ecuaciones planteadas en el paso anterior. En este paso la ortogonalidad para  $E_3$  se plantea con  $E_1$  y simultáneamente con  $E_2$ .

$$\langle E_1, E_3 \rangle = \text{tr} \left( \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^T \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \right)$$

$$0 = a + b + c$$

$$\langle E_2, E_3 \rangle = \text{tr} \left( \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}^T \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \right)$$

$$0 = a - b$$

Al resolver el sistema de ecuaciones

$$\begin{cases} 0 = a + b + c \\ 0 = a - b \end{cases} \Rightarrow \begin{cases} b = a \\ c = -2a \end{cases}$$

Al asignar  $a = 1$ , el tercer vector es  $E_3 = \begin{bmatrix} 1 & 1 \\ 0 & -2 \end{bmatrix}$ . La base ortogonal está completa, por lo que la norma de cada vector es

$$\langle E_1, E_1 \rangle = \text{tr} \left( \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right)$$

$$= 3 \Rightarrow \|E_1\| = \sqrt{3}$$

$$\langle E_2, E_2 \rangle = \text{tr} \left( \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}^T \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \right)$$

$$= 2 \Rightarrow \|E_2\| = \sqrt{2}$$

$$\langle E_3, E_3 \rangle = \text{tr} \left( \begin{bmatrix} 1 & 1 \\ 0 & -2 \end{bmatrix}^T \begin{bmatrix} 1 & 1 \\ 0 & -2 \end{bmatrix} \right)$$

$$= 6 \Rightarrow \|E_3\| = \sqrt{6}$$

$$\therefore B = \left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & 1 \\ 0 & -2 \end{bmatrix} \right\}$$

Para la combinación lineal

$$\begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix} = \alpha_1 \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} + \alpha_2 \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + \alpha_3 \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & 1 \\ 0 & -2 \end{bmatrix}$$

se utilizará el cálculo de escalares mediante el producto interno:

$$\alpha_i = \frac{\langle G, \hat{E}_i \rangle}{\langle \hat{E}_i, \hat{E}_i \rangle}$$

Como se está trabajando con una base ortonormal, el producto  $\langle \hat{E}_i, \hat{E}_i \rangle$  es igual a 1. Por lo tanto los escalares se reducen a calcular el producto interno entre  $G$  y cada vector de la base.

$$\begin{aligned} \alpha_1 &= \langle G, \hat{E}_1 \rangle \\ &= \text{tr} \left( \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}^T \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right) \\ &= \frac{2}{\sqrt{3}} \\ \alpha_2 &= \langle G, \hat{E}_2 \rangle \\ &= \text{tr} \left( \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}^T \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \right) \\ &= \frac{1}{\sqrt{2}} \\ \alpha_3 &= \langle G, \hat{E}_3 \rangle \\ &= \text{tr} \left( \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}^T \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & 1 \\ 0 & -2 \end{bmatrix} \right) \\ &= \frac{5}{\sqrt{6}} \end{aligned}$$

Finalmente, el vector de coordenadas es

$$[G]_B = \frac{1}{\sqrt{6}} \begin{bmatrix} 2\sqrt{2} \\ \sqrt{3} \\ 5 \end{bmatrix}$$

y la expresión por combinación lineal es

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

# Álgebra Lineal

## Lectura 25: Complemento Ortogonal

Ing. Aldo Jiménez Arteaga

Enero 2020

### 1. Complemento Ortogonal

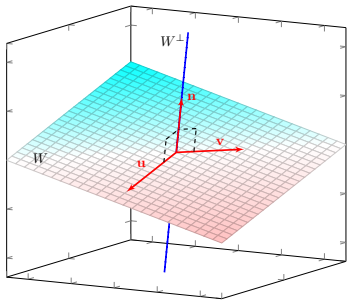


Figura 1. El subespacio  $W$  (plano) tiene a su complemento ortogonal  $W^\perp$  (recta) en el espacio vectorial  $\mathbb{R}^3$  bajo el producto escalar ordinario.

Los conjuntos ortogonales son bases dentro del espacio vectorial. Si se considera ahora la ortogonalidad entre conjuntos, se obtienen no solo unos cuantos vectores sino subespacios completos.

Considerando un plano  $W$  que contiene al origen, la ecuación cartesiana que lo describe algebraicamente se obtiene a partir de su vector normal  $\mathbf{n}$ :

$$\mathbf{n} \cdot \mathbf{w} = 0 \quad (1)$$

donde  $\mathbf{w} \in W$ . Si se considera a ese vector normal como el vector director de la recta, que contiene al origen,  $W^\perp$  cuya ecuación es

$$\mathbf{w}' = t\mathbf{n}$$

entonces ambos subespacios vectoriales son ortogonales entre sí. Los lugares geométricos descritos tienen representación como

$$W = \{\mathbf{w} | \mathbf{n} \cdot \mathbf{w} = 0\} \quad (2)$$

$$W^\perp = \{\mathbf{w}' | \mathbf{w}' = t\mathbf{n}\} \quad (3)$$

Si se aplica el producto escalar ordinario entre los vectores genéricos de (2) y (3), se comprueba que son ortogonales por la ecuación cartesiana (1):

$$\mathbf{w} \cdot \mathbf{w}' = 0$$

$$\mathbf{w} \cdot t\mathbf{n} =$$

$$t(\mathbf{w} \cdot \mathbf{n}) = 0$$

A (2) y (3) se les conoce como complementos ortogonales. La figura 1 muestra la geometría de un subespacio y su complemento ortogonal.

Sean  $V$  un espacio vectorial con producto interno dado, y  $W$  un subespacio de  $V$ . El subespacio  $W^\perp$  se conoce como complemento ortogonal de  $W$  siempre y cuando

$$\langle \mathbf{w}, \mathbf{w}' \rangle = 0 \quad \forall \mathbf{w} \in W, \mathbf{w}' \in W^\perp$$

Estos subespacios se llaman complementos, porque la suma de la dimensión de un subespacio  $W$  más la de su complemento ortogonal  $W^\perp$  devuelven la dimensión del espacio vectorial *padre*:

$$\dim W + \dim W^\perp = \dim V$$

**Ejemplo.** Obtenga el complemento ortogonal de

$$W = \left\{ \left[ \begin{array}{c} x \\ y \\ 2ix - 3y \end{array} \right] \mid x, y \in \mathbb{C} \right\}$$

bajo el producto interno usual en  $\mathbb{C}^3$ .

Para determinar el complemento se designa una base, la cual *representará* a todos los vectores del subespacio.

$$A = \left\{ \left[ \begin{array}{c} 1 \\ 0 \\ 2i \end{array} \right], \left[ \begin{array}{c} 0 \\ 1 \\ -3 \end{array} \right] \right\}$$

Esta base puede o no ser ortogonal. Aplicando el producto interno bajo la condición de ortogonalidad

$$\begin{aligned} \left[ \begin{array}{c} x \\ y \\ z \end{array} \right] \cdot \left[ \begin{array}{c} 1 \\ 0 \\ 2i \end{array} \right] &= 0 \\ x\overline{(1)} + y\overline{(0)} + z\overline{(2i)} &= \\ x - 2iz &= 0 \end{aligned} \tag{4}$$

Como el subespacio  $W$  es de dimensión 2, su complemento ortogonal debe ser de dimensión 1; la ecuación (4) es una de sus condiciones.

$$\begin{aligned} \left[ \begin{array}{c} x \\ y \\ z \end{array} \right] \cdot \left[ \begin{array}{c} 0 \\ 1 \\ -3 \end{array} \right] &= 0 \\ x\overline{(0)} + y\overline{(1)} + z\overline{(-3)} &= \\ y - 3z &= 0 \end{aligned} \tag{5}$$

Al resolver simultáneamente las ecuaciones (4) y (5) se llega a  $x = 2iz$  y  $y = 3z$ , por lo que el complemento ortogonal de  $W$  es

$$W^\perp = \left\{ \left[ \begin{array}{c} 2iz \\ 3z \\ z \end{array} \right] \mid z \in \mathbb{C} \right\}$$

## 2. Teorema de Pitágoras

Cada vector  $\mathbf{u}$  de un espacio vectorial  $V$  tiene una única representación en términos de un vector  $\mathbf{w}$  de un subespacio  $W$  y un vector  $\mathbf{w}'$  del complemento ortogonal  $W^\perp$ :

$$\mathbf{u} = \mathbf{w} + \mathbf{w}' \tag{6}$$

Geoméricamente, la suma algebraica (6) denota un triángulo rectángulo cuyos catetos son  $\mathbf{w}$  y  $\mathbf{w}'$ , en tanto que la hipotenusa es el vector  $\mathbf{u}$ , tal como se muestra en la figura 2. Al ser un triángulo rectángulo puede plantearse el teorema de Pitágoras, donde las longitudes de los lados que

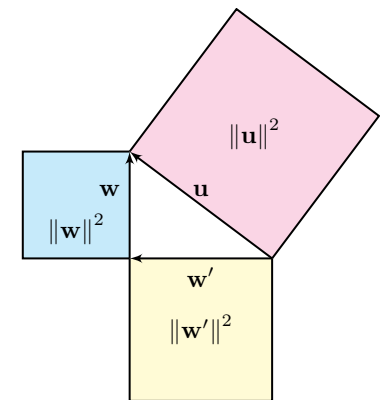


Figura 2. Triángulo rectángulo como suma vectorial. El teorema de Pitágoras se cumple bajo cualquier producto interno.

componen al triángulo son las normas de los vectores involucrados:

$$\|\mathbf{u}\|^2 = \|\mathbf{w}\|^2 + \|\mathbf{w}'\|^2$$

Sin importar el producto interno o el espacio vectorial trabajado, el teorema de Pitágoras en términos de normas siempre se satisface. Debido a esta propiedad, a los espacios vectoriales que poseen un producto interno se les llama euclídeos pues conservan los postulados y teoremas de la Geometría Euclideana.

**Ejemplo.** Verifique que se cumple el teorema de Pitágoras para la matriz  $M = \begin{bmatrix} 1 & 4 \\ -2 & -3 \end{bmatrix}$ , tomando en cuenta el subespacio vectorial  $S = \left\{ \begin{bmatrix} a & b \\ b & d \end{bmatrix} \mid a, b, d \in \mathbb{R} \right\}$  y su complemento ortogonal  $S^\perp = \left\{ \begin{bmatrix} 0 & c \\ -c & 0 \end{bmatrix} \mid c \in \mathbb{R} \right\}$ , bajo el producto interno

$$\langle A, B \rangle = \text{tr } A^T B$$

Primero se determinan los vectores catetos mediante la suma

$$\begin{bmatrix} 1 & 4 \\ -2 & -3 \end{bmatrix} = \begin{bmatrix} a & b \\ b & d \end{bmatrix} + \begin{bmatrix} 0 & c \\ -c & 0 \end{bmatrix}$$

la cual genera el sistema de ecuaciones

$$\begin{aligned} a &= 1 \\ b + c &= 4 \\ b - c &= -2 \\ d &= -3 \end{aligned}$$

cuya solución es  $a = 1$ ,  $b = 1$ ,  $c = 3$  y  $d = -3$ . Por lo tanto, los vectores catetos son

$$W = \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix} \in S \quad W' = \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix} \in S^\perp$$

Ahora se calculan las normas de cada matriz.

$$\begin{aligned} \left\langle \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix} \right\rangle &= \text{tr} \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix} \\ &= 1 + 1 + 1 + 9 \\ &= 12 \quad \Rightarrow \quad \|W\| = \sqrt{12} \end{aligned}$$

$$\begin{aligned} \left\langle \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix} \right\rangle &= \text{tr} \begin{bmatrix} 0 & -3 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix} \\ &= 9 + 9 \\ &= 18 \quad \Rightarrow \quad \|W'\| = \sqrt{18} \end{aligned}$$

$$\begin{aligned} \left\langle \begin{bmatrix} 1 & 4 \\ -2 & -3 \end{bmatrix}, \begin{bmatrix} 1 & 4 \\ -2 & -3 \end{bmatrix} \right\rangle &= \text{tr} \begin{bmatrix} 1 & -2 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ -2 & -3 \end{bmatrix} \\ &= 1 + 4 + 16 + 9 \\ &= 30 \quad \Rightarrow \quad \|M\| = \sqrt{30} \end{aligned}$$

Comprobando el teorema:

$$M = W + W'$$

$$\begin{bmatrix} 1 & 4 \\ -2 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix} + \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix}$$

y en términos de normas

$$\begin{aligned} \|M\|^2 &= \|W\|^2 + \|W'\|^2 \\ 30 &= 12 + 18 \end{aligned}$$



# Álgebra Lineal

## Lectura 26: Proyección Ortogonal

Ing. Aldo Jiménez Arteaga

Enero 2020

### 1. Proyección

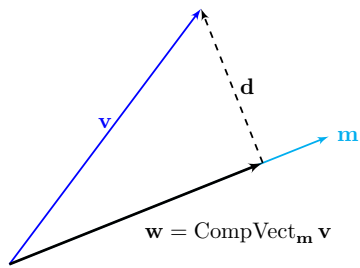


Figura 1. La componente vectorial es la proyección de un vector sobre otro.

La componente vectorial es una herramienta en Geometría Analítica que permite resolver varios problemas como distancias entre lugares geométricos, simétricos o proyecciones sobre un plano o recta. Para obtener la componente vectorial se requiere un vector sobre el cuál proyectar (figura 1). Bajo el producto punto, la componente vectorial es

$$\text{CompVect}_{\mathbf{m}} \mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{m}}{|\mathbf{m}|} \hat{\mathbf{m}}$$

y en términos de producto interno es

$$\text{CompVect}_{\mathbf{m}} \mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{m} \rangle}{\langle \mathbf{m}, \mathbf{m} \rangle} \mathbf{m} \Rightarrow \mathbf{w} \quad (1)$$

La expresión (1) es útil al proyectar sobre un vector. Pero, si extendemos esta definición a la proyección sobre un subespacio, (1) solo

funcionaría con un subespacio de dimensión 1. Para definir la proyección de  $\mathbf{v} \notin W$  en un subespacio  $W$  de mayor dimensión se requiere una base ortogonal  $B = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ . Tomando a  $\mathbf{w} \in W$  como la proyección de  $\mathbf{v}$  sobre  $W$ , su combinación lineal en la base  $B$  es

$$\mathbf{w} = \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \dots + \alpha_n \mathbf{e}_n \quad (2)$$

Al aplicar el producto interno, por ejemplo con el vector  $\mathbf{e}_n$ , a ambos lados de (2), mediante la linealidad y ortogonalidad se obtiene

$$\begin{aligned} \mathbf{w} &= \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \dots + \alpha_n \mathbf{e}_n \\ \langle \mathbf{w}, \mathbf{e}_n \rangle &= \alpha_1 \langle \mathbf{e}_1, \mathbf{e}_n \rangle + \alpha_2 \langle \mathbf{e}_2, \mathbf{e}_n \rangle + \dots + \alpha_n \langle \mathbf{e}_n, \mathbf{e}_n \rangle \\ \langle \mathbf{w}, \mathbf{e}_n \rangle &= \alpha_n \langle \mathbf{e}_n, \mathbf{e}_n \rangle \end{aligned}$$

Por lo que en una base ortogonal los escalares de (2) son

$$\alpha = \frac{\langle \mathbf{w}, \mathbf{e} \rangle}{\langle \mathbf{e}, \mathbf{e} \rangle}$$

La combinación lineal (2) se expresa como

$$\mathbf{w} = \frac{\langle \mathbf{w}, \mathbf{e}_1 \rangle}{\langle \mathbf{e}_1, \mathbf{e}_1 \rangle} \mathbf{e}_1 + \frac{\langle \mathbf{w}, \mathbf{e}_2 \rangle}{\langle \mathbf{e}_2, \mathbf{e}_2 \rangle} \mathbf{e}_2 + \dots + \frac{\langle \mathbf{w}, \mathbf{e}_n \rangle}{\langle \mathbf{e}_n, \mathbf{e}_n \rangle} \mathbf{e}_n \quad (3)$$

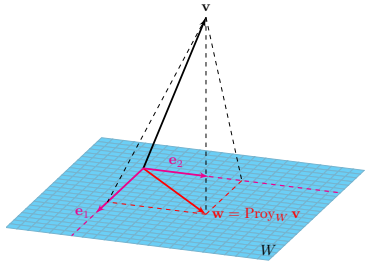


Figura 2. La proyección de un vector  $\mathbf{v}$  sobre un subespacio  $W$  se da a partir de una base ortogonal. Se muestra a  $W$  como un subespacio de dimensión 2.

Como se trata de una distancia, es la norma del vector distancia  $\mathbf{d} = \mathbf{v} - \text{Proj}_W \mathbf{v}$ . Este es el teorema de proyección.

$$\mathbf{w} = \frac{\langle \mathbf{v} - \mathbf{d}, \mathbf{e}_1 \rangle}{\langle \mathbf{e}_1, \mathbf{e}_1 \rangle} \mathbf{e}_1 + \frac{\langle \mathbf{v} - \mathbf{d}, \mathbf{e}_2 \rangle}{\langle \mathbf{e}_2, \mathbf{e}_2 \rangle} \mathbf{e}_2 + \dots + \frac{\langle \mathbf{v} - \mathbf{d}, \mathbf{e}_n \rangle}{\langle \mathbf{e}_n, \mathbf{e}_n \rangle} \mathbf{e}_n$$

$$\mathbf{w} = \frac{\langle \mathbf{v}, \mathbf{e}_1 \rangle}{\langle \mathbf{e}_1, \mathbf{e}_1 \rangle} \mathbf{e}_1 + \frac{\langle \mathbf{v}, \mathbf{e}_2 \rangle}{\langle \mathbf{e}_2, \mathbf{e}_2 \rangle} \mathbf{e}_2 + \dots + \frac{\langle \mathbf{v}, \mathbf{e}_n \rangle}{\langle \mathbf{e}_n, \mathbf{e}_n \rangle} \mathbf{e}_n$$

Esta es la proyección del vector  $\mathbf{v}$  en el subespacio  $W$ . La figura 2 muestra la geometría de la proyección.

Sea  $V$  un espacio vectorial con producto interno definido. Si  $B = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  es una base ortogonal del subespacio  $W$ , entonces la proyección de un vector  $\mathbf{v} \in V$  es

$$\text{Proj}_W \mathbf{v} = \sum_{i=1}^n \frac{\langle \mathbf{v}, \mathbf{e}_i \rangle}{\langle \mathbf{e}_i, \mathbf{e}_i \rangle} \mathbf{e}_i$$

La proyección de un vector  $\mathbf{v}$  sobre un subespacio es como una *sombra*: si se ilumina un vector por encima, en el subespacio se verá una sombra que es la proyección.

Obsérvese que cada sumando de (3) es la expresión de una componente vectorial similar a (1). Como  $\mathbf{w}$  es la proyección de  $\mathbf{v}$  sobre  $W$ , se puede aplicar el teorema de Pitágoras en forma vectorial, tal como lo muestra la figura 1:

$$\mathbf{w} + \mathbf{d} = \mathbf{v}$$

$$\mathbf{w} = \mathbf{v} - \mathbf{d} \quad (4)$$

Aquí se sabe que  $\langle \mathbf{w}, \mathbf{d} \rangle = 0$ . Como  $\mathbf{d}$  pertenece al complemento ortogonal de  $W$ , entonces su producto interno con cada vector de la base  $B$  es nulo. Al sustituir (4) en los escalares de (3) se tiene que

Un uso importante de la proyección es la minimización de la distancia. Como el vector  $\mathbf{v}$  se encuentra fuera del subespacio  $W$ , entonces debe existir una distancia entre los dos. Al hablar de una distancia se hace referencia a la mínima, ya que se pueden medir varias longitudes entre un vector  $\mathbf{v}$  y cualquier elemento de  $W$ , tal como se presenta en la figura 3.

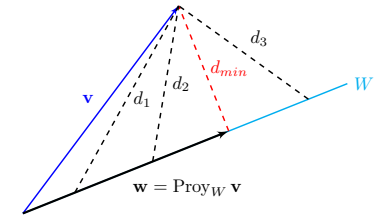


Figura 3. La mínima distancia entre un subespacio y su proyección es mínima.

Como se trata de una distancia, es la norma del vector distancia  $\mathbf{d} = \mathbf{v} - \text{Proj}_W \mathbf{v}$ . Este es el teorema de proyección.

La mínima distancia entre un vector  $\mathbf{v}$  y un subespacio  $W$  es

$$d_{min}(\mathbf{v}, W) = \|\mathbf{v} - \text{Proj}_W \mathbf{v}\|$$

Como el vector distancia es ortogonal al subespacio  $W$  (figura 3), entonces pertenece al complemento ortogonal  $W^\perp$ . Debido al teorema de proyección la mínima distancia entre  $\mathbf{v}$  y el complemento ortogonal es

$$d_{min}(\mathbf{v}, W^\perp) = \|\mathbf{v} - \text{Proj}_{W^\perp} \mathbf{v}\|$$

La figura 4 muestra que las distancias mínimas de  $\mathbf{v}$  y los subespacios  $W$  y  $W^\perp$  forman los catetos del triángulo rectángulo formado por  $\mathbf{v}$ ,  $\mathbf{w}$  y  $\mathbf{w}'$ . Esto permite una nueva representación del teorema de Pitágoras en términos de proyecciones:

$$\mathbf{v} = \mathbf{w} + \mathbf{w}'$$

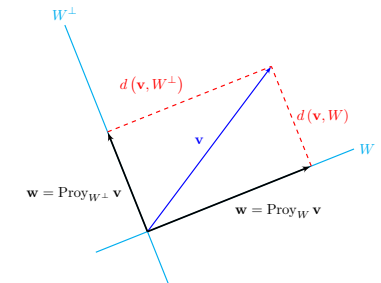


Figura 4. El teorema de Pitágoras puede expresarse en términos de proyecciones sobre los subespacios  $W$  y  $W^\perp$ .

$$\mathbf{v} = \text{Proy}_W \mathbf{v} + \text{Proy}_{W^\perp} \mathbf{v}$$

$$\begin{aligned} \therefore \|\mathbf{v}\|^2 &= \|\text{Proy}_W \mathbf{v}\|^2 + \|\text{Proy}_{W^\perp} \mathbf{v}\|^2 \\ &= \|\mathbf{v} - \text{Proy}_{W^\perp} \mathbf{v}\|^2 + \|\mathbf{v} - \text{Proy}_W \mathbf{v}\|^2 \end{aligned}$$

**Ejemplo.** Obtenga el vector del subespacio

$$W = \left\{ \begin{bmatrix} x \\ y \\ 2x - y \end{bmatrix} \mid x, y \in \mathbb{R} \right\}$$

más próximo a  $\mathbf{u} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ , bajo el producto punto.

Primero debe obtenerse una base ortogonal de  $W$ , y se comienza por

el vector  $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ :

$$\begin{aligned} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 2x - y \end{bmatrix} &= 0 \\ x + 2y &= 0 \\ x &= -2y \end{aligned}$$

Al sustituir la restricción obtenida en el vector genérico del subespacio se obtiene que el segundo vector de la base ortogonal es

$\mathbf{e}_2 = \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix}$ . Ahora puede calcularse la proyección utilizando la base ortogonal  $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ .

$$\begin{aligned} \text{Proy}_W \mathbf{u} &= \frac{\mathbf{u} \cdot \mathbf{e}_1}{\mathbf{e}_1 \cdot \mathbf{e}_1} \mathbf{e}_1 + \frac{\mathbf{u} \cdot \mathbf{e}_2}{\mathbf{e}_2 \cdot \mathbf{e}_2} \mathbf{e}_2 \\ &= \frac{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \frac{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix}}{\begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix}} \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} \\ &= -\frac{3}{5} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} - \frac{9}{30} \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} \end{aligned}$$

La proyección es  $\text{Proy}_W \mathbf{u} = \frac{1}{2} \begin{bmatrix} 0 \\ -3 \\ 3 \end{bmatrix}$ .

## 2. Proceso de Ortogonalización

Las proyecciones son de gran ayuda para construir una base ortogonal a partir de una base cualquiera. En la figura 1 se presentan a los vectores  $\mathbf{d}$ ,  $\mathbf{m}$ ,  $\mathbf{v}$  y  $\mathbf{w}$ . Los vectores  $\mathbf{m}$  y  $\mathbf{v}$  son linealmente independientes, por lo que pueden formar una base  $B$ . Por otro lado,

$$\mathbf{w} = \text{Proy}_{\mathbf{m}} \mathbf{v} \quad (5)$$

y

$$\mathbf{d} = \mathbf{v} - \mathbf{w} \quad (6)$$

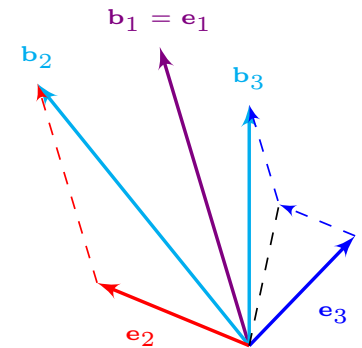


Figura 5. Ortogonalización mediante proyecciones sucesivas.

Al sustituir (5) en (6) se obtiene el vector ortogonal  $\mathbf{d}$  utilizando solo los vectores de la base  $B$ :

$$\mathbf{d} = \mathbf{v} - \text{Proy}_{\mathbf{m}} \mathbf{v}$$

Este proceso, conocido como ortogonalización de Gram-Schmidt, está basado en un algoritmo que crea vectores ortogonales a partir de la resta de proyecciones sucesivas (véase la figura 5). El algoritmo permite calcular una base ortogonal a partir de una base cualquiera  $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ ; consta de dos pasos, uno de los cuales es recursivo:

1.  $\mathbf{e}_1 = \mathbf{v}_1$

Quiere decir que el primer vector de la base  $B$  será el primer vector de la base ortogonal.

2. 
$$\mathbf{e}_i = \mathbf{v}_i - \sum_{j=2}^i \text{Proy}_{\mathbf{w}_{j-1}} \mathbf{v}_i \quad \forall 2 \leq i \leq n$$

Mediante proyecciones sucesivas del vector  $\mathbf{v}_i$  sobre los vectores calculados de la base ortogonal se obtienen los elementos de la nueva base.

**Ejemplo.** Obtenga, bajo el producto punto, una base ortogonal a partir de la base  $B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ .

Primer vector. El vector  $\mathbf{b}_1$  de la base original.

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Segundo vector. Es la resta de  $\mathbf{b}_2$  menos su proyección sobre los vectores anteriores de la nueva base ortogonal ( $\mathbf{e}_1$ ).

$$\begin{aligned} \mathbf{e}_2 &= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \frac{1}{3} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \end{aligned}$$

Tercer vector. Es la resta de  $\mathbf{b}_3$  menos su proyección sobre los vectores anteriores de la nueva base ortogonal ( $\mathbf{e}_1$  y  $\mathbf{e}_2$ ).

$$\begin{aligned} \mathbf{e}_3 &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{6} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \frac{1}{2} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \end{aligned}$$

Por propiedades del producto interno, los escalares  $\frac{1}{3}$  se simplificaron en el segundo sumando. Finalmente, la base ortogonalizada es

$$B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{3} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

# Álgebra Lineal

## Lectura 27: Método de Mínimos Cuadrados

Ing. Aldo Jiménez Arteaga

Enero 2020, Rev. Agosto 2021

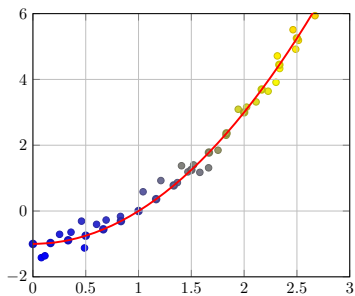


Figura 1. El método de regresión lineal ajusta una curva a un conjunto de puntos.

La aproximación por mínimos cuadrados permite calcular una función que mejor se ajuste a un conjunto de puntos en particular, o bien una solución que se aproxime a satisfacer un sistema de ecuaciones lineales (véase la figura 1). Es sumamente útil en las ciencias experimentales para encontrar modelos matemáticos de fenómenos físicos, o describir comportamientos estadísticos. Es una técnica de análisis numérico que intenta aplicar el criterio de mínimo error cuadrático entre los datos y el modelo matemático.

Un sistema de ecuaciones lineales  $A\mathbf{x} = \mathbf{b}$ , éste puede expresarse como

$$\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{b}$$

Al desarrollar la multiplicación, el vector de términos independientes  $\mathbf{b}$  es una combinación lineal del espacio columna de  $A$ :

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_n\mathbf{u}_n = \mathbf{b} \quad (1)$$

Si la combinación lineal expuesta no se cumple, entonces el sistema de ecuaciones es incompatible. Sin embargo, es posible obtener una solución aproximada. Dicha solución, denotada como  $\mathbf{x}'$ , es aquella que minimizará el error  $e = \|\mathbf{b} - A\mathbf{x}'\|$ ; en otras palabras, es la distancia mínima entre la solución aproximada  $A\mathbf{x}'$  y  $A\mathbf{x} = \mathbf{b}$ .

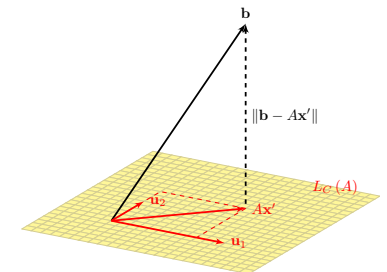


Figura 2. El vector distancia mínima permite ajustar los datos mediante la proyección.  $\mathbf{b} - A\mathbf{x}'$  siempre es perpendicular al espacio columna de la matriz  $A$ .

Para obtener la distancia mínima, es necesario considerar que un punto dado crea una proyección ortogonal sobre la función de ajuste, ya que este tipo de proyección minimiza la distancia entre ambos. Dicha distancia es el error  $e$  mencionado, establecido entre una magnitud real

(por ejemplo, un punto obtenido experimentalmente) y una ideal (el perteneciente a la curva de ajuste).

La curva de ajuste no se conoce, por lo tanto no es posible proyectar sobre ella (en Cálculo Vectorial se estudian las proyecciones sobre vectores tangentes a curvas). La figura 2 muestra cómo el vector de términos independientes  $\mathbf{b}$  se proyecta sobre el espacio columna de  $A$ , pues la combinación lineal (1) es la que plantea el problema de solución aproximada; el vector  $A\mathbf{x}'$  es dicha proyección.

Considerando que  $A\mathbf{w} \neq \mathbf{0}$  pertenece al espacio columna de  $A$ , de acuerdo a la figura 2, éste es ortogonal al vector  $\mathbf{b} - A\mathbf{x}'$ , por lo tanto

$$\langle A\mathbf{w}, \mathbf{b} - A\mathbf{x}' \rangle = 0 \tag{2}$$

por linealidad del producto interno, la ecuación (2) se convierte en

$$\begin{aligned} \langle A\mathbf{w}, \mathbf{b} - A\mathbf{x}' \rangle &= 0 \\ \langle A\mathbf{w}, \mathbf{b} \rangle - \langle A\mathbf{w}, A\mathbf{x}' \rangle &= 0 \end{aligned} \tag{3}$$

Mínimos cuadrados está definido a partir del producto escalar ordinario, es decir el producto punto. Para este caso, se expresará como  $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T \mathbf{v}$ . La ecuación (3) se desarrollará según el producto punto.

$$\begin{aligned} \langle A\mathbf{w}, \mathbf{b} \rangle - \langle A\mathbf{w}, A\mathbf{x}' \rangle &= 0 \\ (A\mathbf{w})^T \mathbf{b} - (A\mathbf{w})^T A\mathbf{x}' &= \\ \mathbf{w}^T A^T \mathbf{b} - \mathbf{w}^T A^T A\mathbf{x}' &= \\ \mathbf{w}^T (A^T \mathbf{b} - A^T A\mathbf{x}') &= 0 \end{aligned} \tag{4}$$

Como  $\mathbf{w}$  no es un vector nulo, entonces de (4) se concluye que

$$A^T \mathbf{b} - A^T A\mathbf{x}' = 0 \tag{5}$$

La expresión (5) es conocida como la ecuación normal de Gauss, y es la representación matricial del método de mínimos cuadrados.

El método de mínimos cuadrados es una técnica que devuelve la solución  $\mathbf{x}'$  que mejor se ajusta a un sistema de ecuaciones lineales incompatible. Para calcular la solución se aplica la ecuación

$$A^T \mathbf{b} = A^T A\mathbf{x}'$$

Al ser desarrollada esta definición matricial se obtiene la llamada regresión lineal, que permite ajustar un conjunto de puntos a una curva mediante cambios de variable.

**Ejemplo.** Obtenga la recta que mejor se ajuste a los puntos

$$\begin{array}{c|cccc} \mathbf{x} & 0 & 1 & 2 & 3 \\ \mathbf{y} & -2 & 0 & 3 & 5 \end{array}$$

Esta recta debe tener la forma  $y = mx + b$ . Para formar las matrices, se usan los coeficientes de la recta propuesta evaluada con los datos:

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

El sistema completo es

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 3 \\ 5 \end{bmatrix}$$

que al aplicarle la ecuación normal, se obtiene un sistema de dos ecuaciones con dos incógnitas que puede resolverse como cualquier sistema determinado.

$$\begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \\ 3 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} 14 & 6 \\ 6 & 4 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 21 \\ 6 \end{bmatrix} \quad (6)$$

La solución de (6) dará como resultado la pendiente y la ordenada al origen de la recta que mejor se ajusta al conjunto de puntos propuesto. Dicha recta es

$$y = \frac{12}{5}x - \frac{21}{10}$$

Cada vez que se realiza la regresión lineal se aplica este proceso, el cual es usado en las calculadoras modernas mediante la ecuación matricial

$$\begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{bmatrix} \quad (7)$$

La ecuación (7) es la expresión general de la ecuación normal de Gauss, y que se presenta en el ejemplo anterior mediante (6).