Espacios con Producto Interno

Definición de producto interno y sus propiedades elementales

Supóngase que se tiene un espacio vectorial V sobre un campo K. Se puede definir en V una función que a cada par de vectores $\bar{u}, \bar{v} \in V$ le asigna un escalar de K, denotado por $\langle \bar{u} | \bar{v} \rangle$. Dicha función se llama producto interno en V.

Una función $F: V \times V \rightarrow K$ es un producto interno o escalar si satisface las siguientes propiedades

- ✓ $\langle \alpha\bar{u} + \beta\bar{v} | \bar{w} \rangle = \alpha\langle \bar{u} | \bar{w} \rangle + \beta\langle \bar{v} | \bar{w} \rangle$ (Linealidad).
- ✓ $\langle \bar{u} | \bar{v} \rangle = \langle \bar{v} | \bar{u} \rangle$ (Simetría).
- ✓ $\langle \bar{u} | \bar{u} \rangle > 0, \bar{u} \neq \bar{0}$ (Positividad).

La propiedad de linealidad puede expresarse como dos propiedades independientes:

$$\langle \bar{u} + \bar{v} | \bar{w} \rangle = \langle \bar{u} | \bar{w} \rangle + \langle \bar{v} | \bar{w} \rangle \qquad \qquad \langle \alpha\bar{u} | \bar{w} \rangle = \alpha\langle \bar{u} | \bar{w} \rangle$$

EJEMPLO 4.1. Verifíquese que el producto punto en $\mathbb{R}^2$ es un producto interno. Para ello se tiene que

$$(x_1, y_1) \cdot (x_2, y_2) = x_1 x_2 + y_1 y_2$$

Linealidad:

$$\begin{aligned}
[\alpha(x_1, y_1) + \beta(x_2, y_2)] \cdot (x_3, y_3) &= \alpha(x_1, y_1) \cdot (x_3, y_3) + \beta(x_2, y_2) \cdot (x_3, y_3) \\
(\alpha x_1 + \beta x_2, \alpha y_1 + \beta y_2) \cdot (x_3, y_3) &= \alpha(x_1 x_3 + y_1 y_3) + \beta(x_2 x_3 + y_2 y_3) \\
(\alpha x_1 + \beta x_2) x_3 + (\alpha y_1 + \beta y_2) y_3 &= \alpha x_1 x_3 + \alpha y_1 y_3 + \beta x_2 x_3 + \beta y_2 y_3 \\
\alpha x_1 x_3 + \beta x_2 x_3 + \alpha y_1 y_3 + \beta y_2 y_3 &= \alpha x_1 x_3 + \alpha y_1 y_3 + \beta x_2 x_3 + \beta y_2 y_3
\end{aligned}$$

Por lo que la propiedad se cumple.

Simetría:

$$\begin{aligned}
(x_1, y_1) \cdot (x_2, y_2) &= (x_2, y_2) \cdot (x_1, y_1) \\
x_1 x_2 + y_1 y_2 &= x_2 x_1 + y_2 y_1
\end{aligned}$$

Al aplicar la conmutación en la multiplicación de los números reales, la propiedad se cumple.

Positividad:

$$\begin{aligned}
(x_1, y_1) \cdot (x_1, y_1) &= x_1 x_1 + y_1 y_1 \\
&= x_1^2 + y_1^2
\end{aligned}$$

Como el cuadrado de un número real siempre es positivo (excepto el cero), entonces la suma de dos cuadrados es positiva; por lo tanto, se cumple la propiedad.

Como conclusión, se tiene que el producto punto en el espacio de las parejas ordenadas es un producto interno.

EJEMPLO 4.2. Sea el espacio vectorial

$$V = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| a, b, c, d \in \mathbb{R} \right\}$$

Determínese si la siguiente función es un producto interno

$$f(A, B) = \operatorname{tr} AB$$

Linealidad:

$$\operatorname{tr}(\alpha A + \beta B)C = \alpha \operatorname{tr} AC + \beta \operatorname{tr} BC$$

Por la propiedades de la traza de una matriz

$$\begin{aligned} \operatorname{tr}(\alpha AC + \beta BC) &= \alpha \operatorname{tr} AC + \beta \operatorname{tr} BC \\ \operatorname{tr}(\alpha AC) + \operatorname{tr}(\beta BC) &= \\ \alpha \operatorname{tr}(AC) + \beta \operatorname{tr}(BC) &= \end{aligned}$$

Por lo que la propiedad se cumple.

Simetría:

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)$$

Por propiedades de la traza de matrices, la propiedad se cumple.

Positividad:

$$\operatorname{tr}(AA) > 0$$

donde

$$\begin{aligned} AA &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ &= \begin{pmatrix} a^2 + bc & ab + bd \\ ac + cd & bc + d^2 \end{pmatrix} \end{aligned}$$

Por lo tanto

$$a^2 + 2bc + d^2 > 0$$

Se tiene la suma de dos cuadrados y un producto que puede ser positivo o negativo, dependiendo de los valores de b y c; en general esta condición no se cumple. Por ejemplo, si se tiene la matriz

$$A = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

y al aplicarle la función

$$AA = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -3 & -4 \\ 4 & -3 \end{pmatrix}$$

se llega a $\operatorname{tr} A = -6$, y la propiedad de positividad no se cumple.

En conclusión la función dada no es un producto interno. No todas las funciones que involucran parejas de vectores pueden llegar a ser un producto interno; sin embargo, existen funciones que con una pequeña variación si lo son.

Algo que se debe resaltar, es que un espacio vectorial puede tener una infinidad de productos internos. Así como en el tema anterior se vio que el número de transformaciones lineales no es limitado, en este se puede definir cualquier producto interno en cualquier espacio vectorial, siempre y cuando su cumplan las tres propiedades descritas anteriormente.

Al igual que otras funciones, el producto interno cumple con algunas propiedades adicionales a las de la su definición; dichas propiedades extensivas son:

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

En la primera y tercera propiedades, $\overline{\langle \bar{v} | \bar{u} \rangle}$ y $\bar{\alpha}$ denotan al conjugado del número correspondiente.

EJEMPLO 4.3. Demuéstrese las propiedades uno y tres, para el espacio vectorial

$$\mathbb{C}^2 = \{(x, y) | x, y \in \mathbb{C}\}$$

Si el producto interno definido es

$$\langle \bar{u} | \bar{v} \rangle = x_1 \bar{x}_2 + y_1 \bar{y}_2 \qquad \forall\, \bar{u} = (x_1, y_1), \bar{v} = (x_2, y_2) \in \mathbb{C}^2$$

donde $\bar{y}_i$ denota al conjugado de y_i.

La propiedad uno establece que

$$\begin{gathered} \langle \bar{u} | \bar{v} \rangle = \overline{\langle \bar{v} | \bar{u} \rangle} \\ x_1 \bar{x}_2 + y_1 \bar{y}_2 = \overline{x_2 \bar{x}_1 + y_2 \bar{y}_1} \end{gathered}$$

Por propiedades del conjugado de los números complejos

$$\begin{gathered} x_1 \bar{x}_2 + y_1 \bar{y}_2 = \overline{x_2 \bar{x}_1} + \overline{y_2 \bar{y}_1} \\ x_1 \bar{x}_2 + y_1 \bar{y}_2 = \bar{x}_2 \bar{\bar{x}}_1 + \bar{y}_2 \bar{\bar{y}}_1 \\ x_1 \bar{x}_2 + y_1 \bar{y}_2 = \bar{x}_2 x_1 + \bar{y}_2 y_1 \end{gathered}$$

que satisface correctamente la propiedad.

La segunda propiedad hace uso del escalar $\alpha \in \mathbb{C}$, por lo que

$$\langle \bar{u} | \alpha \bar{v} \rangle = \bar{\alpha} \langle \bar{u} | \bar{v} \rangle$$
$$\langle (x_1, y_1) | \alpha (x_2, y_2) \rangle = \bar{\alpha} \langle (x_1, y_1) | (x_2, y_2) \rangle$$
$$\langle (x_1, y_1) | (\alpha x_2, \alpha y_2) \rangle = \bar{\alpha} (x_1 \bar{x}_2 + y_1 \bar{y}_2)$$
$$x_1 \overline{\alpha x_2} + y_1 \overline{\alpha y_2} = \bar{\alpha} x_1 \bar{x}_2 + \bar{\alpha} y_1 \bar{y}_2$$
$$x_1 \bar{\alpha} \bar{x}_2 + y_1 \bar{\alpha} \bar{y}_2 = x_1 \bar{\alpha} \bar{x}_2 + y_1 \bar{\alpha} \bar{y}_2$$

Finalmente, la propiedad es verdadera.

Definición de norma de un vector y sus propiedades

Dentro de las características geométricas de un vector, existe el concepto de magnitud; es decir, el valor escalar que por sí mismo posee un vector; o en otras palabras, la distancia que existe entre el punto donde inicia el vector, y el punto donde termina.

En la Física, la magnitud representa el valor que un fenómeno físico tiene. Por ejemplo, la aceleración es una cantidad vectorial, la cual se manipula como una cantidad escalar; ese tipo de manipulación utiliza la magnitud del vector aceleración. Otros ejemplos de este tipo son la fuerza, la velocidad, el campo eléctrico o la corriente eléctrica.

La norma de un vector es su magnitud, y ésta puede obtenerse a partir de un producto interno. Como se describió en el inciso anterior, el espacio vectorial puede tener una infinidad de productos internos definidos; por lo tanto, la norma del vector no será la misma para todos los productos internos.

Sea V un espacio vectorial donde se define el producto interno $\langle \bar{u} | \bar{v} \rangle \ \forall \ \bar{u}, \bar{v} \in V$. La norma de un vector se define como

$$\|\bar{v}\| = \sqrt{\langle \bar{v} | \bar{v} \rangle}$$

EJEMPLO 4.4. Sean los productos internos en el espacio vectorial $\mathbb{R}^2$ definidos como

$$\langle (x_1, y_1) | (x_2, y_2) \rangle = x_1 x_2 + y_1 y_2$$

$$\langle (x_1, y_1) | (x_2, y_2) \rangle = x_1 x_2 - 2x_1 y_2 - 2y_1 x_2 + 5y_1 y_2$$

Se desea encontrar las normas del vector $\bar{x} = (-1, -2)$. Para el caso del producto interno ordinario se tiene

$$\begin{aligned} \|(-1,-2)\| &= \sqrt{\langle (-1,-2) | (-1,-2) \rangle} \\ &= \sqrt{(-1)(-1) + (-2)(-2)} \\ &= \sqrt{5} \end{aligned}$$

Para el segundo producto interno

$$\begin{aligned} \|(-1,-2)\| &= \sqrt{\langle (-1,-2) | (-1,-2) \rangle} \\ &= \sqrt{(-1)(-1) - 2(-1)(-2) - 2(-2)(-1) + 5(-2)(-2)} \\ &= \sqrt{13} \end{aligned}$$

Con lo cual se comprueba que la norma de un mismo vector varía según el producto interno que se utilice.

La norma posee ciertas propiedades que permiten establecer simplificaciones al momento de trabajar con ecuaciones con producto interno. Dichas propiedades son

- ✓ $\|\bar{v}\| \geq 0$
- ✓ $\|\alpha\bar{v}\| = \alpha\|\bar{v}\|, \forall\, \alpha \in K$
- ✓ $\|\bar{u} + \bar{v}\| \leq \|\bar{u}\| + \|\bar{v}\|$

La última propiedad es conocida como la desigualdad del triángulo.

Vectores unitarios

Los vectores unitarios son una propiedad que involucra al vector y su norma. En forma general, un vector unitario es aquel que cumple con la ecuación

$$\|\bar{v}\| = 1$$

Supóngase que se tiene un vector unitario $\hat{v}$, el cual puede multiplicarse por un escalar dando como resultado un vector $\bar{v}$; es decir,

$$\alpha\hat{v} = \bar{v} \ldots (1)$$

El vector $\bar{v}$ tiene una norma $\|\bar{v}\|$. Dicha norma puede igualarse al escalar que se utilizó para multiplicar al vector unitario; es decir, si

$$\|\bar{v}\| = \alpha \ldots (2)$$

Si (2) se sustituye en (1) se tiene que

$$\|\bar{v}\|\hat{v} = \bar{v}$$

Por lo que se deduce que cualquier vector puede tener su correspondiente vector unitario. Esto se comprueba al multiplicar la última ecuación por el inverso de la norma $\|\bar{v}\|$.

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

Cuando se obtiene un vector unitario a partir de un vector $\bar{v}$ cualquiera, se dice que $\bar{v}$ fue normalizado.

EJEMPLO 4.5. Sea el espacio vectorial de las matrices reales de orden 2, y el producto interno definido por

$$\langle A|B\rangle = \operatorname{tr}(B^T A)$$

Se desea normalizar los vectores $\bar{u} = \begin{pmatrix} -2 & 0 \\ 1 & 1 \end{pmatrix}$ y $\bar{v} = \begin{pmatrix} -3 & 2 \\ 1 & 2 \end{pmatrix}$. Para ello, se necesita obtener sus respectivas normas y después, dividir cada vector entre la norma respectiva.

$$\left\|\begin{pmatrix} -2 & 0 \\ 1 & 1 \end{pmatrix}\right\| = \sqrt{6}$$

$$\left\|\begin{pmatrix} -3 & 2 \\ 1 & 2 \end{pmatrix}\right\| = 3\sqrt{2}$$

Finalmente, los vectores unitarios serían:

$$\hat{u} = \frac{1}{\sqrt{6}}\begin{pmatrix} -2 & 0 \\ 1 & 1 \end{pmatrix}, \qquad \hat{v} = \frac{1}{3\sqrt{2}}\begin{pmatrix} -3 & 2 \\ 1 & 2 \end{pmatrix}$$

Se puede verificar fácilmente, que las respectivas normas son iguales a 1.

Definición de distancia entre vectores y sus propiedades

El producto interno con normas permite establecer el concepto de distancia entre los vectores. La figura 4.1 muestra de manera gráfica la distancia entre vectores.

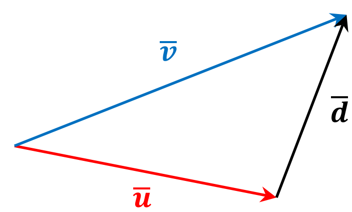

Figura 4.1. Distancia entre vectores.

Puede observarse que la distancia entre vectores se mide desde el punto final de $\bar{u}$ hasta el punto final de $\bar{v}$. La relación entre los vectores mostrados es

$$\bar{u} + \bar{d} = \bar{v}$$
$$\bar{d} = \bar{v} - \bar{u}$$

La distancia entre $\bar{u}$ y $\bar{v}$ se puede calcular como la norma del vector $\bar{d} = \bar{v} - \bar{u}$.

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

Este número es un real positivo, que establece los principios de métrica en un espacio vectorial. Satisface las siguientes propiedades:

- ✓ $d(\bar{u}, \bar{v}) \geq 0$.
- ✓ $d(\bar{u}, \bar{v}) = 0$, si $\bar{u} = \bar{v}$.
- ✓ $d(\bar{u}, \bar{v}) = d(\bar{v}, \bar{u})$.
- ✓ $d(\bar{u}, \bar{w}) \leq d(\bar{u}, \bar{v}) + d(\bar{v}, \bar{w})$.

EJEMPLO 4.6. La desigualdad del triángulo es un teorema geométrico muy importante dentro de la Matemática. La figura 4.2 muestra este problema.

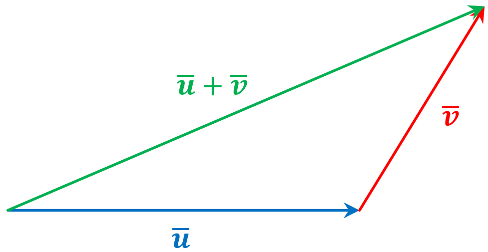

Figura 4.2. Desigualdad triangular.

Las normas son escalares positivos, por lo cual se puede establecer una relación con las propiedades del valor absoluto. Otro argumento para establecer la demostración es tomando al espacio vectorial de los números reales sobre los números reales.

Las propiedades del valor absoluto son:

- ✓ $|x| \leq a$, si, y sólo si $-a \leq x \leq a$.
- ✓ $|x| \geq a$, si, y sólo si, $x \leq -a$ ó $a \leq x$.

Si se tiene en principio

$$-|a| \leq a \leq |a| \ldots (1)$$

$$-|b| \leq b \leq |b| \ldots (2)$$

al sumar las ecuaciones (1) y (2) se tiene que

$$-|a| - |b| \leq a + b \leq |a| + |b|$$

Al utilizar la primera propiedad del valor absoluto se tiene que

$$|a + b| \leq |a| + |b|$$

Si se reescribe en términos de normas, se obtiene

$$\|\bar{u} + \bar{v}\| \leq \|\bar{u}\| + \|\bar{v}\|$$

EJEMPLO 4.7. Sean el espacio vectorial de las matrices cuadradas de orden dos con elementos complejos, y el producto interno definido por

$$\langle A|B\rangle = a_{11}\bar{b}_{11} + a_{12}\bar{b}_{12} + a_{21}\bar{b}_{21} + a_{22}\bar{b}_{22}$$

Calcúlese la distancia entre los vectores $\begin{pmatrix} -i & 2 \\ 1+i & 1-i \end{pmatrix}$ y $\begin{pmatrix} 2-i & -2i \\ 5 & 5i \end{pmatrix}$. Primero se tiene que

$$\begin{pmatrix} -2 & 2+2i \\ -4+i & 1-6i \end{pmatrix} = \begin{pmatrix} -i & 2 \\ 1+i & 1-i \end{pmatrix} - \begin{pmatrix} 2-i & -2i \\ 5 & 5i \end{pmatrix}$$

Por lo que

$$\left\| \begin{pmatrix} -2 & 2+2i \\ -4+i & 1-6i \end{pmatrix} \right\| = \sqrt{(-2)\overline{(-2)} + (2+2i)\overline{(2+2i)} + (-4+i)\overline{(-4+i)} + (1-6i)\overline{(1-6i)}}$$
$$= \sqrt{4 + (4+4) + (16+1) + (1+36)}$$
$$= \sqrt{66}$$

Por lo que la distancia entre las matrices es $\sqrt{66}\ [u]$.

EJEMPLO 4.8. Con el producto interno definido por

$$\langle p(x)|q(x)\rangle = \sum_{i=0}^{3} p(i)q(i)$$

Encuéntrese la distancia entre los polinomios $p(x) = x^3 - x^2 + x - 1$ y $q(x) = -x^2 + 2$.

Se obtiene

$$x^3 + x - 3 = (x^3 - x^2 + x - 1) - (-x^2 + 2)$$

Y, finalmente

$$\|x^3 + x - 3\| = \sqrt{(-3)(-3) + (-1)(-1) + (7)(7) + (27)(27)}$$
$$= \sqrt{788}$$

La distancia entre los polinomios es $\sqrt{788}\ [u]$.

Definición de ángulo entre vectores

El ángulo entre vectores también puede calcularse a partir del producto interno. La figura 4.3 permite establecer ciertas relaciones entre las normas y el ángulo.

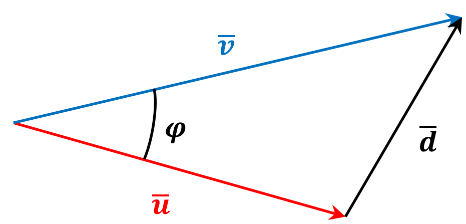

Figura 4.3. Ángulo entre vectores.

Por la ley de cosenos se sabe que

$$\|\bar{d}\|^2 = \|\bar{u}\|^2 + \|\bar{v}\|^2 - 2\|\bar{u}\|\|\bar{v}\|\cos\varphi$$

Pero además $\|\bar{d}\|$ es la norma de $\|\bar{v} - \bar{u}\|$. Por lo tanto, al desarrollar la expresión se tiene

$$\langle \bar{v} - \bar{u}|\bar{v} - \bar{u}\rangle = \|\bar{u}\|^2 + \|\bar{v}\|^2 - 2\|\bar{u}\|\|\bar{v}\|\cos\varphi$$
$$\langle \bar{v}|\bar{v} - \bar{u}\rangle - \langle \bar{u}|\bar{v} - \bar{u}\rangle = \|\bar{u}\|^2 + \|\bar{v}\|^2 - 2\|\bar{u}\|\|\bar{v}\|\cos\varphi$$
$$\langle \bar{v}|\bar{v}\rangle - \langle \bar{v}|\bar{u}\rangle - \langle \bar{u}|\bar{v}\rangle + \langle \bar{u}|\bar{u}\rangle = \|\bar{u}\|^2 + \|\bar{v}\|^2 - 2\|\bar{u}\|\|\bar{v}\|\cos\varphi$$

$$\|\bar{u}\|^2 - 2\langle\bar{u}|\bar{v}\rangle + \|\bar{v}\|^2 = \|\bar{u}\|^2 + \|\bar{v}\|^2 - 2\|\bar{u}\|\|\bar{v}\|\cos\varphi$$
$$-2\langle\bar{u}|\bar{v}\rangle = -2\|\bar{u}\|\|\bar{v}\|\cos\varphi$$

Finalmente, se obtiene que

$$\frac{\langle\bar{u}|\bar{v}\rangle}{\|\bar{u}\|\|\bar{v}\|} = \cos\varphi$$

Y el ángulo entre dos vectores puede obtenerse a partir del producto interno entre sus vectores, y las normas. Hay que destacar que en espacios complejos el producto $\langle\bar{u}|\bar{v}\rangle$ puede ser un número complejo, en cuyo caso sólo se tomará en cuenta la parte real; es decir

$$\frac{\mathrm{Re}\langle\bar{u}|\bar{v}\rangle}{\|\bar{u}\|\|\bar{v}\|} = \cos\varphi$$

EJEMPLO 4.9. Sean los vectores $\bar{u} = (0, 1, -1, 2)$ y $\bar{v} = (0, 0, 1, 1)$. Calcúlese el ángulo entre ellos, utilizando el producto interno usual en $\mathbb{R}^4$.

Se calculan

$$\begin{aligned}\bar{u}\cdot\bar{v} &= (0)(0) + (1)(0) + (-1)(1) + (2)(1)\\ &= 1\end{aligned}$$

$$\begin{aligned}\|\bar{u}\| &= \sqrt{(0)(0) + (1)(1) + (-1)(-1) + (2)(2)}\\ &= \sqrt{6}\end{aligned}$$

$$\begin{aligned}\|\bar{v}\| &= \sqrt{(0)(0) + (0)(0) + (1)(1) + (1)(1)}\\ &= \sqrt{2}\end{aligned}$$

Por lo tanto, el coseno del ángulo entre los vectores es

$$\cos\varphi = \frac{1}{\sqrt{6}\sqrt{2}}$$

Por lo tanto, se obtiene que $\varphi = 73.22°$.

EJEMPLO 4.10. Calcúlese el ángulo entre los vectores $\begin{pmatrix} i & 2 & 1+i \\ 0 & -3i & 1 \end{pmatrix}$ y $\begin{pmatrix} -3 & 4 & -2+i \\ 2i & 0 & i \end{pmatrix}$, si se define el producto interno

$$\langle A|B\rangle = \mathrm{tr}(A^*B)$$

Se calcula

$$\langle A|B\rangle = \mathrm{tr}\left[\begin{pmatrix} -i & 0 \\ 2 & 3i \\ 1-i & 1 \end{pmatrix}\begin{pmatrix} -3 & 4 & -2+i \\ 2i & 0 & i \end{pmatrix}\right]$$

$$= \operatorname{tr}\begin{pmatrix} 3i & -4i & 1+2i \\ -12 & 8 & -7+2i \\ -3+5i & 4-4i & -1+4i \end{pmatrix}$$
$$= 7+7i$$

$$\langle A|A\rangle = \operatorname{tr}\left[\begin{pmatrix} -i & 0 \\ 2 & 3i \\ 1-i & 1 \end{pmatrix}\begin{pmatrix} i & 2 & 1+i \\ 0 & -3i & 1 \end{pmatrix}\right]$$
$$= \operatorname{tr}\begin{pmatrix} 1 & -2i & 1-i \\ 2i & 13 & 2+5i \\ 1+i & 2-5i & 3 \end{pmatrix}$$
$$= 17$$

$$\langle B|B\rangle = \operatorname{tr}\left[\begin{pmatrix} -3 & -2i \\ 4 & 0 \\ -2-i & -i \end{pmatrix}\begin{pmatrix} -3 & 4 & -2+i \\ 2i & 0 & i \end{pmatrix}\right]$$
$$= \operatorname{tr}\begin{pmatrix} 13 & -12 & 8-3i \\ -12 & 16 & -8+4i \\ 8+3i & -8-4i & 6 \end{pmatrix}$$
$$= 35$$

Por lo que el ángulo entre vectores, en este caso, queda definido por

$$\varphi = \arccos\left(\frac{\operatorname{Re}(7+7i)}{\sqrt{17}\sqrt{35}}\right)$$

Obteniéndose que $\varphi = 73.32°$.

Vectores ortogonales

Sea un espacio vectorial V con producto interno. Se dice que dos vectores $\bar{u}, \bar{v} \in V$ son ortogonales si

$$\langle \bar{u}|\bar{v}\rangle = 0$$

Dicha relación es simétrica: si $\bar{u}$ es ortogonal a $\bar{v}$, entonces $\bar{v}$ es ortogonal a $\bar{u}$. Debido a que

$$0 = \cos\frac{\pi}{2}$$

entonces,

$$0 = \frac{\langle \bar{u}|\bar{v}\rangle}{\|\bar{u}\|\|\bar{v}\|}$$

Por lo que se llega a la definición de ortogonalidad entre vectores.

EJEMPLO 4.11. Sean los vectores $\bar{u} = (0,-1,1,2)$ y $\bar{v} = (2,1,1,0)$. Al aplicar el producto escalar ordinario en cuartetos ordenados reales se llega a

$$\bar{u}\cdot\bar{v} = (0,-1,1,2)\cdot(2,1,1,0)$$
$$= (0)(2)+(-1)(1)+(1)(1)+(2)(0)$$
$$= 0$$

Por lo que los vectores son ortogonales.

Si al concepto de vectores ortogonales se le añade el de vector unitario, se establece la definición de vectores ortonormales. Sean $\hat{u}$ y $\hat{v}$ dos vectores unitarios; si se tiene que

$$\langle \hat{u} | \hat{v} \rangle = 0$$

Se dice que los vectores $\hat{u}$ y $\hat{v}$ son ortonormales.

EJEMPLO 4.12. Dados los vectores $p(x) = -x^2 + x - 1$ y $q(x) = ax^2 + x - 2$, encuéntrese los valores $a \in \mathbb{R}$ tales que los dos polinomios sean vectores ortonormales, si se define

$$\langle p(x) | q(x) \rangle = \sum_{i=0}^{2} a_i b_i$$

Si el espacio vectorial es $P_2 = \{a_2x^2 + a_1x + a_0 | a_0, a_1, a_2 \in \mathbb{R}\}$.

Se sabe que el producto entre los dos vectores debe ser nulo, por lo tanto

$$0 = (-1)(a) + (1)(1) + (-1)(-2)$$

y se tiene que $a = 3$. Al obtener las respectivas normas de los polinomios se obtiene que

$$\|p(x)\| = \sqrt{3}$$

$$\|q(x)\| = \sqrt{14}$$

Y los vectores ortonormales son $\hat{p}(x) = -\frac{1}{\sqrt{3}}x^2 + \frac{1}{\sqrt{3}}x - \frac{1}{\sqrt{3}}$ y $\hat{q}(x) = \frac{3}{\sqrt{14}}x^2 + \frac{1}{\sqrt{14}}x - \frac{2}{\sqrt{14}}$.

Conjuntos ortogonales y ortonormales

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. Sea el conjunto

$$B = \{\bar{b}_1, \bar{b}_2, \bar{b}_3 \dots, \bar{b}_n\}$$

Si se cumple que $\langle \bar{b}_i | \bar{b}_j \rangle = 0, \forall\, i \neq j$, entonces se dice que B es un conjunto ortogonal.

EJEMPLO 4.13. El conjunto $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$ es ortogonal bajo el producto interno usual en las ternas ordenadas, ya que

$$(1, 0, 0) \cdot (0, 0, 1) = 0$$

$$(1, 0, 0) \cdot (0, 1, 0) = 0$$

$$(0, 1, 0) \cdot (0, 0, 1) = 0$$

EJEMPLO 4.14. Dado el conjunto de vectores $\{(2a,0,-1),(-1,10,-2),(2,1,b)\}$, ¿qué valores $a,b\in\mathbb{R}$ permiten que el conjunto sea ortogonal?

Se debe plantear que

$$(2a,0,-1)\cdot(-1,10,-2)=0$$

$$(2,1,b)\cdot(2a,0,-1)=0$$

$$(-1,10,-2)\cdot(2,1,b)=0$$

Por lo que surge el sistema de ecuaciones

$$\begin{array}{cccc} -2a & & = & -2 \\ 4a & -b & = & 0 \\ & -2b & = & -8 \end{array}$$

Finalmente se obtiene que $a=1$ y $b=4$.

Para poder establecer un conjunto ortogonal de vectores es necesario considerar el producto interno definido. Entonces, el conjunto puede determinarse a partir de un vector de un subespacio vectorial o incluso a partir de un espacio.

Sea el subespacio vectorial V donde se define un producto interno $\langle\bar{u}|\bar{v}\rangle$. Para obtener un conjunto ortogonal se parte de un vector arbitrario $\bar{b}_1\in V$. Para obtener el segundo vector $\bar{b}_2$ se debe cumplir que

✓ $\langle\bar{b}_1|\bar{b}_2\rangle=0$, que permitirá plantear una ecuación lineal (1).

El tercer vector $\bar{b}_3$ del conjunto ortogonal se obtiene al plantear las condiciones simultáneas

✓ $\langle\bar{b}_1|\bar{b}_3\rangle=0$, que plantea nuevamente la ecuación lineal (1).
✓ $\langle\bar{b}_2|\bar{b}_3\rangle=0$, que añade una nueva ecuación lineal (2)

El vector $\bar{b}_4$ se obtendrá al resolver el sistema de ecuaciones lineales formado por las ecuaciones (1), (2) y una nueva al añadir una ecuación (3), producto de $\langle\bar{b}_3|\bar{b}_4\rangle=0$.

El proceso se continuará de forma análoga al descrito anteriormente, hasta obtener como vector ortogonal al vector nulo. Cabe destacar que cada uno de los vectores obtenidos debe pertenecer al mismo subespacio vectorial en el cual de define el producto interno.

EJEMPLO 4.15. Sea el espacio vectorial

$$M_2=\left\{\begin{pmatrix} a & b \\ c & d\end{pmatrix}\middle| a,b,c,d\in\mathbb{R}\right\}$$

Con el producto interno definido por $\langle A|B\rangle=\operatorname{tr}(AB^T)$. Determínese un conjunto ortogonal a partir del vector $A=\begin{pmatrix} -1 & 0 \\ 1 & -1\end{pmatrix}$.

Primero se plantea $\langle A|B\rangle = 0$, donde $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

$$\langle A|B\rangle = \operatorname{tr}\begin{pmatrix} -a & -c \\ a-b & c-d \end{pmatrix}$$
$$0 = -a + c - d$$

que arroja la condición $c = a + d$; por lo tanto el vector buscado tiene la forma $B = \begin{pmatrix} a & b \\ a+d & d \end{pmatrix}$.

Dando valores, se obtiene $B = \begin{pmatrix} 0 & 1 \\ 2 & 2 \end{pmatrix}$.

El siguiente vector ortogonal se obtiene a partir del planteamiento de las ecuaciones $\langle A|C\rangle = 0$ y $\langle B|C\rangle = 0$, donde $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

$$\langle A|C\rangle = \operatorname{tr}\begin{pmatrix} -a & -c \\ a-b & c-d \end{pmatrix}$$
$$0 = -a + c - d$$

$$\langle B|C\rangle = \operatorname{tr}\begin{pmatrix} b & d \\ 2a+2b & 2c+2d \end{pmatrix}$$
$$0 = b + 2c + 2d$$

Al sustituir las condiciones obtenidas, $a = c - d$ y $b = -2c - 2d$, en el vector C se llega a $C = \begin{pmatrix} c-d & -2c-2d \\ c & d \end{pmatrix}$, que al darle valores establece que $C = \begin{pmatrix} -1 & -6 \\ 1 & 2 \end{pmatrix}$.

El siguiente vector $D = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ deberá cumplir simultáneamente $\langle A|D\rangle = 0$, $\langle B|D\rangle = 0$ y $\langle C|D\rangle = 0$.

$$\langle A|D\rangle = -a + c - d$$

$$\langle B|D\rangle = b + 2c + 2d$$

$$\langle B|C\rangle = \operatorname{tr}\begin{pmatrix} -a-6b & -c-6d \\ a+2b & c+2d \end{pmatrix}$$
$$0 = -a - 6b + c + 2d$$

Las condiciones que se deben satisfacer son $a = c - d$, $b = -2c - 2d$ y $0 = -a - 6b + c + 2d$. Al resolver el sistema ecuaciones se tiene que

$$\begin{array}{cccccc} a & & & +\frac{9}{4}d & = & 0 \\ & & c & +\frac{5}{4}d & = & 0 \\ & b & & -\frac{1}{2}d & = & 0 \end{array}$$

Y el vector tiene la forma $D = \begin{pmatrix} -\frac{9}{4}d & \frac{1}{2}d \\ -\frac{5}{4}d & d \end{pmatrix}$; al dar valores se obtiene $D = \begin{pmatrix} -9 & 2 \\ -5 & 4 \end{pmatrix}$. El siguiente vector que debe ser simultáneamente ortogonal a los cuatro anteriores:

$$\langle A|E\rangle = -a + c - d$$

$$\langle B|E\rangle = b + 2c + 2d$$

$$\langle C|E\rangle = -a - 6b + c + 2d$$

$$\langle D|E\rangle = \operatorname{tr}\begin{pmatrix} -9a + 2b & -9c + 2d \\ -5a + 4b & -5c + 4d \end{pmatrix}$$
$$0 = -9a + 2b - 5c + 4d$$

Al plantear el sistema de ecuaciones

$$\begin{array}{cccccc} -a & & +c & -d & = & 0 \\ & b & +2c & +2d & = & 0 \\ -a & -6b & +c & +2d & = & 0 \\ -9a & +2b & -5c & +4d & = & 0 \end{array}$$

Se descubre que solo acepta la solución trivial; por lo tanto el siguiente vector ortogonal es el vector nulo. Si se sigue el procedimiento nuevamente, se obtendrá siempre el vector nulo; por lo que el conjunto ortogonal es

$$\left\{\begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 2 & 2 \end{pmatrix}, \begin{pmatrix} -1 & -6 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} -9 & 2 \\ -5 & 4 \end{pmatrix}\right\}$$

EJEMPLO 4.16. Para el subespacio $S = \{(w, x, 2w - x, z) | w, x, z \in \mathbb{R}\}$, encuéntrese un conjunto ortogonal.

Se inicia con un vector arbitrario del subespacio $\bar{v}_1 = (1, 1, 1, 1)$:

$$\bar{v}_1 \cdot \bar{v}_2 = (1, 1, 1, 1) \cdot (w, x, 2w - x, z)$$
$$0 = 3w + z$$

El segundo vector es $\bar{v}_2 = (w, x, 2w - x, -3w)$, o bien $\bar{v}_2 = (1, 1, 1, -3)$. El tercer vector es

$$\bar{v}_1 \cdot \bar{v}_3 = 3w + z$$

$$\bar{v}_2 \cdot \bar{v}_3 = (1, 1, 1, -3) \cdot (w, x, 2w - x, z)$$
$$0 = 3w - 3z$$

Al sustituir las dos condiciones, se obtiene el vector $\bar{v}_3 = (0, x, -x, 0)$, que con valores es $\bar{v}_3 = (0, 1, -1, 0)$. Finalmente, el conjunto ortogonal es

$$\{(1, 1, 1, 1), (1, 1, 1, -3), (0, 1, -1, 0)\}$$

Como en el ejemplo anterior, al calcular el siguiente vector ortogonal se obtendrá el vector nulo.

Independencia lineal de un conjunto ortogonal de vectores no nulos

En los ejemplos anteriores puede observarse que el número de vectores de cada conjunto ortogonal coincide con la dimensión del subespacio. Además, se puede comprobar fácilmente que cada conjunto ortogonal es linealmente independiente; por lo tanto, un conjunto ortogonal con el mismo número de vectores que la dimensión del subespacio al que pertenece es una base.

Sea el conjunto ortogonal

$$A = \{\bar{b}_1, \bar{b}_2, \bar{b}_3, \dots, \bar{b}_n\}$$

Si se establece la ecuación de dependencia lineal con los elementos de A se tiene que

$$\bar{0} = \alpha_1\bar{b}_1 + \alpha_2\bar{b}_2 + \alpha_3\bar{b}_3 + \cdots + \alpha_n\bar{b}_n$$

Si se toma un elemento $\bar{b}_i$ del conjunto A y se aplica a la ecuación de dependencia el producto interno para el cual A es ortogonal se obtiene que

$$\begin{aligned}\langle\bar{0}|\bar{b}_i\rangle &= \langle\alpha_1\bar{b}_1|\bar{b}_i\rangle + \langle\alpha_2\bar{b}_2|\bar{b}_i\rangle + \langle\alpha_3\bar{b}_3|\bar{b}_i\rangle + \cdots + \langle\alpha_n\bar{b}_n|\bar{b}_i\rangle \\ 0 &= \alpha_1\langle\bar{b}_1|\bar{b}_i\rangle + \alpha_2\langle\bar{b}_2|\bar{b}_i\rangle + \alpha_3\langle\bar{b}_3|\bar{b}_i\rangle + \cdots + \alpha_n\langle\bar{b}_n|\bar{b}_i\rangle\end{aligned}$$

Como el conjunto es ortogonal, entonces el producto $\langle\bar{b}_j|\bar{b}_i\rangle = 0, \forall\ i \neq j$, y la ecuación de dependencia se reduce a

$$0 = \alpha_i\langle\bar{b}_i|\bar{b}_i\rangle$$

De donde se obtiene que $\alpha_i = 0$. Como se puede realizar este procedimiento con cualquier vector, entonces se concluye que la ecuación de dependencia se cumple solo si todos los escalares son nulos; en consecuencia, el conjunto es linealmente independiente.

EJEMPLO 4.17. El conjunto del ejemplo 4.18 puede establecerse como una matriz, y verificar si es una base del subespacio $S = \{(w, x, 2w - x, z)|w, x, z \in \mathbb{R}\}$.

$$\begin{pmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & -3\\ 0 & 1 & -1 & 0\end{pmatrix} \sim \begin{pmatrix}1 & 1 & 1 & 1\\ 0 & 0 & 0 & 4\\ 0 & 1 & -1 & 0\end{pmatrix} \sim \begin{pmatrix}1 & 1 & 1 & 1\\ 0 & 1 & -1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} \sim \begin{pmatrix}1 & 0 & 2 & 0\\ 0 & 1 & -1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

Se observa que los vectores son linealmente independientes; por lo tanto, el espacio generado es

$$G = \{(a, b, 2a - b, c)|a, b, c \in \mathbb{R}\}$$

que es el mismo subespacio S.

A una base de un subespacio que además es ortogonal se le conoce como base ortogonal. El proceso para obtener una base ortogonal a partir de un vector es el mismo que se describió en el apartado de conjunto ortogonal.

Coordenadas de un vector respecto a una base ortogonal

Al introducir el concepto de base ortogonal, se puede establecer también el de coordenadas de un vector en una base ortogonal. Sea un espacio vectorial V donde se define la base ortogonal

$$A = \{\bar{b}_1, \bar{b}_2, \bar{b}_3, \dots, \bar{b}_n\}$$

y sea un vector $\bar{u} \in V$. Para obtener el vector de coordenadas $[\bar{u}]_A$ se tiene

$$\bar{u} = \alpha_1\bar{b}_1 + \alpha_2\bar{b}_2 + \alpha_3\bar{b}_3 + \cdots + \alpha_n\bar{b}_n$$

Si se aplica el producto interno con un vector $\bar{b}_i$ de la base a ambos lados de la ecuación se llega a

$$\langle \bar{u}|\bar{b}_i\rangle = \langle \alpha_1\bar{b}_1|\bar{b}_i\rangle + \langle \alpha_2\bar{b}_2|\bar{b}_i\rangle + \langle \alpha_3\bar{b}_3|\bar{b}_i\rangle + \cdots + \langle \alpha_n\bar{b}_n|\bar{b}_i\rangle$$

Al igual que en la demostración anterior, se sabe que $\langle \bar{b}_j|\bar{b}_i\rangle = 0, \forall\ i \neq j$; por lo tanto, la expresión anterior se reduce a

$$\langle \bar{u}|\bar{b}_i\rangle = \alpha_i\langle \bar{b}_i|\bar{b}_i\rangle$$

lo cual permite que

$$\frac{\langle \bar{u}|\bar{b}_i\rangle}{\langle \bar{b}_i|\bar{b}_i\rangle} = \alpha_i$$

Entonces, se tiene que el vector de coordenadas de $\bar{u}$ en la base ortogonal A está dado por

$$[\bar{u}]_B = \left(\frac{\langle \bar{u}|\bar{b}_1\rangle}{\langle \bar{b}_1|\bar{b}_1\rangle}, \frac{\langle \bar{u}|\bar{b}_2\rangle}{\langle \bar{b}_2|\bar{b}_2\rangle}, \frac{\langle \bar{u}|\bar{b}_3\rangle}{\langle \bar{b}_3|\bar{b}_3\rangle}, \dots, \frac{\langle \bar{u}|\bar{b}_n\rangle}{\langle \bar{b}_n|\bar{b}_n\rangle}\right)$$

EJEMPLO 4.18. Sea el subespacio $P_2 = \{ax^2 + bx + (-2a + 3b)|a, b \in \mathbb{R}\}$ y la base ortogonal

$$B = \{2x^2 + x - 1, x^2 + 2x + 4\}$$

con respecto al producto interno definido por

$$\langle p(x)|q(x)\rangle = \sum_{i=0}^{2} a_i b_i$$

Calcúlese el vector de coordenadas de $f(x) = x + 3$ en la base ortogonal, utilizando el producto interno. Se sabe que

$$[f(x)]_B = \left(\frac{\langle f|p\rangle}{\langle p|p\rangle}, \frac{\langle f|q\rangle}{\langle q|q\rangle}\right)$$

Por lo que

$$\frac{\langle f|p\rangle}{\langle p|p\rangle} = \frac{(0)(2) + (1)(1) + (3)(-1)}{(2)(2) + (1)(1) + (-1)(-1)} \quad \Rightarrow \quad -\frac{1}{3}$$

$$\frac{\langle f|q\rangle}{\langle q|q\rangle} = \frac{(0)(1) + (1)(2) + (3)(4)}{(1)(1) + (2)(2) + (4)(4)} \quad \Rightarrow \quad \frac{2}{3}$$

Y se obtiene que $[f(x)]_B = \left(-\frac{1}{3}, \frac{2}{3}\right)$.

Proceso de ortogonalización de Gram-Schmidt

Se ha estudiado que a partir de un vector de un subespacio se puede obtener una base ortogonal. Sin embargo, también es posible encontrar una base de este tipo a partir de una base cualquiera del subespacio.

Supóngase que $A = \{\bar{v}_1, \bar{v}_2, \bar{v}_3, \dots, \bar{v}_n\}$ es una base del espacio vectorial V con un producto interno $\langle \bar{u}|\bar{v}\rangle$ definido. Para encontrar una base ortogonal $B = \{\bar{w}_1, \bar{w}_2, \bar{w}_3, \dots, \bar{w}_n\}$ se debe elegir un primer vector de A sobre el cual se operará. En este caso se tiene que

1. $\bar{w}_1 = \bar{v}_1$

Para el segundo vector de la base ortogonal se debe hacer una resta del segundo vector de la base original menos su proyección sobre el vector obtenido en el paso anterior; es decir,

2. $\bar{w}_2 = \bar{v}_2 - \frac{\langle \bar{v}_2 | \bar{w}_1 \rangle}{\langle \bar{w}_1 | \bar{w}_1 \rangle} \bar{w}_1$

Para el tercer vector se realiza la resta del tercer vector de la base original menos la proyección sobre el primer vector menos la proyección sobre el segundo vector de la base ortogonal, ya obtenidos; o sea,

3. $\bar{w}_3 = \bar{v}_3 - \frac{\langle \bar{v}_3 | \bar{w}_1 \rangle}{\langle \bar{w}_1 | \bar{w}_1 \rangle} \bar{w}_1 - \frac{\langle \bar{v}_3 | \bar{w}_2 \rangle}{\langle \bar{w}_2 | \bar{w}_2 \rangle} \bar{w}_2$

El proceso se continúa como en los pasos anteriores hasta obtener el vector n-ésimo de la base ortogonal

n. $\bar{w}_n = \bar{v}_n - \frac{\langle \bar{v}_n | \bar{w}_1 \rangle}{\langle \bar{w}_1 | \bar{w}_1 \rangle} \bar{w}_1 - \frac{\langle \bar{v}_n | \bar{w}_2 \rangle}{\langle \bar{w}_2 | \bar{w}_2 \rangle} \bar{w}_2 - \frac{\langle \bar{v}_n | \bar{w}_3 \rangle}{\langle \bar{w}_3 | \bar{w}_3 \rangle} \bar{w}_3 - \cdots - \frac{\langle \bar{v}_n | \bar{w}_{n-1} \rangle}{\langle \bar{w}_{n-1} | \bar{w}_{n-1} \rangle} \bar{w}_{n-1}$

En forma compacta, este algoritmo crea una base ortogonal a partir de una base cualquiera de un espacio vectorial, donde

$$\bar{w}_1 = \bar{v}_1$$
$$\bar{w}_i = \bar{v}_i - \sum_{k=1}^{i-1} \frac{\langle \bar{v}_i | \bar{w}_k \rangle}{\langle \bar{w}_k | \bar{w}_k \rangle} \bar{w}_k \qquad \forall\, 2 \le i \le n$$

El algoritmo es conocido como el método de ortogonalización de Gram-Schmidt, llamado en honor de sus descubridores, el actuario danés Jørgen Pedersen Gram y el matemático alemán Erhard Schmidt.

EJEMPLO 4.19. Encuéntrese la base ortogonal generada a partir de la base

$$A = \left\{ \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

del espacio de las matrices simétricas de orden 2, si el producto interno que se define es

$$\langle A | B \rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}$$

El proceso se comienzo con la elección del primer vector como inicio de la base, y después se aplica la ecuación

$$\bar{w}_i = \bar{v}_i - \sum_{k=1}^{i-1} \frac{\langle \bar{v}_i | \bar{w}_k \rangle}{\langle \bar{w}_k | \bar{w}_k \rangle} \bar{w}_k$$

Para los vectores siguientes:

$$\bar{w}_1 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$\bar{w}_2 = \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2} & 2 \\ 2 & \frac{1}{2} \end{pmatrix}$$

$$\overline{w}_3 = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{7}{17}\begin{pmatrix} \frac{1}{2} & 2 \\ 2 & \frac{1}{2} \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{12}{17} & \frac{3}{17} \\ \frac{3}{17} & -\frac{12}{17} \end{pmatrix}$$

Y la base ortogonal obtenida es

$$B = \left\{ \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & 2 \\ 2 & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} -\frac{12}{17} & \frac{3}{17} \\ \frac{3}{17} & -\frac{12}{17} \end{pmatrix} \right\}$$

EJEMPLO 4.20. Dada una base de un subespacio de $\mathbb{R}^4$

$$\{(1,1,0,0),(0,1,1,0),(0,0,1,1)\}$$

Encuéntrese una base ortogonal.

Por el proceso de Gram-Schmidt se establece que

$$\overline{w}_1 = (1,1,0,0)$$

El segundo vector es

$$\begin{aligned} \overline{w}_2 &= (0,1,1,0) - \frac{(1,1,0,0)\cdot(0,1,1,0)}{(1,1,0,0)\cdot(1,1,0,0)}(1,1,0,0) \\ &= (0,1,1,0) - \frac{1}{2}(1,1,0,0) \\ &= \left(-\frac{1}{2},\frac{1}{2},1,0\right) \end{aligned}$$

El tercer vector es

$$\begin{aligned} \overline{w}_3 &= (0,0,1,1) - \frac{0}{2}(1,1,0,0) - \frac{\left(-\frac{1}{2},\frac{1}{2},1,0\right)\cdot(0,0,1,1)}{\left(-\frac{1}{2},\frac{1}{2},1,0\right)\cdot\left(-\frac{1}{2},\frac{1}{2},1,0\right)}\left(-\frac{1}{2},\frac{1}{2},1,0\right) \\ &= (0,0,1,1) - \frac{2}{3}\left(-\frac{1}{2},\frac{1}{2},1,0\right) \\ &= \left(\frac{1}{3},-\frac{1}{3},\frac{1}{3},1\right) \end{aligned}$$

Finalmente, la base ortogonal es

$$\left\{(1,1,0,0),\left(-\frac{1}{2},\frac{1}{2},1,0\right),\left(\frac{1}{3},-\frac{1}{3},\frac{1}{3},1\right)\right\}$$

Complemento ortogonal

Las dos opciones que se tienen al manejar conjuntos de vectores y ortogonalidad son:

- ✓ Un conjunto cuyos vectores son ortogonales entre sí.
- ✓ Un conjunto cuyos vectores son ortogonales a otro vector fuera del conjunto.

El primero fue tratado en el tema de bases ortogonales; el segundo se conoce como complemento ortogonal, y se define a continuación.

Sea S un subespacio de un espacio vectorial V con producto interno. Al subespacio de V cuyos vectores son ortogonales a los elementos de S, se le conoce como complemento ortogonal de S, y se denota como

$$S^{\perp} = \{\bar{v} | \langle \bar{v} | \bar{u} \rangle = 0, \bar{u} \in S\}$$

Se debe hacer hincapié que los elementos del complemento ortogonal de S no necesariamente son ortogonales entre sí.

EJEMPLO 4.21. Sea el subespacio $A = \{(w, x, x, w) | w, x \in \mathbb{R}\}$. ¿Cuál es su complemento ortogonal, con respecto al producto escalar ordinario?

En este problema se debe elegir una base de A, la cual planteará la ortogonalidad con otro subespacio de los cuartetos reales ordenados. La base elegida será:

$$\{(1, 0, 0, 1), (0, 1, 1, 0)\}$$

Por lo que se tiene que

$$(1, 0, 0, 1) \cdot (w, x, y, z) = 0$$

$$(0, 1, 1, 0) \cdot (w, x, y, z) = 0$$

De estas dos operaciones se obtienen las condiciones $w + z = 0$ y $x + y = 0$ del complemento ortogonal de A:

$$A^{\perp} = \{(-z, -y, y, z) | y, z \in \mathbb{R}\}$$

Si se realiza el producto escalar entre (w, x, x, w) y $(-z, -y, y, z)$ se obtiene

$$(w, x, x, w) \cdot (-z, -y, y, z) = -wz - xy + xy + wz$$

cuyo resultado es cero.

EJEMPLO 4.22. Encuéntrese el complemento ortogonal al polinomio $g(x) = x^2 - 2x + 1$ bajo el producto interno definido por

$$\langle p(x) | q(x) \rangle = \sum_{i=0}^{2} a_i b_i$$

Si el espacio vectorial es el de los polinomios de grado menor o igual a dos con elementos reales.

Se debe plantear que

$$\langle x^2 - 2x + 1 | ax^2 + bx + c \rangle = 0$$
$$\langle x^2 - 2x + 1 | ax^2 + bx + c \rangle = (1)(a) + (-2)(b) + (1)(c)$$

Por lo que la condición que debe cumplir el complemento ortogonal es $a - 2b + c = 0$. Y se tiene que el complemento ortogonal del polinomio dado es

$$g(x)^{\perp} = \{(2b - c)x^2 + bx + c | y, z \in \mathbb{R}\}$$

EJEMPLO 4.23. Una de las aplicaciones directas del complemento ortogonal es el plano en el espacio. Se sabe que un plano que contiene al origen es un subespacio de $\mathbb{R}^3$; dicho plano puede definirse por la ecuación

$$ax + by + cz = 0$$

Esa ecuación del plano puede descomponerse como $(a, b, c) \cdot (x, y, z) = 0$, donde el vector (x, y, z) representa a un vector que pertenece al plano, y (a, b, c) es un vector perpendicular (conocido en geometría analítica como vector normal) al plano. La figura 4.4 muestra un plano generado por dos vectores, y el vector normal al plano.

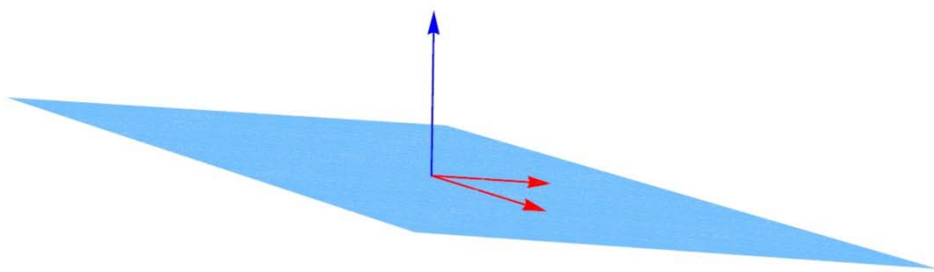

Figura 4.4. Plano generado por dos vectores (rojo) que muestra al vector normal (azul).

Si los vectores que forman el plano son $(-2, 6, 4)$ y $(-3, 3, -3)$, ¿cuál es una ecuación del plano?

El plano es el espacio generado por los vectores dados, es decir

$$\Pi = \{(-x - y, 3x + y, 2x - y) | x, y \in \mathbb{R}\}$$

Para encontrar la ecuación del plano es necesario obtener el vector normal, el cual pertenece al complemento ortogonal del plano. Tomando la base $\{(-1, 3, 2), (-1, 1, -1)\}$, el vector normal se obtiene como

$$(-1, 3, 2) \cdot (x, y, z) = 0$$

$$(-1, 1, -1) \cdot (x, y, z) = 0$$

Que establece las ecuaciones $-x + 3y + 2z = 0$ y $-x + y - z = 0$, de donde se obtiene que

$$\Pi^{\perp} = \left\{\left(-\frac{5}{2}z, -\frac{3}{2}z, z\right) | x, y \in \mathbb{R}\right\}$$

Si se toma el vector $(-5, -3,2)$, que es un vector normal al plano, se obtiene que la ecuación buscada es

$$-5x - 3y + 2y = 0$$

Proyección de un vector sobre un subespacio

De acuerdo al ejemplo anterior, es posible establecer relaciones entre un vector y un subespacio (en el caso del ejemplo, representado por un plano). Sin embargo, ese vector no necesariamente es perpendicular al subespacio; es posible que ese vector sea oblicuo, y por lo tanto, permita establecer una 'sombra' sobre el subespacio. Este concepto se conoce como proyección de un vector sobre un subespacio; la figura 4.5 establece ese concepto.

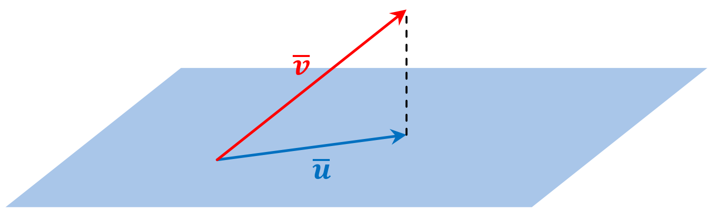

Figura 4.5. Proyección de un vector sobre un subespacio (plano).

Para obtener la proyección de un vector sobre un subespacio se debe considerar el producto interno que se utilizará y el subespacio sobre el cual se proyectará. Sea el conjunto

$$\{\bar{b}_1, \bar{b}_2, \bar{b}_3, \dots, \bar{b}_n\}$$

Una base ortogonal del subespacio H mostrado en la figura 4.5. El vector $\bar{u}$, puede expresarse como

$$\bar{u} = \frac{\langle \bar{u} | \bar{b}_1 \rangle}{\langle \bar{b}_1 | \bar{b}_1 \rangle} \bar{b}_1 + \frac{\langle \bar{u} | \bar{b}_2 \rangle}{\langle \bar{b}_2 | \bar{b}_2 \rangle} \bar{b}_2 + \frac{\langle \bar{u} | \bar{b}_3 \rangle}{\langle \bar{b}_3 | \bar{b}_3 \rangle} \bar{b}_3 + \dots + \frac{\langle \bar{u} | \bar{b}_n \rangle}{\langle \bar{b}_n | \bar{b}_n \rangle} \bar{b}_n \dots (1)$$

Pero el vector $\bar{u}$ puede expresarse como

$$\bar{u} = \bar{v} - \bar{d} \dots (2)$$

Al sustituir (2) en la parte derecha de (1) se obtiene

$$\bar{u} = \frac{\langle \bar{v} - \bar{d} | \bar{b}_1 \rangle}{\langle \bar{b}_1 | \bar{b}_1 \rangle} \bar{b}_1 + \frac{\langle \bar{v} - \bar{d} | \bar{b}_2 \rangle}{\langle \bar{b}_2 | \bar{b}_2 \rangle} \bar{b}_2 + \frac{\langle \bar{v} - \bar{d} | \bar{b}_3 \rangle}{\langle \bar{b}_3 | \bar{b}_3 \rangle} \bar{b}_3 + \dots + \frac{\langle \bar{v} - \bar{d} | \bar{b}_n \rangle}{\langle \bar{b}_n | \bar{b}_n \rangle} \bar{b}_n$$

Al desarrollar la anterior ecuación con propiedades del producto interno se tiene que

$$\bar{u} = \frac{\langle \bar{v} | \bar{b}_1 \rangle}{\langle \bar{b}_1 | \bar{b}_1 \rangle} \bar{b}_1 - \frac{\langle \bar{d} | \bar{b}_1 \rangle}{\langle \bar{b}_1 | \bar{b}_1 \rangle} \bar{b}_1 + \frac{\langle \bar{v} | \bar{b}_2 \rangle}{\langle \bar{b}_2 | \bar{b}_2 \rangle} \bar{b}_2 - \frac{\langle \bar{d} | \bar{b}_2 \rangle}{\langle \bar{b}_2 | \bar{b}_2 \rangle} \bar{b}_2 + \dots + \frac{\langle \bar{v} | \bar{b}_n \rangle}{\langle \bar{b}_n | \bar{b}_n \rangle} \bar{b}_n - \frac{\langle \bar{d} | \bar{b}_n \rangle}{\langle \bar{b}_n | \bar{b}_n \rangle} \bar{b}_n$$

Debido a que el subespacio H es el complemento ortogonal del vector $\bar{d}$, los productos $\langle \bar{d} | \bar{b}_i \rangle$ son iguales a cero, y en consecuencia

$$\frac{\langle \bar{d} | \bar{b}_i \rangle}{\langle \bar{b}_i | \bar{b}_i \rangle} \bar{b}_i = \bar{0}$$

Por lo que el vector $\bar{u}$ se expresa como

$$\bar{u}=\frac{\langle\bar{v}|\bar{b}_1\rangle}{\langle\bar{b}_1|\bar{b}_1\rangle}\bar{b}_1+\frac{\langle\bar{v}|\bar{b}_2\rangle}{\langle\bar{b}_2|\bar{b}_2\rangle}\bar{b}_2+\frac{\langle\bar{v}|\bar{b}_3\rangle}{\langle\bar{b}_3|\bar{b}_3\rangle}\bar{b}_3+\cdots+\frac{\langle\bar{v}|\bar{b}_n\rangle}{\langle\bar{b}_n|\bar{b}_n\rangle}\bar{b}_n$$

Esta expresión es conocida como la proyección de $\bar{v}$ sobre W, denotada como

$$\mathrm{Proy}_W\,\bar{v}=\sum_{i=1}^{n}\frac{\langle\bar{v}|\bar{b}_i\rangle}{\langle\bar{b}_i|\bar{b}_i\rangle}\bar{b}_i$$

EJEMPLO 4.24. Obténgase la proyección de la matriz $A=\begin{pmatrix}1&0\\0&2\end{pmatrix}$ sobre el subespacio

$$X=\left\{\begin{pmatrix}a&b\\-b&a\end{pmatrix}\middle|a,b\in\mathbb{R}\right\}$$

Con el producto interno definido por $\langle A|B\rangle=\mathrm{tr}(A^TB)$.

Una base ortogonal del subespacio es $\left\{B_1=\begin{pmatrix}1&0\\0&1\end{pmatrix},B_2=\begin{pmatrix}0&1\\-1&0\end{pmatrix}\right\}$, por lo tanto, la proyección de A sobre X será

$$\begin{aligned}\mathrm{Proy}_X\,A&=\frac{\langle A|B_1\rangle}{\langle B_1|B_1\rangle}B_1+\frac{\langle A|B_2\rangle}{\langle B_2|B_2\rangle}B_2\\&=\frac{3}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}+\frac{0}{2}\begin{pmatrix}0&1\\-1&0\end{pmatrix}\\\mathrm{Proy}_X\,A&=\frac{3}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}\end{aligned}$$

que es el vector buscado.

En la figura 4.5 se presenta un triángulo rectángulo, cuyos lados son los vectores $\bar{u}$, $\bar{v}$ y $\bar{d}$; la hipotenusa $\bar{v}$ puede expresarse con la ecuación vectorial

$$\bar{v}=\bar{d}+\bar{u}$$

donde $\bar{d}$ es el vector distancia entre $\bar{u}$ y $\bar{v}$, y $\bar{u}$ es la proyección de $\bar{v}$ en el subespacio H; obsérvese que H es el complemento ortogonal de $\bar{d}$; en consecuencia, $\bar{u}$ es ortogonal a $\bar{d}$. Por lo que, un vector $\bar{v}$ cualquiera puede expresarse de la siguiente manera

$$\bar{v}=\bar{u}+\bar{w}\qquad\forall\,\bar{u}\in S,\bar{w}\in S^{\perp}$$

Siendo S un subespacio y $S^{\perp}$ su complemento ortogonal.

EJEMPLO 4.25. Exprésese el vector $(2,-1,0,1)$ como la suma de dos vectores.

$$(2,-1,0,1)=\bar{w}_0+\bar{w}_0^{\perp}$$

Si $\bar{w}_0$ pertenece al subespacio $W=\{(0,x,y,z)|x,y,z\in\mathbb{R}\}$, entonces la proyección se da como

$$\begin{aligned}\bar{w}_0&=\frac{(2,-1,0,1)\cdot(0,1,0,0)}{(0,1,0,0)\cdot(0,1,0,0)}(0,1,0,0)+\frac{(2,-1,0,1)\cdot(0,0,1,0)}{(0,0,1,0)\cdot(0,0,1,0)}(0,0,1,0)+\frac{(2,-1,0,1)\cdot(0,0,0,1)}{(0,0,0,1)\cdot(0,0,0,1)}(0,0,0,1)\\&=-1(0,1,0,0)+0(0,0,1,0)+(0,0,0,1)\end{aligned}$$

$= (0, -1, 0, 1)$

Para encontrar el complemento ortogonal de $\bar{w}_0$:

$$0 = (0, -1, 0, 1) \cdot (w, x, y, z)$$

Por lo tanto, $-x + z = 0$ y el complemento ortogonal es $\{(w, x, y, x) | w, x, y \in \mathbb{R}\}$. Para terminar el cálculo, se debe realizar una ecuación sencilla

$$(2, -1, 0, 1) = (0, -1, 0, 1) + (w, x, y, x)$$

De donde se obtiene que

$$(2, -1, 0, 1) - (0, -1, 0,1) = (w, x, y, x)$$

$$\bar{w}_0^{\perp} = (2, 0, 0, 0)$$

Que son los dos vectores buscados: $(2, -1, 0, 1) = (0, -1, 0, 1) + (2, 0, 0, 0)$

El teorema de proyección

En la figura 4.6 puede observarse la proyección de un vector $\bar{v}$ sobre un vector $\bar{u}$; además, se dibujan las respectivas distancias entre $d(\bar{v}, \bar{u})$ y $d(\bar{v}, \bar{u}_0)$. Es evidente que la distancia entre $\bar{v}$ y su proyección es la menor de las dos.

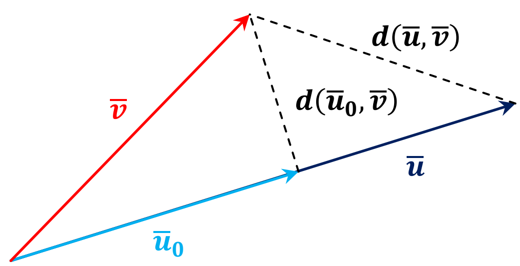

Figura 4.6. Distancia mínima entre vectores.

Esto permite establecer el teorema de proyección, el cual se enuncia a continuación.

Para cada vector $\bar{v} \in V$ existe uno y sólo un vector $\bar{u}_0$ tal que

$$\|\bar{v} - \bar{u}_0\| < \|\bar{v} - \bar{u}\|$$

Este teorema se le conoce como teorema de proyección. Puede observarse que la distancia más corta se da entre el vector $\bar{v}$ y su proyección $\bar{u}_0$.

Ejemplo 4.26. Sea el subespacio $W = \left\{\begin{pmatrix} a & b \\ -b & 2a \end{pmatrix} \middle| a, b \in \mathbb{R}\right\}$ de las matrices cuadradas de orden dos con elementos reales, donde se define el producto interno

$$\langle A|B \rangle = \mathrm{tr}(A^T B)$$

Obténgase la matriz más cercana a $N = \begin{pmatrix} 2 & -1 \\ -1 & 0 \end{pmatrix}$.

Para obtener la proyección, es necesario tener una base ortogonal, la cual es

$$\left\{\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}, \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\right\}$$

y se obtiene la protección de N sobre W, que es

$$\begin{aligned}\mathrm{Proy}_W\,N &= \frac{2}{5}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix} + \frac{0}{2}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\\ &= \frac{2}{5}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}\end{aligned}$$

La matriz más próxima a N es

$$\frac{2}{5}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$$

EJEMPLO 4.27. Dado el subespacio

$$G = \{(a+bi)x^2 + (ai)x + (a+2b) | a, b \in \mathbb{R}\}$$

Encuéntrese el vector $\bar{g}_0 \in G$ más próximo a $p(x) = 2ix + 5$, utilizando el producto interno

$$\langle \bar{p} | \bar{q} \rangle = a_0\bar{b}_0 + a_1\bar{b}_1 + a_2\bar{b}_2 \qquad\qquad \forall\, p(x) = a_0 + a_1x + a_2x^2, q(x) = b_0 + b_1x + b_2x^2$$

donde $\bar{b}_i$ denota al conjugado de b_i. Una base ortogonal de G es $B = \{ix^2 + 2, (5-2i)x^2 + 5ix + 1\}$ y la proyección de $p(x)$ es

$$\begin{aligned}\bar{g}_0 &= \frac{\langle 2ix + 5 | ix^2 + 2\rangle}{\langle ix^2 + 2 | ix^2 + 2\rangle}(ix^2 + 2) + \frac{\langle 2ix + 5 | (5-2i)x^2 + 5ix + 1\rangle}{\langle (5-2i)x^2 + 5ix + 1 | (5-2i)x^2 + 5ix + 1\rangle}[(5-2i)x^2 + 5ix + 1]\\ &= \frac{5(2)}{1+4}(ix^2 + 2) + \frac{2i(-5i) + 5(1)}{29 + 25 + 1}[(5-2i)x^2 + 5ix + 1]\\ &= 2(ix^2 + 2) + \left(\frac{3}{11}\right)[(5-2i)x^2 + 5ix + 1]\\ &= (2ix^2 + 4) + \left[\left(\frac{15}{11} - \frac{6}{11}i\right)x^2 + \frac{15}{11}ix + \frac{3}{11}\right] \Rightarrow \left(\frac{15}{11} + \frac{16}{11}i\right)x^2 + \frac{15}{11}ix + \frac{47}{11}\end{aligned}$$

Por lo que el vector más cercano buscado es $\bar{g}_0 = \left(\frac{15}{11} + \frac{16}{11}i\right)x^2 + \frac{15}{11}ix + \frac{47}{11}$.

Mínimos cuadrados

Es una técnica de análisis numérico encuadrada dentro de la optimización matemática, en la que, dados un conjunto de pares (o ternas, etc.), se intenta encontrar la función que mejor se aproxime a los datos (un mejor ajuste), de acuerdo con el criterio de mínimo error cuadrático. La figura 4.7 muestra el concepto de mínimos cuadrados.

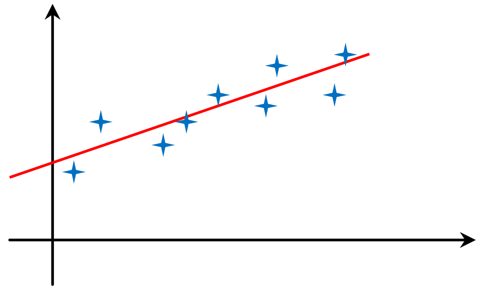

Figura 4.7. Ajuste, por mínimos cuadrados, de un conjunto de puntos a una función.

Tomando como base un sistema de m ecuaciones lineales y n incógnitas $A\bar{x} = \bar{b}$, el vector de términos independientes puede pertenecer al espacio columna de la matriz de coeficientes, dado que el producto $A\bar{x} = \bar{b}$ es una combinación lineal de columnas de A:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n \end{pmatrix} =$$

$$\begin{pmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{pmatrix} + \begin{pmatrix} a_{12}x_2 \\ a_{22}x_2 \\ \vdots \\ a_{m2}x_2 \end{pmatrix} + \begin{pmatrix} a_{13}x_3 \\ a_{23}x_3 \\ \vdots \\ a_{m3}x_3 \end{pmatrix} + \cdots + \begin{pmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{pmatrix} =$$

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \\ \vdots \\ a_{m3} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} =$$

Si la condición expuesta no se cumple, entonces el sistema de ecuaciones no tiene una solución exacta; pero sí es posible obtener una aproximación a ella. Dicha solución, denotada como $\bar{x}'$, es aquélla que minimizará el error $E = \left\|A\bar{x}' - \bar{b}\right\|$; en otras palabras, es la distancia mínima entre la solución aproximada y la solución real. La figura 4.8 establece la relación entre las dos soluciones.

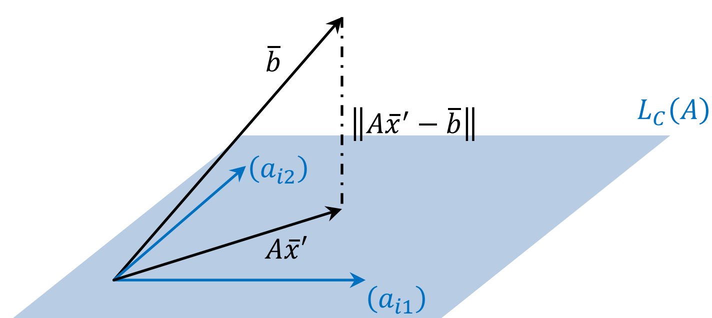

Figura 4.8. Reducción del error entre la solución real $\bar{x}'$ y la aproximada $A\bar{x}'$ de un sistema de ecuaciones.

La figura muestra que la solución aproximada es la proyección del vector de términos independientes sobre el espacio columna de la matriz del sistema; por lo tanto, el vector $A\bar{x}' - \bar{b}$ pertenece al complemento ortogonal de $L_C(A)$.

Supóngase que $\bar{w} \in L_C(A)$. Esto quiere decir, que dicho vector es una combinación lineal de las columnas de A; es decir,

$$\bar{w} = A\bar{c}$$

donde $\bar{c} = (c_1, c_2, c_3, \dots, c_n)$. Entonces, se obtiene que

$$\begin{aligned} \langle \bar{w} | A\bar{x}' - \bar{b} \rangle &= 0 \\ \langle A\bar{c} | A\bar{x}' - \bar{b} \rangle &= \end{aligned}$$

Utilizando el producto interno usual en $\mathbb{R}^n$,

$$\begin{aligned} \langle A\bar{c} | A\bar{x}' - \bar{b} \rangle &= 0 \\ (A\bar{c})^T (A\bar{x}' - \bar{b}) &= \\ \bar{c}^T A^T (A\bar{x}' - \bar{b}) &= \\ \bar{c}^T (A^T A\bar{x}' - A^T\bar{b}) &= \end{aligned}$$

donde, por el concepto de distancia mínima, se puede considerar que

$$\begin{aligned} A^T A\bar{x}' - A^T\bar{b} &= \bar{0} \\ A^T A\bar{x}' &= A^T\bar{b} \end{aligned}$$

Éstas son conocidas como las ecuaciones normales de Gauss, y son la representación matricial del método de mínimos cuadrados.

EJEMPLO 4.28. Regresión Lineal. Sea un conjunto de puntos obtenidos mediante medición

$$\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)\}$$

Se les puede ajustar al comportamiento de una recta, de ecuación

$$y = mx + b$$

Para obtener el valor de los coeficientes m y b, es posible ajustar a un sistema de ecuaciones lineales cuya solución no es exacta

$$\begin{array}{cccc} mx_1 & +b & = & y_1 \\ mx_2 & +b & = & y_2 \\ mx_3 & +b & = & y_3 \\ & & \vdots & \\ mx_n & +b & = & y_n \end{array}$$

cuyo arreglo matricial es

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}\begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_4 \end{pmatrix}$$

Al aplicar las ecuaciones normales $A^T A\bar{x}' = A^T\bar{b}$ se puede reducir la ecuación:

$$\begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix}\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}\begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix}$$

$$\begin{pmatrix} x_1^2 + x_2^2 + \cdots + x_n^2 & x_1 + x_2 + \cdots + x_n \\ x_1 + x_2 + \cdots + x_n & 1 + 1 + \cdots + 1 \end{pmatrix}\begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} x_1y_1 + x_2y_2 + \cdots + x_ny_n \\ y_1 + y_2 + \cdots + y_n \end{pmatrix}$$

$$\begin{pmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{pmatrix}\begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} x_iy_i \\ \sum_{i=1}^{n} y_i \end{pmatrix}$$

que son las ecuaciones conocidas para aplicar la regresión lineal.

EJEMPLO 4.29. Determínese la recta que mejor se ajusta al conjunto de puntos presentados en la tabla 4.1.

X	1	2	3	4	5	6	7
Y	0.5	2.5	2.0	4.0	3.5	6.0	5.5

Tabla 4.1. Datos del ejemplo 4.31.

Con los datos de la tabla 4.1 se puede construir el siguiente sistema de ecuaciones:

$$\begin{array}{rcrcl} m & +b & = & 0.5 \\ 2m & +b & = & 2.5 \\ 3m & +b & = & 2.0 \\ 4m & +b & = & 4.0 \\ 5m & +b & = & 3.5 \\ 6m & +b & = & 6.0 \\ 7m & +b & = & 5.5 \end{array} \Rightarrow \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \\ 6 & 1 \\ 7 & 1 \end{pmatrix}\begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} 0.5 \\ 2.5 \\ 2.0 \\ 4.0 \\ 3.5 \\ 6.0 \\ 5.5 \end{pmatrix}$$

Utilizando las ecuaciones normales $A^T A\bar{x}' = A^T\bar{b}$, se pueden obtener los valores de la pendiente y la ordenada al origen de la recta buscada.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \\ 6 & 1 \\ 7 & 1 \end{pmatrix}\begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0.5 \\ 2.5 \\ 2.0 \\ 4.0 \\ 3.5 \\ 6.0 \\ 5.5 \end{pmatrix}$$

$$\begin{pmatrix} 140 & 28 \\ 28 & 7 \end{pmatrix}\begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} 119.5 \\ 24 \end{pmatrix}$$

Cuya solución es $m = 0.8393$ y $b = 0.0714$. Por lo tanto, la ecuación de la recta buscada es

$$y = 0.8393x + 0.0714$$

El polinomio de primer grado no es el único lugar geométrico al cual se puede ajustar un conjunto de puntos. Cualquier función lineal se puede ajustar satisfactoriamente por el método de mínimos cuadrados; incluso, con cambios de variable adecuados, se puede encontrar una función no-lineal.

EJEMPLO 4.30. Encuéntrese la ecuación de segundo grado que mejor se ajusta al conjunto de puntos mostrado en la tabla 4.2.

X	- 1.1	- 0.4	0.2	0.9	1.8	2.7	3.0	3.8
Y	6.51	3.36	1.44	0.11	- 0.16	1.19	2.00	5.04

Tabla 4.2. Datos del ejemplo 4.32.

El sistema de ecuaciones se construye con el polinomio de segundo grado $y = ax^2 + bx + c$. Eso significa que se debe obtener un nuevo conjunto de datos, utilizando el cuadrado de los datos correspondientes a las abscisas; la tabla 4.3 muestra esos datos.

X	- 1.1	- 0.4	0.2	0.9	1.8	2.7	3.0	3.8
X^2	1.21	0.16	0.04	0.81	3.24	7.29	9.00	14.44
Y	6.51	3.36	1.44	0.11	- 0.16	1.19	2.00	5.04

Tabla 4.3. Datos con la expansión de datos del ejemplo 4.32.

Con estos datos se construye el sistema de ecuaciones lineales

$$\begin{array}{cccccc} 1.21a & -1.1b & +c & = & 6.51 \\ 0.16a & -0.4b & +c & = & 3.36 \\ 0.04a & +0.2b & +c & = & 1.44 \\ 0.81a & +0.9b & +c & = & 0.11 \\ 3.24a & +1.8b & +c & = & -0.16 \\ 7.29a & +2.7b & +c & = & 1.19 \\ 9.00a & +3.0b & +c & = & 2.00 \\ 14.44a & +3.8b & +c & = & 5.04 \end{array} \Rightarrow \begin{pmatrix} 1.21 & -1.1 & 1 \\ 0.16 & -0.4 & 1 \\ 0.04 & 0.2 & 1 \\ 0.81 & 0.9 & 1 \\ 3.24 & 1.8 & 1 \\ 7.29 & 2.7 & 1 \\ 9.00 & 3.0 & 1 \\ 14.44 & 3.8 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 6.51 \\ 3.36 \\ 1.44 \\ 0.11 \\ -0.16 \\ 1.19 \\ 2.00 \\ 5.04 \end{pmatrix}$$

Que se utiliza en las ecuaciones normales

$$A^T = \begin{pmatrix} 1.21 & 0.16 & 0.04 & 0.81 & 3.24 & 7.29 & 9.00 & 14.44 \\ -1.1 & -0.4 & 0.2 & 0.9 & 1.8 & 2.7 & 3.0 & 3.8 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$A^T\begin{pmatrix}1.21 & -1.1 & 1\\ 0.16 & -0.4 & 1\\ 0.04 & 0.2 & 1\\ 0.81 & 0.9 & 1\\ 3.24 & 1.8 & 1\\ 7.29 & 2.7 & 1\\ 9.00 & 3.0 & 1\\ 14.44 & 3.8 & 1\end{pmatrix}\begin{pmatrix}a\\ b\\ c\end{pmatrix}=A^T\begin{pmatrix}6.51\\ 3.36\\ 1.44\\ 0.11\\ -0.16\\ 1.19\\ 2.00\\ 5.04\end{pmatrix}$$

Al realizar las multiplicaciones se obtiene el sistema de ecuaciones

$$\begin{pmatrix}355.3027 & 106.7290 & 36.1900\\ 106.7290 & 36.1900 & 10.9000\\ 36.1900 & 10.9000 & 8.0000\end{pmatrix}\begin{pmatrix}a\\ b\\ c\end{pmatrix}=\begin{pmatrix}107.4957\\ 19.9590\\ 19.4900\end{pmatrix}$$

Cuya solución es $a=1$, $b=-3$ y $c=2$. Por lo tanto, el polinomio que mejor se ajusta al conjunto de puntos dado es

$$y=x^2-3x+2$$