Universidad Nacional de Ingeniería
Escuela Profesional de Matemática
Ciclo 2021-I

[Introducción a los procesos estocásticos]
[J. Ugarte] UNI, 14 de mayo de 2021.

Práctica calificada 2

Tiempo: 2h
Tolerancia 15min

1. *Aplicación de vectores aleatorios*
Los modelos de regresión son variados y muchos poseen distintas generalizaciones del método de mínimos cuadrados aunque en dimensión superior. Comenzamos a estudiar un modelo lineal.

$$Y = X\theta + \xi$$

donde X es una **matriz determinista** de orden $n \times k$ y θ es una matriz columna de k filas, y el **vector aleatorio** ξ satisface $\mathbb{E}(\xi) = 0$ y $\mathbb{E}(\xi\xi^T) = \sigma^2 I_n$ donde I_n es la matriz identidad de orden n.
Recordemos que el método de mínimos cuadrados se basa en encontrar un vector, matriz columna en nuestro caso, $\widehat{\theta}$ tal que $X\widehat{\theta}$ se encuentre lo más próximo a Y. De esta forma definimos:

$$f(\theta) = ||Y - X\theta||^2 + \lambda||\theta||^2$$

Al cual hemos añadido un parámetro de regularización $\lambda > 0$.

a) Determine θ tal que f sea mínimo.
Sugerencia: Desarrolle $f(\theta + \alpha)$ y compare con $f(\theta)$.

b) Sea $\widehat{\theta}_1 = (X^T X + \lambda I_k)^{-1} X^T Y$. Comparar con el θ anterior. Enseguida encuentre $\mathbb{E}(\widehat{\theta}_1)$.
Sugerencia: Y es un vector aleatorio con parte determinista y aleatoria.

c) Encuentre $V(\widehat{\theta}_1)$.

Para comparar este método con el de mínimos cuadrados tradicional, nos enfocamos en el caso cuando $k = 1$ de esta forma X es un vector en $\mathbb{R}^n$. Denotamos por $\widehat{\theta}_{ER} = \widehat{\theta}_1$ y por $\widehat{\theta}_{EMC}$ el obtenido por el método de mínimos cuadrados tradicional así:

$$\widehat{\theta}_{EMC} = \frac{X^T Y}{X^T X}$$

Escriba en este caso, en nuestras condiciones $k = 1$ el valor de:

$$\widehat{\theta}_{ER} = \widehat{\theta}_1 = ...$$

d) Muestre que $\mathbb{E}(\widehat{\theta}_{EMC}) = \theta$, considere Y claramente del modelo lineal inicial.

e) Muestre que

$$\mathbf{EMS}_{EMC} := \mathbb{E}[(\widehat{\theta}_{EMC} - \theta)^2] = \frac{\sigma^2}{||X||^2}$$

.

f) Muestre que

$$\mathbf{EMS}_{ER} := \mathbb{E}[(\widehat{\theta}_{ER} - \theta)^2] = \left(\frac{||X||^2\theta}{||X||^2 + \lambda} - \theta\right)^2 + \frac{\sigma^2||X||^2}{(||X||^2 + \lambda)^2}$$

.

g) Concluir dando un intervalo de valores para λ tal que $\widehat{\theta}_{ER}$ se aproxima más hacia θ que $\widehat{\theta}_{MC}$.

[8 puntos]

2. *Resultados de clase*
   Sea $X \sim \mathcal{N}_p(\mu, \Sigma)$ con Σ simétrica y positiva. Demuestre:

   a) Las combinaciones lineales $a^T X$ para $a \in \mathbb{R}^p$ tenemos que:

   $$a^T X \sim \mathcal{N}(a^T \mu, a^T \Sigma a)$$

   b) Si $Y = BX + c$ donde B es una matriz de orden $q \times p$ y $c \in \mathbb{R}^q$ entonces:

   $$Y \sim \mathcal{N}_q(B\mu + c, B\Sigma B^T)$$

   c) Si $X \sim \mathcal{N}_p(0, \sigma^2 I)$ y Γ una matriz ortogonal entonces $\Gamma X \sim \mathcal{N}_p(0, \sigma^2 I)$

   [3 puntos]

   Solución:

3. *Densidad vía transformaciones*
   Sea $X = (X_1, X_2)^T \sim \mathcal{N}_2(0, I)$.

   a) Recuerde su función de densidad, de X.

   b) Determine la función de densidad de $X^T X$.

   c) Sea $D = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = x_2\}$. Demuestre que la v.a. T definida por:

   $$T = \begin{cases} \left(\frac{X_1+X_2}{X_1-X_2}\right)^2 & \text{si } X \notin D \\ 0 & \text{en otro caso} \end{cases}$$

   admite una densidad y determine dicha densidad.

   [5 puntos]

4. *Cálculo con vectores aleatorios*
   Sea $(X, Y, Z)^T$ un vector aleatorio normal de media cero y cuya matriz de covarianza es:

   $$\Sigma = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

   a) Denotamos $U = -X + Y + Z$, $V = X - Y + Z$, y $W = X + Y - Z$. Determine la distribución del vector aleatorio: $(U, V, W)^T$.

   b) Determine la densidad de $T = U^2 + V^2 + W^2$.

   [4 puntos]