Teorema de Gauss–Markov

En muchas aplicaciones estamos más interesado en estimar funciones lineales de $\boldsymbol{\beta}$ que en estimar $\boldsymbol{\beta}$ en sí mismo.

Estas funciones incluyen el valor esperado de y en una futura observación $\mathbf{x}_o$, por ejemplo.

Si bien puede haber muchos estimadores de una función lineal $\mathbf{c}'\boldsymbol{\beta}$ o $\mathbf{C}\boldsymbol{\beta}$, estudiaremos los estimadores lineales, es decir funciones lineales de las observaciones $y_1, \ldots, y_n$.

Primero veremos cuando una función paramétrica es estimable.

Definición: Una función paramétrica ψ se dice que es una función lineal de los paramétros $\{\beta_1, \ldots, \beta_p\}$ si existen $\{c_1, \ldots, c_p\}$ constantes conocidas tal que

$$\psi = \mathbf{c}'\boldsymbol{\beta} = \sum_{j=1}^{p} c_j\beta_j$$

donde $\mathbf{c} = (c_1, \ldots, c_p)'$.

Definición: Decimos que una función paramétrica $\psi = \mathbf{c}'\boldsymbol{\beta}$ es estimable si tiene un estimador lineal (en $\mathbf{Y}$) insesgado, es decir si existe $\mathbf{a} \in \Re^n$ tal que

$$E(\mathbf{a}'\mathbf{Y}) = \psi = \mathbf{c}'\boldsymbol{\beta} \quad \forall \boldsymbol{\beta} \in \Re^p$$

¿Hay funciones que no son estimables?

Veamos un ejemplo de una función paramétrica no estimable.

Supongamos que queremos comparar la respuesta media de dos tratamientos y un control y que para ello observamos

$$\begin{array}{lll} \text{T1:} & y_{11}, y_{12}, \ldots, y_{1k} & y_{1j} \sim N(\beta_1, \sigma^2) \\ \text{T2:} & y_{21}, y_{22}, \ldots, y_{2k} & y_{2j} \sim N(\beta_2, \sigma^2) \\ \text{Co:} & y_{31}, y_{32}, \ldots, y_{3k} & y_{3j} \sim N(\beta_3, \sigma^2) \end{array}$$

Suponemos igual cantidad de observaciones por tratamiento para simplificar la notación.

Podemos escribir esto como

$$y_{ij} = \beta_i + \epsilon_{ij}$$

Podríamos escribir esto como un modelo lineal:

$$\mathbf{Y} = \begin{pmatrix} y_{11} \\ y_{12} \\ \ldots \\ y_{1k} \\ y_{21} \\ y_{22} \\ \ldots \\ y_{2k} \\ y_{31} \\ y_{32} \\ \ldots \\ y_{3k} \end{pmatrix} ; \mathbf{X} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ . & . & . \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ . & . & . \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ . & . & . \\ 0 & 0 & 1 \end{pmatrix} ; \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

Por ejemplo, T1, T2 y el control podrían ser distintas dosis de una droga de manera que T1 es menor que la dosis del control y T2 mayor que la dosis control. Tendría sentido preguntarse si

$$\beta_3 = \frac{\beta_1 + \beta_2}{2}$$

lo que implicaría cierta linealidad en el efecto medio. En ese caso nos interesaría saber si

$$\left(-\tfrac{1}{2}, -\tfrac{1}{2}, 1 \right) \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = 0$$

Otra manera de escribir el modeo sería

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

donde:

μ es el efecto general

α_i es el efecto del tratamiento i

En ese caso tendríamos

$$\mathbf{Y} = \begin{pmatrix} y_{11} \\ y_{12} \\ \dots \\ y_{1k} \\ y_{21} \\ y_{22} \\ \dots \\ y_{2k} \\ y_{31} \\ y_{32} \\ \dots \\ y_{3k} \end{pmatrix} ; \mathbf{X} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ . & . & . & . \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ . & . & . & . \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ . & . & . & . \\ 1 & 0 & 0 & 1 \end{pmatrix} ; \boldsymbol{\beta} = \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$

¿Son todas las funciones estimables en este modelo?

Consideremos

$$\alpha_1 = \begin{pmatrix} 0, 1, 0, 0 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$

Veremos que α_1 no es estimable.

Veamos el siguiente resultado que caracteriza las funciones paramétricas estimables suponiendo el modelo

$$\Omega : E(\mathbf{Y}) = \mathbf{X}\beta \quad \Sigma_{\mathbf{Y}} = \sigma^2 \mathbf{I}$$

Teorema: La función paramétrica $\psi = \mathbf{c}'\boldsymbol{\beta}$ es estimable si y sólo si $\mathbf{c}$ es una combinación lineal de las filas de $\mathbf{X}$, o sea si existe $\mathbf{a} \in \Re^n$ tal que

$$\mathbf{c}' = \mathbf{a}'\mathbf{X}$$

Lema: Supongamos que vale el modelo Ω. Sean $\psi = \mathbf{c}'\boldsymbol{\beta}$ una función estimable y $\mathcal{V}_r$ el espacio generado por las columnas de $\mathbf{X}$ $(r = rg(\mathbf{X}) \leq p)$. Luego, existe un único estimador lineal insesgado de ψ, digamos $\mathbf{a}^{*'}\mathbf{Y}$ con $\mathbf{a}^* \in \mathcal{V}_r$. Más aún, si $\mathbf{a}'\mathbf{Y}$ es un estimador insesgado de ψ, $\mathbf{a}^*$ es la proyección ortogonal de $\mathbf{a}$ sobre $\mathcal{V}_r$.

Teorema de Gauss–Markov:
Supongamos que vale el modelo $\Omega : E(\mathbf{Y}) = \mathbf{X}\beta \quad \Sigma_{\mathbf{Y}} = \sigma^2\mathbf{I}$.
Toda función estimable $\psi = \mathbf{c}'\boldsymbol{\beta}$ tiene un único estimador $\widehat{\psi}$ lineal insesgado de míinima varianza (BLUE). Este estimador $\widehat{\psi}$ se puede obtener reemplazando a $\boldsymbol{\beta}$ en $\mathbf{c}'\boldsymbol{\beta}$ por $\widehat{\boldsymbol{\beta}}$, el estimador de mínimos cuadrados.

Definición: Dada una función estimable ψ su único estimador lineal insesgado de mínima varianza $\widehat{\psi}$, cuya existencia y cálculo están dados por el Teorema de Gauss–Markov, es el estimador de mínimos cuadrados de ψ.

Tenemos el siguiente resultado:

Corolario: Si $\{\psi_1, \ldots, \psi_q\}$ son q funciones estimables toda combinación lineal $\Psi = \Sigma_{i=1}^q h_i \psi_i$ es estimable y su estimador de mínimos cuadrado está dado por $\Sigma_{i=1}^q h_i \widehat{\psi}_i$.

¿Qué ocurre cuando el rg$(\mathbf{X}) < p$

Si rg$(\mathbf{X}) = r < p$ tenemos que $\widehat{\beta}_1, \ldots, \widehat{\beta}_p$ no son únicos. Esta misma indeterminación afecta a los parámetros $\beta_1, \ldots, \beta_p$, en el sentido de que distintos conjuntos $b_1, \ldots, b_p$ darían origen al mismo η y por lo tanto al mismo modelo

$$\mathbf{Y} = \eta + \boldsymbol{\epsilon} = E(\mathbf{Y}) + \boldsymbol{\epsilon}\,.$$

Sin embargo, tal como vimos si $\mathbf{c}'\boldsymbol{\beta}$ es una función estimable tendrá el mismo valor independientente del $\boldsymbol{\beta}$ que usemos, en tanto

$$\mathbf{c}'\boldsymbol{\beta} = \mathbf{a}'\mathbf{X}\boldsymbol{\beta} = \mathbf{a}'\eta$$

expresión que sólo depende de η, que es único.

¿Cómo podemos eliminar esta indeterminación?

a) Considerar un problema reducido con sólo r parámetros

Podríamos considerar r columnas l.i. de $\mathbf{X}$ que generen a $\mathcal{V}_r$ y mantener en el modelo sólo aquellos β_j asociados a estas columnas.

Así tendríamos una nueva matriz de diseño $\mathbf{X}_1 \in \Re^{n\times r}$ con rango máximo. En este caso tendríamos el modelo

$$\mathbf{Y} = \eta + \boldsymbol{\epsilon} \qquad \text{con } \eta \in \mathcal{V}_r$$

El estimador sería

$$\boldsymbol{\alpha} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{Y}$$

y la matriz de proyección corresponiente $\mathbf{P} = \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'$.

Si asumimos, s.p.g., que las columnas elegidas son las primeras r, tendríamos que

$$\mathbf{X} = [\mathbf{X}_1\mathbf{X}_2]$$

donde $\mathbf{X}_2 \in \Re^{n\times(p-r)}$ y además $\mathbf{X}_2 = \mathbf{X}_1\mathbf{B}$. Por lo tanto

$$\mathbf{X} = \mathbf{X}_1[\mathbf{I}_r \quad \mathbf{B}] = \mathbf{K}\mathbf{L}$$

con $\mathbf{K} \in \Re^{n\times r}, \quad \mathbf{L} \in \Re^{r\times p}$ y $\text{rg}(\mathbf{L}) = r$.

Por lo tanto el modelo original se obtiene como:

$$\mathbf{X}\boldsymbol{\beta} = \mathbf{K}\mathbf{L}\boldsymbol{\beta} = \mathbf{K}\boldsymbol{\alpha}$$

b) Considerar condiciones de contorno adecuadas para los β_j's y sus estimadores

Así podríamos pedir que $\boldsymbol{\beta}_{r+1} = \ldots = \boldsymbol{\beta}_p = 0$ y en este caso obtendríamos el mismo que en la situación a) (suponiendo que las r primeras son las columnas l.i.).

Sin embargo, en otras situaciones, como en la de ANOVA, es frecuente que se impongan otras restricciones lineales de manera de obtener la unicidad.

Consideremos el caso en que imponemos $t \geq p - r$ restricciones lineales a los $\boldsymbol{\beta}_j$, es decir

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{0} \quad \text{con } \mathbf{H} \in \Re^{\mathbf{t}\times\mathbf{p}}$$

Queremos encontrar dentro del conjunto de soluciones de $\mathbf{X}\boldsymbol{\beta} = \boldsymbol{\eta}$ una sola que cumpla $\mathbf{H}\boldsymbol{\beta} = \mathbf{0}$, es decir buscamos $\tilde{\boldsymbol{\beta}}$ que sea única solución de

$$\begin{aligned} \mathbf{X}\tilde{\boldsymbol{\beta}} &= \mathbf{X}\boldsymbol{\beta} \quad (= \boldsymbol{\eta}) \\ \mathbf{H}\tilde{\boldsymbol{\beta}} &= \mathbf{0} \end{aligned}$$

De esta forma las primeras ecuaciones establecen que encontraremos una solución del sistema que nos interesa y las segundas que esta solución será única.

Lo que queremos es que

- toda función estimable del nuevo sistema lo sea en el viejo problema,
- un único conjunto de estimadores de mínimos cuadrados que satisfaga las condiciones de contorno.

El siguiente teorema nos dice como elegir $\mathbf{H}$ para cumplir con este objetivo:

Teorema: Sean $\mathbf{X} \in \Re^{n\times p}$ y $\mathbf{H} \in \Re^{t\times p}$ con $rg(\mathbf{X}) = r$, $p > r$ y $t \geq p - r$. Consideremos $\mathcal{V}_{\mathbf{X}}$ el espacio generado por las columnas de $\mathbf{X}$. El sistema

$$\begin{aligned} \mathbf{Xb} &= \mathbf{z} \\ \mathbf{Hb} &= \mathbf{0} \end{aligned} \qquad (1)$$

tiene solución única $\mathbf{b}$ para todo $\mathbf{z} \in \mathcal{V}_{\mathbf{X}}$ si y sólo si se cumplen las siguientes dos condiciones:

i) si $\text{rg}(\mathbf{G}) = \text{rg}\begin{pmatrix} \mathbf{X} \\ \mathbf{H} \end{pmatrix} = p$

ii) ninguna combinación lineal de las filas de $\mathbf{H}$ es combinación lineal de las de $\mathbf{X}$, excepto el $\mathbf{0}$.

Corolario: Si el sistema (1) cumple la condiciones i) y ii) del Teorema anterior, entonces existe un único conjunto de estimadores de mínimos cuadrados (solución de las ecuaciones normales) $\hat{\beta}_1, \ldots, \hat{\beta}_p$ para el cual $\mathbf{H}\widehat{\boldsymbol{\beta}} = \mathbf{0}$.

Observación:

- En términos estadísticos la condición ii) del Teorema nos dice que si $\mathbf{h}_i$ es la iésima fila de $\mathbf{H}$, entonces no existe $\mathbf{a}$ tal que $\mathbf{h}_i = \mathbf{a}'\mathbf{X}$, por lo tanto las $\mathbf{h}_i'\boldsymbol{\beta}$ no es una función estimable de los parámetros.

Se puede demostrar que:

- Si se cumplen las condiciones i) y ii) del Teorema, entonces los $\tilde{\beta}_j$ son funciones estimables.

  De hecho, si $\mathbf{G}\tilde{\boldsymbol{\beta}} = \begin{pmatrix} \mathbf{X}\boldsymbol{\beta} \\ \mathbf{0} \end{pmatrix}$, entonces

$$\mathbf{G}'\mathbf{G}\tilde{\boldsymbol{\beta}} = \mathbf{G}' \begin{pmatrix} \mathbf{X}\tilde{\boldsymbol{\beta}} \\ \mathbf{0} \end{pmatrix} = (\mathbf{X}' \quad \mathbf{H}') \begin{pmatrix} \mathbf{X}\tilde{\boldsymbol{\beta}} \\ \mathbf{0} \end{pmatrix} = \mathbf{X}'\mathbf{X}\boldsymbol{\beta}\,.$$

  Luego, $(\mathbf{X}'\mathbf{X} + \mathbf{H}'\mathbf{H})\tilde{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{X}\boldsymbol{\beta}$ y como $rg(\mathbf{G}'\mathbf{G}) = rg(\mathbf{G}) = p$ tenemos que

$$\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X} + \mathbf{H}'\mathbf{H})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$$

y tiene un estimador lineal insesgado dado por

$$(\mathbf{X}'\mathbf{X} + \mathbf{H}'\mathbf{H})^{-1}\mathbf{X}'\mathbf{Y}$$

- dada una función estimable ψ, para cualquier $\mathbf{H}$ que elijamos en las condiciones del Teorema anterior, $Var(\widehat{\psi})$ es la misma.

c) Computar una inversa generalizada de $\mathbf{X}'\mathbf{X}$: $(\mathbf{X}'\mathbf{X})^{-}$

En este caso tendríamos que $(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}\mathbf{Y}$ es solución de las ecuaciones normales, por lo tanto otra forma de solucionar nuestro problema. En realidad puede verse que la opción b) y c) quedan ligadas a través del siguiente resultado:

Proposición: Sea $\mathbf{G} = \begin{pmatrix} \mathbf{X} \\ \mathbf{H} \end{pmatrix}$ una matriz que satisface las condiciones i) y ii) del Teorema anterior. Luego $(\mathbf{G}'\mathbf{G})^{-1}$ es una inversa generalizada de $\mathbf{X}'\mathbf{X}$, por lo tanto:

$$(\mathbf{X}'\mathbf{X})(\mathbf{G}'\mathbf{G})^{-1}(\mathbf{X}'\mathbf{X}) = \mathbf{X}'\mathbf{X}$$

En efecto, $\forall \mathbf{Y}$:

$$
\begin{aligned}
(\mathbf{G}'\mathbf{G})(\mathbf{G}'\mathbf{G})^{-1}\mathbf{H}'\mathbf{Y} &= \mathbf{H}'\mathbf{Y} \\
(\mathbf{X}'\mathbf{X}+\mathbf{H}'\mathbf{H})(\mathbf{G}'\mathbf{G})^{-1}\mathbf{H}'\mathbf{Y} &= \mathbf{H}'\mathbf{Y} \\
\mathbf{X}'\mathbf{X}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{H}'\mathbf{Y} &= \mathbf{H}'(\mathbf{I}-\mathbf{H}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{H}')\mathbf{Y}
\end{aligned}
$$

entonces como $\mathbf{X}'\boldsymbol{\alpha} = \mathbf{H}'\boldsymbol{\beta}$ tenemos que

$$\mathbf{X}'\mathbf{X}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{H}'\mathbf{Y} = \mathbf{0}$$

luego

$$\mathbf{X}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{H}'\mathbf{Y} \in \mathcal{V}_r^{\perp}$$

y al mismo tiempo

$$\mathbf{X}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{H}'\mathbf{Y} \in \mathcal{V}_r$$

por lo tanto

$$\mathbf{X}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{H}' = \mathbf{0}$$

Finalmente:

$$(\mathbf{X}'\mathbf{X})(\mathbf{G}'\mathbf{G})^{-1}(\mathbf{X}'\mathbf{X}) = (\mathbf{X}'\mathbf{X} + \mathbf{H}'\mathbf{H})(\mathbf{G}'\mathbf{G})^{-1}(\mathbf{X}'\mathbf{X}) = \mathbf{X}'\mathbf{X}\,,$$

con lo cual es una inversa generalizada.

Mínimos Cuadrados Pesados y Mínimos Cuadrados Generalizados

¿ Qué ocurre cuando $\Sigma_{\mathbf{Y}} = \sigma^2\mathbf{V}$ donde $\mathbf{V} \neq \mathbf{I}$?

Supongamos que $\mathbf{V} \in \Re^{n\times n}$ es una matriz definida positiva de constantes. Podemos entonces escribir: $\mathbf{V} = \mathbf{K}\mathbf{K}'$ con $\mathbf{K}$ una matriz invertible.

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \\ \mathbf{K}^{-1}\mathbf{Y} &= \mathbf{K}^{-1}\mathbf{X}\boldsymbol{\beta} + \mathbf{K}^{-1}\boldsymbol{\epsilon} \end{aligned}$$

donde $E(\mathbf{K}^{-1}\boldsymbol{\epsilon}) = \mathbf{0}$ y $\Sigma_{\mathbf{K}^{-1}\boldsymbol{\epsilon}} = \sigma^2\mathbf{I}$.

Por lo tanto, tenemos un nuevo problema:

$$\tilde{\mathbf{Y}} = \tilde{\mathbf{X}}\boldsymbol{\beta} + \tilde{\boldsymbol{\epsilon}}$$

que satisface las condiciones de Ω.

Hallar el estimador de mínimos cuadrados en el problema transformado equivale a:

$$\begin{aligned}
\min_{\mathbf{b}} \|\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}\mathbf{b}\|^2 &= \min_{\mathbf{b}} (\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}\mathbf{b})'(\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}\mathbf{b}) \\
&= \min_{\mathbf{b}} (\mathbf{Y} - \mathbf{X}\mathbf{b})'\mathbf{K}^{-1'}\mathbf{K}^{-1}(\mathbf{Y} - \mathbf{X}\mathbf{b}) \\
&= \min_{\mathbf{b}} (\mathbf{Y} - \mathbf{X}\mathbf{b})'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\mathbf{b})
\end{aligned}$$

Si $\mathbf{V}$ es una matriz diagonal decimos que tenemos un problema de Mínimos Cuadrados Pesados, mientras que si $\mathbf{V}$ es una matriz definida positiva cualquiera, es de Mínimos Cuadrados Generalizados.

Las ecuaciones normales quedan:

$$\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\mathbf{b} = \tilde{\mathbf{X}}'\tilde{\mathbf{Y}}$$

$$\mathbf{X}'\mathbf{K}^{-1'}\mathbf{K}^{-1}\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{K}^{-1'}\mathbf{K}^{-1}\mathbf{Y}$$
$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$$

Observemos que si $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ tiene inversa, entonces

$$\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$$

y además

- $\tilde{\boldsymbol{\beta}}$ es un estimador insesgado de $\boldsymbol{\beta}$, es decir $E(\tilde{\boldsymbol{\beta}}) = \beta$.
- $\Sigma_{\tilde{\boldsymbol{\beta}}} = \sigma^2(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1} = \sigma^2(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$

Veamos un ejemplo.

Consideremos el caso sencillo de una regresión simple por el origen:

$$\mathbf{Y} = \mathbf{x}\beta + \boldsymbol{\epsilon}$$

donde $\mathbf{Y} = (y_1, \ldots, y_n)'$, $\mathbf{x} = (x_1, \ldots, x_n)'$ y $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)'$ con $E(\boldsymbol{\epsilon}) = \mathbf{0}$ y $\Sigma_{\boldsymbol{\epsilon}} = \sigma^2\mathbf{V} = \sigma^2 diag(w_1, \ldots, w_n)$ con $w_i > 0$.

Probaremos que

$$\boxed{\tilde{\boldsymbol{\beta}} = \frac{\sum_{i=1}^n y_i x_i / w_i}{\sum_{i=1}^n x_i^2 / w_i}}$$

y además

$$\boxed{\Sigma_{\tilde{\boldsymbol{\beta}}} = \sigma^2 (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} = \frac{\sigma^2}{\sum_{i=1}^n x_i^2 / w_i}}$$

Si $\text{rg}(\mathbf{X}) = p$ se puede probar fácilmente que el estimador $\tilde{\boldsymbol{\beta}}$ conserva las propiedades del estimador de mínimos cuadrados: dada una función lineal estimable $\mathbf{c}'\boldsymbol{\beta}$ tenemos que

- $\mathbf{c}'\tilde{\boldsymbol{\beta}}$ es el estimador lineal insesgado de $\mathbf{c}'\boldsymbol{\beta}$ de menor varianza.

Una pregunta muy natural es:

¿ Hay situaciones en las que $\tilde{\beta}$ y $\widehat{\beta}$ coinciden?

Los siguientes resultados nos dan la respuesta

Teorema: Una condición necesaria y suficiente para que $\tilde{\boldsymbol{\beta}}$ y $\widehat{\boldsymbol{\beta}}$ coincidan es que $\mathcal{V}_{\mathbf{V}^{-1}\mathbf{X}} = \mathcal{V}_{\mathbf{X}}$.

Corolario: $\tilde{\boldsymbol{\beta}}$ y $\widehat{\boldsymbol{\beta}}$ coinciden $\Longleftrightarrow \mathcal{V}_{\mathbf{VX}} = \mathcal{V}_{\mathbf{X}}$.

Corolario: Si tenemos un modelo de regresión simple por el origen, $\mathbf{Y} = \mathbf{x}\beta + \boldsymbol{\epsilon}$, entonces

$$\tilde{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}} \quad \forall \mathbf{x} \Longleftrightarrow \mathbf{V} = c\mathbf{I}_n$$