

Ortogonalización y ortonormalización

(Teorema de Gram-Schmidt)

Observación 0.1. Hemos visto ya la utilidad de las bases ortogonales y ortonormales —y, a través de ellas, la interpretación del producto escalar y la norma inducida— sin, hasta el momento, haber justificado su existencia. En este video, mostraremos cómo a partir de un conjunto linealmente dependiente finito se puede producir un conjunto ortogonal linealmente independiente u ortonormal que genere el mismo subespacio vectorial. Este resultado se conoce como el Teorema de Gram-Schmidt, y un corolario de este teorema es que las bases ortogonales y ortonormales siempre existen en espacios vectoriales con producto escalar de dimensión finita.

Observación 0.2. Siguiendo de la observación 0.1, las ideas principales a presentar en este video son:

- (I) En un espacio vectorial real, a partir de cualquier conjunto linealmente independiente de dos vectores se puede obtener —mediante la modificación de uno de ellos con las operaciones de suma vectorial, producto de un vector por un escalar, y producto escalar— un conjunto ortogonal linealmente independiente que genere el mismo subespacio vectorial que el conjunto original (de hecho, se pueden obtener dos, dependiendo de cuál de los vectores modificamos), y lo mismo para conjuntos ortonormales.
- (II) En un espacio vectorial real, a partir de cualquier conjunto linealmente independiente de tres vectores se puede obtener un conjunto ortogonal linealmente independiente que genere el mismo subespacio vectorial que el conjunto original (de hecho, se pueden obtener seis, dependiendo de cuál de ellos no modificamos), y lo mismo para conjuntos ortonormales.
- (III) En un espacio vectorial arbitrario, a partir de cualquier conjunto linealmente independiente finito se pueden obtener conjuntos ortogonales linealmente independientes que generen el mismo subespacio que el conjunto original —y lo mismo para conjuntos ortonormales— mediante procesos algorítmicos finitos. Como consecuencia, todos los espacios vectoriales con producto escalar de dimensión finita tienen bases ortogonales y ortonormales.

1. Primera escena

Supongamos que, en un espacio vectorial real con producto escalar, tenemos un conjunto de dos vectores linealmente independientes¹, pero no ortogonales entre sí, y que queremos modificar a alguno de ellos, utilizando las operaciones disponibles², de tal forma que obtengamos un conjunto de dos vectores linealmente independientes y *ortogonales* que genere el mismo subespacio vectorial que el conjunto original³, el cual mostraremos a continuación.

Por definición, el hecho de que nuestros vectores \vec{a} y \vec{b} no sean ortogonales significa que su producto escalar es distinto de cero⁴, por lo que la proyección vectorial⁵ de \vec{a} sobre \vec{b} es distinta del vector nulo. Observemos que, debido a la propiedad⁶ de simetría conjugada del producto escalar, se sigue que la proyección vectorial de \vec{b} sobre \vec{a} también es no nula y, por ende, que cada uno de los vectores debe tener una componente no nula en el subespacio generado por el otro vector.

¿Qué pasa si le removemos⁷ esta componente a uno de los vectores? Obtenemos un conjunto ortogonal, ¡como queríamos! Veamos que este conjunto ortogonal efectivamente *genera el mismo subespacio vectorial que el conjunto linealmente independiente con el que empezamos*⁸; esto se debe a que obtuvimos al vector \vec{b}' como combinación lineal de los vectores de I y, además, a que los vectores claramente no perdieron su independencia lineal⁹ cuando los ortogonalizamos de esta forma.

Observemos que¹⁰, si partiendo de nuestro conjunto original hubiéramos decidido modificar al *otro* vector de manera análoga, entonces hubiéramos obtenido un conjunto ortogonal *distinto*¹¹ al anterior pero que, nuevamente, genera el *mismo* subespacio¹² que el conjunto original. Si quisiéramos hacer que este conjunto fuese ortonormal, podríamos simplemente normalizar a cada uno de los vectores¹³. Sin embargo, si nuestro objetivo hubiera sido obtener un conjunto ortonormal desde el principio¹⁴, entonces, recordando que es más sencillo calcular proyecciones vectoriales sobre vectores unitarios, pudimos haber empezado¹⁵ el proceso normalizando a uno de los vectores y, luego, removiéndole al *otro* vector su componente a lo largo del eje del vector normalizado, y normalizándolo también. Una vez más, observemos que el subespacio generado¹⁶ por este conjunto *ortonormal* es igual al generado por el conjunto linealmente independiente original.

¹Mostramos brevemente que están en ejes distintos.

²Mostramos brevemente algo como $\vec{u} + \vec{v}$, $a\vec{u}$, $\langle \vec{u}, \vec{v} \rangle$ en pantalla, para que se entienda que nos referimos a las operaciones de suma vectorial, producto de un vector por un escalar y producto escalar.

³Mostramos que $\langle I \rangle$ es todo el plano.

⁴Mostramos $\langle \vec{a}, \vec{b} \rangle \neq 0$ en pantalla.

⁵Hacemos dos proyecciones vectoriales rápidas para mostrar que, en efecto, son no nulas.

⁶Debajo de $\langle \vec{a}, \vec{b} \rangle \neq 0$, mostramos brevemente $\langle \vec{b}, \vec{a} \rangle = \overline{\langle \vec{a}, \vec{b} \rangle}$ y, abajo, $\implies \langle \vec{b}, \vec{a} \rangle \neq 0$ (quedaría como un silogismo); luego, desaparecemos los tres renglones, junto con la proyección vectorial de \vec{a} sobre \vec{b} .

⁷Usamos la proyección vectorial de \vec{b} sobre \vec{a} para ortogonalizar a \vec{b} , obteniendo así a \vec{b}' , y escribimos $\Gamma_1 = \{\vec{a}, \vec{b}'\}$.

⁸Mostramos al generado de Γ_1 (que es todo el plano) y escribimos $\langle \Gamma_1 \rangle = \langle I \rangle$ en pantalla.

⁹Trazamos líneas punteadas para mostrar que los vectores ortogonalizados están en ejes distintos.

¹⁰Regresamos al conjunto I original con las dos proyecciones vectoriales, desaparecemos ahora a la proyección vectorial de \vec{b} sobre \vec{a} y usamos la de \vec{a} sobre \vec{b} para ortogonalizar a \vec{a} y obtener así a \vec{a}' .

¹¹Escribimos $\Gamma_2 = \{\vec{a}', \vec{b}\}$ en pantalla.

¹²Mostramos al generado de Γ_2 (nuevamente, todo el plano) y escribimos $\langle \Gamma_1 \rangle = \langle I \rangle = \langle \Gamma_2 \rangle$ en pantalla.

¹³Tomamos el conjunto ortogonal que ya tenemos y normalizamos a cada uno de los vectores; luego, hacemos *fade out*.

¹⁴Borramos todo y dejamos solamente al conjunto l.i. original y el texto $I = \{\vec{a}, \vec{b}\}$.

¹⁵Esta vez, empezamos normalizando a \vec{b} para definir a \hat{b} , luego ortogonalizamos a \vec{a} proyectándolo sobre \hat{b} , obteniendo así a \vec{a}' , y lo normalizamos para definir a \hat{a}' ; finalmente, escribimos $N = \{\hat{a}', \hat{b}\}$.

¹⁶Mostramos al generado de N y escribimos $\langle I \rangle = \langle N \rangle$.

2. Segunda escena

Consideremos ahora a un espacio vectorial real con producto escalar con un conjunto de tres vectores linealmente independientes pero no ortogonales entre sí. Veamos el subespacio generado por este conjunto. ¿Cómo podríamos obtener un conjunto linealmente independiente y ortogonal de tres vectores a partir del conjunto que tenemos, de tal forma que el subespacio generado no cambie?¹⁷ Sugerimos pausar el video¹⁸ para pensar cómo modificarías a estos tres vectores para cumplir nuestro objetivo, y continuar la reproducción hasta que tengas una idea clara —y, de preferencia, escrita— de cómo hacerlo.

Iniciamos de la misma manera: escogemos a un vector de referencia, por ejemplo, el vector \vec{a} ; proyectamos a otro vector —digamos, a \vec{b} — sobre \vec{a} y restamos esa componente a \vec{b} , obteniendo así un conjunto linealmente independiente y ortogonal de dos vectores

$$\{\vec{a}, \vec{b}'\}.$$

De la misma manera, tomamos al vector restante \vec{c} y le removemos su componente a lo largo del eje del vector \vec{a} ; de esta forma, aseguramos que este nuevo vector sea ortogonal al vector \vec{a} ; sin embargo, esto no necesariamente implica que también sea ortogonal a \vec{b}' . Por ende, además de removerle a \vec{c} su componente a lo largo del eje de \vec{a} , debemos removerle su componente a lo largo del eje de \vec{b}' . De esta forma, aseguraremos que el conjunto compuesto por \vec{a} , \vec{b}' y \vec{c}' sea ortogonal y linealmente independiente¹⁹. Observemos que este nuevo conjunto genera el mismo subespacio vectorial que el conjunto original.

$$\Gamma_1 = \{\vec{a}, \vec{b}', \vec{c}'\}.$$

Hacemos énfasis en que, habiendo elegido un vector de referencia —como el vector \vec{a} ²⁰, en el ejemplo anterior—, si cambiamos el orden en el que modificamos a los demás vectores, obtendremos también un conjunto ortogonal y linealmente independiente, aunque puede ser distinto al que obtuvimos anteriormente. Lo mismo ocurre si elegimos a cualquier otro vector como vector de referencia²¹. No es difícil ver que cualquiera de estos conjuntos ortogonales generará el mismo subespacio vectorial que el conjunto original.

Nuevamente, si desde el principio nuestro objetivo hubiera sido obtener un conjunto *ortonormal* de tres vectores, pudimos haber empezado por normalizar a uno de los vectores, después, proyectar a alguno de los otros vectores sobre el vector normalizado, remover esa componente y normalizar al vector resultante y, finalmente, proyectar al último vector sobre los dos vectores normales que tenemos, removiendo dichas componentes y normalizando al vector resultante²².

¹⁷Tal vez poner un i ? grande y divertido en pantalla jijj

¹⁸Dejamos el espacio rotando unos segundos de tal forma que, cuando empiece el siguiente párrafo, ya se encuentre en una posición conveniente para empezar a hacer las proyecciones/modificaciones.

¹⁹Escribimos “Nota: Ver Ejercicio 3.1.”.

²⁰Mostramos dos copias del mismo conjunto *l.i.* arriba y abajo (sin nombres), resaltamos al mismo vector (de referencia), ortogonalizamos en diferente orden y escribimos Γ_1 y Γ_2 .

²¹Hacemos algo análogo para Γ_3 y Γ_4 y, luego, para Γ_5 y Γ_6 .

²²En este párrafo ya no se hacen referencias a vectores con un nombre específico, pues en el párrafo pasado ya se estableció que tanto el punto como de partida como el orden es arbitrario; la animación debería reforzar este hecho quitándole los nombres a los vectores. Conceptualmente, esto facilita la transición a la escena tres.

3. Tercera escena

Este procedimiento para, a partir de un conjunto linealmente independiente de dos o tres vectores, obtener conjuntos linealmente independientes y *ortogonales* que generen el mismo subespacio que el conjunto original puede ser generalizado para conjuntos finitos de cualquier cardinalidad, y se conoce como el proceso de Gram-Schmidt. Lo mismo ocurre con el procedimiento que nos permite obtener conjuntos *ortonormales* sin cambiar el subespacio generado. Éste es conocido como el proceso de Gram-Schmidt *modificado*²³. A continuación, enunciaremos ambos procesos de forma matemática. Sugerimos comparar esto con la versión escrita vista anteriormente.

④ Sea L un conjunto l.i.

Proceso de Gram-Schmidt — ② —	Gram-Schmidt modificado
<p>③ {</p> <ol style="list-style-type: none"> 1. Tomar a un vector de L y agregarlo a un nuevo conjunto Γ. 2. Tomar a otro de los vectores de L, restarle sus proyecciones sobre todos los vectores de Γ, y después agregarlo a Γ. 3. Repetir el paso 2 hasta que Γ tenga tantos vectores como L. <p style="text-align: center;">¡ Γ es \perp !</p>	<ol style="list-style-type: none"> 1. Tomar a un vector de L, normalizarlo y agregarlo a un nuevo conjunto N. 2. Tomar a otro de los vectores de L, restarle sus proyecciones sobre todos los vectores de N, normalizarlo, y después agregarlo a N. 3. Repetir el paso 2 hasta que N tenga tantos vectores como L. <p style="text-align: center;">¡ N es o.n. !</p> <p style="text-align: center;">④ }</p>

— ⑤ —

En conjunto, este resultado se conoce como el Teorema de Gram-Schmidt. Observemos que, como consecuencia de este teorema, se sigue que la existencia de bases ortogonales y ortonormales siempre está asegurada en los espacios vectoriales con producto escalar de dimensión finita²⁴.

⑧ Teorema de Gram-Schmidt

① $\dim(V) = K < \infty$

<p>② {</p> $\beta = \{\vec{v}_1, \dots, \vec{v}_K\}$ $\langle \beta \rangle = V$ $\beta \text{ es l.i.}$	<p>④ {</p> $\Gamma = \{\vec{g}_1, \dots, \vec{g}_K\}$ $\langle \Gamma \rangle = V$ $\Gamma \text{ es l.i.}$	<p>⑥ {</p> $N = \{\vec{n}_1, \dots, \vec{n}_K\}$ $\langle N \rangle = V$ $N \text{ es l.i.}$
	<p>③ {</p> $\vec{g}_1 := \vec{v}_1$ $\vec{g}_j := \vec{v}_j - \sum_{i=1}^{j-1} \frac{\langle \vec{v}_j, \vec{g}_i \rangle}{\ \vec{g}_i\ ^2} \vec{g}_i, \quad 1 < j \leq K$	<p>⑤ {</p> $\vec{n}_1 := \frac{\vec{v}_1}{\ \vec{v}_1\ }$ $\vec{n}_j := \frac{\vec{v}_j - \sum_{i=1}^{j-1} \langle \vec{v}_j, \vec{n}_i \rangle \vec{n}_i}{\ \vec{v}_j - \sum_{i=1}^{j-1} \langle \vec{v}_j, \vec{n}_i \rangle \vec{n}_i\ }, \quad 1 < j \leq K$

⑦ { β es base de V \Rightarrow Γ es base ortogonal de V , N es base ortonormal de V }

²³En esta imagen faltó agregar $\langle \Gamma \rangle = \langle I \rangle = \langle N \rangle$ centrado debajo de donde dice 5 (aparecería hasta el final).

²⁴Mejor poner los vectores “n”s con gorritos ya que, por la forma en que los estamos definiendo, son unitarios, y el objetivo es que la base N sea ortonormal.

4. Escena final

Ejercicio 3.1 Sea V un espacio vectorial con producto escalar. Demuestra que

$$\left\langle \vec{u}, \vec{v} - \frac{\langle \vec{v}, \vec{u} \rangle}{\langle \vec{u}, \vec{u} \rangle} \vec{u} \right\rangle = 0 \quad \text{y} \quad \left\langle \left\{ \vec{u}, \vec{v} - \frac{\langle \vec{v}, \vec{u} \rangle}{\langle \vec{u}, \vec{u} \rangle} \vec{u} \right\} \right\rangle = \langle \{\vec{u}, \vec{v}\} \rangle$$

para $\vec{u}, \vec{v} \in V$ con $\vec{u}, \vec{v} \neq \vec{0}$, y dibuja un ejemplo en \mathbb{R}^2 .²⁵

Ejercicio 3.2 Usa la intuición geométrica generada por este video para demostrar el Teorema de Gram-Schmidt.

Pregunta 3.3 ¿Qué sucedería si aplicáramos el proceso de Gram-Schmidt a un conjunto de vectores linealmente *dependiente*?²⁶

Pregunta 3.4 ¿Qué sucedería si, en el proceso de Gram-Schmidt modificado, normalizáramos a alguno de los vectores *antes* de ortogonalizarlo, y no *después*?²⁷

Pregunta 3.5 El Teorema de Gram-Schmidt **no** asegura la existencia de bases ortogonales y ortonormales en espacios vectoriales con producto escalar de dimensión **infinita**, ¿por qué?²⁸ (Sugerencia: piensa en la Pregunta 1.2.)

²⁵En la descripción del video o en los comentarios podemos poner la siguiente “Pista: Para la primera parte, utiliza el Ejercicio 1.1.”

²⁶Respuesta: obtendríamos un conjunto ortogonal linealmente independiente con tantos vectores como vectores linealmente independientes haya tenido el conjunto original.

²⁷Respuesta: los vectores obtenidos no serían unitarios, por lo que el conjunto resultante no sería ortonormal.

²⁸Respuesta: porque las combinaciones lineales infinitas, en general, no están bien definidas.