

Clase 14. Vectores y Subespacios Ortogonales.

Curso ‘Linear Algebra’ del MIT.

Resumen

La **ortogonalidad** entre dos vectores y dos subespacios es otra información relevante que vamos a considerar desde ahora. En esta ocasión estudiaremos cuándo lo son en ambos casos, incluyendo a los cuatro subespacios fundamentales. Para terminar, veremos cómo resolver sistemas de ecuaciones lineales que, en principio, no tienen solución.

14.1 Ortogonalidad.

La **ortogonalidad** es la noción geométrica de la **perpendicularidad** en el Álgebra Lineal. En otras palabras, son conceptos intercambiables (o sinónimos) en esta rama de la matemática.

14.1.1 Vectores Ortogonales.

Geométricamente, dos vectores son ortogonales si se forma un **ángulo de 90°** entre ellos.

Recordemos que el **producto punto** (o producto escalar) entre dos vectores es la suma de la multiplicación entre sus componentes que resulta en un valor escalar. Es decir, sean \vec{x} e \vec{y} dos vectores columna en \mathbb{R}^n dimensiones, calculamos esta operación algebraica como:

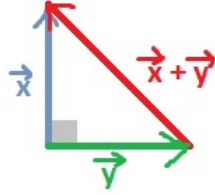
$$\vec{x}^T \cdot \vec{y} = x_1 \cdot y_1 + x_2 \cdot y_2 + \cdots + x_n \cdot y_n$$

Señalamos lo del **producto punto** porque se cumple que **dos vectores son ortogonales** (o perpendiculares) si, en esta operación algebraica, el escalar resultante **es cero**.

$$\vec{x}^T \vec{y} = 0$$

La igualdad de arriba podemos comprobarla usando el Teorema de Pitágoras. Partamos con

el supuesto de que los vectores \vec{x} e \vec{y} son perpendiculares. En ese sentido, entre ambos y el vector $\vec{x} + \vec{y}$ podemos construir un triángulo rectángulo.



Debido a lo anterior, podemos aplicar el Teorema de Pitágoras:

$$\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$$

donde, por ejemplo:

$$\|\vec{x}\|^2 = \left(\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \right)^2 = x_1^2 + x_2^2 + \cdots + x_n^2$$

Algo interesante es que:

$$\vec{x}^T \vec{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1^2 + x_2^2 + \cdots + x_n^2 \end{bmatrix}$$

la cual es una matrix de 1×1 dimensiones.

En ese sentido, podemos establecer de forma conveniente¹ que $\vec{x}^T \vec{x} = \|\vec{x}\|^2$, lo cual también podemos aplicar a los otros vectores de la ecuación de arriba, reescribiéndola como:

$$(\vec{x} + \vec{y})^T (\vec{x} + \vec{y}) = \vec{x}^T \vec{x} + \vec{y}^T \vec{y}$$

Además, $(\vec{x} + \vec{y})^T = \vec{x}^T + \vec{y}^T$:

$$\begin{aligned} (\vec{x}^T + \vec{y}^T)(\vec{x} + \vec{y}) &= \vec{x}^T \vec{x} + \vec{y}^T \vec{y} \\ \vec{x}^T \vec{x} + \vec{x}^T \vec{y} + \vec{y}^T \vec{x} + \vec{y}^T \vec{y} &= \vec{x}^T \vec{x} + \vec{y}^T \vec{y} \\ \vec{x}^T \vec{y} + \vec{y}^T \vec{x} &= 0 \end{aligned}$$

Y $\vec{x}^T \vec{y} = \vec{y}^T \vec{x}$ (prop. conmutativa). En consecuencia:

¹Aunque no es técnicamente cierto, ya que $\vec{x}^T \vec{x}$ es una matrix, no un escalar, pero para la explicación lo consideraremos así.

$$\vec{x}^T \vec{y} + \vec{y}^T \vec{x} = 0$$

$$2\vec{x}^T \vec{y} = 0$$

$$\vec{x}^T \vec{y} = 0$$

donde esta igualdad es verdadera sí y solo sí \vec{x} e \vec{y} son perpendiculares/ortogonales.

Como es de suponer, el $\vec{0}$ **siempre será ortogonal a cualquier otro vector** porque su producto punto será igual a cero en todos los casos.

14.1.2 Subespacios Vectoriales Ortogonales.

Dos **subespacios vectoriales** S y T también pueden ser **ortogonales**, la condición es que **todos los $\vec{s} \in S$ deben ser perpendiculares a todos los $\vec{t} \in T$.**

Si **dos subespacios se intersectan** o, en otras palabras, si uno o todos los vectores pertenecen a ambos conjuntos, de inmediato podemos concluir que **no son ortogonales** porque un vector no puede ser perpendicular a sí mismo, salvo que sea el $\vec{0}$.

Lo anterior podemos verlo con el producto punto. Digamos que $S \cap T = \{\vec{s}\}$. Es decir, $\vec{s} \in S$ y $\vec{s} \in T$. Si \vec{s} es distinto de cero, siempre se dará que $\vec{s}^T \cdot \vec{s} \neq 0$, pero si $\vec{s} = \vec{0}$, entonces $\vec{s}^T \cdot \vec{s} = 0$

14.1.2.1 Ortogonalidad en los Cuatro Subespacios Fundamentales.

Un caso de dos subespacio que se intersectan en el origen y que son perpendiculares entre sí, son el **Espacio Nulo** y el **Espacio Fila**. Recordemos que el primero consiste de todas las soluciones a la ecuación $Ax = 0$.

$$Ax = 0$$

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Si nos damos cuenta, Ax también corresponde a los **productos punto** entre los vectores fila de A y \vec{x} , que resultan en el escalar cero y, por consiguiente, son ortogonales.

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = 0 \quad \cdots \quad \begin{bmatrix} a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = 0$$

Esto implica, a su vez, que los productos punto entre los vectores que resultan de la combinación lineal de los vectores fila de A y el vector \vec{x} también serán iguales a cero, implicando que **el $C(A^T)$ es ortogonal al $N(A)$** (y viceversa).

Lo mismo ocurre entre el **Espacio Columna** y el **Espacio Nulo Izquierdo**. A partir de $A^T y = 0$ o de $y^T A = 0^T$ podemos demostrar que **el $C(A)$ es ortogonal al $N(A^T)$** .

Así, agreguemos esta información a la tabla que hicimos en la clase 10.

Subesp	Esp. Vectorial	Base	Dimensión	Ortogonal
$C(A)$	\mathbb{R}^n	Columnas Pivotes A	r	$N(A^T)$
$C(A^T)$	\mathbb{R}^m	Filas Pivotes A	r	$N(A)$
$N(A)$	\mathbb{R}^m	x en $Ax = 0$	$m - r$	$C(A^T)$
$N(A^T)$	\mathbb{R}^n	y en $A^T y = 0$	$n - r$	$C(A)$

Algo a tener en cuenta, además, es que **la suma de las dimensiones de dos subespacios que son ortogonales siempre es menor o igual a la dimensión de su espacio vectorial**. Esto podemos verlo con los cuatros subespacios especiales. Por ejemplo:

$$\dim(C(A)) + \dim(N(A^T)) = \dim(\mathbb{R}^n)$$

$$r + (n - r) = n$$

$$n = n$$

Es decir, **los vectores que conforman el $C(A)$ y el $N(A^T)$** , que son ortogonales entre sí, **cubren todo su espacio vectorial \mathbb{R}^n** .

14.1.2.2 Complemento Ortogonal.

Cuando dos subespacios son ortogonales, **el conjunto de todos los vectores de uno de ellos que son perpendiculares a todos los del otro**, se conoce como **Complemento Ortogonal**.

En ese sentido podemos decir, por ejemplo, que el $N(A)$ y el $C(A^T)$ son **complementos ortogonales** en el espacio \mathbb{R}^m , que es lo mismo a señalar que el espacio nulo contiene todos los vectores que son perpendiculares a los del espacio fila.

14.2 Resolviendo un sistema $Ax = b$ sin solución.

Un problema que empezaremos a abordar, es intentar resolver un sistema de ecuaciones lineales $Ax = b$ que no tiene solución. Otra forma de expresarlo, es que \vec{b} no está en el $C(A)$. El propósito es, al menos, hacerlo de la mejor forma posible.

Si bien puede ser raro intentar resolver un sistema que no tiene solución, en realidad es bastante común. Esto lo vemos cuando intentamos modelar un fenómeno o un proceso. Muchas veces solo nos interesan algunas variables o solo una, pero tenemos que considerar otras porque de algún modo influyen en lo que estudiamos, aunque solo generen ruido en \vec{b} .

El caso habitual en el que $Ax = b$ no tiene solución, con A siendo de $n \times m$, es cuando esta matriz es **rectangular**. En particular, en el caso de que:

$$n \text{ (ecuaciones)} > m \text{ (incógnitas/variables)}$$

Esto significa, además, que tendremos más componentes en \vec{b} de los que necesitamos.

La idea es evitar tener que eliminar ecuaciones y dejar “las que sirvan”, porque no hay motivo para aquello. Todas nos entregan información. Lo mismo sería hacerlo tal que quedemos con una matriz invertible y, por tanto, un sistema resoluble.

En vez de aquello, vamos a trabajar con la siguiente matriz:

$$A^T A$$

donde A es **rectangular**.

Como recordaremos de la clase 5, el producto entre la transpuesta de una matriz y ella misma, $A^T A$, **siempre resulta en una matriz simétrica y cuadrada**.

Por ejemplo, si A es de $n \times m$, entonces A^T es de $m \times n$. Por lo tanto, $A^T A$ es de $m \times m$. Por otra parte, la simetría de $A^T A$ la comprobamos calculando su transpuesta:

$$(A^T A)^T = A^T (A^T)^T = A^T A$$

En ese sentido, para resolver el sistema $Ax = b$ que no tiene solución, lo multiplicaremos por A^T obteniendo lo siguiente:

$$A^T A \hat{x} = A^T b$$

Denotamos al vector de soluciones x como \hat{x} para diferenciarla de la primera, porque ésta tendrá las soluciones (i.e, será la mejor solución) que no tiene \vec{x} .

Por otra parte, $A^T A$ será **invertible** sí y solo sí los **vectores columna** de A son **independientes** entre sí. Cuando este hecho ocurre, se da como consecuencia que:

$$N(A^T A) = N(A) \quad \text{y} \quad \text{rango}(A^T A) = \text{rango}(A)$$

Y como los vectores columnas A son independientes entre sí y $A^T A$ es simétrica, cuando $(A^T A)^{-1}$ existe, $A^T A$ es de rango completo, lo que implica que el vector $\vec{0}$ es el único que generan el $N(A^T A)$ y también el $N(A)$.