CAPÍTULO 3 - ORTOGONALIDAD (2DA. PARTE) [1]

Facultad de Cs. Exactas, Ingeniería y Agrimensura
Universidad Nacional de Rosario

[1] Siguiendo *Linear Algebra and its applications*, G. Strang.

Outline

V: espacio vectorial con producto interno $\langle \perp, \perp \rangle$.

- **Definición**: $u \perp v$ si y sólo si $\langle u, v \rangle = 0$.
- **Definición**: $\|w\|^2 = \langle w, w \rangle$.

El Teorema de Pitágoras vale en cualquier espacio vectorial con producto interno.

Esto es, $u \perp v$ si y solo si

$$\|u\|^2 + \|v\|^2 = \ \|u - v\|^2 = \|u + v\|^2$$

EJERCICIO

V: espacio vectorial con producto interno,
W_1, W_2: subespacios vectoriales de V:

- W_1 y W_2 son *ortogonales* si, para todo $u \in W_1$ y todo $v \in W_2$, se verifica $u \perp v$. Notamos: $W_1 \perp W_2$.
- *Condición necesaria*: $W_1 \perp W_2 \Longrightarrow W_1 \cap W_2 = \{0\}$.
- *Condición suficiente*: Sean U_1, U_2 tales que $W_1 = \langle U_1 \rangle$ y $W_2 = \langle U_2 \rangle$. Entonces, $\forall u \in U_1, v \in U_2, u \perp v \Longrightarrow W_1 \perp W_2$.
- $W_2 = \{u \in V : u \perp v$ para todo $v \in W_1\} = W_1^{\perp}$, *complemento ortogonal de* W_1.
- $W_1^{\perp}$ es un subespacio de V, $W_1^{\perp} \perp W_1$ y $(W_1^{\perp})^{\perp} = W_1$.

Ejemplo: Sea W el subespacio de $\mathbb{R}^{2\times 2}$ generado por $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Obtener $W^{\perp}$ respecto al producto interno de $\mathbb{R}^{2\times 2}$ definido por

$$\langle A,B\rangle = \sum_{i=1}^{2}\sum_{j=1}^{2} A_i^j B_i^j.$$

Solución: Queremos caracterizar las matrices $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in W^{\perp}$.

Sabemos que es condición suficiente que $M \perp A$ y $M \perp B$, o equivalentemente, $\langle M,A\rangle = 0$ y $\langle M,B\rangle = 0$.

Planteando esas dos condiciones obtenemos

$$\left\langle \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} \right\rangle = a+2b-d = 0 \text{ y } \left\langle \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\rangle = a = 0.$$

Solución (continuación)
Resolviendo el sistema, nos queda $a = 0,\ d = 2b$ y c libre. O sea,

$$W^{\perp} = \{ \begin{bmatrix} 0 & b \\ c & 2b \end{bmatrix} : b, c \in \mathbb{R} \}.$$

Si queremos una base de $W^{\perp}$ podemos elegir las *soluciones especiales* del sistema, esto es, las soluciones correspondientes a $b = 1,\ c = 0$ y $b = 0,\ c = 1$. En efecto, es fácil verificar que, para todo $b, c \in \mathbb{R}$,

$$\begin{bmatrix} 0 & b \\ c & 2b \end{bmatrix} = b \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

- $W_1 \perp W_2 \Longrightarrow dim(W_1) + dim(W_2) \leq dim(V)$.
- $W_1 \perp W_2$, $dim(W_1) + dim(W_2) = dim(V) \Longrightarrow W_2 = W_1^{\perp}$.

  Entonces

  Teorema Fundamental del Álgebra Lineal (Parte II): Sea A una matriz $m \times n$. Entonces,

  1. El espacio fila y el espacio nulo de A son complementos ortogonales en $\mathbb{R}^n$.
  2. El espacio columna y el espacio nulo a izquierda de A son complementos ortogonales en $\mathbb{R}^m$.

- **Teorema** (*de descomposición*): Para todo W subespacio de V, W y $W^{\perp}$ *descomponen* al espacio: para todo $v \in V$ existen únicos $w \in W$ y $w^{\perp} \in W^{\perp}$ tales que $v = w + w^{\perp}$.

  Idea en la Prueba:

  1. $\mathscr{B}_1 = \{w^1, \ldots, w^k\}$ base de W y $\mathscr{B}_2 = \{z^1, \ldots, z^{n-k}\}$ base de $W^{\perp}$ entonces $\mathscr{B}_1 \cup \mathscr{B}_2$ base de V.
  2. Para todo $v \in V$, $v = \sum_{i=1}^{k} \alpha_i w^i + \sum_{i=1}^{n-k} \beta_i z^i = w + w^{\perp}$.

- $W_1 \perp W_2$, $dim(W_1) + dim(W_2) = dim(V) \Longrightarrow W_2 = W_1^{\perp}$.

¿Vale la recíproca?

Teorema: Sea V un espacio vectorial con producto interno, W_1, W_2 subespacios ortogonales de V. Entonces, $W_2 = W_1^{\perp}$ si y solo si $dim(W_1) + dim(W_2) = dim(V)$

Prueba: Sólo resta probar $W_2 = W_1^{\perp} \Longrightarrow dim(W_1) + dim(W_2) = dim(V)$.

Sea $W_2 = W_1^{\perp}$ y, para $i = 1, 2$, sea $\mathscr{B}_i$ una base de W_i. Como $W_2 \perp W_1$, ya sabemos que $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2$ es un conjunto de vectores l.i. y además $|\mathscr{B}| = |\mathscr{B}_1| + |\mathscr{B}_2| = dim(W_1) + dim(W_2)$ (justificar). Basta probar que $V = < \mathscr{B} >$ (justificar).

Por el Teorema de descomposición, sabemos que para todo $v \in V$ existen $w \in W_1$ y $w^{\perp} \in W_1^{\perp} = W_2$ tales que $v = w + w^{\perp}$. Como $w \in < \mathscr{B}_1 >$ y $w^{\perp} \in < \mathscr{B}_2 >$, claramente $v \in < \mathscr{B} >$. Por lo tanto, $V = < \mathscr{B} >$. $\square$

Recordemos que dado V un espacio vectorial y U_1, U_2, dos subespacios de V, si para todo $v \in V$, existen únicos $u_i \in U_i$, $i = 1, 2$ tales que $v = u_1 + u_2$ entonces $V = U_1 \oplus U_2$ (suma directa).

Así,

Teorema(*de descomposición*): Sea V un espacio vectorial con producto interno y W un subespacio vectorial de V. Entonces $V = W \oplus W^{\perp}$.

y

Teorema Fundamental del Álgebra Lineal (Parte II): Sea A una matriz $m \times n$. Entonces,

$$\mathbb{R}^n = C(A^T) \oplus N(A) \text{ y } \mathbb{R}^m = C(A) \oplus N(A^T).$$

Veamos qué nos dice este resultado respecto al *efecto* de la transformación lineal definida por una matriz $m \times n$ sobre los espacios $\mathbb{R}^n$ y $\mathbb{R}^m$.

Sea $x \in \mathbb{R}^n$. Sabemos que existen únicos $x_F \in C(A^T)$ y $x_N \in N(A)$ tal que $x = x_F + x_N$. Por lo tanto

$$Ax = Ax_F + Ax_N = Ax_F + 0 = Ax_F \in C(A) \subset \mathbb{R}^m.$$

La tranformación lineal A lleva a todo el espacio $\mathbb{R}^n$ en su espacio columna. En particular su espacio fila *se transforma* en su espacio columna (y su espacio nulo, en el $\{0\} \subset \mathbb{R}^m$).

Lema: Para toda matriz, A $m \times n$, A define un isomorfismo entre su espacio fila y su espacio columna.
Prueba: Ejercicio.

Volvamos al Teorema de descomposición en espacios vectoriales generales.

Teorema (*de descomposición*): Sea V un espacio vectorial con producto interno y W un subespacio vectorial de V. Entonces, para todo $v \in V$ existen únicos $w \in W$ y $w^{\perp} \in W^{\perp}$ tales que $v = w + w^{\perp}$.

¿Cómo obtenemos las "componentes" en cada uno de los subespacios?

Siguiendo la demostración del teorema de *descomposición*, si $\mathscr{B}$ y $\mathscr{B}^{\perp}$ son bases de W y $W^{\perp}$, respectivamente, $\mathscr{B}' = \mathscr{B} \cup \mathscr{B}^{\perp}$ es una base de V y, para todo $v \in V$, los vectores en w y $w^{\perp}$ tales que $v = w + w^{\perp}$ se obtienen fácilmente a partir de **la representación de** v en $\mathscr{B}'$.

Sabemos que, en general, no es fácil obtener la representación de un vector en una base. Salvo que la base sea ortogonal.

Observación: Si $\mathscr{B}, \mathscr{B}^{\perp}$ son bases ortogonales de $W, W^{\perp}$ respectivamente, entonces $\mathscr{B} \cup \mathscr{B}^{\perp}$ es una base ortogonal de V.

Ejercicio: Si $\mathscr{B} = \{w^i : i = 1, \dots, p\}$ y $\mathscr{B}^\perp = \{z^i : i = 1, \dots, k\}$ son bases ortogonales de W y $W^\perp$, respectivamente, entonces, para todo $v \in V$, $v = w + w^\perp$ con

$$w = \sum_{i=1}^{p} \frac{\langle v, w^i \rangle}{\|w^i\|} w^i \quad \text{y} \quad w^\perp = \sum_{i=1}^{k} \frac{\langle v, z^i \rangle}{\|z^i\|} z^i.$$

¿Qué interpretación geométrica tienen estas *componentes ortogonales*?

Recordemos que dados dos vectores no nulos $x, y \in V$, definimos el *coseno de ángulo que forman* x *e* y como $cos(\hat{xy}) = \frac{\langle x,y \rangle}{\|x\|\|y\|}$, entonces si W y $W^\perp$ tienen las bases ortogonales de antes, w y $w^\perp$ pueden reescribirse como

$$w = \sum_{i=1}^{p} \|v\| \, cos(v\hat{w^i}) \, w^i \quad \text{y} \quad w^\perp = \sum_{i=1}^{k} \|v\| \, cos(v\hat{z^i}) \, z^i.$$

Analicemos en $\mathbb{R}^2$ y $\mathbb{R}^3$, donde podemos *ver* los subespacios.

En $\mathbb{R}^2$, *el eje* x y *el eje* y son subespacios ortogonales y $\mathscr{B}_1 = \{e^1\}$ $\mathscr{B}_2 = \{e^2\}$ son bases ortonormales de cada uno de ellos, respectivamente.

El teorema de *descomposición* dice que cualquier vector $v \in \mathbb{R}^2$ puede expresarse como $v = v_x + v_y$, con

$$v_x = \|v\|\, cos(v\hat{e}^1)\, e^1 \quad \text{y} \quad v_y = \|v\|\, cos(v\hat{e}^2)\, e^2.$$

(ver dibujo en archivo ProyeccionR2 o hacer uno más lindo.)

Observar que v_x y v_y son otra cosa que los *vectores proyección de* v *sobre el eje* x *y el eje* y, respectivamente. O sea, cada una de las componentes son las proyecciones del vector sobre cada uno de los subespacios complementos ortogonales.

Pensemos ahora en $\mathbb{R}^3$, W_1 el subespacio vectorial correspondiente al plano xy con base ortonormal $\mathscr{B}_1 = \{e^1, e^2\}$ y W_2, el correspondiente al eje z, con base ortonormal $\mathscr{B}_1 = \{e^3\}$.

Tenemos que $W_2 = W_1^{\perp}$ y el teorema de *descomposición* nos dice que cualquier vector $v \in \mathbb{R}^3$, $v = v_{xy} + v_z$ con

$$v_{xy} = \|v\|\, cos(v\hat{\ }e^1)\, e^1 + \|v\|\, cos(v\hat{\ }e^2) e^2 \qquad v_z = \|v\|\, cos(v\hat{\ }e^3)\, e^3.$$

¿Quiénes son v_{xy} y v_z? Son las proyecciones de v sobre W_1 y W_2, respectivamente.

(ver dibujo en el archivo ProyeccionR3 o hacer uno más lindo.)

Definición: Sea V un espacio vectorial con producto interno y W un subespacio de W con base ortogonal $\{w^1, \ldots, w^p\}$. Para todo $v \in V$, *el vector proyección de v sobre W* es el vector

$$proy_{s/W}\, v = \sum_{i=1}^{p} \|v\|\, cos(v\hat{\ }w^i)\, \frac{w^i}{\|w^i\|} = \sum_{i=1}^{p} \frac{\langle v, w^i \rangle}{\|w^i\|^2}\, w^i.$$

Así, si W y $W^{\perp}$ *tienen bases ortogonales*, la descomposición de $v \in V$ resulta:

$$v = proy_{s/W}\, v + proy_{s/W^{\perp}}\, v.$$

Más importante aún es la *recíproca* de esta propiedad:

Lema: Sea W y $W^{\perp}$ subespacios vectoriales de V *con bases ortogonales*. Sea $v \in V$ y $z = v - proy_{s/W} v$. Entonces, $z \in W^{\perp}$.

Prueba: Sabemos que $v = proy_W v + proy_{W^{\perp}} v$. Además, por definición de z, $v = proy_W v + z$. Por lo tanto $z = proy_{W^{\perp}} v$ y $z \in W^{\perp}$. □

Otra prueba: Si $\mathscr{B} = \{w^i : i = 1, \ldots, p\}$ es base ortogonal de W, basta probar que $\langle z, w^j \rangle = 0$ para $j = 1, \ldots, p$.

$$\langle z, w^j \rangle = \langle v - \sum_{i=1}^{p} \frac{\langle v, w^i \rangle}{\|w^i\|^2} w^i, w^j \rangle = \langle v, w^j \rangle - \sum_{i=1}^{p} \frac{\langle v, w^i \rangle}{\|w^i\|^2} \langle w^i, w^j \rangle =$$

$$= \langle v, w^j \rangle - \frac{\langle v, w^j \rangle}{\|w^j\|^2} \langle w^j, w^j \rangle = \langle w^j, v \rangle - \frac{\langle v, w^j \rangle}{\|w^j\|^2} \|w^j\|^2 = 0.$$

Recordemos:

Definición: Sea V un espacio vectorial con producto interno y W un subespacio de W con base ortogonal $\{w^1, \ldots, w^p\}$. Para todo $v \in V$, *el vector proyección de* v *sobre* W es el vector

$$proy_{s/W}\, v = \sum_{i=1}^{p} \|v\|\, cos(v\hat{\ }w^i)\, \frac{w^i}{\|w^i\|} = \sum_{i=1}^{p} \frac{\langle v, w^i \rangle}{\|w^i\|^2}\, w^i.$$

¿Sólo podemos proyectar sobre subespacios con bases ortogonales?

Veremos que todos los espacios vectoriales (dimensión finita, con producto interno) tienen bases ortogonales.

V: espacio vectorial con producto interno.

Veremos un algoritmo que, a partir de p vectores l.i. de V, genera p vectores ortogonales no nulos en V.

Proceso de Ortogonalización de Gram-Schmidt:
ENTRADA: $\{v^i : i = 1, \ldots, p\}$: vectores l.i. de V.
SALIDA: $\{w^i : i = 1, \ldots, p\}$: vectores ortogonales no nulos de V.

PROCESO:
Para $j = 1, \ldots, p$ hacer

$$w^j = v^j - \sum_{i=1}^{j-1} \frac{\left\langle v^j, w^i \right\rangle}{\left\| w^i \right\|^2} w^i.$$

Teorema: Sean $\{v^i : i = 1, \ldots, p\}$ vectores l.i. de un espacio vectorial V con producto interno y $\{w^i : i = 1, \ldots, p\}$ los vectores resultantes de aplicar sobre ellos el Proceso de Ortogonalización de Gram-Schmidt. Entonces, para todo $i = 1, \ldots, p$, $w^i \neq 0$ y $w^i \perp w^j$ para todo $j = 1, \ldots, p$, $j \neq i$.

Prueba: Sea $i \in \{1, \ldots, p\}$.
Observemos primero que w^i es una combinación lineal de $\{v^j : j = 1, \ldots, i\}$ (Justificar: usar inducción en i). Por lo tanto, como por hipótesis los vectores $\{v^j : j = 1, \ldots, i\}$ son l.i., obtenemos que $w^i \neq 0$. (Justificar).

Por otro lado, si $W^i = \langle \{w^j : j = 1, \ldots, i\} \rangle$, entonces $w^i = v^i - proy_{s/W^{i-1}} v^i$. De acuerdo al lema anterior, $w^i \in (W^{i-1})^{\perp}$ y $w^i \perp w^j$, para todo $j = 1, \ldots, i-1$. Por lo tanto, $w^i \perp w^j$ para todo $i, j = 1, \ldots, p, \;\; i \neq j$. $\square$

Observación: $W^i = \langle \{w^j : j = 1, \ldots, i\} \rangle = \langle \{v^j : j = 1, \ldots, i\}$. Justificar.

El Proceso de Ortogonalización de Gram-Schmidt nos permite probar el siguiente importante resultado:

Teorema: Todo espacio vectorial con producto interno y de dimensión finita tiene una base ortonormal.
Prueba: Ejercicio.

También, gracias a Gram-Schmidt (GS, a partir de ahora) el Teorema de descomposición (por subespacios complementos ortonormales) puede ser enunciado:

Teorema(de descomposición): Sea V un espacio vectorial de dimensión finita, con producto interno y W un subespacio de V. Entonces, $W = W \oplus W^{\perp}$ y, para todo $v \in V$,

$$v = proy_{s/W}\, v + proy_{s/W^{\perp}}\, v.$$

V: espacio vectorial con producto interno
W: subespacio de V con base ortogonal $\{w^1, \dots, w^p\}$.

El *vector proyección ortogonal de* $v \in V$ *sobre* W es

$$proy_{s/W}\, v = \sum_{i=1}^{p} \|v\|\, cos(v\hat{\,}w^i)\, \frac{w^i}{\|w^i\|} = \sum_{i=1}^{p} \frac{\left\langle v, w^i \right\rangle}{\|w^i\|^2}\, w^i.$$

Veamos un ejemplo sencillo y simplificado del rol fundamental de las proyecciones en aplicaciones.

Ejemplo: Cálculo de la velocidad de un satélite.
Para conocer la velocidad (constante) v con la que un satélite viaja hacia Marte, hacemos $k+1$ mediciones en tiempos $t_0 = 0 < t_1 < \ldots < t_k$ y obtenemos las distancias b_i, $i = 0, \ldots k$. Sabemos que v debe verificar,

$$t_i v = b_i - b_0, \;\; i = 1, \dots, k \;\; \text{(sistema con } k \text{ ecuaciones y una variable).}$$

Los errores de medición no nos aseguran que tenga solución. ¿Qué valor de v consideramos?

Sabemos que v debería verificar $t_i v = b_i - b_0, \; i = 1, \dots, k$. Sean

$T = (t_1, \dots, t_k)^T$ (matriz $k \times 1$) y $b = (b_1 - b_0, \dots, b_k - b_0)^T \in \mathbb{R}^k$. El sistema tiene la forma $Tv = b$, $v \in \mathbb{R}$.
Si $b \in C(T)$, el sistema tiene solución (única). Caso contrario,

¿qué valor de v tomamos?

Buscamos *la mejor aproximación de b en $C(T)$* (esto es, buscamos $\tilde{b} \in C(T)$ tal que la distancia $\left\|b - \tilde{b}\right\|$ sea mínima). Luego elegimos v tal que $Tv = \tilde{b}$.

¿Cómo conseguimos $\tilde{b}$?

Como $\tilde{b} \in C(T)$, sabemos que $\tilde{b} = \alpha T$, para algún $\alpha \in \mathbb{R}$.

Si $k=3$, $C(T)$ es una recta en $\mathbb{R}^3$ que pasa por el origen y $b \in \mathbb{R}^3$ un punto por fuera de la recta. Es claro que *el punto más cercano a* b en la recta, se encuentra en el pie de la perpendicular a la recta que pasa por b. O sea, buscamos $\tilde{b}$ tal que $(b-\tilde{b}) \perp T$. Como $\tilde{b} = \alpha T$ para algún $\alpha \in \mathbb{R}$, tenemos:

$$\left\langle b-\tilde{b}, T\right\rangle = \langle b-\alpha T, T\rangle = \langle b, T\rangle - \langle \alpha T, T\rangle = \langle b, T\rangle - \alpha \langle T, T\rangle = 0$$

y por lo tanto $\alpha = \frac{\langle b,T\rangle}{\|T\|^2}$. Así,

$$\tilde{b} = \alpha T = \frac{\langle b,T\rangle}{\|T\|^2}\, T = proy_{s/C(T)}\, b.$$

Si tuviéramos un sistema inconsistente $Ax=b$, con A matriz 3×2 y $b \in \mathbb{R}^2$ y queremos encontrar *la mejor solución* $\hat{x}$ del sistema, buscaríamos $\tilde{b}$ *la mejor aproximación de* b sobre $C(A)$.

Si el rango de A es 2, $C(A)$ puede pensarse como un plano en $\mathbb{R}^3$ que pasa por el origen y b un vector fuera de él. Nuevamente, para encontrar *la mejor aproximación* de b en el plano, buscaríamos el pie de la perpendicular al plano que pasa por b.

Ejercicio: Sea W un subespacio de $\mathbb{R}^3$ con base ortonormal $\{w_1, w_2\}$ y $b \in \mathbb{R}^3 \setminus W$. Sea $\tilde{b} \in W$ tal que $b - \tilde{b} \perp W$. Probar que $\tilde{b} = proy_{s/W} b$.

Esta idea intuitiva respecto a la relación entre *la mejor aproximación* y la proyección son correctas en todos los espacios de dimensión finita (y para algunos de dimensión infinita).

Formalizamos definiciones:

Definición: Dado un espacio vectorial V con producto interno, W un subespacio de V y $v \in V$. Decimos que $\bar{w}$ es *una mejor aproximación de* v *en* W si, para todo $w \in W$,

$$\|\bar{w} - v\| \leq \|w - v\|.$$

Recordemos: El Teorema de Pitágoras vale en cualquier espacio vectorial con producto interno. Esto es, $u \perp v$ si y solo si

$$\|u\|^2 + \|v\|^2 = \ \|u - v\|^2 = \|u + v\|^2$$

Teorema: Sea W un subespacio vectorial de V, con producto interno, $v \in V$ y $\bar{w} = proy_{s/W} v$. Entonces, $\bar{w}$ es la única mejor aproximación de v en W.

Prueba: Como $v - \bar{w} \in W^{\perp}$, para todo $w \in W$, resulta $v - \bar{w} \perp \bar{w} - w$. Entonces, por Pitágoras, tenemos:

$$\|v - \bar{w}\|^2 + \|\bar{w} - w\|^2 = \|(v - \bar{w}) + (\bar{w} - w)\|^2 = \|v - w\|^2 .$$

Si $w \neq \bar{w}$ entonces $\|\bar{w} - w\|^2 > 0$ y resulta

$$\|v - w\|^2 = \|v - \bar{w}\|^2 + \|\bar{w} - w\|^2 > \|v - \bar{w}\|^2 .$$

Entonces, $\|v - \bar{w}\| < \|v - w\|$ para todo $w \in W \setminus \{\bar{w}\}$ y por lo tanto $\bar{w}$ es la única mejor aproximación de v en W. $\square$

Recordemos que el algoritmo de GS, en su paso j ($j = 1, \ldots, p-1$), obtiene $w^{j+1} = v^{j+1} - proy_{s/W^j} v^{j+1}$, donde $W^{(j)}$ es el subespacio generado por los vectores (ortogonales) $w^1, \ldots, w^j$. Esto es, w^j es el *vector error* en la aproximación de v^j en $W^{(j-1)}$.

Veremos en la práctica, a través de ejemplos, cómo estos procesos se aplican en espacios vectoriales de funciones.

Los polinomios son las únicas funciones que puede calcular una computadora. Por lo tanto, cuando una computadora calcula el valor de una función (continua) f no polinómica, en realidad está calculando el valor de *la mejor aproximación de* f en subespacio de polinomios.

Para calcular proyecciones fácilmente, necesitamos bases ortogonales. Si consideramos la integral del producto de la funciones en el intervalo $[-1, 1]$ como producto interno, las potencias enteras (base canónica de los polinomios) no resultan ortogonales. Aplicando G-S a las funciones potencias enteras, obtenemos los denominados *polinomios de Legendre*.

A partir de ahora, trabajamos en $\mathbb{R}^n$ con suma, producto por escalar y producto interno habituales.

En el ejemplo que vimos para aproximar la velocidad del satélite, proyectábamos el vector b correspondiente a las distancias observadas sobre el subespacio generado por el vector de tiempos de medición T. La velocidad aproximada es $v \in \mathbb{R}$ tal que $Tv = proy_{s/<T>}\, b$.

En general, la proyección de un vector $b \in \mathbb{R}^n$ sobre la recta generada por otro vector $a \in \mathbb{R}^n$ es

$$proy_{s/<a>}\, b = \frac{a^T b}{a^T a}\, a.$$

Observar que $\underbrace{(a^T b)}_{\alpha\in\mathbb{R}}\; a = a\,(a^T b) = (a\,a^T)\, b$. Si consideramos la matriz $P = \frac{1}{a^T a}\, aa^T$, resulta

$$proy_{s/<a>}\, b = Pb.$$

Así, la proyección de un vector de $\mathbb{R}^n$ sobre una recta puede ser pensada como una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^n$.

Observaciones:

- P es simétrica. Prueba: $(a\,a^T)^T = (a^T)^T\ a^T = aa^T$
- $P^2 = P$. Prueba: $(i)\ Pb \in <a>$. $\quad (ii)\, d \in <a> \Longrightarrow proy_{s/<a>} d = d$.

Ejemplo: Proyección sobre la recta en $\mathbb{R}^2$ que forma un ángulo θ con el eje x. La recta es el subespacio generado por el vector $a = (cos\theta, sen\theta)$. Como $\|a\| = 1$,

$$P = a\,a^T = \begin{bmatrix} cos^2\theta & cos\theta sen\theta \\ cos\theta sen\theta & sen^2\theta \end{bmatrix}$$

Esta transformación lineal ya la habíamos visto en el Capítulo 2.

Veamos ahora las proyecciones sobre un subespacio general W de $\mathbb{R}^n$.

Dado $b \in \mathbb{R}^n$, buscamos $\hat{b} = proy_{s/W} b$. Sabemos que $\hat{b}$ debe verificar:

- $\hat{b} \in W$ y
- $(b - \hat{b}) \perp W$ o, equivalentemente, $(b - \hat{b}) \in W^{\perp}$.

Si $\{a^i : i = 1, \ldots, k\}$ es una base de W y A es la matriz $n \times k$ cuya columna i-ésima $A^i = a^i$, tenemos que $W = C(A)$.

Por lo tanto, $\hat{b}$ debe verificar: $\hat{b} \in C(A)$ y $(b - \hat{b}) \in (C(A))^{\perp} = N(A^T)$.
Equivalentemente,

$$\hat{b} = A\hat{x} \;\; \text{para algún } \hat{x} \in \mathbb{R}^k \quad \text{y} \quad A^T(b - \hat{b}) = A^T(b - A\hat{x}) = 0.$$

Como

$$A^T(b - A\hat{x}) = 0 \Longrightarrow (A^T A)\hat{x} = A^T b,$$

para obtener $\hat{b} = proy_{s/W} b$:

1. Resolvemos el *sistema de ecuaciones normales* $(A^T A)\hat{x} = A^T b$,
2. Calculamos $\hat{b} = A\hat{x}$.

Observación: Si tenemos un sistema $Ax = b$ inconsistente, la *mejor solución* $\hat{x}$ (que minimiza el error) es solución del sistema de ecuaciones normales:

$$(A^T A)\,\hat{x} = A^T b.$$

Si las columnas de A son l.i., $A^T A$ es inversible (Ejercicio). En ese caso,
$\hat{x} = (A^T A)^{-1} A^T b \quad$ y $\quad \hat{b} = A\hat{x} = A(A^T A)^{-1} A^T b$.

Ejemplo: Sean

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}, \; b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}.$$

Aquí el subespacio $W = C(A)$ resulta ser el plano xy en $\mathbb{R}^3$ y es fácil intuir que $\hat{b} = (4,5,0)^T$. No es difícil chequear que $\hat{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y finalmente,

$$\hat{b} = A\hat{x} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}.$$

Hemos visto entonces que, dada A una matriz $n \times k$ con columnas l.i. y $b \in \mathbb{R}^n$,

$$\hat{b} = proy_{s/C(A)} b = \left(A(A^T A)^{-1} A^T\right) b = Pb.$$

Nuevamente, la proyección sobre $C(A)$ puede ser vista como una tranformación lineal definida por la *matriz de proyección (sobre $C(A)$)*

$$P = A(A^T A)^{-1} A^T.$$

Veamos algunas propiedades que caracterizan a las matrices de proyección:

Teorema: Sea A una matriz $n \times k$ con columnas l.i. y P la matriz proyección (sobre $C(A)$). Entonces P es idempotente ($P^2 = P$) y simétrica.
Recíprocamente, toda matriz P simétrica e idempotente es una matriz proyección (sobre $C(P)$).
Prueba: Ejercicio.

Observación: Si P es la matriz de proyección sobre $C(A)$ y $b \in \mathbb{R}^n$ entonces $Pb \in C(A)$, $e = b - Pb \in N(A^T)$ y $b = Pb + (b - Pb)$. Por lo tanto, Pb y e son *la descomposición de* b en $C(A)$ y $C(A)^\perp = N(A^T)$.

Observaciones:

- Si A es inversible, $P = A(A^TA)^{-1}A^T = I$. ¿Tiene sentido esto?
- Si A es inversible, sus columnas son una base de $\mathbb{R}^n$. Si la base es una base ortonormal (sabemos que existe, por G-S), ¿quién es A^TA?.
  En ese caso, $A^TA = I$.
- Si las columas de A son ortonormales, A no necesariamente cuadrada, $A^TA = I$.
- Sea A una matriz de tamaño $n \times k$ con columnas ortonormales y $b \in \mathbb{R}^n$. Entonces, la matriz P proyección (sobre $C(A)$) es $P = AA^T$. (Justificar).

Definición: Una matriz ortogonal es una matriz cuadrada con columnas ortonormales.

Ejemplos:

- Las matrices de rotación.
- Las matrices de permutación.

Lema: Sea Q es una matriz $n \times n$. Entonces:

1. Q es ortogonal si y solo si $Q^T = Q^{-1}$.
2. Si Q es ortogonal, entonces $\|Qx\| = \|x\|$ para todo $x \in \mathbb{R}^n$. (Q preserva longitud).

Prueba Ejercicio.

Sea A una matriz $n \times k$ cuyas columnas son l.i. y Q^i $i = 1, \ldots, k$, los vectores ortonormales obtenidos al aplicar GS a los vectores columna de A y luego normalizarlos (dividiendo por su norma). Sea Q la matriz $n \times k$ de columnas Q^i, $i = 1, \ldots, k$.

Se puede probar que, en general, $A = QR$, donde R es una matriz triangular superior de tamaño $k \times k$.

Esto es lo que se denomina *factorización* QR de una matriz.

Teorema: Toda matriz A de tamaño $n \times k$ y con columnas l.i. puede ser factorizada $A = QR$ con Q una matriz $n \times k$ cuyas columnas son ortonormales y R, triangular superior e inversible de tamaño $k \times k$.

Prueba: Ver última sección.

Contar con la factorización QR de una matriz simplifica la resolución de sistemas inconsistentes por aproximación.
En efecto, si A es una matriz de columnas l.i. y $A = QR$ es su factorización QR, entonces la matriz el sistema de ecuaciones normales para obtener $\hat{b} = proy_{s/C(A)}b$ resulta:

$$A^T A\hat{x} = A^T b \Longleftrightarrow R^T Q^T QR\hat{x} = R^T Q^T b \Longleftrightarrow R^T R\hat{x} = R^T Q^T b \Longleftrightarrow R\hat{x} = Q^T b.$$

Por lo tanto $\hat{x}$ se obtiene resolviendo un sistema triangular, por sustitución hacia atrás.

Teorema: Toda matriz A de tamaño $n \times k$ y con columnas l.i. puede ser factorizada $A = QR$ con Q una matriz $n \times k$ cuyas columnas son ortonormales y R, triangular superior e inversible de tamaño $k \times k$.

Prueba: Sea Q la matriz cuyas columnas son los vectores obtenidos aplicando el proceso de $G-S$ a las columnas de A, normalizadas. Debemos encontrar R tal que $QR = A$.

En primer lugar observemos que, por ser R triangular superior, la columna i-ésima de QR es de la forma:

$$(QR)^i = \sum_{j=1}^{i} R_j^i Q^j, \, i = 1, \dots, n.$$

Como queremos probar que $QR = A$, debemos encontrar las entradas R_j^i, $i = 1, \dots, n, j = 1 \dots, i$ tales que

$$A^i = \sum_{j=1}^{i} R_j^i Q^j, \, i = 1, \dots, n.$$

Por otro lado, recordemos que si $\tilde{Q}^i$, $i = 1, \ldots, n$ son los vectores que se obtienen de aplicar el proceso de ortogonalización de $G-S$ a las columnas de A (sin normalizar), entonces $Q^i = \frac{\tilde{Q}^i}{\|\tilde{Q}^i\|}$ y

$$\tilde{Q}^i = A^i - \sum_{j=1}^{i-1} \frac{\left\langle A^i, \tilde{Q}^j \right\rangle}{\left\|\tilde{Q}^j\right\|^2} \tilde{Q}^j = A^i - \sum_{j=1}^{i-1} \left\langle A^i, \frac{\tilde{Q}^j}{\left\|\tilde{Q}^j\right\|} \right\rangle \frac{\tilde{Q}^j}{\left\|\tilde{Q}^j\right\|} = A^i - \sum_{j=1}^{i-1} \left\langle A^i, Q^j \right\rangle Q^j.$$

Así,

$$A^i = \tilde{Q}^i + \sum_{j=1}^{i-1} \left\langle A^i, Q^j \right\rangle Q^j = \left\|\tilde{Q}^i\right\| Q^i + \sum_{j=1}^{i-1} \left\langle A^i, Q^j \right\rangle Q^j.$$

Es fácil verificar que $\left\|\tilde{Q}^i\right\| = \left\langle A^i, Q^i \right\rangle$. Por lo tanto, $A^i = \sum_{j=1}^{i} \left\langle A^i, Q^j \right\rangle Q^j$.

Recordemos que buscamos R tal que, para $i = 1, \ldots, n$, $A^i = \sum_{j=1}^{i} R^i_j Q^j$.
Por lo tanto, basta considerar R tal que $R^i_j = \left\langle A^i, Q^j \right\rangle$ para $i = 1, \ldots, n$ y $j = 1, \ldots, i$. □