TEMA 3: ORTOGONALIDAD Y MÍNIMOS CUADRADOS.

5.1. Producto escalar y ortogonalidad.

Producto escalar y norma de vectores

Producto escalar. Consideremos dos vectores $u = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}, v = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \in \mathbb{R}^n$,

se denomina *producto escalar de los vectores u y v*, al número real

$$u \cdot v = u^T v = c_1 d_1 + c_2 d_2 + \cdots + c_n d_n.$$

Norma de un vector. Se denomina *norma del vector $v \in \mathbb{R}^n$* al número real no-negativo

$$\|v\| = \sqrt{v \cdot v} \geq 0,$$

5.1. Producto escalar y ortogonalidad.

Distancia entre vectores.

Distancia entre dos vectores. Se denomina *distancia entre* $u, v \in \mathbb{R}^n$ al número real no-negativo

$$d(u, v) = \|u - v\| .$$

Ángulo entre dos entre dos vectores. El ángulo determinado por dos vectores no nulos $u, v \in \mathbb{R}^n$ puede caracterizarse mediante la igualdad

$$u \cdot v = \|u\| \, \|v\| \cos(\widehat{u, v})$$

Propiedades del producto escalar. Sean $u, v \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$, se verifican:

(1) $\|v\| = 0$ si y sólo si $v = 0$.

(2) $\|\alpha v\| = |\alpha| \, \|v\|$.

(3) Desigualdad triangular: $\|u + v\| \leq \|u\| + \|v\|$.
Como consecuencia $\|u - v\| \leq \|u\| + \|v\|$.

(4) Desigualdad de Cauchy-Schwartz: $|u \cdot v| \leq \|u\| \, \|v\|$.

5.1. Producto escalar y ortogonalidad.

Ortogonalidad

Vectores ortogonales. Se dice que dos vectores $u, v \in \mathbb{R}^n$ son *ortogonales*, y se denota por $u \perp v$, si su producto escalar vale 0, esto es, si $u \cdot v = 0$.

En general, el conjunto de vectores $\{v_1, v_2, \ldots, v_p\}$ de $\mathbb{R}^n$ es ortogonal si cada uno de los vectores v_k es ortogonal a todos los demás, esto es, $v_k \cdot v_j = 0$ para $j \neq k$. Si además, cada uno de los vectores v_k tiene norma igual a 1, esto es,

$$v_k \cdot v_j = 0 \text{ y } \|v_k\| = 1 \text{ para } j \neq k \text{ y } k, j \in \{1, 2, \ldots, p\}.$$

se dice que el conjunto es *ortonormal*.

5.1. Producto escalar y ortogonalidad.

Ortogonalidad

Propiedades de la ortogonalidad.

(1) Los vectores $u, v \in \mathbb{R}^n$ son ortogonales si y sólo si forman un ángulo de 90 grados.

(2) *Teorema de Pitágoras.* Los vectores $u, v \in \mathbb{R}^n$ son ortogonales si y sólo si

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

(3) Si $\{v_1, v_2, \ldots, v_p\}$ es un conjunto de vectores no nulos ortogonales dos a dos, entonces son linealmente independientes.

Cuestión: Demostrar (2) (cuidado, se trata de una doble implicación).

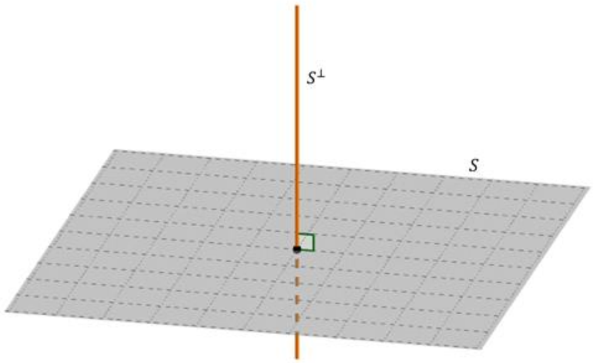
5.2. El subespacio ortogonal a un subespacio dado.

Subespacio ortogonal (Complemento ortogonal)

Subespacio ortogonal a uno dado. Dado un subespacio vectorial S de $\mathbb{R}^n$ se denomina *subespacio ortogonal de* S al conjunto $S^\perp$ formado por todos los vectores de $\mathbb{R}^n$ que son ortogonales a todos los de S,

$$S^\perp = \{v \in \mathbb{R}^n : v \perp u \text{ para todo } u \in S\}.$$

El subespacio ortogonal al subespacio nulo $\{0\}$ es $\mathbb{R}^n$ y viceversa.



5.2. El subespacio ortogonal a un subespacio dado.

Subespacio vectorial

Propiedades del subespacio ortogonal. Dado un subespacio S de $\mathbb{R}^n$ se verifica:

(1) $S^\perp$ es un subespacio vectorial.

(2) $\left(S^\perp\right)^\perp = S$.

(3) El único vector que está en S y en $S^\perp$ es el vector nulo.

(4) Si $S = Gen\{v_1, v_2, \ldots, v_p\}$ entonces

$$v \in S^\perp \text{ si y sólo si } v \perp v_1,\ v \perp v_2,\ \ldots,\ v \perp v_p,$$

esto es, para probar que un vector es ortogonal a todos los vectores de un subespacio vectorial S basta ver que es ortogonal a los vectores que generan a S.

(5) Si A es una matriz real $m \times n$. Se verifica:

$$(Col(A))^\perp = Nul(A^T), \qquad (Nul(A))^\perp = Col(A^T).$$

El espacio $Col(A^T)$ se suele denominar *espacio fila de la matriz* A.

(6) $\dim(S^\perp) = n - \dim(S)$.

5.3. Bases ortonormales de un subespacio.

Base ortonormal

Base ortonormal de un subespacio. Sea S un subespacio vectorial de $\mathbb{R}^n$ y sea $\{v_1, v_2, \ldots, v_p\}$ un conjunto de vectores de S. Decimos que $\{v_1, v_2, \ldots, v_p\}$ constituyen una *base ortogonal* de S si son base de S y, además, conjunto ortogonal. Decimos que $\{v_1, v_2, \ldots, v_p\}$ constituyen una *base ortonormal* de S si es una base y conjunto ortonormal.

Desarrollo de Fourier de un vector. Sea $\{v_1, v_2, \ldots, v_p\}$ una base ortogonal de un subespacio S de $\mathbb{R}^n$. Entonces, las coordenadas de un vector $v \in S$ respecto de dicha base vienen dadas por $\dfrac{v \cdot v_k}{\|v_k\|^2}$, es decir, se verifica que

$$v = \frac{v \cdot v_1}{\|v_1\|^2} v_1 + \frac{v \cdot v_2}{\|v_2\|^2} v_2 + \cdots + \frac{v \cdot v_p}{\|v_p\|^2} v_p.$$

Como caso particular tenemos que si $\{v_1, v_2, \ldots, v_p\}$ es una base ortonormal de S, entonces $v = (v \cdot v_1)\, v_1 + (v \cdot v_2)\, v_2 + \cdots + (v \cdot v_p)\, v_p.$

5.3. Bases ortonormales de un subespacio.

Base ortonormal

Matriz ortogonal. Se denomina *matriz ortogonal* a toda matriz Q real cuadrada cuyas columnas son ortonormales.

Propiedades de las matrices ortogonales. Sea Q una matriz ortogonal. Se verifica:

(1) Q es no singular y $Q^{-1} = Q^T$.

(2) $\det(Q) = \pm 1$.

(3) Q^T es ortogonal. Por tanto, las filas de Q son ortonormales.

(4) Si Q' es otra matriz ortogonal entonces QQ' es ortogonal.

(5) Si T es la transformación lineal asociada a la matriz Q entonces T conserva ángulos y distancias, es decir:

a) $\|Qx\| = \|x\|$.

b) $(Qx) \cdot (Qy) = x \cdot y$.

c) $Qx \perp Qy$ si y sólo si $x \perp y$.

5.4. La proyección ortogonal.

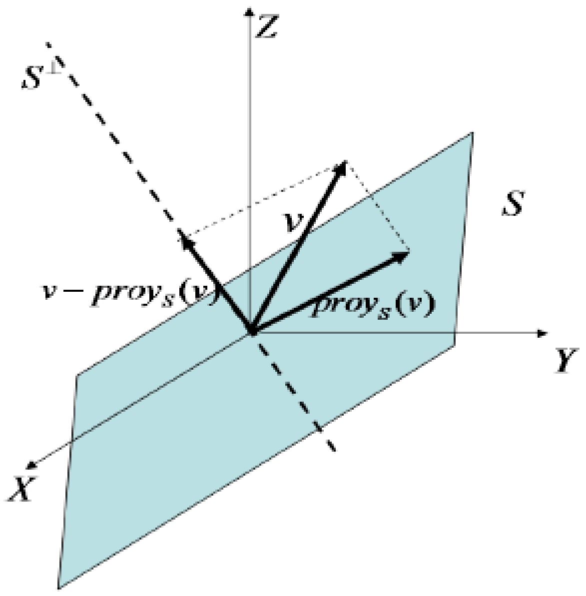
Proyección ortogonal

Proyección ortogonal. Sea S un subespacio vectorial de $\mathbb{R}^n$. Dado cualquier vector $v \in \mathbb{R}^n$ existe un único vector $u \in S$ llamado *proyección ortogonal de v sobre S* tal que $v - u \in S^{\perp}$.

De hecho, si $\{u_1, u_2, \ldots, u_p\}$ es una base ortogonal de S, entonces la proyección ortogonal de v sobre S es

$$u := proy_S(v) = \frac{v \cdot u_1}{\|u_1\|^2} u_1 + \frac{v \cdot u_2}{\|u_2\|^2} u_2 + \cdots + \frac{v \cdot u_p}{\|u_p\|^2} u_p$$

y la proyección ortogonal de v sobre $S^{\perp}$ es $proy_{S^{\perp}}(v) = v - u$.



5.4. La proyección ortogonal.

Proyección ortogonal

Propiedades de la proyección ortogonal. Sea S un subespacio vectorial de $\mathbb{R}^n$.

(1) Si $v \in S$, entonces $proy_S(v) = v$ y $proy_{S^\perp}(v) = 0$.

(2) Se verifica que

$$proy_S(v) + proy_{S^\perp}(v) = v.$$

Por tanto, todo vector $v \in \mathbb{R}^n$ se puede expresar de forma única como suma de un vector de S y otro de $S^\perp$.

(3) Si $\{u_1, u_2, \ldots, u_p\}$ es una base ortonormal de S, entonces la proyección ortogonal de un vector $v \in \mathbb{R}^n$ sobre S es

$$u := proy_S(v) = (v \cdot u_1)\, u_1 + (v \cdot u_2)\, u_2 + \cdots + (v \cdot u_p)\, u_p.$$

(4) La transformación que a cada $v \in \mathbb{R}^n$ le hace corresponder $proy_S(v) \in S$ es una transformación lineal.

5.4. La proyección ortogonal.

Proyección ortogonal

Matriz de la proyección ortogonal. Sea S un subespacio vectorial de $\mathbb{R}^n$. Si U es una matriz cuyas columnas forman una base ortonormal de S, entonces la matriz de la transformación proyección ortogonal sobre S es $P_S = UU^T$, esto es,

$$proy_S(v) = UU^T v \quad \text{para cualquier } v \in \mathbb{R}^n.$$

Dicha matriz verifica las siguientes propiedades:

(1) $(P_S)^2 = P_S$

(2) P_S es simétrica.

(3) $P_S + P_{S^\perp} = I$.

Teorema de la mejor aproximación. Sea S un subespacio vectorial de $\mathbb{R}^n$ y consideremos un vector $v \in \mathbb{R}^n$ y un vector $u \in S$. Son equivalentes:

(1) u es la proyección ortogonal de v sobre S, esto es, $u \in S$ y $v - u \in S^\perp$.

(2) u es el vector de S más próximo a v, esto es, $u \in S$ y para todo $w \in S$ se tiene que $\|v - u\| \leq \|v - w\|$.

5.4. La proyección ortogonal.

El Método de ortogonalización de Gram-Schmidt

El método de ortogonalización de Gram-Schmidt. El método de ortogonalización de Gram-Schmidt permite construir de manera progresiva una base ortogonal de un subespacio vectorial a partir de una base de dicho subespacio e incluso a partir de un conjunto de vectores que genere dicho subespacio. Consideremos una base $\{v_1, v_2, \ldots, v_p\}$ de un subespacio vectorial S de $\mathbb{R}^n$. Entonces los siguientes vectores

$$\begin{aligned}
u_1 &= v_1 \\
u_2 &= v_2 - \frac{v_2 \cdot u_1}{\|u_1\|^2} u_1 \\
u_3 &= v_3 - \frac{v_3 \cdot u_1}{\|u_1\|^2} u_1 - \frac{v_3 \cdot u_2}{\|u_2\|^2} u_2 \\
&\vdots \\
u_p &= v_p - \frac{v_p \cdot u_1}{\|u_1\|^2} u_1 - \frac{v_p \cdot u_2}{\|u_2\|^2} u_2 - \cdots - \frac{v_p \cdot u_{p-1}}{\|u_{p-1}\|^2} u_{p-1}
\end{aligned}$$

5.4. La proyección ortogonal.

El Método de ortogonalización de Gram-Schmidt

están bien definidos, son no nulos y ortogonales dos a dos. Además:

(1) $\{u_1, u_2, \ldots, u_p\}$ es una base ortogonal de $S = Gen\{v_1, v_2, \ldots, v_p\}$.

(2) Para cada $k = 1, 2, ..., p$, $\{u_1, u_2, \ldots, u_k\}$ es una base ortogonal del subespacio $Gen\{v_1, v_2, \ldots, v_k\}$.

Notas:

(1) Si el objetivo es conseguir una base ortonormal de S, una vez que se ha obtenido una base ortogonal basta normalizar los vectores obtenidos.

(2) En cada paso del método de Gram-Schmidt que acabamos de describir podríamos multiplicar o dividir el vector obtenido por un coeficiente no nulo y seguir los cálculos con dicho vector.

(3) Si no se parte de una base sino que, por ejemplo, el vector v_k es combinación lineal de los anteriores $v_1, v_2, \ldots, v_{k-1}$, al aplicar el método de Gram-Schmidt obtenemos $u_k = 0$. Es decir, el método de Gram-Schmidt devuelve el vector nulo cuando se aplica a un conjunto de vectores linealmente dependientes.

5.5. Problemas de mínimos cuadrados.

Mínimos cuadrados

Solución en el sentido de los mínimos cuadrados. Sea $Ax = b$ un sistema de ecuaciones lineales, con A matriz $m \times n$. Encontrar una solución en el sentido de mínimos cuadrados consiste en encontrar un vector $x_0 \in \mathbb{R}^n$ para el cual $\|Ax_0 - b\|$ sea mínima, es decir,

$$\|Ax_0 - b\| \leq \|Ax - b\| \text{ para todo } x \in \mathbb{R}^n.$$

Ecuaciones normales de Gauss. Consideremos un sistema $Ax = b$ con A una matriz real $m \times n$ y $b \in \mathbb{R}^m$ y sea $x_0 \in \mathbb{R}^n$. Son equivalentes:

(1) x_0 es solución en el sentido de mínimos cuadrados del sistema $Ax = b$.

(2) x_0 verifica $Ax_0 = proy_{Col(A)}(b)$.

(3) x_0 verifica $A^T A x_0 = A^T b$, es decir, es solución del sistema lineal

$$A^T A x = A^T b$$

llamado *ecuaciones normales de Gauss.*

5.5. Problemas de mínimos cuadrados.

Mínimos cuadrados

Notas:

(1) Las ecuaciones normales de Gauss constituyen un sistema compatible.

(2) El sistema de ecuaciones $Ax = proy_{Col(A)}(b)$ es equivalente a las ecuaciones normales de Gauss.

(3) La ecuaciones normales de Gauss forman un sistema compatible determinado si y sólo si el sistema homogéneo $Ax = 0$ tiene solución única.