Índice general

Capítulo 6

Espacios vectoriales con producto interior

Durante las primeras semanas de clases trabajamos con vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$ y calculamos, con ayuda del teorema de Pitágoras, la longitud de vectores de $\mathbb{R}^2$ y $\mathbb{R}^3$ (observa las figuras 6.1 y 6.2), a esta longitud le llamamos *norma euclidiana*: la norma euclidiana de un vector $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ es

$$\|x\| = \sqrt{x_1^2 + x_2^2},$$

mientras que la norma euclidiana de $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$ es

$$\|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

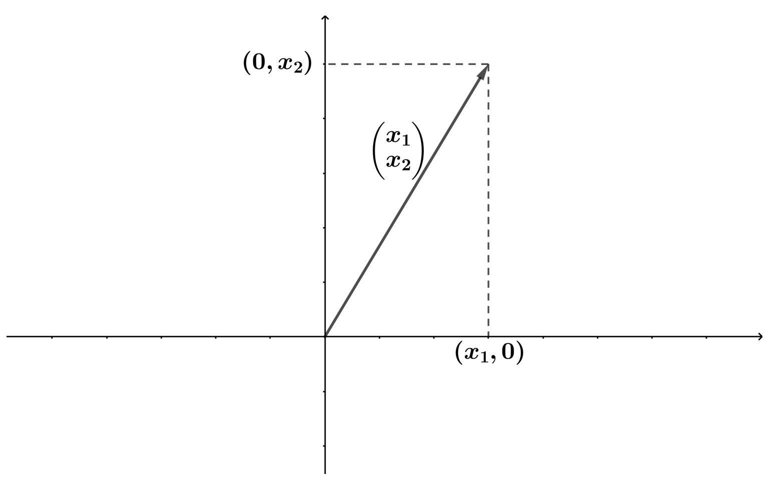

Figura 6.1: Norma de vectores de $\mathbb{R}^2$.

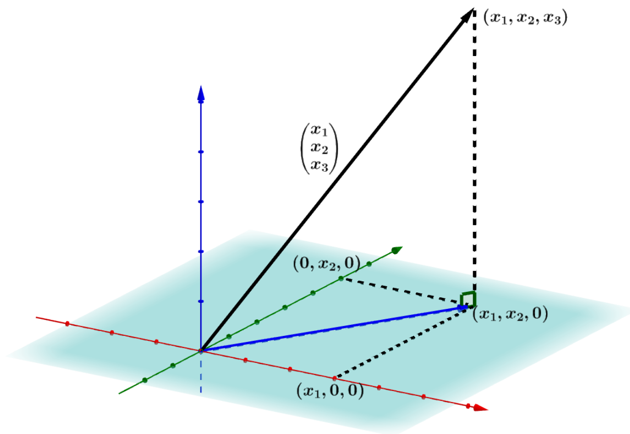

Figura 6.2: Norma de vectores de $\mathbb{R}^3$.

Volviendo a lo que aprendimos en Álgebra 1, observa ahora la figura 6.3. El ángulo α es el ángulo entre el vector u y la parte positiva del eje X, β es el ángulo entre v y la parte positiva del eje X, mientras que θ es el ángulo entre u y v, $\theta = \beta - \alpha$.

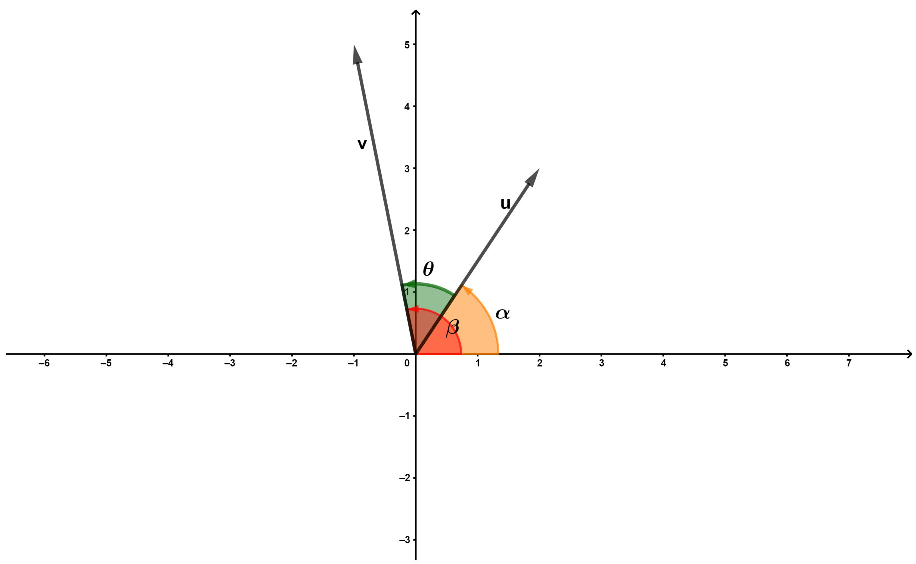

Figura 6.3: Ángulo entre vectores de $\mathbb{R}^2$.

Supongamos que $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$,

$$\|u\| = \sqrt{u_1^2 + u_2^2}, \qquad \|v\| = \sqrt{v_1^2 + v_2^2}.$$

Recordando la forma en que definimos las funciones seno y coseno se tiene que

$$\cos\alpha = \frac{u_1}{\|u\|}, \quad \sin\alpha = \frac{u_2}{\|u\|}, \quad \cos\beta = \frac{v_1}{\|v\|}, \quad \sin\beta = \frac{v_2}{\|v\|}.$$

Entonces

$$\cos\theta = \cos(\beta - \alpha) = \cos\beta\cos\alpha + \sin\beta\sin\alpha = \frac{v_1}{\|v\|}\frac{u_1}{\|u\|} + \frac{v_2}{\|v\|}\frac{u_2}{\|u\|} = \frac{u_1v_1 + u_2v_2}{\|u\|\|v\|}.$$

El ángulo entre los vectores u y v es el cociente entre la suma $u_1v_1 + u_2v_2$ y el producto de las normas de u y v. A la suma $u_1v_1 + u_2v_2$ le llamamos *producto interior usual* entre los vectores u y v de $\mathbb{R}^2$.

En este último capítulo del curso comenzaremos definiendo qué es un *espacio vectorial con producto interior* y veremos que $\mathbb{R}^n$ es, con el producto interior usual entre vectores de $\mathbb{R}^n$ (generalización del producto interior entre vectores de $\mathbb{R}^2$), un espacio vectorial con producto interior.

Espacios vectoriales con producto interior son importantes porque en ellos es posible definir: ángulos entre vectores (generalizando lo que hicimos antes para calcular el ángulo entre u y v), normas (”longitud” o ”magnitud”) de vectores, distancias entre vectores y, lo más importante, proyecciones de vectores del espacio vectorial sobre un subespacio vectorial del mismo.

Comencemos entonces definiendo qué es un espacio vectorial con producto interior.

Observación 6.1. *En este capítulo solo trabajaremos con espacios vectoriales reales. Algunas definiciones las escribiremos para espacios vectoriales V sobre un cuerpo $\mathbb{K}$, pero luego mencionaremos cómo ella se simplifica si $\mathbb{K} = \mathbb{R}$ y en todos los ejemplos solo trabajaremos con e.v. sobre $\mathbb{R}$.*

Definición 6.2. Un espacio vectorial V sobre un cuerpo $\mathbb{K}$ es un *espacio vectorial con producto interior* si y solo si existe una función, que denotaremos mediante $\bullet$, que a cada par de vectores de V hace corresponder un escalar en $\mathbb{K}$ y tiene las siguientes propiedades:

(i) para todo par de vectores $u, v \in V$ se cumple que $u \bullet v = \overline{v \bullet u}$, donde $\overline{v \bullet u}$ es el conjugado del número complejo $v \bullet u$.
Si V es un e.v. real esta propiedad es $u \bullet v = v \bullet u$, es decir, en espacios vectoriales reales el producto interior es una operación conmutativa.
Por otro lado, si $u = v$, la propiedad $u \bullet v = \overline{v \bullet u}$ se escribe como

$$v \bullet v = \overline{v \bullet v},$$

lo que significa que, aún en espacios vectoriales complejos, el producto interior de un vector por sí mismo tiene que ser un número real,

$$\forall\, v \in V \ : \ v \bullet v \in \mathbb{R}.$$

Entonces tiene sentido escribir la siguiente propiedad que debe cumplir $\bullet$ para ser un producto interior.

(ii) el producto interior de todo vector por sí mismo es un número real mayor o igual que cero,

$$\forall u \in V \ : \ u \bullet u \geq 0.$$

Además el único vector cuyo producto por sí mismo es cero es el vector nulo de V,

$$\forall\, u \in V \ : \ u \bullet u = 0 \quad \Leftrightarrow \quad u = \theta_V.$$

(iii) Por último, para todo par de escalares $\alpha, \beta \in \mathbb{K}$ y vectores $u, v, w \in V$ debe cumplirse

$$(\alpha u + \beta v) \bullet w = \alpha\,(u \bullet w) + \beta\,(v \bullet w)\,.$$

Un producto interior muy usado es el llamado *producto interior usual* en $\mathbb{R}^n$.

Dados $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ e $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, el producto interior usual entre ellos es el número real

$$x \bullet y = x_1\, y_1 + x_2\, y_2 + \ldots + x_n\, y_n = \sum_{i=1}^{n} x_i\, y_i.$$

No es difícil demostrar que éste cumple las propiedades mencionadas en la definición anterior.

Ejemplo 6.3. *Sean* $x = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$, $y = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}$, $z = \begin{pmatrix} 1 \\ -5 \\ 2 \end{pmatrix}$, $u = \begin{pmatrix} -1 \\ 0 \\ \frac{1}{2} \end{pmatrix}$, *entonces*

$$x \bullet y = 0 + (-2) + 6 = 4, \qquad y \bullet z = 0 + (-10) + 6 = -4, \qquad z \bullet u = -1 + 1 = 0.$$

Observa que:

1. si V es un $\mathbb{K}$-espacio vectorial con producto interior $\bullet$ y θ_V es el elemento neutro de V, entonces el producto interior de cualquier vector de V por θ_V es cero y esto es cierto porque

$$u \bullet \theta_V = u \bullet (u - u) = \overline{(u - u) \bullet u} = \overline{u \bullet u - u \bullet u} = \overline{0} = 0.$$

2. Además el único vector de V cuyo producto interior con cualquier otro vector de V es cero es θ_V, es decir, si $u \in V$ es tal que $\forall\, v \in V \ :\ u \bullet v = 0$, entonces $u = \theta_V$. Esto es cierto pues si u es tal que el para todo $v \in V$, el producto $u \bullet v = 0$, entonces

$$u \bullet u = 0 \qquad \Rightarrow \qquad u = \theta_V.$$

3. Por último, en espacios vectoriales complejos es importante tener cuidado al calcular productos interiores pues, a diferencia de espacios vectoriales reales, en e.v. complejos el producto interior no es una operación conmutativa.
Nota que, según la definición de producto interior, se cumple que para vectores u, v, w y escalares α, β,

$$(\alpha u + \beta v) \bullet w = \alpha(u \bullet w) + \beta(v \bullet w),$$

pero, ¿a qué es igual $w \bullet (\alpha u + \beta v)$? En e.v. reales $w \bullet (\alpha u + \beta v) = \alpha(w \bullet u) + \beta(w \bullet v)$, pero esta igualdad no se cumple en e.v. complejos, veamos por qué

$$\begin{aligned} w \bullet (\alpha u + \beta w) &= \overline{(\alpha u + \beta v) \bullet w} = \overline{\alpha(u \bullet w) + \beta(v \bullet w)}, \\ &= \overline{\alpha}\, \overline{u \bullet w} + \overline{\beta}\, \overline{v \bullet w} = \overline{\alpha}(w \bullet u) + \overline{\beta}(w \bullet v). \end{aligned}$$

Si α y β son números reales, $\overline{\alpha} = \alpha$ y $\overline{\beta} = \beta$ y $w \bullet (\alpha u + \beta v) = \alpha(w \bullet u) + \beta(w \bullet v)$, pero si α y β son complejos, entonces $w \bullet (\alpha u + \beta v) = \overline{\alpha}(w \bullet u) + \overline{\beta}(w \bullet v)$.

Ejemplo 6.4. *Demostremos que si $A \in \mathcal{M}_2(\mathbb{R})$ es invertible, entonces la función $\bullet_A$ que a $x, y \in \mathbb{R}^2$ le hace corresponder el número real*

$$x \bullet_A y = (Ax) \bullet (Ay)$$

es un producto interior en $\mathbb{R}^2$.

Si A es la matriz identidad, entonces $x \bullet_I y$ es igual al producto interior usual en $\mathbb{R}^2$.

Si

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

entonces

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \bullet_A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_1 + x_2 \end{pmatrix} \bullet \begin{pmatrix} y_1 - y_2 \\ y_1 + y_2 \end{pmatrix} = (x_1 - x_2)(y_1 - y_2) + (x_1 + x_2)(y_1 + y_2).$$

Demostremos que la función $\bullet_A$ cumple las propiedades de un producto interior, pero no para la matriz A anterior, sino suponiendo que A es una matriz invertible de 2 filas y 2 columnas.

1. *La primera propiedad que debe cumplirse es: para cualquier par de vectores $x, y \in \mathbb{R}^2$, $x \bullet_A y$ debe ser igual a $y \bullet_A x$. Pero,*

   $$x \bullet_A y = (Ax) \bullet (Ay)$$

   y como $\bullet$ es un producto interior en $\mathbb{R}^2$, se cumple que $(Ax) \bullet (Ay) = (Ay) \bullet (Ax)$, por tanto, $x \bullet_A y = (Ax) \bullet (Ay) = (Ay) \bullet (Ax) = y \bullet_A x$ y con esto hemos demostrado la primera propiedad.
2. *Ahora debemos demostrar que para cualquier $x \in \mathbb{R}^2$ se cumple que $x \bullet_A x \geq 0$. Esto también se demuestra utilizando que $\bullet$ es un producto interior pues*

   $$x \bullet_A x = (Ax) \bullet (Ax)$$

   y como $\bullet$ es un producto interior, se tiene que $(Ax) \bullet (Ax) \geq 0$.
   La siguiente propiedad es un poco más difícil de demostrar y es la única en la que es importante nuestra suposición de que A es invertible. Debemos demostrar que $x \bullet_A x = 0$ si y solo si $x = \theta$. Supongamos que $x \in \mathbb{R}^2$ es tal que

   $$x \bullet_A x = (Ax) \bullet (Ax) = 0$$

   *como $\bullet$ es un producto interior, $(Ax) \bullet (Ax) = 0$ si y solo si $Ax = \theta$ y, **como A es invertible**, $Ax = \theta$ si y solo si $x = \theta$.*
   Nota, por ejemplo, que si

   $$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

   que no es invertible, entonces

   $$x \bullet_A x = \begin{pmatrix} x_1 + x_2 \\ x_1 + x_2 \end{pmatrix} \bullet \begin{pmatrix} x_1 + x_2 \\ x_1 + x_2 \end{pmatrix} = 2(x_1 + x_2)^2$$

   y, por ejemplo, $\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix} \bullet_A \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right) = 0$ sin ser $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ el vector nulo de $\mathbb{R}^2$. Para que $\bullet_A$ sea un producto interior en $\mathbb{R}^2$ es de vital importancia que A sea una matriz invertible.

3. *Solo nos resta demostrar la última propiedad: que para escalares reales cualesquiera α, β y vectores $x, y, z \in \mathbb{R}^2$ se cumple $(\alpha x + \beta y) \bullet_A z = \alpha(x \bullet_A z) + \beta(y \bullet_A z)$.*
   Utilizaremos las propiedades del producto matriz - vector y el producto escalar - vector para demostrar esta propiedad.

$$(\alpha x + \beta y) \bullet_A z = (A(\alpha x + \beta y)) \bullet (Az) = (\alpha(Ax) + \beta(Az)) \bullet (Az).$$

   Como $\bullet$ es producto interior

$$(\alpha(Ax) + \beta(Az)) \bullet (Az) = \alpha\left((Ax) \bullet (Az)\right) + \beta\left((Ax) \bullet (Az)\right) = \alpha(x \bullet_A z) + \beta(y \bullet_A z).$$

La demostración anterior es la misma si quisiéramos probar que, dada una matriz invertible $A \in \mathcal{M}_n(\mathbb{R})$, la función $\bullet_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ que a cada par de vectores $x, y \in \mathbb{R}^n$ hace corresponder el número real $x \bullet_A y = (Ax) \bullet (Ay)$ (siendo $\bullet$ el producto interior usual en $\mathbb{R}^n$), es un producto interior en $\mathbb{R}^n$.

Más adelante volveremos a trabajar con este producto interior y con el producto interior usual en $\mathbb{R}^n$.

Como mencionamos antes, una de las ventajas de trabajar con espacios vectoriales con producto interior es que en ellos es posible definir normas (longitudes) de vectores. Veamos primero qué propiedades debe cumplir una norma de vectores para luego ver por que en espacios vectoriales con producto interior siempre es posible definir una función con estas propiedades.

Definición 6.5. Si V es un espacio vectorial sobre cierto cuerpo $\mathbb{K}$, una función

$$\|\cdot\| : V \to \mathbb{R},$$

es una *norma* en V si y solo si

(i) $\forall\, u \in V \;:\; \|u\| \geq 0 \;\wedge\; \|u\| = 0 \Leftrightarrow u = \theta$,
(ii) $\forall\, \alpha \in \mathbb{K} \;:\; \forall\, u \in V \;:\; \|\alpha\, u\| = |\alpha|\|u\|$,
(iii) $\forall\, u, v \in V \;:\; \|u + v\| \leq \|u\| + \|v\|$.

Si V es un espacio vectorial con producto interior (e.v.p.i.), entonces la función que a cada vector $u \in V$ hace corresponder el número real $\|u\| = \sqrt{u \bullet u}$ es una norma en V y se denomina *norma inducida por el producto interior* $\bullet$.

Si $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, la *norma inducida por el producto interior usual* en $\mathbb{R}^n$ es entonces

$$\|x\| = \sqrt{x \bullet x} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}.$$

Esta norma es la *norma euclidiana* o *norma 2* de vectores de $\mathbb{R}^n$.

Ejemplo 6.6. *La norma euclidiana del vector* $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ *es* $\left\|\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}\right\| = \sqrt{1 + 1 + 4} = \sqrt{6}$.

Ejemplo 6.7. *¿Cuál es la norma inducida por el producto anterior $\bullet_A$ que definimos antes?*

Llamemos $\|\cdot\|_A$ a la norma inducida por este producto interior,

$$\|x\|_A = \sqrt{x \bullet_A x} = \sqrt{(Ax) \bullet (Ax)} = \|Ax\|.$$

Si, por ejemplo,

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

entonces

$$\left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\|_A = \|Ax\| = \left\| \begin{pmatrix} x_1 - x_2 \\ x_1 + x_2 \end{pmatrix} \right\| = \sqrt{(x_1 - x_2)^2 + (x_1 + x_2)^2}.$$

Mientras

$$\left\| \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\| = \sqrt{5},$$

se tiene que

$$\left\| \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\|_A = \sqrt{10}.$$

Demostremos que si $\bullet$ es un producto interior en un e.v. real V, entonces la función que a cada u de V le hace corresponder el número real $\sqrt{u \bullet u}$ cumple todas las propiedades de una norma.

1. Como $\bullet$ es un producto interior, para todo $u \in V$ se cumple que $u \bullet u \geq 0$ y además $u \bullet u = 0$ si y solo si $u = \theta_V$.
   Estas propiedades de $\bullet$ implican que

   $$\|u\| = \sqrt{u \bullet u} \geq 0$$

   y

   $$\|u\| = 0 \;\Leftrightarrow\; \sqrt{u \bullet u} = 0 \;\Leftrightarrow\; u \bullet u = 0 \;\Leftrightarrow\; u = \theta_V.$$

2. Además si $\alpha \in \mathbb{R}$ y $u \in V$

   $$\|\alpha u\| = \sqrt{(\alpha u) \bullet (\alpha u)} = \sqrt{\alpha (u) \bullet (\alpha u)}.$$

   Como V es un e.v. real, $u \bullet (\alpha u) = (\alpha u) \bullet u = \alpha (u \bullet u)$. Entonces,

   $$\|\alpha u\| = \sqrt{\alpha^2 (u \bullet u)} = |\alpha| \sqrt{u \bullet u} = |\alpha| \|u\|.$$

Nota que en la demostración anterior $\bullet$ no es necesariamente el producto interior usual en $\mathbb{R}^n$, sino cualquier producto interior en un e.v. real, $\bullet$ pudo ser, por ejemplo, el producto interior $\bullet_A$.

La tercera y última propiedad que debe cumplir la norma inducida por un producto interior es la desigualdad triangular. La demostración de esta propiedad es más sencilla si utilizamos una desigualdad, denominada desigualdad de Cauchy-Schwarz, que la cumplen cualquier producto interior y la norma inducida por él.

Lema 6.8 (Desigualdad de Cauchy – Schwarz)**.** Sea V un e.v. real. Para cualquier par de vectores $u, v \in V$, se cumple que

$$|u \bullet v| \leq \|u\| \, \|v\|,$$

donde $\|u\| = \sqrt{u \bullet u}$ y $\|v\| = \sqrt{v \bullet v}$.

Demostración. Sean $u, v \in V$ vectores distintos del vector nulo. Observa que si uno de ellos es el vector nulo la desigualdad de Cauchy-Schwarz es cierta, por tanto, nos resta demostrarla para el caso en que ambos son distintos del vector nulo.

Dado que el producto interior de cualquier vector por sí mismo es un número mayor o igual que cero, entonces para todo número real λ se cumple que

$$(u - \lambda v) \bullet (u - \lambda v) \geq 0.$$

Teniendo en cuenta las propiedades del producto interior se tiene que

$$(u - \lambda v) \bullet (u - \lambda v) = \lambda^2 \|v\|^2 - 2\,\lambda\,(u \bullet v) + \|u\|^2$$

y, por tanto, completando cuadrados se cumple que para todo par de vectores u, v distintos del vector nulo y todo $\lambda \in \mathbb{R}$,

$$\lambda^2 \|v\|^2 - 2\lambda (u \bullet v) + \|u\|^2 \geq 0 \quad \Leftrightarrow \quad \left(\lambda - \frac{u \bullet v}{\|v\|^2}\right)^2 + \frac{\|u\|^2 \|v\|^2 - (u \bullet v)^2}{\|v\|^4} \geq 0.$$

Dado que esta última expresión es válida para cualquier $\lambda \in \mathbb{R}$, si se considera en particular el escalar $\lambda = \dfrac{u \bullet v}{\|v\|^2}$, el primero de los términos a la izquierda de la desigualdad anterior es cero y ella se reduce a

$$\begin{aligned} \frac{\|u\|^2 \|v\|^2 - (u \bullet v)^2}{\|v\|^4} \geq 0 \quad & \Leftrightarrow \quad \|u\|^2 \|v\|^2 - (u \bullet v)^2 \geq 0, \\ & \Leftrightarrow \quad \|u\|^2 \|v\|^2 \geq (u \bullet v)^2 , \\ & \Leftrightarrow \quad \|u\| \|v\| \geq |(u \bullet v)| , \end{aligned}$$

lo cual demuestra la desigualdad de Cauchy-Schwarz. □

Demostremos ahora que cualquier norma inducida por un producto interior satisface la desigualdad triangular (la última de las propiedades de una norma en la definición de norma). Si $u, v \in V$, entonces

$$\begin{aligned} \|u + v\|^2 &= (u + v) \bullet (u + v), & \\ &= u \bullet (u + v) + v \bullet (u + v), & \text{propiedad de producto interior} \\ &= u \bullet u + u \bullet v + v \bullet u + v \bullet v, & \\ &= \|u\|^2 + 2\, u \bullet v + \|v\|^2, & \text{definición de norma inducida por producto interior.,} \\ & & \text{En e.v. reales } u \bullet v = v \bullet u, \\ &\leq \|u\|^2 + 2\, |u \bullet v| + \|v\|^2, & \text{cualquier número real es menor o igual que su valor absoluto} \\ &\leq \|u\|^2 + 2\, \|u\| \, \|v\| + \|v\|^2, & \text{hemos utilizado la desigualdad de Cauchy-Schwarz} \\ &= (\, \|u\| + \|v\| \,)^2, & \end{aligned}$$

de donde se concluye que
$$\|u+v\| \leq \|u\| + \|v\|.$$

La norma inducida por un producto interior satisface la siguiente propiedad, que se denomina *Regla del Paralelogramo*: para todo par de vectores u, v en un e.v. real V con producto interior se cumple que
$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2,$$
si $\|\cdot\|$ es la norma inducida por el producto interior en V. Demostremos esta igualdad

$$\begin{aligned}\|u+v\|^2 + \|u-v\|^2 &= (u+v) \bullet (u+v) + (u-v) \bullet (u-v),\\ &= \|u\|^2 + u \bullet v + v \bullet u + \|v\|^2 + \|u\|^2 - u \bullet v - v \bullet u + \|v\|^2,\\ &= 2\|u\|^2 + 2\|v\|^2.\end{aligned}$$

Si V es un e.v. con producto interior, éste puede utilizarse para definir una norma de vectores, pero también podemos definir normas de vectores sin necesidad de tener un producto interior. Por ejemplo, la función que a cada $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ hace corresponder el número

$$\|x\|_1 = |x_1| + |x_2| + \ldots + |x_n| \tag{6.1}$$

es una norma en $\mathbb{R}^n$.

Una norma de vectores puede utilizarse para definir una distancia entre ellos.

Veamos primero qué es una distancia entre vectores de un espacio vectorial. La definición siguiente es una generalización de la que escribimos a comienzos de semestre para $\mathbb{R}^n$.

Definición 6.9. Sea V un e.v. sobre $\mathbb{K}$.

Una *distancia* en V es una función que a un par de vectores de V hace corresponder un número real mayor o igual que cero, usualmente denotada como $\mathrm{d} : V \times V \longrightarrow \mathbb{R}$, que cumple las siguientes propiedades:

1. $\forall\, x, y \in V \ :\ \mathrm{d}(x,y) \geq 0,$
2. $\mathrm{d}(x,y) = 0 \quad \Leftrightarrow \quad x = y,$
3. $\forall\, x, y \in V \ :\ \mathrm{d}(x,y) = \mathrm{d}(y,x),$
4. $\forall\, x, y, z \in V \ :\ \mathrm{d}(x,y) \leq \mathrm{d}(x,z) + \mathrm{d}(z,y).$

Si $\|\cdot\|$ es una norma en V, la función que a cada par de vectores $x, y \in V$ le hace corresponder la norma de la diferencia entre ellos cumple con todas las propiedades mencionadas anteriormente, es decir,
$$\mathrm{d}(x,y) = \|x-y\| = \sqrt{(x-y) \bullet (x-y)}$$

satisface las propiedades necesarias para ser una distancia entre vectores de un e.v. real V.

Con la norma euclidiana definida antes tendríamos que si $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ e $y = (y_1, y_2, \ldots, y_n)^{\mathrm{T}}$, entonces

$$\mathrm{d}(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

Esta distancia se denomina *distancia euclidiana.*

Ejemplo 6.10. *La distancia euclidiana entre los vectores* $x = (1, 2, 1)^T$ *y* $y = (-1, 2, 2)^T$ *está dada por,*

$$d(x, y) = \|x - y\| = \left\| \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} \right\| = \sqrt{(2)^2 + 0^2 + (-1)^2} = \sqrt{5}.$$

Con la norma $\|\cdot\|_1$, definida en (6.1), también podemos definir una distancia entre vectores de $\mathbb{R}^n$: si $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ e $y = (y_1, y_2, \ldots, y_n)^{\mathrm{T}}$, entonces

$$\mathrm{d}_1(x, y) = \|x - y\|_1 = \sum_{i=1}^{n} |x_i - y_i|.$$

Ejemplo 6.11. *Con la distancia definida por* $\|\cdot\|_1$, *la distancia entre los vectores* $x = (1, 2, 1)^T$ *y* $y = (-1, 2, 2)^T$ *está dada por,*

$$d_1(x, y) = \|x - y\|_1 = \left\| \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} \right\| = |2| + |0| + |-1| = 3.$$

6.1. Ortogonalidad

Definición 6.12. Dos vectores u, v de un espacio vectorial con producto interior se denominan *ortogonales* si y solo si $u \bullet v = 0$. La notación $u \perp v$ indica que los vectores u y v son ortogonales.

Si u y v son ortogonales, ellos cumplen la igualdad de Pitágoras,

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

La demostración se basa en las propiedades de producto interior:

$$\|u + v\|^2 = (u + v) \bullet (u + v) = u \bullet (u + v) + v \bullet (u + v) = \|u\|^2 + u \bullet v + v \bullet u + \|v\|^2 = \|u\|^2 + \|v\|^2.$$

Definición 6.13. Un conjunto de vectores $\{v_1, v_2, \ldots, v_n\}$ de un espacio vectorial V se dice *ortogonal* si y solo si todos los vectores del conjunto son distintos del vector nulo de V y todo par de vectores del conjunto son ortogonales, es decir,

$$(\forall i, j \in \{1, 2, \ldots, n\} \; : \; i \neq j \to v_i \bullet v_j = 0) \; \wedge \; (\forall\, i \in \{1, 2, \ldots, n\} \; : \; \|v_i\| \neq 0).$$

Un conjunto de vectores $\{v_1, v_2, \ldots, v_n\}$ de un espacio vectorial V se denomina *ortonormal* si y solo si él es ortogonal y además $\forall\, i \in \{1, 2, \ldots, n\} \; : \; \|v_i\| = 1$.

Ejemplo 6.14. *La base canónica de $\mathbb{R}^2$ es, con respecto al producto interior usual en $\mathbb{R}^2$ y la norma inducida por él, un conjunto ortonormal de vectores.*

Sin embargo, ella no es un conjunto ortonormal de vectores si consideramos el producto interior $\bullet_A$ con

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

pues

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \bullet_A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \bullet \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1 \neq 0.$$

El conjunto $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{2} \\ 1 \end{pmatrix} \right\}$ *sí es un conjunto ortogonal de vectores de $\mathbb{R}^2$ con respecto a $\bullet_A$ pues*

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \bullet_A \begin{pmatrix} -\frac{1}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \bullet \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = 0.$$

Nota que este conjunto es ortogonal (con respecto a $\bullet_A$), pero no ortonormal pues

$$\left\| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\|_A = \left\| \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\| = \sqrt{2} \neq 1.$$

No es difícil comprobar que el conjunto $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{2} \\ 1 \end{pmatrix} \right\}$ *es li y, por tanto, base de $\mathbb{R}^2$.*

En el siguiente lema veremos que cualquier conjunto ortogonal es li, es decir, si tenemos un conjunto ortogonal de n vectores de $\mathbb{R}^n$, podemos asegurar que este conjunto es una base de $\mathbb{R}^n$.

Lema 6.15. Sea V un e.v. sobre $\mathbb{K}$, con producto interior. Suponga que $\{v_1, v_2, \ldots, v_n\}$ es un conjunto ortogonal de vectores de V. Entonces

1. $\{v_1, v_2, \ldots, v_n\}$ es li.
2. $\left\{ \frac{1}{\|v_1\|} v_1, \frac{1}{\|v_2\|} v_2, \ldots, \frac{1}{\|v_n\|} v_n \right\}$ es un conjunto ortonormal.
3. Si $\mathcal{A} = \{v_1, v_2, \ldots, v_n\}$ es un conjunto ortogonal de vectores de V y V tiene dimensión n, entonces $\{v_1, v_2, \ldots, v_n\}$ es una base de V y para todo $v \in V$,

$$v = \sum_{j=1}^{n} \frac{v \bullet v_j}{\|v_j\|^2} v_j,$$

es decir, para todo $v \in V$ se cumple que

$$[v]_{\mathcal{A}} = \begin{pmatrix} \dfrac{v \bullet v_1}{\|v_1\|^2} \\ \dfrac{v \bullet v_2}{\|v_2\|^2} \\ \vdots \\ \dfrac{v \bullet v_n}{\|v_n\|^2} \end{pmatrix}.$$

Si $\mathcal{A}$ fuera ortonormal, entonces $\mathcal{A}$ es ortogonal y también es una base de V, pero además para todo $v \in V$ se cumple que

$$[v]_{\mathcal{A}} = \begin{pmatrix} v \bullet v_1 \\ v \bullet v_2 \\ \vdots \\ v \bullet v_n \end{pmatrix}.$$

Demostración. Demostremos cada uno de los puntos anteriores.

1. $\{v_1, v_2, \ldots, v_n\}$ es li.
   Para probarlo supongamos que existe una combinación lineal de ellos que es igual al vector nulo de V, es decir, existen $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ tales que $\sum_{i=1}^{n} \alpha_i v_i = \theta_V$.
   Entonces, como el producto de cualquier vector de V por $\sum_{i=1}^{n} \alpha_i v_i$ es igual a cero, podemos asegurar que para cada $j \in \{1, 2, \ldots, n\}$ se tiene que

$$0 = \left(\sum_{i=}^{n} \alpha_i v_i \right) \bullet v_j = \sum_{i=1}^{n} \alpha_i (v_i \bullet v_j) = \alpha_j (v_j \bullet v_j) = \alpha_j \|v_j\|^2.$$

   Dado que $\forall\, j \in \{1, 2, \ldots, n\} \,:\, \|v_j\| \neq 0$, la igualdad anterior se cumple si y solo si $\alpha_j = 0$.
   Con esto se tiene entonces que los únicos valores de $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ para los que $\sum_{i=1}^{n} \alpha_i v_i = \theta_V$ son $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, lo que significa que $\{v_1, v_2, \ldots, v_n\}$ es un conjunto li.
2. El conjunto

$$\left\{ \frac{1}{\|v_1\|} v_1, \frac{1}{\|v_2\|} v_2, \ldots, \frac{1}{\|v_n\|} v_n \right\}$$

   es ortonormal si y solo si él es ortogonal y además la norma de todos sus vectores es 1.
   Dado que $\{v_1, v_2, \ldots, v_n\}$ es ortogonal, para cada par de índices $i, j \in \{1, 2, \ldots, n\}$ distintos entre sí, se cumple que

$$\left(\frac{1}{\|v_i\|} v_i \right) \bullet \left(\frac{1}{\|v_j\|} v_j \right) = \frac{1}{\|v_i\| \|v_j\|} (v_i \bullet v_j) = 0,$$

   además como $\{v_1, v_2, \ldots, v_n\}$ es ortogonal, ninguno de los vectores en el conjunto es el vector nulo de V y, por tanto, los vectores en $\left\{ \frac{1}{\|v_1\|} v_1, \frac{1}{\|v_2\|} v_2, \ldots, \frac{1}{\|v_n\|} v_n \right\}$ también son todos distintos del vector nulo de V.
   Por último, para cada $i \in \{1, 2, \ldots, n\}$ se tiene que

$$\left\| \frac{1}{\|v_i\|} v_i \right\| = \frac{1}{\|v_i\|} \|v_i\| = 1.$$

3. Como $\mathcal{A}$ es ortogonal, $\mathcal{A}$ es li. Como V tiene dimensión n y $\mathcal{A}$ es li, $\mathcal{A}$ es base de V.
   Entonces para todo $v \in V$ existen $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tales que

$$v = \sum_{i=1}^{n} \alpha_i v_i$$

Supongamos que $j \in \{1, 2, \ldots, n\}$. Entonces el producto interior de v por v_j es tal que

$$v \bullet v_j = \left(\sum_{i=1}^{n} \alpha_i v_i \right) \bullet v_j = \alpha_j (v_j \bullet v_j) = \alpha_j \|v_j\|^2.$$

De la igualdad $v \bullet v_j = \alpha_j \|v_j\|^2$, podemos despejar α_j y obtenemos

$$\alpha_j = \frac{v \bullet v_j}{\|v_j\|^2}$$

y, con ello,

$$v \in V \;\Rightarrow\; v = \frac{v \bullet v_1}{\|v_1\|^2} v_1 + \frac{v \bullet v_2}{\|v_2\|^2} v_2 + \ldots + \frac{v \bullet v_n}{\|v_n\|^2} v_n.$$

Si $\mathcal{A}$ es ortonormal, entonces la los vectores en $\mathcal{A}$ tienen norma 1 y

$$v = (v \bullet v_1) v_1 + (v \bullet v_2) v_2 + \ldots + (v \bullet v_n) v_n.$$

□

6.2. Proyección ortogonal de $u \in V$ sobre $W \subset V$

Si V es un espacio vectorial con producto interior y W es un s.e.v. de V, entonces es posible, como veremos en esta sección, determinar cuál es el vector en W más cercano a un cierto vector $v \in V$ que no esté en W. Esto tiene múltiples aplicaciones, al final de esta sección veremos una de ellas.

Definición 6.16. Sea V un espacio vectorial real con producto interior.

Supongamos que W es un subespacio vectorial de V y que $\{w_1, w_2, \ldots, w_r\}$ es una base de W.

Un vector $v \in V$ es ortogonal a W si y solo si v es ortogonal a cada vector de W y esto, a su vez, se cumple si y solo si v es ortogonal a los vectores $w_1, w_2, \ldots, w_r$.

Se denomina *proyección ortogonal de un elemento $u \in V$ sobre W* al vector $w^* \in W$ tal que $u - w^*$ es ortogonal a W.

Consideremos primero $W = \langle \{w_1\} \rangle$. Veamos cómo determinar la proyección ortogonal de $u \in V$ sobre W. Llamémosle w^* a este vector. Como él debe pertenecer a W, entonces $w^* = \alpha w_1$ con $\alpha \in \mathbb{R}$, es decir,

$$w^* \in W \;\Rightarrow\; \exists \alpha \in \mathbb{R} \;:\; w^* = \alpha\, w_1.$$

El vector w^* además debe cumplir que $u - w^*$ es ortogonal a W, esto se cumple si y solo si $u - w^*$ es ortogonal a w_1

$$\begin{aligned}
u - w^* \perp W \;&\Leftrightarrow\; u - w^* \perp w_1, \\
&\Leftrightarrow\; (u - w^*) \bullet w_1 = 0, \\
&\Leftrightarrow\; (u - \alpha\, w_1) \bullet w_1 = 0, \\
&\Leftrightarrow\; u \bullet w_1 - \alpha(w_1 \bullet w_1) = 0, \\
&\Leftrightarrow\; \alpha = \frac{u \bullet w_1}{\|w_1\|^2}.
\end{aligned}$$

La proyección ortogonal de u sobre W es entonces el vector $w^* = \dfrac{u \bullet w_1}{\|w_1\|^2} w_1$.

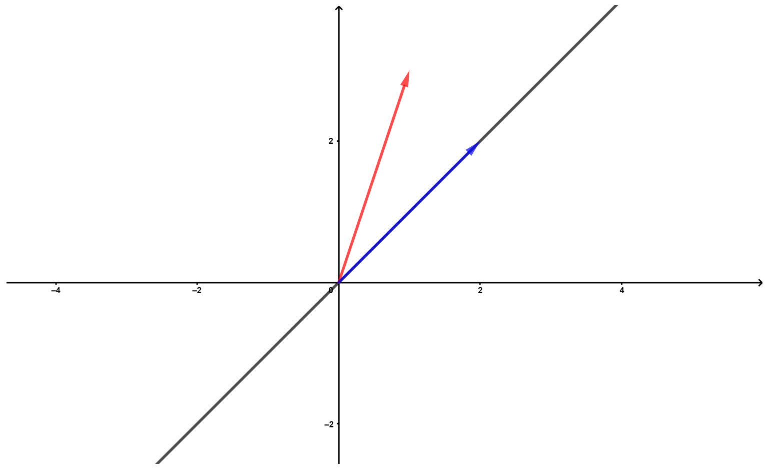

Figura 6.4: La proyección del vector rojo sobre el s.e.v. de $\mathbb{R}^2$ representado en color negro es el vector azul.

Ejemplo 6.17. *Consideremos el espacio vectorial $\mathbb{R}^2$, con el producto interior usual, y llamemos W al subespacio vectorial de él generado por el vector $\begin{pmatrix}1\\1\end{pmatrix}$.*

Calculemos la proyección ortogonal del vector $\begin{pmatrix}1\\3\end{pmatrix}$, que no pertenece a W, sobre W, es decir, determinemos cuál es el vector de la forma $\begin{pmatrix}\alpha\\\alpha\end{pmatrix}$, $\alpha \in \mathbb{R}$ para el que se cumple que $\begin{pmatrix}1\\3\end{pmatrix} - \alpha\begin{pmatrix}1\\1\end{pmatrix}$ es ortogonal a $\begin{pmatrix}1\\1\end{pmatrix}$.

$$\begin{aligned}
\left(\begin{pmatrix}1\\3\end{pmatrix} - \alpha\begin{pmatrix}1\\1\end{pmatrix}\right) \perp \begin{pmatrix}1\\1\end{pmatrix} \;&\Leftrightarrow\; \left(\begin{pmatrix}1\\3\end{pmatrix} - \alpha\begin{pmatrix}1\\1\end{pmatrix}\right) \bullet \begin{pmatrix}1\\1\end{pmatrix} = 0,\\
&\Leftrightarrow\; \begin{pmatrix}1\\3\end{pmatrix} \bullet \begin{pmatrix}1\\1\end{pmatrix} - \alpha\begin{pmatrix}1\\1\end{pmatrix} \bullet \begin{pmatrix}1\\1\end{pmatrix} = 0,\\
&\Leftrightarrow\; 4 - \alpha\left\|\begin{pmatrix}1\\1\end{pmatrix}\right\|^2 = 0,\\
&\Leftrightarrow\; \alpha = 2.
\end{aligned}$$

Ocurre entonces que la proyección ortogonal de $\begin{pmatrix}1\\3\end{pmatrix}$ sobre W es el vector $\begin{pmatrix}2\\2\end{pmatrix}$.

En la figura 6.4 están representados W, $\begin{pmatrix}1\\3\end{pmatrix}$ y su proyección ortogonal sobre W con respecto al producto interior usual.

¿Qué ocurre si escogemos otro producto interior en $\mathbb{R}^2$? La proyección ortogonal de $\begin{pmatrix}1\\3\end{pmatrix}$ sobre W

es un vector diferente. Calculémoslo. Tomemos, por ejemplo, el producto interior $\bullet_A$ *con*

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

$$\begin{aligned}
\left(\begin{pmatrix} 1 \\ 3 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) \perp \begin{pmatrix} 1 \\ 1 \end{pmatrix} &\Leftrightarrow \left(\begin{pmatrix} 1 \\ 3 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) \bullet_A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0, \\
&\Leftrightarrow \begin{pmatrix} 1 \\ 3 \end{pmatrix} \bullet_A \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} \bullet_A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0, \\
&\Leftrightarrow 9 - \alpha \left\| \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\|_A^2 = 0, \\
&\Leftrightarrow 9 - 5\alpha = 0 \Leftrightarrow \alpha = \frac{9}{5}.
\end{aligned}$$

Ocurre entonces que la proyección ortogonal de $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ *sobre* W*, considerando el producto interior* $\bullet_A$ *en* $\mathbb{R}^2$*, en lugar del producto interior usual, es el vector* $\begin{pmatrix} \frac{9}{5} \\ \frac{9}{5} \end{pmatrix}$.

Calculemos ahora la proyección de $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ *sobre* W *con respecto al producto interior usual. Nota (figura 6.5) que* $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ *es ortogonal a* W *pues*

$$\begin{pmatrix} -1 \\ 1 \end{pmatrix} \bullet \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0.$$

¿Qué vector es la proyección ortogonal de $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ *sobre* W*. Antes de seguir leyendo, observa la figura 6.5, ¿qué vector crees que resulte?*

La proyección ortogonal de $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ *sobre* W *es el vector de la forma* $\begin{pmatrix} \alpha \\ \alpha \end{pmatrix}$, $\alpha \in \mathbb{R}$ *para el que se cumple que* $\begin{pmatrix} -1 \\ 1 \end{pmatrix} - \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}$ *es ortogonal a* W*:*

$$\begin{aligned}
\left(\begin{pmatrix} -1 \\ 1 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) \perp \begin{pmatrix} 1 \\ 1 \end{pmatrix} &\Leftrightarrow \left(\begin{pmatrix} -1 \\ 1 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) \bullet \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0, \\
&\Leftrightarrow \begin{pmatrix} -1 \\ 1 \end{pmatrix} \bullet \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} \bullet \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0, \\
&\Leftrightarrow 0 - \alpha \left\| \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\|_A^2 = 0, \\
&\Leftrightarrow 0 - 2\alpha = 0 \Leftrightarrow \alpha = 0.
\end{aligned}$$

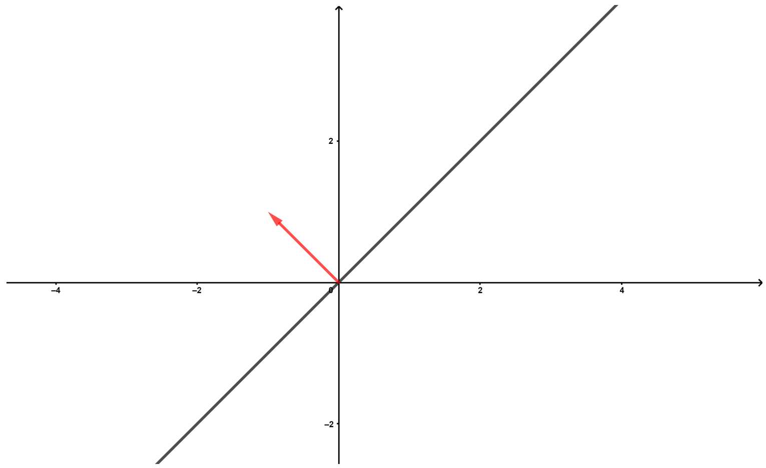

Figura 6.5: ¿Cuál debe ser la proyección ortogonal del vector rojo sobre s.e.v. de $\mathbb{R}^2$ representado con una recta negra?

La proyección ortogonal de $\begin{pmatrix}-1\\1\end{pmatrix}$ *(que es ortogonal a* W*) sobre* W *es el vector nulo de* $\mathbb{R}^2$*.*

Nota que $\begin{pmatrix}-1\\1\end{pmatrix}$ *no es ortogonal a* W *con respecto a* $\bullet_A$ *pues*

$$\begin{pmatrix}-1\\1\end{pmatrix}\bullet_A\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}\bullet\begin{pmatrix}1\\2\end{pmatrix}\neq 0,$$

por tanto, la proyección ortogonal de $\begin{pmatrix}-1\\1\end{pmatrix}$ *sobre* W*, con respecto a* $\bullet_A$*, no es el vector nulo de* $\mathbb{R}^2$*.*

¿Cuál crees que sea la proyección ortogonal de un vector en W *sobre* W*? Debería ser el propio vector, ¿cierto? Demostrémoslo.*

Supongamos que queremos proyectar el vector $\begin{pmatrix}a\\a\end{pmatrix}\in W$ *sobre* W*. El resultado es el vector de la forma* $\begin{pmatrix}\alpha\\\alpha\end{pmatrix}$ *para el que se cumple que*

$$\left(\begin{pmatrix}a\\a\end{pmatrix}-\begin{pmatrix}\alpha\\\alpha\end{pmatrix}\right)\bullet\begin{pmatrix}1\\1\end{pmatrix}=0.$$

Como

$$\left(\begin{pmatrix}a\\a\end{pmatrix}-\begin{pmatrix}\alpha\\\alpha\end{pmatrix}\right)=(a-\alpha)\begin{pmatrix}1\\1\end{pmatrix},$$

el producto anterior es cero si y solo si

$$(a-\alpha)\left\|\begin{pmatrix}1\\1\end{pmatrix}\right\|^2=0$$

y esto ocurre si y solo si $\alpha=a$*.*

Sea ahora $W = \langle\{w_1, w_2, \ldots, w_r\}\rangle$. La proyección ortogonal de $u \in V$ sobre W es el vector $w^* \in W$ para el cual se cumple

$$u - w^* \perp W.$$

Si $\{w_1, w_2, \ldots, w_r\}$ es una base de W, entonces $u - w^* \perp W$, es decir, de modo que

$$u - w^* \perp w_1 \ \wedge \ u - w^* \perp w_2 \ \wedge \ \cdots \ \wedge \ u - w^* \perp w_r.$$

Como $w^* \in W$, w^* es combinación lineal de $w_1, w_2, \ldots, w_r$ con escalares $\alpha_1, \alpha_2, \ldots, \alpha_r$ y nuestra tarea consiste en determinar $\alpha_1, \alpha_2, \ldots, \alpha_r \in \mathbb{R}$ de modo que

$$u - w^* \perp w_1 \ \Leftrightarrow \qquad \alpha_1\|w_1\|^2 + \alpha_2(w_2 \bullet w_1) + \ldots \alpha_r(w_r \bullet w_1) = u \bullet w_1, \tag{6.2a}$$

$$u - w^* \perp w_2 \ \Leftrightarrow \qquad \alpha_1(w_1 \bullet w_2) + \alpha_2\|w_2\|^2 + \ldots \alpha_r(w_r \bullet w_2) = u \bullet w_2, \tag{6.2b}$$

$$\vdots$$

$$u - w^* \perp w_r \ \Leftrightarrow \qquad \alpha_1(w_1 \bullet w_r) + \alpha_2(w_2 \bullet w_r) + \ldots \alpha_r\|w_r\|^2 = u \bullet w_r. \tag{6.2c}$$

Resolviendo ese sistema de r ecuaciones (es un sistema compatible determinado, aunque no lo demostraremos en este curso) podemos determinar la proyección ortogonal de u sobre W.

Ejemplo 6.18. *Consideremos el e.v. $\mathbb{R}^3$, con el producto interior usual y sea*

$$W = \left\{ \begin{pmatrix} x \\ y \\ 2x+y \end{pmatrix} \ : \ x, y \in \mathbb{R} \right\} = \left\langle \left\{ \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\} \right\rangle.$$

El vector $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ no pertenece a W. Calculemos su proyección ortogonal sobre W. Éste es el vector

$$\alpha \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \ \alpha, \beta \in \mathbb{R}$$

tal que

$$\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} - \beta \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right) \bullet \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = 0 \quad y \quad \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} - \beta \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right) \bullet \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 0.$$

En la figura 6.6 se muestran W (en color celeste) y el vector $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \notin W$.

Los valores de $\alpha, \beta \in \mathbb{R}$ que nos permiten determinar la proyección de $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ sobre W son entonces la solución del sistema de ecuaciones lineales

$$5\alpha + 2\beta = 3,$$
$$2\alpha + 2\beta = 2,$$

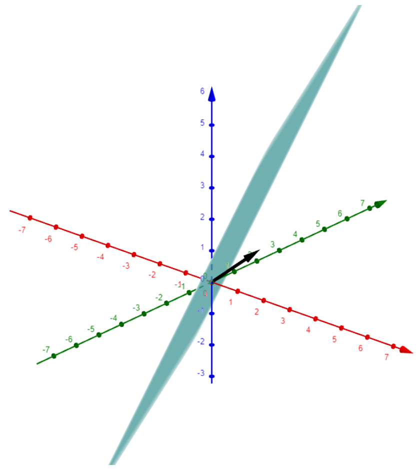

Figura 6.6: El vector $\begin{pmatrix}1\\1\\1\end{pmatrix}$ no es elemento del s.e.v. de $\mathbb{R}^3$ dibujado en color celeste

es decir, $\alpha = \frac{1}{3}$, $\beta = \frac{2}{3}$.

La proyección de $\begin{pmatrix}1\\1\\1\end{pmatrix}$ *sobre* W *es entonces el vector*

$$\frac{1}{3}\begin{pmatrix}1\\0\\2\end{pmatrix} + \frac{2}{3}\begin{pmatrix}0\\1\\1\end{pmatrix} = \begin{pmatrix}\frac{1}{3}\\\frac{2}{3}\\\frac{4}{3}\end{pmatrix}.$$

En la figura 6.7 puedes observar al vector $\begin{pmatrix}1\\1\\1\end{pmatrix}$ *y su proyección ortogonal sobre* W.

Si $\{w_1, \ldots, w_r\}$ hubiera sido una base ortogonal de W (dada una base de un e.v.p.i es posible encontrar una base ortogonal para él, pero no veremos cómo en este curso, si te interesa, puedes buscar el algoritmo de ortogonalización de Gramm-Schmidt), el sistema (6.3) se reduce a

$$u - w^* \perp w_1 \Leftrightarrow \qquad \alpha_1\|w_1\|^2 = u \bullet w_1, \tag{6.3a}$$
$$u - w^* \perp w_2 \Leftrightarrow \qquad \alpha_2\|w_2\|^2 = u \bullet w_2, \tag{6.3b}$$
$$\vdots$$
$$u - w^* \perp w_r \Leftrightarrow \qquad \alpha_r\|w_r\|^2 = u \bullet w_r, \tag{6.3c}$$

pues para cada par de índices $i, j \in \{1, 2, \ldots, r\}$ distintos entre sí el producto $w_i \bullet w_j = 0$.

El vector w^*, proyección ortogonal de u sobre W es entonces el vector

$$w^* = \sum_{i=1}^{r} \frac{u \bullet w_i}{\|w_i\|^2} w_i.$$

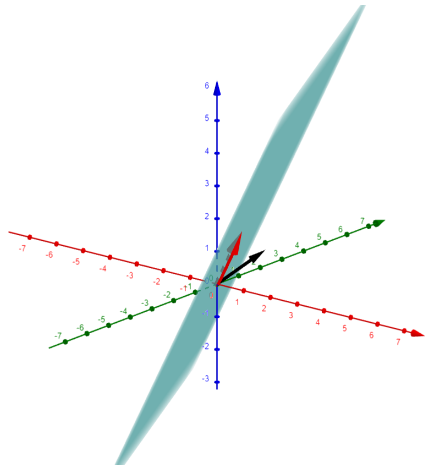

Figura 6.7: El vector $\begin{pmatrix}1\\1\\1\end{pmatrix}$ no es elemento del s.e.v. de $\mathbb{R}^3$ dibujado en color celeste. Su proyección sobre el s.e.v. es el vector rojo.

Ejemplo 6.19. *El conjunto*

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

es base del s.e.v. de $\mathbb{R}^3$ *que llamamos* W *en el ejemplo anterior, pero ésta no es una base ortogonal de* W *pues*

$$\begin{pmatrix}1\\0\\2\end{pmatrix}\cdot\begin{pmatrix}0\\1\\1\end{pmatrix}=2\neq 0.$$

Sin embargo, si reemplazamos el segundo vector del conjunto anterior por el vector

$$\begin{pmatrix}0\\1\\1\end{pmatrix}-\frac{2}{5}\begin{pmatrix}1\\0\\2\end{pmatrix}=\begin{pmatrix}-\frac{2}{5}\\1\\\frac{1}{5}\end{pmatrix},$$

obtenemos una nueva base para W

$$\left\{\begin{pmatrix}1\\0\\2\end{pmatrix},\begin{pmatrix}-\frac{2}{5}\\1\\\frac{1}{5}\end{pmatrix}\right\},$$

que es un conjunto ortogonal.

La proyección ortogonal de $\begin{pmatrix}1\\1\\1\end{pmatrix}$ *sobre* W *es más sencilla de calcular con esta nueva base pues ésta es el vector*

$$\alpha\begin{pmatrix}1\\0\\2\end{pmatrix}+\beta\begin{pmatrix}-\frac{2}{5}\\1\\\frac{1}{5}\end{pmatrix}$$

con

$$\alpha = \frac{\begin{pmatrix}1\\1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\0\\2\end{pmatrix}}{\left\|\begin{pmatrix}1\\0\\2\end{pmatrix}\right\|^2} = \frac{3}{5}, \quad \beta = \frac{\begin{pmatrix}1\\1\\1\end{pmatrix}\cdot\begin{pmatrix}-\frac{2}{5}\\1\\\frac{1}{5}\end{pmatrix}}{\left\|\begin{pmatrix}-\frac{2}{5}\\1\\\frac{1}{5}\end{pmatrix}\right\|^2} = \frac{2}{3}.$$

La proyección ortogonal de $\begin{pmatrix}1\\1\\1\end{pmatrix}$ *sobre* W *es el mismo vector que calculamos antes,*

$$\frac{3}{5}\begin{pmatrix}1\\0\\2\end{pmatrix} + \frac{2}{3}\begin{pmatrix}-\frac{2}{5}\\1\\\frac{1}{5}\end{pmatrix} = \begin{pmatrix}\frac{1}{3}\\\frac{2}{3}\\\frac{4}{3}\end{pmatrix}.$$

¿Por qué es importante calcular la proyección ortogonal de un vector u en un cierto e.v. V sobre un s.e.v. W de V? Porque ese vector es, de entre los vectores en W, el más cercano a u. Si w es un vector cualesquiera de W y w^* es la proyección ortogonal de u sobre W, entonces podemos calcular la distancia entre u y w (o entre u y w^*) como $\|u-w\|$ (ésta es la norma inducida por el producto interior). Observa que entonces

$$\begin{aligned}\|u-w\|^2 &= \|u-w^*+w^*-w\|^2 = (u-w^*+w^*-w)\cdot(u-w^*+w^*-w)\,,\\ &= (u-w^*)\cdot(u-w^*) + (u-w^*)\cdot(w^*-w) + (w^*-w)\cdot(u-w^*) + (w^*-w)\cdot(w^*-w)\,.\end{aligned}$$

El segundo y tercer términos en la suma anterior son iguales a cero porque $u-w^* \perp W$ y $w^*-w \in W$. Por tanto,

$$\|u-w\|^2 = \|u-w^*\|^2 + \|w^*-w\|^2$$

y, con esto se cumple que $\forall\, w \in W \;:\; \|u-w\| \geq \|u-w^*\|$.

Una de las aplicaciones de lo que hemos aprendido en este capítulo es la solución, *en el sentido de los mínimos cuadrados*, de un sistema sobredeterminado de ecuaciones lineales.

Un *sistema sobredeterminado de ecuaciones lineales* es uno en el que el número de incógnitas es inferior al número de ecuaciones que, en forma matricial, puede escribirse como: dados $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ y $b \in \mathbb{R}^m$, con $m > n$, determinar $x \in \mathbb{R}^n$ de modo que $Ax = b$.

En un sistema de este tipo, si las columnas de A son un conjunto li de vectores de $\mathbb{R}^m$, ellas son una base de un s.e.v. de $\mathbb{R}^m$ de dimensión n. Si $b \notin \operatorname{im}(A)$, el sistema es incompatible, es decir, no tiene solución.

La solución de un problema así, *en el sentido de los mínimos cuadrados* consiste en

1. determinar el vector $b^* \in \operatorname{im}(A)$ más cercano a b, es decir, determinar la proyección ortogonal de b sobre $\operatorname{im}(A)$,

2. resolver $Ay = b^*$, que sí tiene solución, en lugar de $Ay = b$. Nota que y es el vector cuyas componentes son los escalares que permiten escribir a b^* como combinación lineal de las columnas de A y éstos los determinamos en el paso anterior.

De este modo, como b^* es el vector en $\operatorname{im}(A)$ más cercano a b se cumple que para todo $w \in \operatorname{im}(A)$

$$\|b^* - b\| \leq \|w - b\|,$$

y como $b^* = Ay$, entonces el vector y determinado es tal que para todo $x \in \mathbb{R}^n$ se cumple que

$$\|Ay - b\| \leq \|Ax - b\|.$$

Ejemplo 6.20. *Encontremos la solución del siguiente sistema de ecuaciones lineales en el sentido de los mínimos cuadrados,*

$$\begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

La matriz ampliada del sistema anterior es tal que

$$\left(\begin{array}{cc|c} 1 & -2 & 4 \\ 1 & -1 & 2 \\ 1 & 0 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 1 \end{array}\right) \begin{array}{c} f_2 \leftarrow f_2 - f_1 \\ f_3 \leftarrow f_3 - f_1 \\ \sim \\ f_4 \leftarrow f_4 - f_1 \\ f_5 \leftarrow f_5 - f_1 \end{array} \left(\begin{array}{cc|c} 1 & -2 & 4 \\ 0 & 1 & -2 \\ 0 & 2 & -3 \\ 0 & 4 & -3 \\ 0 & 5 & -3 \end{array}\right) \begin{array}{c} f_3 \leftarrow f_3 - 2f_2 \\ \sim \\ f_4 \leftarrow f_4 - 4f_2 \\ f_5 \leftarrow f_5 - 5f_2 \end{array} \left(\begin{array}{cc|c} 1 & -2 & 4 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \\ 0 & 0 & 5 \\ 0 & 0 & 7 \end{array}\right).$$

El vector $b = \begin{pmatrix} 4 \\ 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ *no pertenece al espacio generado por las columnas de A.*

Determinemos su proyección ortogonal sobre el espacio generado por las columnas de A con respecto al producto interior usual. La proyección ortogonal de b sobre $\operatorname{im}(A)$ *es el vector en* $\operatorname{im}(A)$, *es decir, el vector de la forma*

$$\alpha \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} -2 \\ -1 \\ 0 \\ 2 \\ 3 \end{pmatrix}, \quad \alpha, \beta \in \mathbb{R}$$

para el que se cumple que $b - b^*$ es ortogonal a $\operatorname{im}(A)$, *es decir, α y β deben cumplir*

$$\left(\begin{pmatrix} 4 \\ 2 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \beta \begin{pmatrix} -2 \\ -1 \\ 0 \\ 2 \\ 3 \end{pmatrix} \right) \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = 0 \quad y \quad \left(\begin{pmatrix} 4 \\ 2 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \beta \begin{pmatrix} -2 \\ -1 \\ 0 \\ 2 \\ 3 \end{pmatrix} \right) \cdot \begin{pmatrix} -2 \\ -1 \\ 0 \\ 2 \\ 3 \end{pmatrix} = 0$$

Los valores de α *y* β *son, por tanto, la solución del sistema de ecuaciones lineales*

$$\begin{aligned} 5\alpha + 2\beta &= 9, \\ 2\alpha + 18\beta &= -5, \end{aligned}$$

es decir, $\alpha = 2$ *y* $\beta = -\frac{1}{2}$.

Nota que ahora el vector y *tal que* $Ay = b^*$ *es precisamente* $y = \begin{pmatrix} 2 \\ -\frac{1}{2} \end{pmatrix}$ *pues*

$$Ay = 2\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} -2 \\ -1 \\ 0 \\ 2 \\ 3 \end{pmatrix} \quad \Leftrightarrow \quad y = \begin{pmatrix} 2 \\ -\frac{1}{2} \end{pmatrix}.$$

Este vector y *es tal que para todo* $x \in \mathbb{R}^2$ *se cumple que* $\|Ay - b\|$ *es menor o igual que* $\|Ax - b\|$, *donde* $\|\cdot\|$ *es la norma inducida por el producto interior usual que es el que hemos utilizado para calcular la proyección ortogonal de* b *sobre* $\mathrm{im}(A)$.