

Norma inducida y bases ortonormales

Observación 0.1. *Las ideas principales a presentar en este video son:*

- (I) *La norma es una operación que, de estar presente en un espacio vectorial, nos da una noción de magnitud de un vector, y nos permite hacer comparaciones entre vectores en este sentido. Los vectores con norma uno son llamados normales o unitarios y todo vector no nulo puede ser reescalado de tal forma que el vector resultante sea normal.*
- (II) *Por definición, un producto escalar en un espacio vectorial induce una norma en ese espacio. Aunque no todas las normas provienen de un producto escalar¹, en esta serie de videos nos enfocaremos en este tipo de normas.*
- (III) *La proyección vectorial de un vector arbitrario sobre un vector no nulo puede ser reescrita considerando la norma inducida. Una base en la que todos los vectores son normales y ortogonales entre sí se conoce como una base ortonormal.*
- (IV) *Si un espacio vectorial con producto escalar tiene dimensión finita y existe una base ortonormal, entonces el problema de encontrar los coeficientes necesarios para expresar a un vector arbitrario como combinación lineal de esta base tiene una solución extremadamente simple, en donde cada coeficiente se calcula a través de un producto escalar.*

Observación 0.2. *Considerando que la serie está enfocada en abordar la descomposición espectral de operadores lineales, creo que **no** es necesario incluir en ella una discusión sobre métrica ni normas distintas de aquellas inducidas por un producto escalar positivo definido; sin embargo, sería bueno buscar referencias para estos temas y ponerlas en la descripción.*

¹Podríamos colocar ejemplos de normas no inducidas por productos escalares positivo definidos en la descripción del video.

1. Primera escena

1.1. SE1 (Norma)

Por definición, *todos* los espacios vectoriales deben tener una operación llamada *suma* o *adición vectorial* y otra llamada *producto de un vector por un escalar* o, simplemente, *reescalamiento*². Por ello, decimos que estas son las operaciones *esenciales* de los espacios vectoriales. Sin embargo, puede que un espacio vectorial tenga, además, *otras* operaciones, que doten al espacio de una estructura adicional. Un ejemplo es la operación de producto escalar, la cual introduce nociones de ortogonalidad y proyecciones, como vimos en el video anterior. En este video veremos otra operación conocida como *norma*, así como las nociones que introduce en un espacio vectorial, y la forma en que se relaciona con el producto escalar.

Una norma en un espacio vectorial es una operación que a cada vector del espacio le asigna número real, tal que cumple las siguientes propiedades:

$$\begin{array}{c} \text{Norma} \\ || \cdot || : V \rightarrow \mathbb{R} \end{array}$$

$$\forall \vec{u} \in V, \forall a \in K,$$

$$||a\vec{u}|| = |a| ||\vec{u}||$$

$$||\vec{u}|| = 0 \iff \vec{u} = \vec{0}$$

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

La primera propiedad, conocida como *escalabilidad absoluta*, nos dice que la norma de un vector reescalado es directamente proporcional al *valor absoluto* del escalar en cuestión. En particular, se sigue que la norma de un vector no cambia si lo reescalamos por cualquier escalar con valor absoluto uno. Recordando la interpretación geométrica del producto de un vector por un escalar se sigue que, en espacios vectoriales reales, la norma de un vector es invariante bajo inversiones de sentido y, más generalmente, que en espacios vectoriales complejos, la norma es invariante bajo rotaciones. Por ende, la norma de un vector debe estar *de alguna forma* relacionada con la longitud de la flecha que lo representa.

La siguiente propiedad se conoce como *distinción del vector nulo*, y nos dice que el vector nulo es el único vector del espacio que tiene norma igual a cero. Observemos que el vector nulo es el único vector representado geoméricamente por un punto, que podemos considerar como una *flecha de longitud cero*, reforzando nuestra sospecha anterior de que la norma de un vector debe estar relacionada con su longitud.

La última propiedad se conoce como la *desigualdad del triángulo*, dado que nos recuerda a la desigualdad existente entre los lados de un triángulo aplicada al triángulo que se obtiene mediante la *Ley del paralelogramo* cuando sumamos dos vectores.

²Podríamos aprovechar la mención de ambas operaciones para implícitamente introducir su interpretación geométrica en el plano real y complejo, siguiendo las notas del curso.

$$\begin{array}{l} \text{Norma} \\ || \cdot || : V \rightarrow K \end{array}$$

$$\forall \vec{u} \in V, \forall a \in K,$$

$$||a\vec{u}|| = |a| ||\vec{u}|| \quad \text{Escalabilidad absoluta}$$

$$||\vec{u}|| = 0 \iff \vec{u} = \vec{0} \quad \text{Distinción del vector nulo}$$

$$||\vec{u} + \vec{v}|| \leq ||\vec{u}|| + ||\vec{v}|| \quad \text{Desigualdad del triángulo}$$

Por último, no es difícil demostrar, a partir de las tres propiedades anteriores, que la norma es positivo definida³. Una vez más, esto refuerza nuestra sospecha de que la norma de un vector debe estar relacionada con la longitud de la flecha que lo representa, pues no existen flechas con *longitud negativa*.

A un espacio vectorial con una función que cumpla estas propiedades se le conoce como un *espacio normado*.

1.2. SE2 (Vectores y conjuntos normales)

Supongamos que en un espacio vectorial normado tenemos un vector no nulo, y lo queremos reescalar de tal forma que su norma sea igual a uno, pero sin cambiar su dirección ni su sentido. Dado que la norma es positivo definida y nuestro vector es distinto al vector nulo, sabemos que su norma es positiva. Observemos que, si reescalamos al vector por el inverso multiplicativo de su norma entonces, por la propiedad de escalabilidad absoluta de la norma, se sigue que el vector reescalado tiene norma igual a uno.

$$\vec{v} \in V, \vec{v} \neq \vec{0} \Rightarrow ||\vec{v}|| > 0.$$

$$\frac{1}{||\vec{v}||} \vec{v}$$

$$\left\| \frac{1}{||\vec{v}||} \vec{v} \right\| = \left| \frac{1}{||\vec{v}||} \right| ||\vec{v}||$$

$$= \frac{1}{||\vec{v}||} ||\vec{v}||$$

$$= 1.$$

³Escribir “Ver Ejercicio 2.1.” como nota al pie.

Este proceso de reescalar a un vector no nulo por el inverso multiplicativo de su norma para convertirlo en un vector de norma uno se conoce como *normalización*, y a los vectores con norma igual a uno se les conoce como vectores *normales* o *unitarios*, y se les suele denotar de esta forma

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

Notemos que, dado que la norma de \vec{v} es positiva, el vector normal que hemos obtenido efectivamente apunta en la misma dirección y sentido que el vector original.

Si multiplicamos ambos lados de la igualdad anterior por la norma de \vec{v} , obtenemos la siguiente ecuación

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

Esta igualdad nos dice que, en un espacio normado, cualquier vector no nulo \vec{v} es igual al vector normal en su misma dirección y sentido reescalado⁴ por su norma⁵. Por ende, podemos interpretar geoméricamente a la norma de un vector como la longitud de la flecha que lo representa. De esta manera, la norma nos da una noción de la *magnitud* de un vector y, además, nos sirve como una escala con la cual comparar a diferentes vectores. Más aún, a partir de cualquier norma se puede definir una *métrica inducida*, lo que nos da una noción de *distancia* entre vectores⁶, aunque esto vas más allá del enfoque de esta serie de videos.

Decimos que un conjunto de vectores en un espacio normado es *normal* si todos sus vectores son normales

$$U = \{\vec{u}_1, \dots, \vec{u}_k\} \subseteq V \text{ es normal si } \|\vec{u}_i\| = 1 \text{ para } 1 \leq i \leq k.$$

Notemos que cualquier conjunto que *no* contenga al vector nulo puede ser convertido en un conjunto normal, simplemente normalizando a cada uno de sus vectores:

$$\vec{0} \notin \{\vec{v}_1, \dots, \vec{v}_k\} \subseteq V \Rightarrow \{\hat{v}_1, \dots, \hat{v}_k\} \subseteq V \text{ es un conjunto normal.}$$

⁴Enfatizar a \hat{v} .

⁵Enfatizar a $\|\vec{v}\|$.

⁶Escribir “Ver Ejercicio 2.2.” como nota al pie.

2. Segunda escena

Hasta ahora, en esta serie de videos hemos visto que tanto la norma como el producto escalar son operaciones positivo definidas que distinguen al vector nulo⁷. Sin embargo, estas no son las únicas relaciones existentes entre estas dos operaciones. Observemos lo siguiente: supongamos que tenemos un espacio vectorial con producto escalar y que definimos una función de la siguiente manera

$$(V, K)$$

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow K \text{ es un producto escalar en } V$$

$$\|\vec{v}\| := \sqrt{\langle \vec{v}, \vec{v} \rangle} \quad \forall \vec{v} \in V$$

Entonces, se sigue de las propiedades del producto escalar que nuestra función asigna a cada vector del espacio un número real y, además, que cumple las siguientes propiedades

$$(V, K)$$

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow K \text{ es un producto escalar en } V$$

$$\|\vec{v}\| := \sqrt{\langle \vec{v}, \vec{v} \rangle} \quad \forall \vec{v} \in V$$

$$\|\cdot\| : V \rightarrow K$$

$$\forall \vec{u} \in V, \forall a \in K,$$

$$\|a\vec{u}\| = \sqrt{\langle a\vec{u}, a\vec{u} \rangle} = \sqrt{a\bar{a}\langle \vec{u}, \vec{u} \rangle} = \sqrt{a\bar{a}}\sqrt{\langle \vec{u}, \vec{u} \rangle} = |a| \|\vec{u}\|$$

$$\|\vec{u}\| = 0 \iff \sqrt{\langle \vec{u}, \vec{u} \rangle} = 0 \iff \langle \vec{u}, \vec{u} \rangle = 0 \iff \vec{u} = \vec{0}$$

En otras palabras, nuestra función ya cumple dos de las propiedades de norma. Más aún, se puede demostrar que también cumple la desigualdad del triángulo⁸, por lo que la función que definimos es una norma. En general, siempre que tengamos un espacio vectorial con un producto escalar, ¡podemos definir una norma en ese espacio *a partir* del producto escalar de la misma manera en que lo hicimos anteriormente! A esto se le conoce como una *norma inducida por un producto escalar*. A pesar de que no todas las normas en un espacio vectorial tengan que ser inducidas por un producto escalar, nos enfocaremos principalmente en este tipo de normas durante el resto de la serie⁹.

⁷Escribir “Ver Ejercicio 1.2.” como nota al pie.

⁸Colocar referencias de desigualdad de Cauchy-Schwarz y la desigualdad del triángulo en la descripción.

⁹Podríamos colocar referencias a normas no inducidas por productos escalares en la descripción del video y añadir una

3. Tercera escena

3.1. SE1 (Proyecciones vectoriales y vectores normales/unitarios)

Ahora, volvamos a la expresión de la proyección vectorial de un vector \vec{u} sobre un vector no nulo \vec{v} que vimos al final del video anterior

$$\frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v}$$

Observemos que, considerando la norma inducida, podemos reescribir esta expresión como sigue:

$$\begin{aligned} \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v} &= \frac{\langle \vec{u}, \vec{v} \rangle}{\left(\sqrt{\langle \vec{v}, \vec{v} \rangle}\right)^2} \vec{v} \\ &= \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|^2} \vec{v} \\ &= \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|} \left(\frac{1}{\|\vec{v}\|} \vec{v} \right) \\ &= \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|} \hat{v}. \end{aligned}$$

Es decir, la componente del vector \vec{u} que vive en el subespacio generado por el vector no nulo \vec{v} se obtiene reescalando al vector unitario en la misma dirección y sentido que \vec{v} por el producto escalar de \vec{u} con \vec{v} entre la norma de \vec{v} .

$$\frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v} = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|} \hat{v}$$

3.2. SE2 (Conjuntos y bases ortonormales)

En un espacio vectorial con producto escalar y normado, decimos que un conjunto es *ortonormal* si es ortogonal y normal. Es decir, un conjunto ortonormal es un conjunto de vectores de norma uno que son ortogonales entre sí.

$N = \{\vec{n}_1, \dots, \vec{n}_k\} \subseteq V$ es *ortonormal* si $\langle \vec{n}_i, \vec{n}_j \rangle = 0$ para $i \neq j$ con $1 \leq i, j \leq k$ y $\|\vec{n}_i\| = 1$ para $1 \leq i \leq k$.

Si consideramos la norma inducida por el producto escalar, entonces el que la norma de un vector sea uno es equivalente a que su producto escalar consigo mismo sea igual a uno

$N = \{\vec{n}_1, \dots, \vec{n}_k\} \subseteq V$ es *ortonormal* si $\langle \vec{n}_i, \vec{n}_j \rangle = 0$ para $i \neq j$ con $1 \leq i, j \leq k$ y $\langle \vec{n}_i, \vec{n}_i \rangle = 1$ para $1 \leq i \leq k$.

por lo que podemos caracterizar a un conjunto ortonormal como sigue.

$$N = \{\vec{n}_1, \dots, \vec{n}_k\} \subseteq V \text{ es } \textit{ortonormal} \text{ si } \langle \vec{n}_i, \vec{n}_j \rangle = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \text{ para } 1 \leq i, j \leq k.$$

nota al pie que diga “Ver referencias [N-1], [N-2], etc. en la descripción del video sobre normas no inducidas por productos escalares.”.

En particular, si una base cumple con ser un conjunto ortonormal, decimos que es una *base ortonormal*. Dado que se puede demostrar que todo conjunto ortonormal finito es linealmente independiente¹⁰ se sigue que, en un espacio vectorial de dimensión finita con producto escalar, cualquier conjunto ortonormal con tantos vectores como la dimensión del espacio es una base ortonormal del espacio.

$N = \{\vec{n}_1, \dots, \vec{n}_k\}$ es base *ortonormal* de V si

$$\dim(V) = k, \langle \vec{n}_i, \vec{n}_j \rangle = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

Este tipo de bases son de gran utilidad, como veremos a continuación.

¹⁰Escribir “Ver Ejercicio 2.3.” como nota al pie.

4. Cuarta escena

Regresemos ahora al problema por el cual introdujimos la operación de producto escalar, y veamos cómo entra la norma en nuestra discusión: tenemos un espacio vectorial de dimensión finita y queremos encontrar los coeficientes necesarios para expresar a un vector no nulo arbitrario como combinación lineal de una base. En el video anterior vimos que, si suponemos que nuestro espacio tiene producto escalar y que existe una base ortogonal, entonces los coeficientes se obtienen de la siguiente manera. Ahora, si consideramos la norma inducida por este producto escalar y suponemos que existe una base *ortonormal* entonces, dado que dicha base en particular es ortogonal, podemos aplicar el mismo resultado. Más aún, como ahora todos los vectores de la base son normales, el resultado anterior se simplifica de la siguiente manera. Es decir, obtenemos que el cálculo de los coeficientes buscados se efectúa simplemente a través de un producto escalar.

Por esto es que las bases ortonormales son tan útiles en espacios vectoriales con producto escalar. Algunos ejemplos de este tipo de bases incluyen las bases canónicas de \mathbb{R}^2 y \mathbb{R}^3 .

En el siguiente video, veremos cómo podemos obtener conjuntos ortogonales y ortonormales a partir de un conjunto linealmente independiente. Como consecuencia de esto, veremos además que las bases ortogonales y ortonormales siempre existen en espacios vectoriales de dimensión finita con producto escalar, lo que justifica *a posteriori* toda la discusión contenida en este video y en el video previo.

5. Escena final

Ejercicio 2.1 Sean V un espacio vectorial normado y $\vec{u} \in V$. Demuestra que $\|\vec{u}\| > 0$ si $\vec{u} \neq \vec{0}$, es decir, que la norma es positivo definida.

Ejercicio 2.2 Una *métrica* o *función de distancia* en un conjunto X es una función $f(\cdot, \cdot) : X \times X \rightarrow [0, \infty)$ que cumple las siguientes propiedades:

- (I) $d(x, y) = 0$ si, y sólo si, $x = y$;
- (II) $d(x, y) = d(y, x)$ para cualesquiera $x, y \in X$;
- (III) $d(x, y) \leq d(x, z) + d(z, y)$ para cualesquiera $x, y, z \in X$.

Demuestra que si (V, K) es un espacio vectorial con norma $\|\cdot\|$, entonces la función dada por $d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|$ para todo $\vec{x}, \vec{y} \in V$ es una métrica en V . Por ende, toda norma *induce* una métrica y, en particular, todo producto escalar positivo definido induce una métrica; sin embargo, existen métricas que no son inducidas por normas ni productos escalares.

Ejercicio 2.3 Demuestra que todo conjunto ortonormal finito es linealmente independiente.