TAREA 3

INDICACIONES:

- Para resolver el problema 1, crear un archivo en `Jupyter Notebook`, y llámelo hmw3_1. Para escribir las respuestas a estas preguntas usen MarkDown.
- Para resolver el problema 2, crear un archivo en `Jupyter Notebook`, y llámelo hmw3_2. Para escribir las respuestas a estas preguntas usen MarkDown.
- Para resolver el problema 3, crear un archivo en `Jupyter Notebook`, y llámelo hmw3_3. Para escribir las respuestas a estas preguntas usen MarkDown.
- Para resolver el problema 4, crear un archivo en `Jupyter Notebook`, y llámelo hmw3_4. Para escribir las respuestas a estas preguntas usen MarkDown.
- Usted deberá realiar el PUSH a la carpeta github ECOP2037_NN, antes de la fecha indicada en esta tarea 3.

1. Método Delta

Tome el modelo

$$
\begin{aligned}
Y &= X_1\beta_1 + X_2\beta_2 + e_i \\
\mathbb{E}\left[X_i e_i\right] &= 0
\end{aligned}
$$

con $\beta_1 \in \mathbb{R}$ y $\beta_2 \in \mathbb{R}$, y defina el parámetro $\theta = \beta_1\beta_2$.

(a) ¿Cuál es el estimador apropiado $\hat{\theta}$ para θ?

(b) Encuentre la distribución asintótica de $\widehat{\theta}$ bajo condiciones de regularidad estándar.

(c) Muestre cómo calcular un intervalo de confianza asintótico del 95 %.

2. MCO Ponderados

Las variables $\{Y_i, X_i, W_i\}$ son una muestra aleatoria. El parámetro β se estima minimizando la función criterio.

$$
S(\beta) = \sum_{i=1}^{n} W_i \left(Y_i - X_i'\beta\right)^2
$$

Es decir, $\widehat{\beta} = \operatorname{argmin}_\beta S(\beta)$.

(a) Encuentre una expresión explícita para $\widehat{\beta}$.

(b) ¿Qué parámetro poblacional β está estimando $\widehat{\beta}$? (Sea explícito sobre cualquier suposición que deba imponer. Pero no haga más suposiciones de las necesarias).

(c) Encuentre el límite de probabilidad para $\widehat{\beta}$ como $n \to \infty$.

(d) Encuentre la distribución asintótica de $\sqrt{n}(\widehat{\beta} - \beta)$ como $n \to \infty$.

3. Consistencia y eficiencia

El modelo simple sin constante es $Y = \beta X + e$ con $\mathbb{E}[e \mid X] = 0$ y $X \in \mathbb{R}$. Considere los dos estimadores.

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}$$
$$\widetilde{\beta} = \frac{1}{n} \sum_{i=1}^{n} \frac{Y_i}{X_i}$$

(a) Bajo los supuestos planteados, ¿son ambos estimadores consistentes para β?

(b) ¿Existen condiciones bajo las cuales alguno de los estimadores sea eficiente?

Ridge y Lasso Regression: Introducción

En contextos de regresión lineal, el estimador de Mínimos Cuadrados Ordinarios (MCO) es insesgado y consistente bajo supuestos clásicos. Sin embargo, cuando el número de variables explicativas es grande, o cuando existe alta colinealidad entre ellas, el estimador MCO puede presentar alta varianza, problemas de inestabilidad numérica, o incluso no ser identificable si $\mathbf{X}'\mathbf{X}$ no es invertible.

Para enfrentar estos problemas, se han desarrollado métodos de estimación penalizada como **Ridge** y **Lasso**, que modifican el criterio de mínimos cuadrados incorporando un término de regularización que penaliza la magnitud de los coeficientes estimados.

- **Ridge Regression** agrega una penalización cuadrática de la forma $\lambda\|\beta\|^2 = \lambda \sum_{j=1}^{k} \beta_j^2$, lo cual reduce la varianza del estimador, pero introduce un pequeño sesgo. Es especialmente útil en presencia de colinealidad.
- **Lasso Regression** agrega una penalización lineal de la forma $\lambda \sum_{j=1}^{k} |\beta_j|$, lo cual no solo reduce la varianza, sino que también puede forzar algunos coeficientes a ser exactamente cero, permitiendo así la **selección automática de variables**.

Ambos métodos requieren elegir un parámetro de penalización $\lambda > 0$ que controla el grado de regularización. Dependiendo de cómo λ se comporta con el tamaño muestral n, los estimadores pueden o no ser consistentes.

A continuación, se presenta un ejercicio que compara el comportamiento asintótico de los estimadores Ridge y Lasso bajo distintas elecciones de λ_n.

4. Comparación: Ridge vs. Lasso Regression

Considere el modelo lineal clásico:

$$Y_i = X_i'\beta + e_i, \quad \text{con } \mathbb{E}[X_i e_i] = 0$$

Parte A: Ridge Regression

El estimador ridge está definido como:

$$\widehat{\beta}^{\text{ridge}} = \left(\sum_{i=1}^{n} X_i X_i' + \lambda_n I_k \right)^{-1} \left(\sum_{i=1}^{n} X_i Y_i \right)$$

(a) Suponga que $\lambda_n = \lambda > 0$ es constante. Encuentre el límite en probabilidad de $\widehat{\beta}^{\text{ridge}}$ cuando $n \to \infty$. ¿Es consistente para β?

(b) Suponga ahora que $\lambda_n = cn$, con $c > 0$ constante. ¿Cuál es el límite en probabilidad de $\widehat{\beta}^{\text{ridge}}$? ¿Es consistente?

Parte B: Lasso Regression

El estimador lasso está definido como:

$$\widehat{\beta}^{\text{lasso}} = \arg\min_{\beta} \left\{ \sum_{i=1}^{n} (Y_i - X_i'\beta)^2 + \lambda_n \sum_{j=1}^{k} |\beta_j| \right\}$$

(c) Suponga que $\lambda_n = \lambda > 0$ es constante. ¿Cuál es el límite en probabilidad de $\widehat{\beta}^{\text{lasso}}$? ¿Es consistente?

(d) Suponga que $\lambda_n = cn$, con $c > 0$. ¿Cuál es el límite en probabilidad de $\widehat{\beta}^{\text{lasso}}$? ¿Es consistente?