

Índice general

1	De dónde venimos y qué añadimos hoy	2
2	El coseno mide la abertura del ángulo	3
3	Ortogonalidad en $\mathbb{R}^n$	5
4	Teorema de Pitágoras en $\mathbb{R}^n$	5
5	Desigualdad de Cauchy–Schwarz	7
6	Ángulo, alineación, normalización	8
7	Desigualdad triangular y distancia	9
8	Recapitulación y guiño a la lección 4	10

Lección 3. El lenguaje geométrico del curso (parte II)

Marcos Bujosa

19 de septiembre de 2026

Resumen

Ángulo, ortogonalidad y teorema de Pitágoras en $\mathbb{R}^n$.
“La ortogonalidad es la piedra angular de la regresión.”

- ([slides](#)) — ([html](#)) — ([pdf](#))

1 De dónde venimos y qué añadimos hoy

- En la lección 2: vectores, producto escalar, norma —justificada con Pitágoras en $\mathbb{R}^2$ y $\mathbb{R}^3$ (dado por conocido).
- Hoy añadimos la otra magnitud geométrica: el *ángulo*.
- Y su caso estrella para nosotros: la *ortogonalidad* (ángulo recto).
- Programa:
 1. El coseno del ángulo: $\cos \theta = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|}$.
 2. Ortogonalidad: $\langle \mathbf{a}, \mathbf{b} \rangle = 0$.
 3. *Teorema de Pitágoras en $\mathbb{R}^n$* —demostración seria; con rigor y sin dibujo.
 4. Cauchy–Schwarz: lo que *legitima* la fórmula del coseno en $\mathbb{R}^n$.
 5. El ángulo, ya con todo rigor; y la desigualdad triangular.

De dónde venimos y qué añadimos hoy

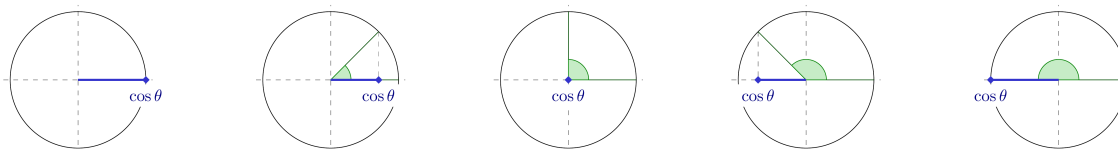
En la lección 2, se definieron los vectores en $\mathbb{R}^n$ como listas ordenadas de números reales y se presentaron sus operaciones básicas: suma y multiplicación por escalares. Además, se introdujeron el producto escalar $\langle \mathbf{x}, \mathbf{y} \rangle$ y la norma $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$; así como su particularización al caso usual en el espacio euclídeo $\mathbb{R}^n$: el producto $\langle \mathbf{x}, \mathbf{y} \rangle_e = \mathbf{x} \cdot \mathbf{y} = \sum_i x_i y_i$ y la norma $\|\mathbf{x}\|_e = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_e} = \sqrt{\sum_i x_i^2}$. Se explicó que esta última mide la “longitud” de un vector; su justificación en $\mathbb{R}^2$ y $\mathbb{R}^3$ se basa en el teorema de Pitágoras (dado por conocido), extendiéndose por analogía a $\mathbb{R}^n$.

Licencia: Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

Hoy añadimos la otra magnitud geométrica fundamental: el *ángulo* entre dos vectores. Y con el ángulo aparece su caso particular más importante para el curso, la *ortogonalidad* (ángulo recto), que es la piedra angular de toda la regresión. Además, la ortogonalidad nos permitirá por fin *demostrar* el teorema de Pitágoras en $\mathbb{R}^n$ —la primera demostración sería del curso, imposible en la lección 2 porque requería precisamente este concepto—. Cerraremos con un par de desigualdades (Cauchy–Schwarz y la triangular) que completan el aparato geométrico y dan sentido a la distancia entre vectores.

Todo el contenido de esta lección está desarrollado, con más detalle y figuras, en el capítulo 11 de mi *Curso de Álgebra Lineal*: [libro online](#). Le remitiré a él en los puntos donde convenga ver los detalles completos.

2 El coseno mide la abertura del ángulo



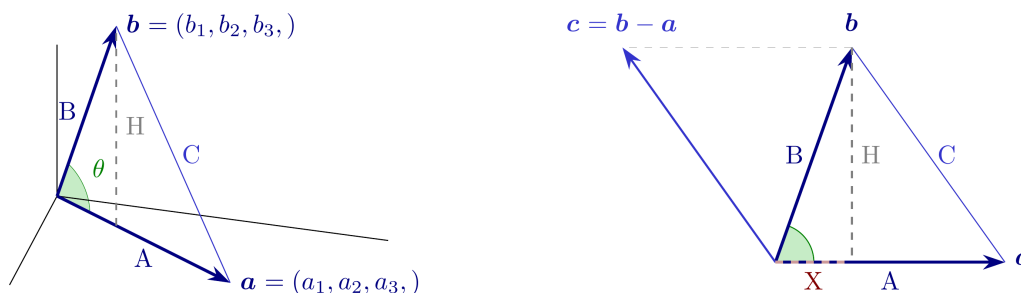
- $\cos \theta = 1$: ángulo nulo (misma dirección y sentido).
- $\cos \theta = 0$: ángulo recto.
- $\cos \theta = -1$: ángulo de 180° (sentidos opuestos).
- El coseno *decrece* de 1 a -1 a medida que el ángulo se abre.

¿Cómo medimos el ángulo entre dos vectores a partir de sus componentes?

La respuesta es una de las fórmulas más bellas y útiles del curso:

$$\cos \theta = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|}.$$

Caso unitario ($\|\mathbf{a}\| = \|\mathbf{b}\| = 1$): aplicando Pitágoras en los dos triángulos rectángulos:



$$x^2 + h^2 = 1 \quad \text{y} \quad (1 - x)^2 + h^2 = c^2 \quad \implies \quad x = \frac{2 - c^2}{2}.$$

Y como $c^2 = \|\mathbf{a} - \mathbf{b}\|^2 = 2 - 2\langle \mathbf{a}, \mathbf{b} \rangle$: $\cos \theta = x = \langle \mathbf{a}, \mathbf{b} \rangle$.

Caso general: normalizando $\mathbf{a}$ y $\mathbf{b}$ tenemos:

$$\cos \theta = \left\langle \frac{\mathbf{a}}{\|\mathbf{a}\|}, \frac{\mathbf{b}}{\|\mathbf{b}\|} \right\rangle = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|}$$

¡La fórmula es consecuencia de T. de Pitágoras!

Cálculo del coseno del ángulo entre dos vectores

El *coseno* es una función que mide el grado de abertura de un ángulo: vale 1 cuando el ángulo es nulo (vectores que apuntan en la misma dirección), 0 cuando el ángulo es recto (vectores perpendiculares) y -1 cuando el ángulo es de 180° (vectores opuestos).

El coseno del ángulo entre dos vectores $\mathbf{a}$ y $\mathbf{b}$ se puede calcular a partir de la siguiente fórmula:

$$\cos \theta = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|}.$$

Para deducir la fórmula, formamos un triángulo con dos vectores $\mathbf{a}$ y $\mathbf{b}$ y dividimos el problema con Pitágoras.

Caso unitario (cuando $\|\mathbf{a}\| = \|\mathbf{b}\| = 1$). Sea x la longitud de la proyección de $\mathbf{b}$ sobre la dirección de $\mathbf{a}$ (la longitud del segmento X de la figura de la derecha). Por definición de coseno en el triángulo rectángulo, $\cos \theta = x$. Llamando h a la altura del segmento H y c a la longitud del lado opuesto del triángulo (es decir, a la longitud del segmento C , que coincide con la norma del vector diferencia: $c = \|\mathbf{b} - \mathbf{a}\|$); aplicando Pitágoras en los dos triángulos rectángulos obtenemos las ecuaciones:

$$x^2 + h^2 = 1, \quad (1 - x)^2 + h^2 = c^2.$$

Restando la primera de la segunda:

$$(1 - 2x + x^2) - x^2 = c^2 - 1 \implies 2 - 2x = c^2 \implies x = \frac{2 - c^2}{2}.$$

Ahora desarrollamos c^2 con las propiedades del producto escalar:

$$c^2 = \|\mathbf{a} - \mathbf{b}\|^2 = \langle \mathbf{a}, \mathbf{a} \rangle - 2\langle \mathbf{a}, \mathbf{b} \rangle + \langle \mathbf{b}, \mathbf{b} \rangle = 2 - 2\langle \mathbf{a}, \mathbf{b} \rangle,$$

donde hemos usado $\|\mathbf{a}\| = \|\mathbf{b}\| = 1$. Sustituyendo:

$$\cos \theta = x = \frac{2 - (2 - 2\langle \mathbf{a}, \mathbf{b} \rangle)}{2} = \langle \mathbf{a}, \mathbf{b} \rangle.$$

Es decir, para vectores unitarios, *el coseno del ángulo es directamente su producto escalar*.

Caso general (vectores no nulos cualesquiera). Basta normalizar: los vectores $\frac{\mathbf{a}}{\|\mathbf{a}\|}$ y $\frac{\mathbf{b}}{\|\mathbf{b}\|}$ son unitarios, así que

$$\cos \theta = \left\langle \frac{\mathbf{a}}{\|\mathbf{a}\|}, \frac{\mathbf{b}}{\|\mathbf{b}\|} \right\rangle = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|}.$$

Esta demostración está sacada del capítulo 11 del libro. Lo importante es que la fórmula *no sale de la nada*: es consecuencia del Teorema de Pitágoras.

Para profundizar:

- Bujosa, M. *Curso de Álgebra Lineal*, capítulo 11, sección “Ángulo formado por dos vectores”: [libro online](#).
- Vídeo: [Dot products and duality](#) de 3Blue1Brown (capítulo 9 de *Essence of Linear Algebra*).

3 Ortogonalidad en $\mathbb{R}^n$

El ángulo es recto cuando $\cos \theta = 0$, es decir, cuando el numerador $\langle \mathbf{a}, \mathbf{b} \rangle$ se anula:

Definición. $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ son *ortogonales* (perpendiculares) si

$$\langle \mathbf{a}, \mathbf{b} \rangle = 0.$$

Entonces se escribe: $\mathbf{a} \perp \mathbf{b}$.

- El vector $\mathbf{0}$ es ortogonal a *todos* los vectores.
- *La ortogonalidad es una noción central del curso:* en la regresión, el residuo (datos menos ajuste) será ortogonal a los *regresores* (los vectores con que se construye el ajuste).

Ortogonalidad en $\mathbb{R}^n$

De la fórmula del coseno extraemos el caso particular más importante del curso. El ángulo es recto cuando $\cos \theta = 0$, y como $\cos \theta = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|}$, esto ocurre exactamente cuando el numerador se anula. Esto sugiere la definición:

Definición. Dos vectores $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ son *ortogonales* (o *perpendiculares*; usaremos ambos términos indistintamente, aunque preferiremos “ortogonal”) cuando $\langle \mathbf{a}, \mathbf{b} \rangle = 0$. Se escribe $\mathbf{a} \perp \mathbf{b}$.

Obsérvese una consecuencia inmediata: el vector $\mathbf{0}$ es ortogonal a *todos* los vectores, incluido él mismo, porque $\langle \mathbf{0}, \mathbf{b} \rangle = 0$ siempre.

La ortogonalidad es la noción que vertebra el curso. Cuando lleguemos a la regresión, la condición que define el ajuste por mínimos cuadrados será precisamente una condición de ortogonalidad: el vector de residuos¹ será ortogonal a los regresores.² Por eso conviene fijar bien esta definición ahora.

Para profundizar:

- Bujosa, M. *Curso de Álgebra Lineal*, capítulo 11, sección “Generalizando al espacio euclídeo $\mathbb{R}^n$ ”: [libro online](#).

4 Teorema de Pitágoras en $\mathbb{R}^n$

Teorema: $\mathbf{x} \perp \mathbf{y} \iff \|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$

Demostración:

¹La diferencia entre los datos originales objeto de estudio y el ajuste que el modelo logra para dichos datos.

²Los vectores que se combinan para construir el ajuste. Los definiremos con precisión al llegar a la regresión.

$$\begin{aligned}\|\mathbf{x} + \mathbf{y}\|^2 &= \langle (\mathbf{x} + \mathbf{y}), (\mathbf{x} + \mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ &= \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2.\end{aligned}$$

La igualdad pitagórica se cumple $\Leftrightarrow$ el término central $2\langle \mathbf{x}, \mathbf{y} \rangle = 0 \Leftrightarrow \mathbf{x} \perp \mathbf{y}$. ■

El cuadrado del binomio vectorial:

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2$$

aparecerá más veces en el curso.

Teorema de Pitágoras en $\mathbb{R}^n$

Esta es la primera demostración seria del curso. Conviene seguirla paso a paso, porque ilustra cómo las propiedades del producto escalar (simetría, linealidad) se usan para demostrar resultados geométricos.

En la lección 2 justificamos —con Pitágoras clásico en $\mathbb{R}^2$ y $\mathbb{R}^3$ — que la norma mide la longitud. Ahora, con la noción de ortogonalidad que acabamos de definir, podemos ir más lejos: demostrar el teorema en $\mathbb{R}^n$ sin recurrir a ningún dibujo.

Teorema (Pitágoras en $\mathbb{R}^n$). Dos vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ son ortogonales si y solo si

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$

Demostración. Partimos de la norma al cuadrado de la suma y la desarrollamos usando que $\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v} \rangle$ y la linealidad del producto escalar:

$$\begin{aligned}\|\mathbf{x} + \mathbf{y}\|^2 &= \langle (\mathbf{x} + \mathbf{y}), (\mathbf{x} + \mathbf{y}) \rangle && \text{(definición de norma)} \\ &= \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle && \text{(linealidad en ambos argumentos)} \\ &= \langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle && \text{(simetría: } \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle) \\ &= \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2. && \text{(definición de norma)}\end{aligned}$$

De esta igualdad —que es válida *siempre*, sean o no ortogonales los vectores— se lee directamente el teorema: la igualdad $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ se cumple si y solo si el término central se anula, es decir, si y solo si $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, esto es, si y solo si $\mathbf{x} \perp \mathbf{y}$. ■

Aquí $\mathbf{x}$ e $\mathbf{y}$ son los catetos y $\mathbf{x} + \mathbf{y}$ la hipotenusa. La igualdad intermedia

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2$$

conviene recordarla por sí misma: es el análogo vectorial del cuadrado del binomio, y la usaremos repetidamente (en Cauchy–Schwarz, en la desigualdad triangular y en la descomposición de la varianza en la regresión).

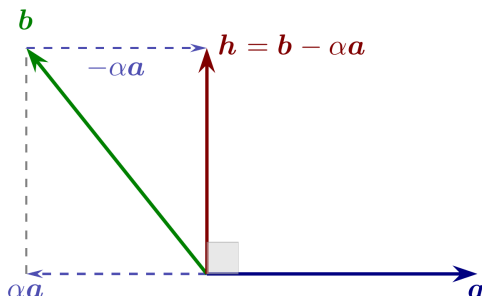
Para profundizar:

- Bujosa, M. *Curso de Álgebra Lineal*, capítulo 11, "Teorema de Pitágoras en $\mathbb{R}^n$ ": [libro online](#).

5 Desigualdad de Cauchy–Schwarz

¿Qué asegura que $\cos \theta = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|} \in [-1, 1]$ en $\mathbb{R}^n$?

Lo hace la desigualdad de *Cauchy–Schwarz*: $|\langle \mathbf{a}, \mathbf{b} \rangle| \leq \|\mathbf{a}\| \|\mathbf{b}\|$.



Idea de la demostración: construir $\mathbf{h} = \mathbf{b} - \alpha \mathbf{a}$ ortogonal a $\mathbf{a}$:

$$\langle \mathbf{a}, \mathbf{h} \rangle = 0 \implies \boxed{\alpha = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\langle \mathbf{a}, \mathbf{a} \rangle}}.$$

Por Pitágoras, $\|\mathbf{b}\|^2 = \|\mathbf{h}\|^2 + \alpha^2 \|\mathbf{a}\|^2 \geq \alpha^2 \|\mathbf{a}\|^2$, de donde sale la desigualdad.

- Igualdad $\Leftrightarrow$ vectores alineados; ortogonales $\Leftrightarrow \langle \mathbf{a}, \mathbf{b} \rangle = 0$.

Desigualdad de Cauchy–Schwarz

Hay un detalle que hemos pasado por alto. Hemos *definido* el coseno del ángulo en $\mathbb{R}^n$ como $\cos \theta = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|}$. Pero un coseno solo puede tomar valores entre -1 y 1 . ¿Qué nos garantiza que ese cociente nunca se sale de ese intervalo en $\mathbb{R}^n$? La respuesta es un resultado clásico:

Proposición (desigualdad de Cauchy–Schwarz). Para todo $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ se cumple

$$|\langle \mathbf{a}, \mathbf{b} \rangle| \leq \|\mathbf{a}\| \|\mathbf{b}\|.$$

La demostración es bonita y *siembra una idea que reaprovecharemos en la lección 4*: la de descomponer un vector en dos partes; donde la primera está alineada con otro vector y la segunda es ortogonal a la primera.

Demostración (caso en que $\mathbf{a}$ y $\mathbf{b}$ son linealmente independientes, es decir, no están alineados). Construimos un vector $\mathbf{h}$ que sea ortogonal a $\mathbf{a}$ restando a $\mathbf{b}$ un múltiplo adecuado de $\mathbf{a}$:

$$\mathbf{h} = \mathbf{b} - \alpha \mathbf{a}.$$

Es decir, elegimos α para forzar la ortogonalidad $\mathbf{a} \perp \mathbf{h}$:

$$0 = \langle \mathbf{a}, \mathbf{h} \rangle = \langle \mathbf{a}, (\mathbf{b} - \alpha \mathbf{a}) \rangle = \langle \mathbf{a}, \mathbf{b} \rangle - \alpha \langle \mathbf{a}, \mathbf{a} \rangle \implies \boxed{\alpha = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\langle \mathbf{a}, \mathbf{a} \rangle}}.$$

Como $\mathbf{b} = \mathbf{h} + \alpha \mathbf{a}$ con $\mathbf{h} \perp \alpha \mathbf{a}$, aplicamos Pitágoras:

$$\|\mathbf{b}\|^2 = \|\mathbf{h}\|^2 + \alpha^2 \|\mathbf{a}\|^2.$$

Como $\mathbf{a}$ y $\mathbf{b}$ son linealmente independientes, $\mathbf{h} \neq \mathbf{0}$, luego $\|\mathbf{h}\|^2 > 0$ y por tanto

$$\|\mathbf{b}\|^2 > \alpha^2 \|\mathbf{a}\|^2 = \frac{\langle \mathbf{a}, \mathbf{b} \rangle^2}{\langle \mathbf{a}, \mathbf{a} \rangle^2} \|\mathbf{a}\|^2 = \frac{\langle \mathbf{a}, \mathbf{b} \rangle^2}{(\|\mathbf{a}\|^2)^2} \|\mathbf{a}\|^2 = \frac{\langle \mathbf{a}, \mathbf{b} \rangle^2}{\|\mathbf{a}\|^2}.$$

Multiplicando por $\|\mathbf{a}\|^2$ obtenemos $\|\mathbf{a}\|^2 \|\mathbf{b}\|^2 > \langle \mathbf{a}, \mathbf{b} \rangle^2$, y tomando raíces, $\|\mathbf{a}\| \|\mathbf{b}\| > |\langle \mathbf{a}, \mathbf{b} \rangle|$.

En un extremo tenemos $\langle \mathbf{a}, \mathbf{b} \rangle = 0$ cuando $\mathbf{a} \perp \mathbf{b}$.

En el otro, si $\mathbf{b} = \lambda \mathbf{a}$ (si ambos vectores están alineados) entonces $\mathbf{h} = \mathbf{0}$, es decir, $\|\mathbf{h}\|^2 = 0$; y entonces $\|\mathbf{b}\|^2 = \alpha^2 \|\mathbf{a}\|^2$. Consecuentemente $|\langle \mathbf{a}, \mathbf{b} \rangle| = \|\mathbf{a}\| \|\mathbf{b}\|$. Así, la desigualdad es estricta cuando los vectores **no** están alineados, y se convierte en igualdad cuando lo están. (Si $\mathbf{a} = \mathbf{0}$ no hay nada que probar: ambos lados de la desigualdad valen 0. Ese caso queda fuera de la construcción anterior porque α exige dividir por $\langle \mathbf{a}, \mathbf{a} \rangle$, que entonces se anula.)

Idea para recordar. El vector $\alpha \mathbf{a}$, con $\alpha = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\langle \mathbf{a}, \mathbf{a} \rangle}$, es la “sombra” de $\mathbf{b}$ sobre la dirección de $\mathbf{a}$. En la lección 4 reconoceremos en esta construcción la *proyección ortogonal*, y ese α resultará ser la media (cuando $\mathbf{a} = \mathbf{1}$); más adelante será también la pendiente de la regresión lineal simple (cuando $\mathbf{a}$ son los datos centrados de la variable explicativa).

Para profundizar:

- Bujosa, M. *Curso de Álgebra Lineal*, capítulo 11, “Desigualdad de Cauchy–Schwarz”: [libro online](#).
- Para quien quiera otra demostración (la del parámetro t): cualquier manual de álgebra lineal; es la versión que considera $\|\mathbf{b} - t\mathbf{a}\|^2 \geq 0$ como polinomio en t .

6 Ángulo, alineación, normalización

Dado que se verifica Cauchy–Schwarz en $\mathbb{R}^n$, es legítima la definición:

$$\cos \theta = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|}, \quad -1 \leq \cos \theta \leq 1.$$

- *Vector unitario*: norma 1.
 - Un vector $\mathbf{a}$, no nulo, se *normaliza* dividiendo por su longitud: $\frac{1}{\|\mathbf{a}\|} \mathbf{a}$.
- *Vectores alineados* ($\mathbf{b} = \lambda \mathbf{a}$): $\cos \theta = \frac{\lambda}{|\lambda|} = \pm 1$.
 - Caso opuesto a la ortogonalidad.
 - *Avance*: regresores muy correlacionados $\approx$ vectores casi alineados $\rightarrow$ lo veremos en la lección 14, al hablar de colinealidad.

El ángulo en $\mathbb{R}^n$, vectores alineados, normalización

Ahora que Cauchy–Schwarz nos garantiza que el cociente está en $[-1, 1]$, podemos definir con todo rigor:

Definición. El *coseno del ángulo* θ formado por dos vectores no nulos $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ es

$$\cos \theta = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|}, \quad -1 \leq \cos \theta \leq 1.$$

Dos conceptos auxiliares que reaparecerán:

- *Vector unitario*: un vector de norma 1. Para obtener un vector unitario en la dirección de $\mathbf{a} \neq \mathbf{0}$, se *normaliza* dividiéndolo por su longitud: $\frac{1}{\|\mathbf{a}\|}\mathbf{a}$ tiene norma 1, porque $\left\| \frac{1}{\|\mathbf{a}\|}\mathbf{a} \right\| = \frac{1}{\|\mathbf{a}\|}\|\mathbf{a}\| = 1$ (por la homogeneidad de la norma vista en la lección 2).

Ejemplo: si $\mathbf{x} = (1, 1, 1, 1)$, entonces $\|\mathbf{x}\|_e = \sqrt{4} = 2$, y el vector normalizado $\frac{1}{2}\mathbf{x} = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ es unitario.

- *Vectores alineados*: cuando $\mathbf{b} = \lambda\mathbf{a}$ con $\lambda \neq 0$. En ese caso el ángulo es 0 (si $\lambda > 0$, "apuntan" en el mismo sentido) o 180° (si $\lambda < 0$), porque

$$\cos \theta = \frac{\lambda \|\mathbf{a}\|^2}{|\lambda| \|\mathbf{a}\|^2} = \frac{\lambda}{|\lambda|} = \pm 1.$$

Los vectores alineados son el caso opuesto a los ortogonales: máxima dependencia frente a máxima independencia direccional. Esta imagen será útil en la lección 14, cuando hablemos de *colinealidad*: dos regresores muy correlacionados son vectores *casi* alineados (ángulo pequeño, coseno cercano a ± 1), y eso hará inestable la regresión.

Para profundizar:

- Bujosa, M. *Curso de Álgebra Lineal*, capítulo 11, secciones sobre el coseno del ángulo y vectores alineados: [libro online](#).

7 Desigualdad triangular y distancia

Desigualdad triangular (aplicación de Cauchy–Schwarz):

$$\|\mathbf{a} + \mathbf{b}\|^2 = \|\mathbf{a}\|^2 + 2\langle \mathbf{a}, \mathbf{b} \rangle + \|\mathbf{b}\|^2 \leq (\|\mathbf{a}\| + \|\mathbf{b}\|)^2.$$

$$\implies \|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|.$$

Distancia entre vectores:

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$

La desigualdad triangular y la distancia entre vectores

Como aplicación inmediata de Cauchy–Schwarz, demostramos un resultado intuitivo: el camino directo nunca es más largo que el indirecto.

Proposición (desigualdad triangular). Para todo $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$,

$$\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|.$$

Demostración. Partimos del cuadrado del binomio vectorial (la identidad que destacamos en la sección 4):

$$\begin{aligned}\|\mathbf{a} + \mathbf{b}\|^2 &= \|\mathbf{a}\|^2 + 2\langle \mathbf{a}, \mathbf{b} \rangle + \|\mathbf{b}\|^2 \\ &\leq \|\mathbf{a}\|^2 + 2\|\mathbf{a}\|\|\mathbf{b}\| + \|\mathbf{b}\|^2 \quad (\text{por Cauchy–Schwarz: } \langle \mathbf{a}, \mathbf{b} \rangle \leq |\langle \mathbf{a}, \mathbf{b} \rangle| \leq \|\mathbf{a}\|\|\mathbf{b}\|) \\ &= (\|\mathbf{a}\| + \|\mathbf{b}\|)^2.\end{aligned}$$

Tomando raíces cuadradas (pues ambos lados son no negativos), $\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$. ■

Esta desigualdad permite definir la *distancia* entre dos vectores como la longitud de su diferencia:

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$

Es simétrica ($d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$, porque $\|\mathbf{x} - \mathbf{y}\| = \|\mathbf{y} - \mathbf{x}\|$), se anula solo cuando los vectores coinciden ($d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$), y verifica su propia desigualdad triangular:

$$d(\mathbf{a}, \mathbf{c}) = \|\mathbf{a} - \mathbf{c}\| = \|(\mathbf{a} - \mathbf{b}) + (\mathbf{b} - \mathbf{c})\| \leq \|\mathbf{a} - \mathbf{b}\| + \|\mathbf{b} - \mathbf{c}\| = d(\mathbf{a}, \mathbf{b}) + d(\mathbf{b}, \mathbf{c}).$$

Esta noción de distancia es, literalmente, el corazón de la regresión: en la lección 6 buscaremos el vector $\hat{\mathbf{y}}$ —de entre todos los de un subespacio— que *minimiza la distancia* $\|\mathbf{y} - \hat{\mathbf{y}}\|$ al vector de datos $\mathbf{y}$. Ajustar por mínimos cuadrados será, exactamente, minimizar una distancia $\|\mathbf{y} - \hat{\mathbf{y}}\|$, donde $\hat{\mathbf{y}}$ es el vector ajustado (aún no definido formalmente; lo haremos en la lección 6).

Para profundizar:

- Bujosa, M. *Curso de Álgebra Lineal*, capítulo 11, "Desigualdad triangular" y "distancia entre vectores": [libro online](#).

8 Recapitulación y guiño a la lección 4

Concepto	Fórmula matemática / definición algebraica
Coseno del ángulo	$\cos \theta = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\ \mathbf{a}\ \ \mathbf{b}\ }$
Ortogonalidad	$\langle \mathbf{a}, \mathbf{b} \rangle = 0$
Pitágoras	$\mathbf{x} \perp \mathbf{y} \iff \ \mathbf{x} + \mathbf{y}\ ^2 = \ \mathbf{x}\ ^2 + \ \mathbf{y}\ ^2$
Cauchy–Schwarz	$ \langle \mathbf{a}, \mathbf{b} \rangle \leq \ \mathbf{a}\ \ \mathbf{b}\ $
Distancia	$d(\mathbf{x}, \mathbf{y}) = \ \mathbf{x} - \mathbf{y}\ $

Hoy, $\alpha = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\langle \mathbf{a}, \mathbf{a} \rangle}$ fue un truco auxiliar; pero en la lección 4 será la *proyección ortogonal* que, con $\mathbf{a} = \mathbf{1}$, dará lugar a la *media aritmética*.

Recapitulación y guiño a la lección 4

Lo esencial de hoy:

1. La *longitud*, justificada en la lección 2 con Pitágoras clásico en $\mathbb{R}^2$ y $\mathbb{R}^3$, se extiende por analogía a $\mathbb{R}^n$. Hoy hemos demostrado el *teorema de Pitágoras en $\mathbb{R}^n$* con rigor: $\mathbf{x} \perp \mathbf{y} \Leftrightarrow \|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$.
2. El *coseno del ángulo* es $\cos \theta = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|}$; se deduce en $\mathbb{R}^2$ con Pitágoras y se extiende a $\mathbb{R}^n$ como definición.
3. Dos vectores son *ortogonales* cuando $\langle \mathbf{a}, \mathbf{b} \rangle = 0$. Esta noción es central en el curso.
4. *Teorema de Pitágoras en $\mathbb{R}^n$* : $\mathbf{x} \perp \mathbf{y} \Leftrightarrow \|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$. Demostrado con las propiedades del producto escalar.
5. *Cauchy-Schwarz* ($|\langle \mathbf{a}, \mathbf{b} \rangle| \leq \|\mathbf{a}\| \|\mathbf{b}\|$) garantiza que el coseno está en $[-1, 1]$. Su demostración construye la "sombra" (la proyección) $\alpha \mathbf{a}$ de $\mathbf{b}$ sobre $\mathbf{a}$, con $\alpha = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\langle \mathbf{a}, \mathbf{a} \rangle}$.
6. La *desigualdad triangular* da sentido a la *distancia* $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$.

Guiño a la lección 4. Dos piezas de hoy son la semilla de toda la econometría que viene. Primero, la *distancia*: la regresión consistirá en encontrar el vector de un subespacio más cercano a los datos, es decir, en minimizar $\|\mathbf{y} - \hat{\mathbf{y}}\|$. Segundo, la *construcción* $\mathbf{h} = \mathbf{b} - \alpha \mathbf{a}$ de la demostración de Cauchy-Schwarz: ese $\alpha = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\langle \mathbf{a}, \mathbf{a} \rangle}$, que hoy ha aparecido como truco auxiliar, resultará ser la *proyección ortogonal*. Cuando $\mathbf{a} = \mathbf{1}$ (el vector de unos), α será la *media* de los datos. No es casualidad: es el primer indicio de que la media, la regresión y la geometría son la misma cosa vista desde ángulos distintos.