

Análisis Matemático

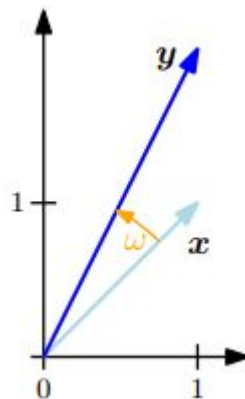
Clase 2

Algunos comentarios sobre la clase pasada

Ángulo entre elementos de un e.v.

Def: A partir de un p.i. se puede definir el **ángulo** ω entre dos vectores x , y como

$$\cos(\omega) = \frac{\langle x, y \rangle}{\|x\| \|y\|}$$



Obs: si x , y son ortogonales, entonces

$$\cos(\omega) = 0, \text{ y } \omega = \pm\pi$$

Base ortonormal

Def: Una **base ortonormal (BON)** de un espacio vectorial V es una base $B = \{v_1, \dots, v_n\}$ que satisface

$$\begin{aligned}\langle v_i, v_j \rangle &= 0 & i \neq j \\ \langle v_i, v_i \rangle &= 1\end{aligned}$$

Obs: Si sólo se cumple que $\langle v_i, v_j \rangle = 0 \quad i \neq j$, B se dice una **base ortogonal**

Para transformar una base en una base ortonormal usamos el proceso de **Gram-Schmidt**

Complemento ortogonal

Def: Sea V un EV de dimensión n y $S \subseteq V$ un SEV de V de dimensión $m \leq n$. El **complemento ortogonal** S^\perp es un SEV de V de dimensión $(n - m)$ que satisface

$$\begin{aligned} S \cup S^\perp &= V \\ S \cap S^\perp &= \{0\} \end{aligned}$$

Matrices definidas positivas

Def: Una matriz $A \in \mathbb{R}^{n \times n}$ se dice **definida positiva** si es simétrica y vale que

$$x^T A x > 0 \quad \forall x \in \mathbb{R}^n$$

Si vale que $x^T A x \geq 0 \quad \forall x \in \mathbb{R}^n$ se la llama **semi definida positiva**

Matrices def. positivas y p.i.

Teor: Sea V un EV de dimensión finita, y B una base de V , vale que $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ es un p.i. **sii** existe una matriz definida positiva $A \in \mathbb{R}^{n \times n}$ tal que

$$\langle x, y \rangle = \hat{x}^T A \hat{y}$$

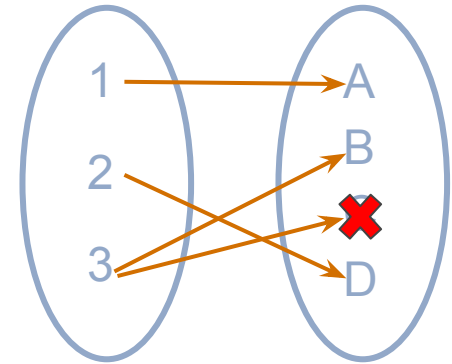
donde, \hat{x} , \hat{y} son las representaciones de x , y en la base B

Transformaciones

Transformaciones

Def: Sea una transformación $\Phi : V \rightarrow W$, donde V, W son dos conjuntos arbitrarios. Φ se dice:

- **Inyectiva:** si $\forall x, y \in V : \Phi(x) = \Phi(y) \Rightarrow x = y$
- **Surjectiva:** si $\Phi(V) = W$
- **Bijectiva:** si es inyectiva y surjectiva



Transformaciones lineales

Def: Sean V y W dos espacios vectoriales.

$\Phi : V \rightarrow W$ es una transformación lineal si:

$$\forall \lambda, \psi \in K, x, y \in V : \Phi(\lambda x + \psi y) = \lambda \Phi(x) + \psi \Phi(y)$$

Corolario: toda transformación lineal se puede representar de forma matricial.

Transformaciones lineales especiales

- Isomorfismo: $\Phi : V \rightarrow W$ es lineal y biyectiva
- Endomorfismo: $\Phi : V \rightarrow V$ es lineal
- Automorfismo: $\Phi : V \rightarrow V$ es lineal y biyectiva
- $id_V : V \rightarrow V, x \mapsto x$ (transformación identidad)

Representaciones

Teorema: V y W , dos espacios vectoriales de dim. finita son un isomorfismo sii $\dim(V) = \dim(W)$

Corolario: Todo e.v. V de dimensión finita (n) tiene un isomorfismo con \mathbb{R}^n . Si consideramos la base $B = (v_1, \dots, v_n)$ todo $v \in V$ puede escribirse como $v = \alpha_1 v_1 + \dots + \alpha_n v_n$. Luego las coordenadas de v en la base B resulta

$$\alpha = [\alpha_1, \dots, \alpha_n]^T \in \mathbb{R}^n$$

Espacio nulo e Imagen.

Def: Sea una transformación $\Phi : V \rightarrow W$. Se define

Espacio nulo (kernel): $Nul(\Phi) := \{v \in V : \Phi(v) = 0_W\}$

Imagen (rango): $rg(\Phi) := \Phi(V) = \{w \in W \mid \exists v \in V : w = \Phi(v)\}$

Si Φ es una transformación lineal, con matriz asociada A ,
el kernel y rango de la transformación se asocian con
los espacios nulo y columna de A

Sea $A \in \mathbb{R}^{n \times m}$ se definen:

Espacio nulo: $x \in \mathbb{R}^m \mid Ax = 0$

Espacio columna: $gen(a_1, \dots, a_m)$, con a_i la columna i de A

Proyección ortogonal

Proyección ortogonal: motivación

Las proyecciones ortogonales son un caso particular (importantísimo!) de las transformaciones lineales.

Algunos usos:

- Reducción de dimensiones (proyecto sobre un subconjunto más pequeño)
- Regresión/Clasificación
- Compresión

Proyección

Def: Sea V un EV y $U \subseteq V$ un SEV de V . Una transformación lineal $\pi : V \rightarrow U$ es la **proyección** si

$$\pi^2 = \pi \circ \pi = \pi$$

Esto significa que la matriz de transformación asociada a la proyección debe satisfacer:

$$P_{\pi}^2 = P_{\pi} \text{ (idempotencia)}$$

Proyección ortogonal

Dado V un EV con **p.i.** y $S \subset V$ un SEV de V , el objetivo es dado $v \in V$ hallar $\hat{v} \in S$ que sea “lo más parecido posible” a v .

$$\hat{v} \in S : \hat{v} = \arg \min_{s \in S} \|v - s\|.$$

Además vale que $\langle v - \hat{v}, s \rangle = 0 \quad \forall s \in S$

Obs: La ortogonalidad de la proyección tiene que ver con el p.i. que se esté usando

Teorema de proyección

Teorema: Sea V un EV de dimensión finita con p.i. $\langle \cdot, \cdot \rangle$, S un SEV de V . Dado $x \in V$ existe un **único** $\hat{x} \in S$ tal que

$$\|x - \hat{x}\| \leq \|x - z\| \quad \forall z \in S$$

Más aún: $\langle x - \hat{x}, s \rangle = 0 \quad \forall s \in S$

Cómo hallar la proyección

Sea V un EV de dimensión n con p.i. $\langle \cdot, \cdot \rangle$, y $S \subset V$ un subespacio de V de dimensión $m \geq 1$, y sea $B = \{s_1, \dots, s_m\}$ una BON de S . Buscamos encontrar la proyección $\hat{v} \in S$ de $v \in V$ ($\hat{v} = \pi_S(v)$)

Como $\hat{v} \in S$, $\hat{v} = \sum_{i=1}^m \alpha_i s_i \Rightarrow$ Busco los coeficientes que minimizan $\|v - \sum_{i=1}^m \alpha_i s_i\|$. El problema puede escribirse como:

$$\pi_S(v) = \sum_{i=1}^m \alpha_i s_i = B\alpha, \quad B = [s_1, \dots, s_m], \quad \alpha = [\alpha_1, \dots, \alpha_m]^T$$

Cómo hallar la proyección.

Como por definición $\langle v - \pi_S(v), s \rangle = 0 \quad \forall s \in S$, debo resolver el siguiente sistema de ecuaciones:

$$\begin{array}{ccc} \langle v - \pi_S(v), s_1 \rangle = 0 & s_1^T (v - B\alpha) = 0 & \\ \vdots & \vdots & \\ \langle v - \pi_S(v), s_m \rangle = 0 & s_m^T (v - B\alpha) = 0 & \end{array} \quad \Rightarrow \quad \begin{bmatrix} s_1^T \\ \vdots \\ s_m^T \end{bmatrix} [v - B\alpha] = B^T [v - B\alpha] = 0$$

Ecuaciones normales

$$B^T [v - B\alpha] = 0 \Leftrightarrow \boxed{B^T v = B^T B \alpha} \Rightarrow \alpha = (B^T B)^{-1} B^T v$$

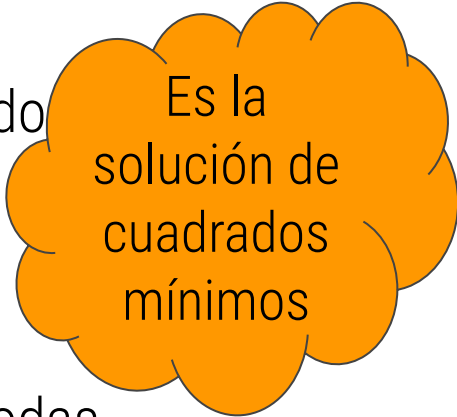
Obs: Si B es
BON $P_{\pi_S} = BB^T$

$$P_{\pi_S} = B(B^T B)^{-1} B^T$$

Una aplicación: Cuadrados mínimos

Supongamos que tenemos un sistema sobredeterminado de la forma

$$Xb = y, \quad X \in \mathbb{R}^{n \times m}, \quad b \in \mathbb{R}^m, \quad y \in \mathbb{R}^n, \quad m > n$$



Es la
solución de
cuadrados
mínimos

Como $m > n$, puede que no exista b que satisfaga todas las m ecuaciones, entonces busco la solución que más se acerque (busco la proyección de y sobre $Col(X)$)

$$b = (X^T X)^{-1} X^T y \quad \longrightarrow \quad P = X(X^T X)^{-1} X^T$$

Aplicación: predicción

Modelo

Problema general:

Tenemos un sistema que sigue el siguiente comportamiento:

$$\begin{aligned}x_{k+1} &= A_k x_k + B_k w_k \\ y_k &= C_k x_k + v_k\end{aligned}$$

donde y_k son las observaciones del sistema, w_k y v_k son el ruido de proceso y de medición respectivamente

El **objetivo** es obtener una estimación $\hat{x}_{k|k-1}$ a partir de las mediciones de los instantes y_1, \dots, y_k

Suposiciones generales

1. $\{v_k\}, \{w_k\}$ son procesos gaussianos de media nula y varianza conocida, e independientes.

$$\mathbb{E}[v_k v_l^*] = R_k \delta_{kl}, \quad \mathbb{E}[w_k w_l^*] = Q_k \delta_{kl}$$

2. Condiciones iniciales del problema:
 - a. Para $k = 0$ partimos de un estado inicial x_0 descorrelacionado de los proceso de ruido.

Estimador óptimo

Dos problemas de interés:

- Hallar $\hat{x}_{k|k}$: la mejor estimación de x_k dada la información hasta el instante k ($Y_k = [y_0, \dots, y_k]$) [estim. a posteriori]
- Hallar $\hat{x}_{k|k-1}$: la mejor estimación de x_k dada la información hasta el instante $k-1$ ($Y_{k-1} = [y_0, \dots, y_{k-1}]$) [estim. a priori]

Esto equivale a resolver el problema de proyectar x_k sobre Y_k y Y_{k-1} respectivamente.

Estimador óptimo

1. $\langle x_k - \hat{x}_{k|k}, y_n \rangle, \quad n = 0, \dots, k$
2. $\langle x_{k+1} - \hat{x}_{k+1|k}, y_n \rangle, \quad n = 0, \dots, k$

La pregunta clave es: Si conozco $\hat{x}_{k|k} \in Y_k$,
¿puedo hallar $\hat{x}_{k+1|k} \in Y_k$?

Conjetura a partir del modelo: $\hat{x}_{k+1|k} = A_k \hat{x}_{k|k}$

Estimador óptimo

Demostramos que

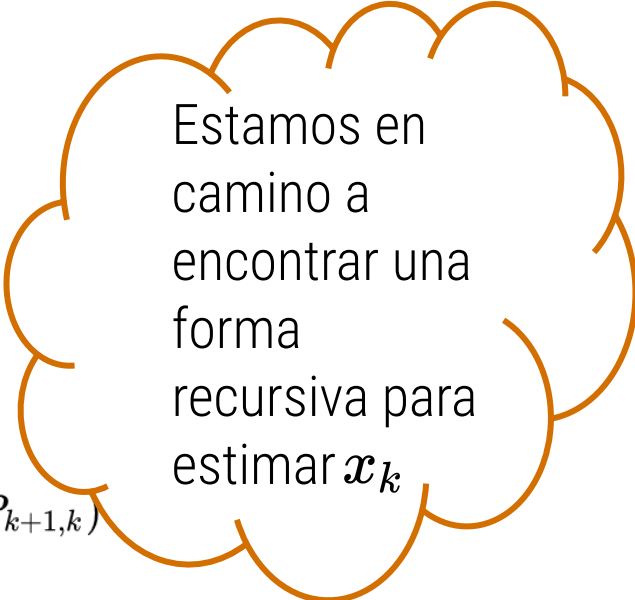
$$\hat{x}_{k+1|k} = A_k \hat{x}_{k|k}$$

y además tenemos una forma de hallar el error:

$$e_{k|k}^2 = \langle x_k - \hat{x}_{k|k}, x_k - \hat{x}_{k|k} \rangle = \text{Tr}(P_{k,k})$$

$$e_{k+1|k}^2 = \langle x_{k+1} - \hat{x}_{k+1|k}, x_{k+1} - \hat{x}_{k+1|k} \rangle = \text{Tr}(P_{k+1,k})$$

$$P_{k+1|k} = A_k P_{k|k} A_k^H + B_k Q_k B_k^H$$



Estamos en camino a encontrar una forma recursiva para estimar x_k

Filtro de Kalman

Sistema de ecuaciones recursivas:

$$\begin{aligned}\hat{x}_{k|k-1} &= A_{k-1} \hat{x}_{k-1|k-1} \\ P_{k|k-1} &= A_{k-1} P_{k-1|k-1} A_{k-1}^T + B_{k-1} Q_{k-1} B_{k-1}^T \\ K_k &= P_{k|k-1} C_k^T (R_k + C_k P_{k|k-1} C_k^T)^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k (y_k - C_k \hat{x}_{k|k-1}) \\ P_{k|k} &= (I - K_k C_k) P_{k|k-1}\end{aligned}$$

Condiciones iniciales: $\hat{x}_0 = \mathbb{E}[x_0]$, $P_{0|0} = \mathbb{E}[x_0 x_0^T]$ conocidas
(si mi estimación inicial es muy mala, poner la cov. grande)

Bibliografía extra

- “[Kalman Filtering](#)”, Charles K. Chui and Guanrong Chen, Springer
- “Optimal Filtering”, Brian D.O. Anderson and John B. Moore, Dover