

Capítulo 4

El producto escalar

4.1. Definición y propiedades

Todos conocemos el *producto escalar* en \mathbb{R}^n que hemos usado en cursos pasados con $n = 2$ y $n = 3$. Para $n > 3$ se extiende de la forma obvia:

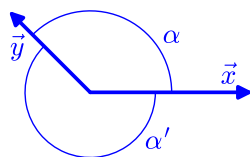
$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \quad \text{con} \quad \vec{x} = (x_1, x_2, \dots, x_n)^t, \quad \vec{y} = (y_1, y_2, \dots, y_n)^t.$$

Pensando los vectores como matrices columna, se tiene $\vec{x} \cdot \vec{y} = \vec{x}^t \vec{y}$. Otra notación medianamente común para el producto escalar es $\langle \vec{x}, \vec{y} \rangle$, sobre todo para generalizaciones de $\vec{x} \cdot \vec{y}$ como las que veremos después.

Esta operación tan sencilla de multiplicar las coordenadas respectivas y sumar los resultados es interesante porque está relacionada con medir longitudes y ángulos. Concretamente, la *norma* (*longitud*) de un vector \vec{x} y el *ángulo* α entre dos vectores \vec{x} e \vec{y} vienen dados por las fórmulas

$$(4.1) \quad \|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}} \quad \text{y} \quad \cos \alpha = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|} \quad \text{con} \quad 0 \leq \alpha \leq \pi.$$

Geoméricamente, dos vectores en el plano dibujados como flechas determinan dos ángulos (en sentido antihorario) que suman 2π , el que va de \vec{x} a \vec{y} y el que va de \vec{y} a \vec{x} , por convenio se toma siempre el más pequeño y de ahí la restricción $0 \leq \alpha \leq \pi$.



$$\alpha' = 2\pi - \alpha, \quad \cos \alpha = \cos \alpha'.$$

Dicho sea de paso, en matemáticas medianamente avanzadas siempre se usan radianes, así que π son los 180° que la gente de a pie conoce (si le pides al pizzero 45° de *pizza* se reirá de ti pero seguramente te entenderá. Prueba sin embargo a pedirle $\pi/4$).

En cuanto uno tenga un poco de curiosidad científica se plantea por qué las fórmulas (4.1) están de acuerdo con nuestro concepto intuitivo de longitudes y ángulos,

al menos en el caso $n = 2$. Consideremos $\vec{x} = (x_1, x_2)^t$, $\vec{y} = (y_1, y_2)^t$. La primera fórmula es solo una manifestación del teorema de Pitágoras. Si dibujamos la flecha que representa \vec{x} , esta será la hipotenusa de un triángulo rectángulo con catetos de longitudes x_1 y x_2 , salvo el signo:

$$\vec{x} = \sqrt{x_1^2 + x_2^2} = \sqrt{\vec{x} \cdot \vec{x}}.$$

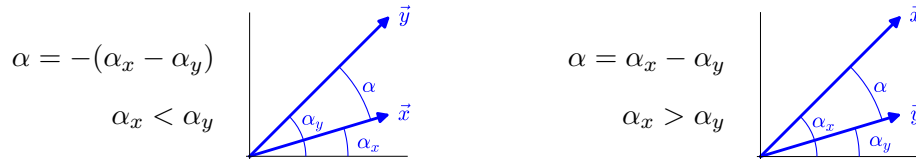
Para la segunda fórmula digamos que los ángulos que van del eje OX a los vectores \vec{x} e \vec{y} son α_x y α_y , respectivamente. Esto significa, empleando trigonometría elemental,

$$x_1 = \|\vec{x}\| \cos \alpha_x, \quad x_2 = \|\vec{x}\| \operatorname{sen} \alpha_x, \quad y_1 = \|\vec{y}\| \cos \alpha_y, \quad y_2 = \|\vec{y}\| \operatorname{sen} \alpha_y.$$

De aquí

$$\frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|} = \cos \alpha_x \cos \alpha_y + \operatorname{sen} \alpha_x \operatorname{sen} \alpha_y = \cos(\alpha_x - \alpha_y)$$

donde se ha usado la fórmula de adición del coseno. La observación final es que el ángulo α entre \vec{x} e \vec{y} es $\pm(\alpha_x - \alpha_y)$ donde el signo no afecta al valor del coseno y depende de si α_x es mayor o menor que α_y .



Veamos ahora algunos indicios de por qué es conveniente a veces considerar generalizaciones del producto escalar a espacios bien diferentes de \mathbb{R}^n o definidos por fórmulas distintas de la habitual.

Consideremos una señal que dura un segundo representada por una función $f : [0, 1] \rightarrow \mathbb{R}$. En la práctica, sobre todo en nuestro mundo digital, la *muestreamos*, lo que significa que la sustituimos por el vector $(f(1/N), f(2/N), \dots, f(N/N))^t \in \mathbb{R}^N$ correspondiente a medir la señal cada $1/N$ segundos con N grande (en señales de audio es habitual $N = 44100$). El producto escalar correspondiente a dos señales f y g sería $\sum_{k=1}^N f(k/N)g(k/N)$. Para que esto no resulte enorme se suele normalizar dividiendo entre N . En el límite $N \rightarrow \infty$, cuando el muestreo es infinitamente fino, se obtiene $\int_0^1 fg$. Por esta razón para señales que duran entre los tiempos $t = a$ y $t = b$ es natural definir los productos escalares

$$\langle f, g \rangle = \int_a^b fg \quad \text{o} \quad \langle f, g \rangle = \frac{1}{b-a} \int_a^b fg$$

que solo difieren en decidir si se normaliza también con la duración de la señal.

Por supuesto, uno puede definir lo que le venga en gana, otra cosa es que sirva para algo. Resulta que estos productos escalares sirven y mucho por varias razones, entre otras que en física e ingeniería $\langle f, f \rangle$ está relacionado con la energía. Por dar

un ejemplo simple y un poco tonto (con la excusa de que este no es un curso de teoría de señales), seguro que sabes que el voltaje de los enchufes de nuestra casa es actualmente¹ de 230 voltios y seguramente también sepas que es una onda sinusoidal que oscila una vez cada $T = 1/50$ segundos, matemáticamente $f(t) = A \operatorname{sen}(2\pi t/T)$. Con esta información estarás tentado a creer que lo que sale de tu enchufe tiene $A = 230$. No es así, cuando se habla del voltaje de los enchufes lo que se da es $\|f\| = \sqrt{\langle f, f \rangle}$ con el último producto escalar. Es decir,

$$230 = \left(\frac{1}{T} \int_0^T A^2 \operatorname{sen}^2(2\pi t/T) dt \right)^{1/2} = \left(\frac{1}{T} \cdot \frac{A^2 T}{2} \right)^{1/2} = \frac{1}{\sqrt{2}} A.$$

En resumidas cuentas, lo que sale del enchufe es $230\sqrt{2}\operatorname{sen}(2\pi t/T)$ que tiene picos de tensión de más de 300 voltios. ¿Por qué se habla entonces de 230? Porque esta es la tensión que se aprovecha para crear energía, la que correspondería a una corriente continua con igual rendimiento.

Incluso si no nos salimos de nuestro confortable \mathbb{R}^2 o \mathbb{R}^3 hay razones para considerar productos escalares diferentes del habitual. Supongamos que alguien viniera con la siguiente fórmula bajo el brazo:

$$\langle \vec{x}, \vec{y} \rangle = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2 \quad \text{para } \vec{x} = (x_1, x_2)^t, \vec{y} = (y_1, y_2)^t \in \mathbb{R}^2.$$

Pensaríamos que es muy raro porque sus longitudes y ángulos a través de (4.1) serían bien diferentes de las nuestras, por ejemplo $(1, 0)^t$ mediría $\sqrt{2}$ y el ángulo entre $(1, 0)^t$ y $(0, 1)^t$, es decir, entre los ejes, sería $\alpha = \pi/4$. Sin embargo es posible mostrar que podría desarrollar con ello una geometría perfectamente coherente, aunque distinta de la habitual. La razón para esta inesperada coherencia, y lo que justifica que esta nueva fórmula para el producto escalar no sea tan loca, es que si empleamos la base $\{\vec{u}_1 = (1, 1)^t, \vec{u}_2 = (0, 1)^t\}$ entonces lo que llamamos vectores de coordenadas $(x_1, x_2)^t$ e $(y_1, y_2)^t$ son los vectores $x_1\vec{u}_1 + x_2\vec{u}_2$ e $y_1\vec{u}_1 + y_2\vec{u}_2$ cuyo producto escalar usual es

$$\begin{aligned} (x_1\vec{u}_1 + x_2\vec{u}_2) \cdot (y_1\vec{u}_1 + y_2\vec{u}_2) &= x_1y_1\vec{u}_1 \cdot \vec{u}_1 + x_1y_2\vec{u}_1 \cdot \vec{u}_2 + x_2y_1\vec{u}_2 \cdot \vec{u}_1 + x_2y_2\vec{u}_2 \cdot \vec{u}_2, \\ &= 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2. \end{aligned}$$

Es decir, $\langle \vec{x}, \vec{y} \rangle$ con su extraña fórmula no es más que el producto escalar de toda la vida cuando usamos cierta base distinta de la canónica, un sistema de referencia deformado.

La existencia de productos escalares raros útiles motivó en el desarrollo del álgebra lineal dar una definición abstracta muy general de producto escalar, una declaración de mínimos requeridos para que algo se pueda considerar que lejanamente tiene que ver con longitudes y ángulos.

¹Incluso si eres mucho más talludito que tus compañeros quizá no sepas que en 2003 se cambió oficialmente de 220 voltios a 230 para unificar con Europa. Si eres tan viejo como yo, sí recordarás los 125 voltios de la infancia y que en el cambio a los 220 el gobierno distribuyó transformadores en los hogares para que los electrodomésticos funcionaran.

Si V es un espacio vectorial sobre \mathbb{R} , se dice que una función $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ es un producto escalar (generalizado) si tiene las tres siguientes propiedades:

- 1) Es lineal en el primer argumento²: Para todo $\vec{x}, \vec{y}, \vec{z} \in V$, $\lambda, \mu \in \mathbb{R}$ se cumple $\langle \lambda\vec{x} + \mu\vec{y}, \vec{z} \rangle = \lambda\langle \vec{x}, \vec{z} \rangle + \mu\langle \vec{y}, \vec{z} \rangle$.
- 2) Es *simétrica*: $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$ para todo $\vec{x}, \vec{y} \in V$.
- 3) Es *definida positiva*: $\langle \vec{x}, \vec{x} \rangle > 0$ para todo $\vec{x} \in V - \{\vec{0}\}$.

Los espacios vectoriales sobre \mathbb{R} en los que hemos definido un producto escalar con estas propiedades se dice que son *espacios euclídeos*.

Por ejemplo, en la línea de lo dicho antes, $\langle P, Q \rangle = \int_{-1}^1 PQ$ define un producto escalar en el espacio de polinomios $\mathbb{R}_n[x]$ convirtiéndolo en un espacio euclídeo. La linealidad se sigue de la linealidad de la integral:

$$\langle \lambda P_1 + \mu P_2, Q \rangle = \int_{-1}^1 (\lambda P_1 + \mu P_2)Q = \lambda \int_{-1}^1 P_1Q + \mu \int_{-1}^1 P_2Q = \lambda \langle P_1, Q \rangle + \mu \langle P_2, Q \rangle.$$

La simetría es evidente por la conmutativa del producto de polinomios. Finalmente, $\langle P, P \rangle > 0$ para todo polinomio P no nulo porque $\langle P, P \rangle = \int_{-1}^1 P^2$ es el área limitada por la gráfica de P^2 sobre el intervalo $[-1, 1]$ del eje X .

La política que usaremos en este capítulo será preferir la notación $\langle \vec{x}, \vec{y} \rangle$ para los resultados válidos con cualquier producto escalar y reservar $\vec{x} \cdot \vec{y}$ para el producto escalar usual. Se puede probar, aunque no lo haremos aquí, que todo producto escalar generalizado en \mathbb{R}^n se reduce al de toda la vida empleando una base adecuada, como vimos en un ejemplo anterior.

La última propiedad exigida a un producto escalar generalizado asegura que las longitudes definidas con la primera fórmula de (4.1) tienen sentido, no dan números imaginarios. Mostrar que los ángulos también lo tienen requiere mostrar que en la segunda fórmula el resultado está en $[-1, 1]$. Este resultado no es de ningún modo inmediato y tiene nombre propio (normalmente dos y a veces tres).

Proposición 4.1.1 (Desigualdad de Cauchy-Schwarz). *En cualquier espacio euclídeo V se cumple*

$$|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \|\vec{y}\| \quad \text{para todo } \vec{x}, \vec{y} \in V.$$

Aquí las barras que rodean $\langle \vec{x}, \vec{y} \rangle$ significan valor absoluto, es decir, que pongas siempre el signo positivo.

Demostración. Si \vec{y} es un múltiplo de \vec{x} entonces el resultado es evidente, incluso con igualdad. En otro caso, por la última propiedad $\langle \lambda\vec{x} - \vec{y}, \lambda\vec{x} - \vec{y} \rangle > 0$. Usando la linealidad (la primera propiedad) esto es lo mismo que

$$0 < \lambda^2 \langle \vec{x}, \vec{x} \rangle - \lambda \langle \vec{x}, \vec{y} \rangle - \lambda \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle = \lambda^2 \|\vec{x}\|^2 - 2\lambda \langle \vec{x}, \vec{y} \rangle + \|\vec{y}\|^2$$

²Tradicionalmente se pide la linealidad en ambos argumentos pero por la segunda propiedad, basta hacerlo en uno.

donde se ha usado la definición de norma y la simetría en la igualdad. La última expresión es un polinomio cuadrático en λ que no tiene raíces reales (es positivo) por tanto su discriminante $b^2 - 4ac = 4\langle \vec{x}, \vec{y} \rangle^2 - 4\|\vec{x}\|^2\|\vec{y}\|^2$ es negativo y esto equivale al resultado deseado. \square

Estrechamente relacionada con esta desigualdad hay otra bien conocida acerca de longitudes.

Proposición 4.1.2 (Desigualdad triangular). *En un espacio euclídeo V se cumple*

$$\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| \quad \text{para todo } \vec{x}, \vec{y} \in V.$$

Demostración. Elevando al cuadrado y recordando la definición de la norma en (4.1), hay que probar

$$\langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle \leq \langle \vec{x}, \vec{x} \rangle + 2\|\vec{x}\|\|\vec{y}\| + \langle \vec{y}, \vec{y} \rangle.$$

Cuando expandimos el primer miembro por la linealidad y cancelamos términos iguales, esto equivale a $\langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle \leq 2\|\vec{x}\|\|\vec{y}\|$ que es consecuencia de la desigualdad de Cauchy-Schwarz. \square

Del último resultado general que destacaremos es posible que te sorprenda el nombre, ciertamente el aludido se extrañará si en su metempsicosis ve esto.

Proposición 4.1.3 (Teorema de Pitágoras). *En cualquier espacio euclídeo V se verifica*

$$\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2 \quad \text{para cualesquiera } \vec{x}, \vec{y} \in V \text{ con } \langle \vec{x}, \vec{y} \rangle = 0.$$

El nombre proviene de que, según la segunda fórmula de (4.1), $\langle \vec{x}, \vec{y} \rangle = 0$ corresponde a $\alpha = \pi/2$. Con ello, al poner \vec{y} al final de \vec{x} forman los catetos de un triángulo rectángulo cuya hipotenusa mide $\|\vec{x} + \vec{y}\|$.

Demostración. Utilizando, como antes, la linealidad del producto escalar

$$\langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle$$

y la hipótesis y la simetría aseguran que los dos términos intermedios son nulos. \square

Después de toda esta sección centrada en \mathbb{R} cabe preguntarse qué ocurre en \mathbb{C} . Es muy posible que nunca te hayan definido en \mathbb{C}^n ni siquiera el producto escalar “usual”. La fórmula es muy similar a la de \mathbb{R}^n con una pequeña salvedad. Para $\vec{x}, \vec{y} \in \mathbb{C}^n$ se define³

$$\vec{x} \cdot \vec{y} = x_1\bar{y}_1 + x_2\bar{y}_2 + \cdots + x_n\bar{y}_n \quad \text{con } \vec{x} = (x_1, x_2, \dots, x_n)^t, \quad \vec{y} = (y_1, y_2, \dots, y_n)^t$$

³Aquí pongo, con bastante desconfianza, la definición habitual en álgebra lineal que es la que aparece en fuentes bibliográficas principales del curso como [31] y [20], pero en física y en otras partes de las matemáticas las x_j y las y_j están intercambiadas de modo que la linealidad que falla es la del primer argumento.

donde la barra sobre las y_j significa el conjugado. Esto se hace para que la norma $\|\vec{x}\|$ definida como la raíz cuadrada de $\vec{x} \cdot \vec{x}$ sea un número mayor o igual que cero. Como ejemplo, si $\vec{v}_1 = (2 + i, 1 + i)^t$, $\vec{v}_2 = (-1 + i, 2 - i)^t$ y $\vec{v}_3 = (3, i)^t$, se tiene

$$\vec{v}_1 \cdot \vec{v}_2 = 0, \quad \vec{v}_1 \cdot \vec{v}_3 = 7 + 2i, \quad \vec{v}_3 \cdot \vec{v}_1 = 7 - 2i, \quad \vec{v}_2 \cdot \vec{v}_3 = -4 + i.$$

También hay una definición generalizada de producto escalar para espacios vectoriales sobre \mathbb{C} igual a la de \mathbb{R} excepto que la simetría de la segunda propiedad hay que cambiarla por

$$\langle \vec{x}, \vec{y} \rangle = \overline{\langle \vec{y}, \vec{x} \rangle}.$$

Con ello la desigualdad de Cauchy-Schwarz, la triangular y el teorema de Pitágoras se siguen cumpliendo en este contexto, entendiéndose que en la Proposición 4.1.1, $|\langle \vec{x}, \vec{y} \rangle|$ significa el módulo de $\langle \vec{x}, \vec{y} \rangle$. Los espacios vectoriales sobre \mathbb{C} con un producto escalar en este sentido generalizado se llaman *espacios hermiticos*. Esto es, son los análogos sobre \mathbb{C} de los espacios euclídeos.

Solo para practicar, veamos que los vectores $\vec{x}, \vec{y} \in \mathbb{C}^3$ (con el producto escalar usual) dados por $\vec{x} = (2 + i, 1 + i, 1)^t$, $\vec{y} = (1 - 3i, 1 - i, 1 + 9i)^t$ están bajo las hipótesis del teorema de Pitágoras y cumplen su conclusión. Las cuentas, con un poco de detalle son:

$$\vec{x} \cdot \vec{y} = (2 + i)(1 + 3i) + (1 + i)(1 + i) + (1 - 9i) = (-1 + 7i) + 2i + (1 - 9i) = 0.$$

La suma es $\vec{x} + \vec{y} = (3 - 2i, 2, 2 + 9i)^t$ y se tienen las normas

$$\begin{aligned} \|\vec{x}\| &= \sqrt{|2 + i|^2 + |1 + i|^2 + 1^2} = \sqrt{5 + 2 + 1} = \sqrt{8}, \\ \|\vec{y}\| &= \sqrt{|1 - 3i|^2 + |1 - i|^2 + |1 + 9i|^2} = \sqrt{10 + 2 + 82} = \sqrt{94}, \\ \|\vec{x} + \vec{y}\| &= \sqrt{|3 - 2i|^2 + 2^2 + |2 + 9i|^2} = \sqrt{13 + 4 + 85} = \sqrt{102}. \end{aligned}$$

En consonancia con el teorema, se tiene $8 + 94 = 102$, ¡Pitágoras nunca falla!

Exprimiendo el silicio [opcional]. En `matlab/octave` el apóstrofo indica la operación \dagger que coincide con la trasposición en el caso real, así que si v y w son vectores columna $v' * w$ da el producto escalar de v por w en el caso real mientras que si queremos que esto funcione también en el caso complejo hay que cambiar el orden⁴ escribiendo $w' * v$ o conjugar el resultado anterior, lo cual se consigue mediante `conj`. Para comprobar los ejemplos anteriores en \mathbb{C}^2 podríamos escribir

```

1  % Definición de los vectores:
2  v1 = [2+i; 1+i];
3  v2 = [-1+i; 2-i];
4  v3 = [3; i];
5
6  % Productos escalares:
7  conj(v1' * v2)
8  conj(v1' * v3)
9  conj(v3' * v1)
10 conj(v2' * v3)

```

⁴Una razón más para que se cambie el convenio en álgebra lineal poniendo el conjugado en el primer argumento.

y si esto de conjugar nos parece enrevesado, sería equivalente

```

1  % Productos escalares:
2  v2'*v1
3  v3'*v1
4  v1'*v3
5  v3'*v2

```

Hay un comando `norm` que funciona correctamente con números complejos. Por ejemplo, `norm(v1)` mostraría una aproximación de $\sqrt{7}$.

En `sagemath` existe una instrucción `dot_product` que se limita a calcular $\sum x_k y_k$ y por tanto funciona en \mathbb{R}^n . Para que sirva también en \mathbb{C}^n hay que conjugar el segundo argumento con `conjugate`. Así el análogo del código anterior sería:

```

1  # Definición de los vectores:
2  v1 = vector(CC, [2+i, 1+i])
3  v2 = vector(CC, [-1+i, 2-i])
4  v3 = vector(CC, [3, i])
5
6  # Productos escalares:
7  print( v1.dot_product(v2.conjugate()) )
8  print( v1.dot_product(v3.conjugate()) )
9  print( v3.dot_product(v1.conjugate()) )
10 print( v2.dot_product(v3.conjugate()) )

```

Hay también un comando `norm` como en `matlab/octave`. El código anterior produce resultados con decimales. La razón para ello es que estamos trabajando en \mathbb{C} , indicado mediante `CC`. Una forma de conseguir resultados exactos forzando el cálculo simbólico es omitir `CC` (y la coma) o sustituirlo por `SR`. Otra forma más artificiosa es cambiar al cuerpo $K = \{a + bi : a, b \in \mathbb{Q}\}$ reemplazando las cuatro primeras líneas por

```

1  # Definición del nuevo cuerpo (redefine i)
2  K.<i> = QuadraticField(-1)
3  # Definición de los vectores:
4  v1 = vector(K, [2+i, 1+i])
5  v2 = vector(K, [-1+i, 2-i])
6  v3 = vector(K, [3, i])

```

4.2. Ortogonalidad

En el teorema de Pitágoras (Proposición 4.1.3) apareció la condición $\langle \vec{x}, \vec{y} \rangle = 0$ que en \mathbb{R}^2 o en \mathbb{R}^3 no es otra cosa que la perpendicularidad. Dentro de la jerga del álgebra lineal se prefiere el término *ortogonalidad*. Es decir, se dice que dos vectores \vec{x} e \vec{y} son *ortogonales* si $\langle \vec{x}, \vec{y} \rangle = 0$. Si además \vec{x} e \vec{y} son *unitarios* se dice que son *ortonormales*. Por extensión, diremos que una base $\mathcal{B} = \{\vec{u}_1, \dots, \vec{u}_n\}$ de un espacio vectorial con producto escalar es *base ortogonal* si los \vec{u}_j son ortogonales dos a dos y que es *base ortonormal* si son ortonormales. Nada impide hablar en general de conjuntos finitos ortogonales u ortonormales aunque, como veremos en breve, excluyendo vectores nulos, estos conjuntos son base del espacio que generan, gracias a que la independencia lineal está asegurada.

Por ejemplo $\mathcal{B} = \{\vec{u}_1 = (6, 4)^t, \vec{u}_2 = (2, -3)^t\}$ es base ortogonal de \mathbb{R}^2 . Es base porque está compuesta por $2 = \dim \mathbb{R}^2$ vectores linealmente independientes de \mathbb{R}^2