

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
 FACULTAD DE CIENCIAS SOCIALES
 ESPECIALIDAD DE ECONOMÍA

ESTADÍSTICA INFERENCIAL
 EXAMEN PARCIAL

Fecha: 16 de mayo de 2015

Clave del curso: EST241

Horario: 522

Duración: 2 h y 30 min.

Profesor: José Flores Delgado

M = A + B \Delta

Ejercicio 1.

(5 puntos)

Considere el modelo de regresión lineal sin intercepto:

$$Y_j = \beta x_j + \epsilon_j, \text{ para } j = 1, \dots, n,$$

donde x_1, \dots, x_n son constantes conocidas, β es un parámetro por estimar y los errores $\epsilon_1, \dots, \epsilon_n$ son variables aleatorias independientes con $E(\epsilon_j) = 0$ y $V(\epsilon_j) = \sigma^2$, para $j = 1, \dots, n$.

- a) Determine el vector de medias y la matriz de varianzas-covarianzas del vector aleatorio $\mathbf{Y} = (Y_1, \dots, Y_n)^t$. (1 punto)
- b) Determine la distribución del vector aleatorio $(Y_1, \dots, Y_n)^t$. (1 punto)
- c) A continuación se muestra el estimador usual de β y otros dos más: (2 puntos)

$$\hat{\beta}_1 = \sum_{j=1}^n b_j Y_j, \text{ con } b_j = \frac{x_j}{\sum_{i=1}^n x_i^2}, \quad j = 1, \dots, n;$$

$$\hat{\beta}_2 = \sum_{j=1}^n c_j Y_j, \text{ con } c_j = \frac{x_j - \bar{X}}{\sum_{i=1}^n (x_i - \bar{X})^2}, \quad j = 1, \dots, n;$$

$$\hat{\beta}_3 = \sum_{j=1}^n d_j Y_j, \text{ con } d_j = n\bar{X}, \quad j = 1, \dots, n.$$

Determine el vector de medias y la matriz de varianzas-covarianzas del vector aleatorio $(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)^t$, mediante propiedades matriciales del valor esperado y de la varianza.

Note que $(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)^t = B\mathbf{Y}$, donde $\mathbf{Y} = (Y_1, \dots, Y_n)^t$ y $B = \begin{pmatrix} b_1 & \dots & b_n \\ c_1 & \dots & c_n \\ d_1 & \dots & d_n \end{pmatrix}$.

Obtenga expresiones simplificadas.

- d) A partir de los resultados de las partes anteriores, determine la distribución del vector aleatorio $(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)^t$. (1 punto)

(4 puntos)

Ejercicio 2.

Para estimar la varianza σ^2 de una variable X cuya media $E(X) = 0$, a partir de una muestra aleatoria X_1, \dots, X_n , considere los estimadores siguientes:

$$\hat{\sigma}_1^2 = \frac{\sum_{j=1}^n X_j^2}{n}, \quad \hat{\sigma}_2^2 = S^2 \text{ y } \hat{\sigma}_3^2 = n\bar{X}^2.$$

$$\hat{\sigma}_1^2 = \frac{\sum_{j=1}^n X_j^2}{n}, \quad \hat{\sigma}_2^2 = S^2, \quad \hat{\sigma}_3^2 = n\bar{X}^2.$$

(2,5 puntos)

a) Deducir cuáles de estos estimadores son insesgados.

b) Si además, X tiene distribución normal, determine cuál de estos estimadores es más eficiente. Tenga en cuenta que, cuando X tiene distribución normal: (1,5 puntos)

$$\frac{\sum_{j=1}^n X_j^2}{\sigma^2} \sim \chi^2(n); \quad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \quad \text{y} \quad \frac{n\bar{X}^2}{\sigma^2} \sim \chi^2(n); \quad \text{donde } \chi^2(n) = G\left(\frac{n}{2}; \frac{1}{2}\right).$$

(6 puntos)

Ejercicio 3.

Sea X_1, \dots, X_n una muestra aleatoria de X , tal que $E(X^k) = m_k, \forall k \in \mathbb{N}^+$.

a) Determine $Cov(\bar{X}; \frac{1}{n} \sum_{i=1}^n X_i^2)$. (2 puntos)

b) Halle $E(X_1 + \dots + X_n | X_1 = 1)$. (1,5 puntos)

c) Halle $V(\sum_{j=1}^n (X_j - \bar{X})^2)$ (2,5 puntos)

Ejercicio 4.

Sean X e Y tales que $X \sim G(\alpha; \beta)$ y $Y|X = x \sim G(2; x), \forall x > 0$.

a) Halle $E(X^2 Y | X)$ y $E(X^2 Y)$. (2 puntos)

b) Halle el modelo condicional $X|Y = y, \forall y > 0$. (3 puntos)

Recordatorio

$$V(X) = E((X - E(X))^2) = E(X^2) - E^2(X). \quad cov(X, Y) = E(XY) - E(X)E(Y).$$

$$Y_{m \times 1} = A + BX \Rightarrow \mu_Y = A + B\mu_X \text{ y } \Sigma_Y = B\Sigma_X B^t.$$

$$X \sim G(\alpha, \beta); \quad \alpha > 0, \beta > 0 : \quad f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0; \quad E(X) = \frac{\alpha}{\beta}, \quad V(X) = \frac{\alpha}{\beta^2}.$$

$$\text{Si } Y \text{ es discreta: } f_x(x) = \sum_{y=-\infty}^{\infty} f_{x,y}(y). \quad \text{Si } Y \text{ es continua: } f_x(x) = \int_{-\infty}^{\infty} f_{x,y}(y) dy.$$

$$f_{x|Y=y}(y) = \frac{f_{x,y}(y)}{f_y(y)}.$$

$$E(X|Y) = g(Y), \quad \text{donde } g(y) = E(X|Y = y). \quad E(X) = E(E(X|Y)).$$

$$X_1, \dots, X_n \text{ independientes: } V(a_0 + \sum_{i=1}^n a_i X_i) = \sum_{i=1}^n a_i^2 V(X_i).$$

$$V(a_0 + \sum_{i=1}^n a_i X_i) = \sum_{i=1}^n a_i^2 V(X_i) + 2 \sum_{i=1}^{n-1} \sum_{j>i}^n a_i a_j Cov(X_i, X_j).$$

$$Cov\left(a_0 + \sum_{i=1}^n a_i X_i, b_0 + \sum_{j=1}^m b_j Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j Cov(X_i, Y_j).$$

Ex. Parcial

16 mayo 2015

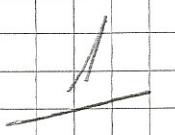
$$(1) \quad Y_j = Bx_j + \varepsilon_j$$

$$E(\varepsilon_j) = 0 \quad ; \quad V(\varepsilon_j) = \sigma^2$$

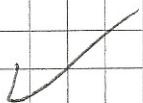
a) Determinar vector medio y matriz varianzas de los errores

vector medio $Y_j = Bx_j + \varepsilon_j$

$$\begin{pmatrix} Y_{j1} \\ \vdots \\ Y_{jn} \end{pmatrix}_{n \times 1} = \begin{pmatrix} Bx_{j1} \\ \vdots \\ Bx_{jn} \end{pmatrix}_{n \times 1} + \begin{pmatrix} \varepsilon_{j1} \\ \vdots \\ \varepsilon_{jn} \end{pmatrix}_{n \times 1}$$



$$\begin{pmatrix} E(Y_{j1}) \\ \vdots \\ E(Y_{jn}) \end{pmatrix}_{n \times 1} = \begin{pmatrix} Bx_{j1} \\ \vdots \\ Bx_{jn} \end{pmatrix}_{n \times 1} + \underbrace{\begin{pmatrix} E(\varepsilon_{j1}) \\ \vdots \\ E(\varepsilon_{jn}) \end{pmatrix}_{n \times 1}}_{= 0}$$



$$\begin{pmatrix} M_{y1} \\ \vdots \\ M_{yn} \end{pmatrix}_{n \times 1} = \begin{pmatrix} Bx_{j1} \\ \vdots \\ Bx_{jn} \end{pmatrix}_{n \times 1} \quad \boxed{M_y = (Bx_j)_{n \times 1}}$$



matriz varianzas

covarianzas

$$Y_j = Bx_j + \varepsilon_j$$

$$\Sigma_y = I_m \sum_{\varepsilon_j} I_m^T$$

I_m = matriz identidad

$$\sum_{\varepsilon_j} = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix}_{n \times n}$$

$$\Sigma_y = \sum_{\varepsilon_j} = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix}_{n \times n}$$

My

sin

b) Determinar distribución $(y_1, \dots, y_n)^T \sim \text{Cov}$

No

$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \sim N_n(\bar{y}; \Sigma^2) \rightarrow$ modelo normal con media \bar{y}

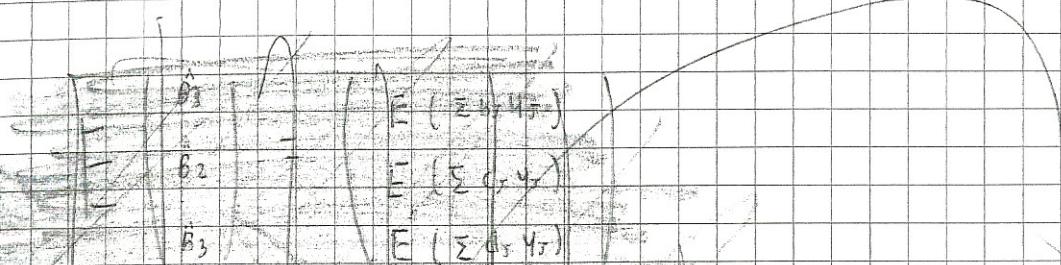
$N \sim \Sigma$

y variancia Σ

0,5

$\text{Sums } \sum y$

c) Determinar variancia media y variancia de $(\hat{\beta}_1; \hat{\beta}_2; \hat{\beta}_3)^T$



$$\textcircled{1} \quad E\left(\sum_{j=1}^m b_j y_j\right) = E\left(\sum_{j=1}^m \frac{x_j}{\sum(x_i - \bar{x})^2} (Bx_j + \varepsilon_j)\right)$$

$$= E\left(B \frac{\sum (x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right) = E(B) = B$$

$$= E\left(B \frac{\sum (x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right) = E(B) = B$$

$$\textcircled{2} \quad E\left(\sum_{j=1}^n d_j M_j\right) = E\left(\sum_{j=1}^n m \bar{x} (Bx_j + \varepsilon_j)\right)$$

$$= E\left(m \bar{x} B \left(\frac{\sum x_j}{n}\right)\right) = E(m^2 \bar{x}^2 B) = m^2 \bar{x}^2 B$$

Up

$\frac{1}{n} \sum x_j^2$

No vale

ACA está RPTA a matriz covarianza.

Vector medias

en otro ordenamiento

$$\underline{\beta} = \begin{pmatrix} \alpha \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

Sí vale

C) matriz varianza

covarianza

$$\underline{\beta}^T = B Y \rightarrow \Sigma_{\beta} = B \Sigma_Y B^T$$

$$\Sigma_{\beta} = \begin{pmatrix} b_1 & \dots & b_n \\ c_1 & \dots & c_n \\ d_1 & \dots & d_n \end{pmatrix}_{3 \times n} \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{pmatrix}_{n \times n} \begin{pmatrix} b_1 & c_1 & d_1 \\ \vdots & \vdots & \vdots \\ b_n & c_n & d_n \end{pmatrix}_{n \times 3}$$

$$\Sigma_{\beta} = \begin{pmatrix} b_1 \sigma^2 & b_2 \sigma^2 & \dots & b_n \sigma^2 \\ c_1 \sigma^2 & \dots & c_n \sigma^2 \\ d_1 \sigma^2 & \dots & d_n \sigma^2 \end{pmatrix}_{3 \times n} \begin{pmatrix} b_1 & c_1 & d_1 \\ \vdots & \vdots & \vdots \\ b_n & c_n & d_n \end{pmatrix}_{n \times 3}$$

$$\Sigma_{\beta} = \begin{pmatrix} b_1 \sigma^2 \sum_{i=1}^n b_i & b_2 \sigma^2 \sum_{i=1}^n c_i & b_3 \sigma^2 \sum_{i=1}^n d_i \\ c_1 \sigma^2 \sum_{i=1}^n b_i & c_2 \sigma^2 \sum_{i=1}^n c_i & c_3 \sigma^2 \sum_{i=1}^n d_i \\ d_1 \sigma^2 \sum_{i=1}^n b_i & d_2 \sigma^2 \sum_{i=1}^n c_i & d_3 \sigma^2 \sum_{i=1}^n d_i \end{pmatrix}_{3 \times 3}$$

$$④ X \sim G(x|B) \quad ; \quad \forall x \in \mathbb{R} \quad Y|x=x \sim G(z|x)$$

$$a) \textcircled{1} E(X^2 y/x) \quad \textcircled{2} E(X^2 y/x=x) = E(X^2 y/x=x)$$

$$\begin{aligned} &= x^2 E(y/x=x) \quad \cancel{\int_0^{\infty} y^2 e^{-y/x} dy / x^2} = 2x \\ &= x^2 \cdot \frac{2}{x} = 2x \quad \Rightarrow \text{en variable de integración} \end{aligned}$$

En variable aleatoria:

$$E(X^2 y/x) = 2X$$

$$\textcircled{3} E(X^2 y) = \textcircled{4} E(E(X^2 y/x)) = E(2X) = 2E(X)$$

$$= 2 \cdot \frac{\infty}{B}$$

$$E(X^2 y) = \frac{2X}{B}$$

2

b) Hullen models conditional $X|Y=y$ $\forall y > 0$

$$f(x) = \frac{B^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-Bx}$$

$$f(y) = \frac{x^2}{\Gamma(2)} \cdot y \cdot e^{-xy}$$

$$f_{X,Y}(x,y) = \frac{B^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-Bx} \cdot \frac{x^2}{\Gamma(2)} y e^{-xy}$$

proportional

$$f_{X|Y=y}(x) \propto x^{\alpha-1} e^{-Bx} x^2 e^{-xy}$$

$$f_{X|Y=y}(x) \propto x^{\alpha+1} e^{-(B+y)x} \Rightarrow X|y \sim G(\alpha+2, B+y)$$

$$f_{X|Y=y}(x) = \frac{(B+y)^{\alpha+2}}{\Gamma(\alpha+2)} x^{\alpha+1} e^{-(B+y)x}$$

(3) x_1, \dots, x_n miteinander unkorreliert $\mathbb{E}(x_k) = m_k$; $\forall k \in \mathbb{N}$

$$\text{Var}\left(\bar{x}; \frac{1}{n} \sum_{i=1}^n x_i^2\right)$$

$$\text{Var}\left(\frac{1}{n} \sum_{j=1}^n x_j, \frac{1}{n} \sum_{i=1}^n x_i^2\right) = \frac{1}{n^2} \text{Var}\left(x_1 + x_2 + \dots + x_n, x_1^2 + \dots + x_n^2\right)$$

$$\begin{aligned} &= \frac{1}{n^2} \sum_{i=1}^n \text{Var}(x_i; x_i^2) \quad \text{real justified} \\ &= \frac{1}{n^2} \sum_{i=1}^n [E(x_i x_i^2) - E(x_i) E(x_i^2)] \end{aligned}$$

$$\begin{aligned} &= \frac{1}{n^2} \sum_{i=1}^n [m_3 - m_1 m_2] \\ &= \frac{1}{m^2} [m_3 - m_1 m_2] \end{aligned}$$

$$\text{Var}\left(\bar{x}; \frac{1}{n} \sum_{i=1}^n x_i^2\right) = \frac{[m_3 - m_1 m_2]}{n}$$

Faltó

prueba
cálculo

b) $E(x_1 + \dots + x_n | x_1 = 1) = E(1 + \dots + x_n | x_1 = 1)$
 $= 1 + E(x_2 + \dots + x_n)$

$E(x_2 + \dots + x_n | x_1 = 1) = 1 + (n-1) m_2$

c) $V\left(\sum_{j=1}^m (x_j - \bar{x})^2\right) = V\left(\sum_{j=1}^m (x_j^2 - 2x_j \bar{x} + \bar{x}^2)\right)$

$= V\left(\sum_{j=1}^m x_j^2 - 2x_j \sum_{j=1}^m \bar{x} + m \bar{x}^2\right)$

No cumple

$V(x+4+2) \\ V(x) + V(4) + V(2)$

$= V\left(\sum_{j=1}^m x_j^2\right) + 4 V(x_1 + x_2 + \dots + x_m) + m^2 V(x_1^2)$

$= \sum_{j=1}^m V(x_j^2) + 4 \left[V(x_1) + \sum_{j=2}^m V(x_j x_1) \right] + m^2 V(x_1^2)$

(*) $V(x_i^2) = E(x_i^2) - E^2(x_i) = [m_2 - m_3^2]$

$= m[m_2 - m_3^2] + 4 \left[m_2 - m_3^2, \sum_{j=2}^m V(x_j x_1) \right] + m^2 [m_2 - m_3^2]$

(**) $V(x_i x_j) = E(x_i^2 x_j^2) - E^2(x_i x_j) = E(x_i^2) E(x_j^2) - [E(x_i) E(x_j)]^2$

$\boxed{i \neq j}$
 $= m_2 m_3 - [m_2 m_3]^2 = m_2^2 - m_3^4$

$= m[m_2 - m_3^2] + 4 \left[m_2 - m_3^2 + (m-2)(m_2^2 - m_3^4) \right] + m^2 [m_2 - m_3^2]$

$V\left(\sum_{j=1}^m (x_j - \bar{x})^2\right) = [m^2 + m + 4][m_2 - m_3^2] + 4(m-1)(m_2^2 - m_3^4)$

② Estimar variancia σ^2 de X ; cuando medimos \bar{x} $E(\bar{x}) = \mu$

$$\hat{\sigma}_1^2 = \frac{1}{m} \sum_{j=1}^n x_j^2 \quad ; \quad \hat{\sigma}_2^2 = S^2 = \frac{\sum_{j=1}^n (x_j - \bar{x})^2}{m-1}$$

$$\hat{\sigma}_3^2 = m \bar{x}^2$$

a) ¿ Cuales son los sesgos?

$$\begin{aligned} \text{b)} E(\hat{\sigma}_3^2) &= E\left(\frac{\sum_{j=1}^n x_j^2}{m}\right) = \frac{1}{m} \sum_{j=1}^n E(x_j^2) \\ &= \frac{1}{m} \sum_{j=1}^n (V(x_j) + E^2(x_j)) \\ &= \frac{1}{m} \sum_{j=1}^m (\sigma^2 + 0) \\ &= \frac{1}{m} [m \sigma^2] \end{aligned}$$

$$E(\hat{\sigma}_3^2) = \sigma^2$$

(\Rightarrow Es im sesgado)

f_1

$$\begin{aligned}
 0) \quad E(\hat{\sigma}_s^2) &= E(S^2) = E\left(\frac{\sum_{j=1}^n (x_j - \bar{x})^2}{n-1}\right) \\
 &= \left(\frac{1}{n-1}\right) E\left(\sum_{j=1}^n x_j^2 - n\bar{x}^2\right) \\
 &= \left(\frac{1}{n-1}\right) \left[E\left(\sum_{j=1}^n x_j^2\right) - nE(\bar{x}^2)\right] \\
 &= \left(\frac{1}{n-1}\right) \left[\sum_{j=1}^n E(x_j^2) - n(V(\bar{x}) + E^2(\bar{x})) \right] \\
 2) \quad &= \left(\frac{1}{n-1}\right) \left[\sum_{j=1}^n (\sigma^2 + 0) - n\left(\frac{V(x)}{n} + E^2(x)\right) \right] \\
 &= \left(\frac{1}{n-1}\right) \left[n\sigma^2 - \cancel{n\sigma^2 + 0} \right]
 \end{aligned}$$

$$\boxed{E(\hat{\sigma}_s^2) = \frac{(n-1)}{n} \sigma^2 = \sigma^2} \rightarrow \text{also interessant!}$$

$$\begin{aligned}
 0) \quad E(\hat{\sigma}_3^2) &= E(n\bar{x}^2) = nE(\bar{x}^2) = n[V(\bar{x}) + E^2(\bar{x})] \\
 &= n\left[\frac{V(x)}{n} + E^2(x)\right]
 \end{aligned}$$

$$\begin{aligned}
 &= V(x) + n(0) \\
 \boxed{E(\hat{\sigma}_3^2) = \sigma^2} \quad &\rightarrow \text{also interessant}
 \end{aligned}$$

b) X tiene distribución normal. Ver cuál es más eficiente.

Para ver cuál es más eficiente basta ver cuál tiene menor varianza.
Como son estimadores, su error estándar medio es su varianza.

①

$$\text{Var}(\bar{x}) = \frac{1}{n} \sum_{i=1}^n \text{Var}(x_i) = \frac{\sigma^2}{n}$$

②

$$\text{Var}\left(\frac{\sum x_i^2}{n}\right)$$

③

$$\text{Var}\left(\frac{\hat{\sigma}_3^2}{\sigma^2}\right) = \text{Var}\left(\frac{\frac{1}{n} \sum x_i^2 - \bar{x}^2}{\sigma^2}\right) = \text{Var}\left(\frac{\sum x_i^2}{n\sigma^2}\right) \cdot \frac{(\sigma^2)^2}{n^2}$$

$$\text{Var}\left(\frac{\hat{\sigma}_1^2}{\sigma^2}\right) = \boxed{\frac{2(\sigma^2)^2}{n}}$$

④

$$\text{Var}\left(\frac{\hat{\sigma}_2^2}{\sigma^2}\right) = \text{Var}\left(\frac{s^2 \cdot (n+1)}{(n-1)} \cdot \frac{\sigma^2}{\sigma^2}\right) = \frac{(n-1)^2}{(n-1)^2} \cdot 2(n-1)$$

$$\text{Var}\left(\frac{\hat{\sigma}_2^2}{\sigma^2}\right) = \boxed{\frac{2(\sigma^2)^2}{(n-1)}}$$

⑤

$$\text{Var}\left(\frac{\hat{\sigma}_3^2}{\sigma^2}\right) = \text{Var}\left(\frac{n^{-2}}{n} \frac{\sigma^2}{\sigma^2}\right) = \boxed{(\sigma^2)^2 \cdot 2n}$$

•

Ver:

$$\frac{1}{n} < \frac{1}{n-1} < 1n$$

Cada sección es de longitud igual.

→ El más eficiente es

$$\boxed{\hat{\sigma}_1^2}$$

Si vale la respuesta ítem (a)
 y parte ítem (c) de
 pregunta

Proyecto 1

distribución del vector $(\hat{B}_1; \hat{B}_2; \hat{B}_3)^T \sim N_3(\mu_B; \Sigma_B)$

es normal ya que si después de una combinación lineal de variables aleatorias ~~son~~ normales con ~~desviación~~ medias 0 y varianza σ^2

c) vector medias $(\bar{B}_1; \bar{B}_2; \bar{B}_3)^T = \bar{B}$

$$\bar{B} = \frac{1}{n} \sum_{i=1}^n B_i$$

$$\mu_B = B \mu_Y$$

$$\mu_B = B (\underbrace{B' Y}_{\text{Sistema de ecuaciones}} | \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}}_{\text{columnas}} | \underbrace{\begin{pmatrix} B X_1 \\ \vdots \\ B X_n \end{pmatrix}}_{\text{columnas}})$$

$$\text{Sistema de ecuaciones } \underbrace{\begin{pmatrix} b_1 & \cdots & b_m \\ \vdots & \ddots & \vdots \\ b_n & \cdots & b_n \end{pmatrix}}_{m \times n} \quad \underbrace{\begin{pmatrix} B X_1 & \cdots & B X_n \end{pmatrix}}_{n \times 1}.$$