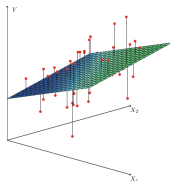


Modelos Lineales de Rango Completo

Teoría y Práctica

Christian Amao Suxo

Escuela Profesional de Ingeniería Estadística
Universidad Nacional de Ingeniería



Semestre I - 2020

Temario

1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- Propiedades estadísticas del método MCO
- Medidas de Bondad de Ajuste del modelo
- Modelo Lineal Centrado

2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- Propiedades estadísticas del método MCO
- Medidas de Bondad de Ajuste del modelo
- Modelo Lineal Centrado

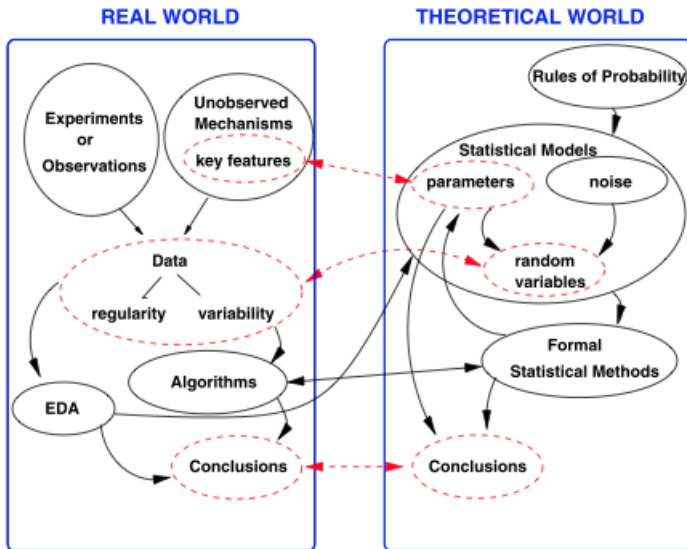
2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

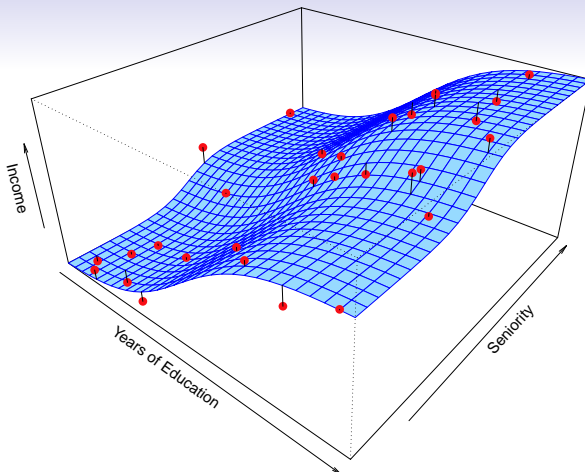
3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

¿Qué explican los modelos lineales?



¿Qué son los modelos lineales?



$$\text{income} \approx f(\text{education}, \text{seniority})$$

1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- Propiedades estadísticas del método MCO
- Medidas de Bondad de Ajuste del modelo
- Modelo Lineal Centrado

2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

¿En qué consiste el modelo lineal general?

Planteamiento del problema

Supongamos que se tiene interés en explicar la variabilidad de un cierto constructo Y (**variable dependiente**) con respecto a la influencia de ciertas variables X_1, X_2, \dots, X_p (**variables regresoras**). Esta relación se puede representar mediante el modelo:

$$Y = f(X_1, X_2, \dots, X_p) + e \quad (1)$$

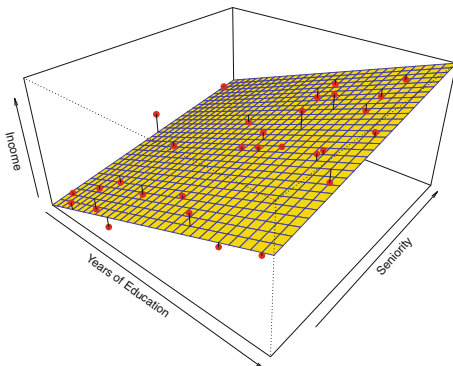
donde $f: \mathbb{R}^p \rightarrow \mathbb{R}$ es una función y e es una variable aleatoria^a con $\mathbb{E}(e) = 0$ y $Var(e) = \sigma^2$.

^aEsta variable aleatoria es conocida como **término de perturbación aleatorio**

¿En qué consiste el modelo lineal general?

Interrogante: ¿Cuál sería la función f más sencilla para el modelo?

Respuesta: La función lineal $f(X_1, \dots, X_p) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$.



$$\text{income} = \beta_0 + \beta_1 \times \text{education} + \beta_2 \times \text{seniority} + e$$

¿En qué consiste modelo lineal general?

Modelo Lineal General

Dada una variable dependiente Y , p variables regresoras X_1, X_2, \dots, X_p y una muestra de tamaño n , el modelo lineal general es aquella relación

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + e_i \quad (i = 1, 2, \dots, n) \quad (2)$$

donde:

- y_i : Valor observado de la variable dependiente Y en el i -ésimo individuo.
- x_{ij} : Valor observado de la variable X_j en el i -ésimo individuo.
- e_i : Error aleatorio producto de las mediciones hechas al i -ésimo individuo.

¿Cómo plantear el modelo lineal general?

Representación matricial

Extendiendo los resultados muestrales de y_i para todo $i = 1, 2, \dots, n$; es posible obtener una representación más compacta de las n ecuaciones:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

$$\underset{n \times 1}{\mathbf{Y}} = \underset{n \times (p+1)}{\mathbf{X}} \underset{(p+1) \times 1}{\boldsymbol{\beta}} + \underset{n \times 1}{\boldsymbol{\varepsilon}} \quad (3)$$

La ecuación (3) es conocida como la **representación matricial** del modelo lineal general.

¿Cuáles son los supuestos de un modelo lineal general?

Supuestos del modelo lineal general

Para el estudio estadístico del modelo lineal general, se supone que:

1. $\mathbb{E}(e_i) = 0, \forall i = 1, 2, \dots, n$
2. $Var(e_i) = \sigma^2, \forall i = 1, 2, \dots, n$ y $Cov(e_i, e_j) = 0$ para $i \neq j$.
3. Las variables regresoras son **determinísticas** y $n \gg p$.

Los supuestos anteriores se pueden expresar también haciendo uso de la notación matricial de la ecuación (3):

1. $\mathbb{E}(\boldsymbol{\varepsilon}) = 0$ y $Cov(\boldsymbol{\varepsilon}) = \sigma^2 \mathbb{I}_n$
2. \mathbf{X} es **no estocástica**.

Observación: Se supondrá que la matriz \mathbf{X} es de **rango completo**. Es por ello que a estos modelos se denominan de rango completo.

1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- Propiedades estadísticas del método MCO
- Medidas de Bondad de Ajuste del modelo
- Modelo Lineal Centrado

2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

¿En qué consiste el método de Mínimos Cuadrados Ordinarios (MCO)?

Estimación de β mediante el método MCO

Dado el modelo lineal general

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon, \quad (4)$$

el método Mínimos Cuadrados Ordinarios (MCO) consiste en buscar aquel β que minimice $\sum_{i=1}^n e_i^2$, i.e.

$$\hat{\beta}_{MCO} = \underset{\beta \in \mathbb{R}^{p+1}}{\operatorname{argmin}} \varepsilon' \varepsilon \quad (5)$$

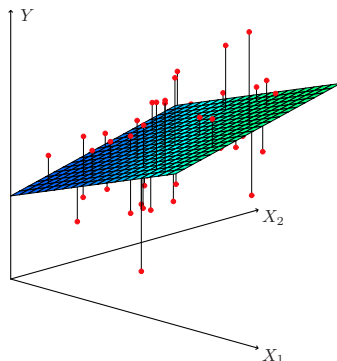
Al resolver la ecuación (5), se obtiene que $\hat{\beta}_{MCO}$ forma parte de un **sistema de ecuaciones normales**:

$$\mathbf{X}'\mathbf{X}\hat{\beta}_{MCO} = \mathbf{X}'\mathbf{Y}. \quad (6)$$

¿En qué consiste el método de Mínimos Cuadrados Ordinarios (MCO)?

Desde que \mathbf{X} es una matriz de rango completo se tiene que:

$$\hat{\beta}_{MCO} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \quad (7)$$



1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- **Propiedades estadísticas del método MCO**
- Medidas de Bondad de Ajuste del modelo
- Modelo Lineal Centrado

2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

¿Cuáles son las propiedades del método MCO?

Propiedades de los estimadores MCO

- 1 $\hat{\beta}_{MCO}$ es un estimador insesgado de β .
- 2 $Var(\hat{\beta}_{MCO}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$.
- 3 Las variables explicativas son ortogonales al vector de residuos MCO, i.e. $\mathbf{X}'\hat{\varepsilon} = \mathbf{0}$.
- 4 El predictor es ortogonal al vector de residuos MCO, i.e. $\hat{\mathbf{Y}}'\hat{\varepsilon} = \mathbf{0}$.
- 5 $\hat{\beta}_{MCO}$ y $\hat{\varepsilon}$ están incorrelacionados.
- 6 El vector de residuos MCO es una transformación lineal del término de perturbación aleatoria.
- 7 $E(\hat{\varepsilon}) = \mathbf{0}$ y $Var(\hat{\varepsilon}) = \sigma^2 M_{\mathbf{X}}$, donde $M_{\mathbf{X}} = \mathbb{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$
- 8 Si el modelo posee término independiente, la suma de residuos es cero.

¿Cuáles son las propiedades del método MCO?

Propiedades importantes

① $s^2 = \frac{\hat{\epsilon}'\hat{\epsilon}}{n-p-1}$ es un estimador insesgado para σ^2

② Dadas las siguientes sumas:

- Suma de Cuadrados Total (SCT) $:= \sum_{i=1}^n (y_i - \bar{y})^2$
- Suma de Cuadrados Explicada (SCE) $:= \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$
- Suma de Cuadrados Residual (SCR) $:= \sum_{i=1}^n (y_i - \hat{y}_i)^2$

Si el modelo lineal general posee término independiente, entonces

$$SCT = SCE + SCR \quad (8)$$

Teorema de Gauss - Markov

El estimador $\hat{\beta}_{MCO}$ es el mejor estimador lineal insesgado para β , i.e. cualquier otro estimador tiene matriz de covarianzas mayor que la del estimador MCO.

1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- Propiedades estadísticas del método MCO
- **Medidas de Bondad de Ajuste del modelo**
- Modelo Lineal Centrado

2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

¿Qué medida de bondad de ajuste podemos usar para ajustar el modelo?

Coeficiente de determinación

Dado un modelo lineal general de la forma $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ con término independiente, el coeficiente de determinación es una medida de bondad de ajuste para un modelo lineal estimado. Este se define como

$$R^2 = 1 - \frac{SCR}{SCT} = \frac{SCE}{SCT} \quad (9)$$

Si el modelo no posee término independiente ($\beta_0 = 0$) entonces se utiliza una modificación del R^2 de la forma

$$R_0^2 = 1 - \frac{SCR}{\sum_{i=1}^n y_i^2} \quad (10)$$

Propiedad: Al agregar una variable explicativa al modelo, el R^2 siempre mejora.

¿Cómo comparar modelos?

Interrogante: Imagina que se tienen dos modelos candidatos con el mismo R^2 , pero el primero tiene más variables regresoras que el otro.

¿Qué modelo escogerías? ¿Por qué?

Coefficiente de determinación ajustado

El coeficiente de determinación ajustado es un indicador de bondad de ajuste que penaliza el uso de más variables regresoras en el modelo.

$$R_{adj}^2 = 1 - \frac{SCR/(n-p-1)}{SCT/(n-1)} = 1 - (1 - R^2) \frac{n-1}{n-p-1} \quad (12)$$

Observaciones:

- El R_{adj}^2 puede ser negativo y además $R_{adj}^2 \leq R^2$.
- El R_{adj}^2 es una medida para **comparar modelos**.

1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- Propiedades estadísticas del método MCO
- Medidas de Bondad de Ajuste del modelo
- **Modelo Lineal Centrado**

2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

¿En qué consiste el modelo lineal centrado?

Algunas veces es usual estimar los parámetros del modelo centrando las observaciones con respecto a sus medias. Para ello supóngase el siguiente modelo lineal

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i \quad (i = 1, 2, \dots, n)$$

Centrando las observaciones con respecto a sus medias respectivas

$$y_i = \alpha + \beta_1(x_{i1} - \bar{x}_1) + \beta_2(x_{i2} - \bar{x}_2) + \cdots + \beta_p(x_{ip} - \bar{x}_p) + e_i \quad (14)$$

donde $\alpha = \beta_0 + \beta_1 \bar{x}_1 + \cdots + \beta_p \bar{x}_p$.

Observación: El modelo lineal centrado es solo un cambio de forma del modelo lineal original. En esencia el modelo sigue siendo el mismo porque los **parámetros son los mismos**.

¿En qué consiste el modelo lineal centrado?

El modelo lineal centrado se puede presentar de forma matricial

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \dots & x_{1p} - \bar{x}_p \\ 1 & x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \dots & x_{2p} - \bar{x}_p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 & \dots & x_{np} - \bar{x}_p \end{pmatrix} \begin{pmatrix} \alpha \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

$$\mathbf{Y} = \mathbf{X}^* \boldsymbol{\beta}^* + \boldsymbol{\varepsilon} \quad (17)$$

donde $\mathbf{X}^* = (\mathbf{1}_n \mid \mathbf{X}_c)$ y $\boldsymbol{\beta}^* = (\alpha \mid \boldsymbol{\beta}_1')'$. Si calculamos el estimador MCO para $\boldsymbol{\beta}^*$ se obtiene que:

$$\hat{\alpha} = \bar{y} \text{ y } \hat{\boldsymbol{\beta}}_1 = (\mathbf{X}_c' \mathbf{X}_c)^{-1} \mathbf{X}_c' \mathbf{Y} \quad (18)$$

Pregunta: ¿Los estimadores MCO usando el modelo centrado son los mismos estimadores MCO del modelo inicial? ¡Demuéstrelo!

¿Qué tanto has aprendido?

¡Hora de ejercitarse!

- 1 Pruebe que dado un modelo lineal general con intercepto de la forma $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, entonces la suma de cuadrados explicada se puede representar como

$$SCE = \hat{\boldsymbol{\beta}}_1 \mathbf{X}'_c \mathbf{X}_c \hat{\boldsymbol{\beta}}_1 = \hat{\boldsymbol{\beta}}_1 \mathbf{X}'_c \mathbf{Y}$$

- 2 Pruebe que la suma de cuadrados residual de un modelo lineal general no cambia si usamos la versión centrada del mismo modelo.
- 3 Sea el modelo lineal general $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ con p variables explicativas y n observaciones. Suponiendo que $\boldsymbol{\varepsilon}$ posee distribución normal multivariada, halle la distribución de

$$W = \frac{\hat{\boldsymbol{\beta}}_1 \mathbf{X}'_c \mathbf{Y} / p}{(\mathbf{Y}'(\mathbb{I}_n - \frac{1}{n}\mathbb{J}_n)\mathbf{Y} - \hat{\boldsymbol{\beta}}_1 \mathbf{X}'_c \mathbf{Y}) / (n - p - 1)}$$

- 4 ¿Qué ocurre con el R^2_{adj} al agregar una variable explicativa al modelo?

1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- Propiedades estadísticas del método MCO
- Medidas de Bondad de Ajuste del modelo
- Modelo Lineal Centrado

2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

¿Cómo plantear un modelo lineal clásico?

Modelo Lineal Clásico

El modelo lineal clásico trabaja con los mismos supuestos del modelo lineal general pero ahora se agrega el supuesto de que los términos de perturbación aleatoria se **distribuyen normalmente**, i.e. $e_i \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$. Entonces podemos plantear el modelo lineal clásico de la forma matricial

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon \quad (19)$$

bajo los supuestos de que:

- $\varepsilon \sim \mathcal{N}_n(\mathbf{0}, \sigma^2 \mathbb{I}_n)$.
- \mathbf{X} es no estocástica y $\text{ran}(\mathbf{X}) = p + 1$.

1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- Propiedades estadísticas del método MCO
- Medidas de Bondad de Ajuste del modelo
- Modelo Lineal Centrado

2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

¿En qué consiste el Método de Máxima Verosimilitud?

Estimación Máximo Verosímil

Para estimar los parámetros del modelo mediante este método, definimos primero la **función de verosimilitud**

$$L(\beta, \sigma^2) = p(\mathbf{Y}|\beta, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{Y} - \mathbf{X}\beta)' (\mathbf{Y} - \mathbf{X}\beta) \right\}$$

Luego, el estimador máximo verosímil para $\theta = (\beta, \sigma^2)$ es

$$(\hat{\beta}_{MV}, \hat{\sigma}_{MV}^2) = \underset{\beta, \sigma^2}{\operatorname{argmax}} L(\beta, \sigma^2) = \underset{\beta, \sigma^2}{\operatorname{argmax}} \ln\{L(\beta, \sigma^2)\} \quad (20)$$

Resolviendo la ecuación (16) se obtiene que

$$\hat{\beta}_{MV} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \quad y \quad \hat{\sigma}_{MV}^2 = \frac{(\mathbf{Y} - \mathbf{X}\hat{\beta}_{MV})'(\mathbf{Y} - \mathbf{X}\hat{\beta}_{MV})}{n} \quad (21)$$

1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- Propiedades estadísticas del método MCO
- Medidas de Bondad de Ajuste del modelo
- Modelo Lineal Centrado

2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

¿Cuáles son las propiedades estadísticas en un MLC?

Propiedades de los estimadores en un Modelo Lineal Clásico

- 1 Los estimadores para β obtenidos mediante el método MCO y MV son los mismos.
- 2 Los estimadores para σ^2 obtenidos mediante el método MCO y MV son distintos. De hecho $\hat{\sigma}_{MV}^2 = \frac{n-p-1}{n} \hat{\sigma}_{MCO}^2$.
- 3 $\hat{\beta}_{MV} = \hat{\beta}_{MCO} \sim \mathcal{N}_{p+1}(\beta, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$.
- 4 Los estimadores $\hat{\beta}_{MCO}$ y $\hat{\sigma}_{MCO}^2$ son independientes. También lo son $\hat{\beta}_{MV}$ y $\hat{\sigma}_{MV}^2$.
- 5 $\frac{n-p-1}{\sigma^2} \hat{\sigma}_{MCO}^2 = \frac{n}{\sigma^2} \hat{\sigma}_{MV}^2 \sim \chi^2(n-p-1)$.
- 6 $\hat{\sigma}_{MV}^2$ es un estimador más eficiente para σ^2 que $\hat{\sigma}_{MCO}^2$.
- 7 $\hat{\sigma}_{MV}^2$ es un estimador sesgado para σ^2 ; sin embargo, es un estimador asintóticamente insesgado para σ^2 .

1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- Propiedades estadísticas del método MCO
- Medidas de Bondad de Ajuste del modelo
- Modelo Lineal Centrado

2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

¿Cómo se plantea un modelo de Regresión Lineal Múltiple?

Interrogante: ¿Cómo desarrollaríamos un modelo si las variables regresoras son estocásticas? ¿Los resultados estimados serían los mismos?

Sea y variable dependiente y x_1, x_2, \dots, x_p un conjunto de p variables regresoras estocásticas tal que $(y, x_1, \dots, x_p)' = (y \mid \mathbf{x}')' \sim \mathcal{N}_{p+1}(\mu, \Sigma)$ donde μ y Σ son los parámetros particionados como:

$$\mu = \begin{pmatrix} \mu_y \\ \mu_{\mathbf{x}} \end{pmatrix} \quad y \quad \Sigma = \begin{pmatrix} \sigma_{yy} & \sigma'_{\mathbf{x}y} \\ \sigma_{\mathbf{x}y} & \Sigma_{\mathbf{xx}} \end{pmatrix}$$

El problema consiste en hallar la combinación lineal $\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$ que tenga una máxima correlación (en magnitud) con y . Esto es

$$\hat{\beta} = \underset{\beta \in \mathbb{R}^{p+1} \setminus \{0\}}{\operatorname{argmax}} \rho^2(y, \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p) \quad (23)$$

¿Cómo se plantea un modelo de Regresión Lineal Múltiple?

Con un poco de álgebra se demuestra que

$$\max_{\beta \in \mathbb{R}^{p+1} \setminus \{0\}} \rho^2(y, \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p) = \frac{\sigma'_{xy} \Sigma^{-1} \sigma_{xy}}{\sigma_{yy}} \quad (28)$$

Entonces, queda la interrogante: *¿Cuál es la combinación lineal $\beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$ que hace alcanzar esta máxima correlación?*

La respuesta está en $\mathbb{E}(y|\mathbf{x})$. De la propiedad de la esperanza condicional de la normal multivariada se obtiene que:

$$\mathbb{E}(y|\mathbf{x}) = \mu_y + \sigma'_{xy} \Sigma_{xx}^{-1} (\mathbf{x} - \mu_{\mathbf{x}}) = \beta_0 + \beta_1' \mathbf{x} \quad (29)$$

donde $\beta_0 = \mu_y - \sigma'_{xy} \Sigma_{xx}^{-1} \mu_{\mathbf{x}}$ y $\beta_1 = \Sigma_{xx}^{-1} \sigma_{xy}$.

¿Cómo se plantea un modelo de Regresión Lineal Múltiple?

En efecto, se puede probar que:

$$\rho^2(y, \mathbb{E}(y|\mathbf{x})) = \frac{\sigma'_{\mathbf{x}y} \Sigma^{-1} \sigma_{\mathbf{x}y}}{\sigma_{yy}} \quad (30)$$

Observación: Si se calcula la varianza condicional bajo los supuestos del MRL, $var(y | \mathbf{x}) = \sigma_{yy} - \sigma'_{\mathbf{x}y} \Sigma_{\mathbf{xx}}^{-1} \sigma_{\mathbf{x}y}$, se observa que esta no depende de los valores de \mathbf{x} .

1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- Propiedades estadísticas del método MCO
- Medidas de Bondad de Ajuste del modelo
- Modelo Lineal Centrado

2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

¿Cómo estimamos los parámetros en este tipo de modelos?

Sea y, x_1, x_2, \dots, x_p un conjunto de p variables regresoras estocásticas. Al medir n observaciones muestrales de estas $p + 1$ variables obtenemos

$$Z = \begin{pmatrix} y_1 & x_{11} & x_{12} & \dots & x_{1p} \\ y_2 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_n & x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} = \begin{pmatrix} \mathbf{z}_1' \\ \mathbf{z}_2' \\ \vdots \\ \mathbf{z}_n' \end{pmatrix}$$

donde $\mathbf{z}_i = (y_i, x_{i1}, x_{i2}, \dots, x_{ip})' = (y_i \mid \mathbf{x}_i')' \stackrel{\text{iid}}{\sim} \mathcal{N}_{p+1}(\mu, \Sigma)$ tal que

$$\mu = \begin{pmatrix} \mu_y \\ \mu_{\mathbf{x}} \end{pmatrix} \quad y \quad \Sigma = \begin{pmatrix} \sigma_{yy} & \sigma'_{\mathbf{x}y} \\ \sigma_{\mathbf{x}y} & \Sigma_{\mathbf{x}\mathbf{x}} \end{pmatrix}$$

¿Cómo estimamos los parámetros en este tipo de modelos?

Proposición

Sea $B_{n \times n}$ una matriz simétrica y definida positiva y $b > 0$. Entonces

$$\frac{1}{2b}B = \operatorname{argmax}_{\Sigma > 0} \left(\frac{1}{|\Sigma|^b} e^{-\frac{\operatorname{tr}(\Sigma^{-1}B)}{2}} \right) \quad (31)$$

Teorema

Sea $(y_1 | \mathbf{x}_1')', (y_2 | \mathbf{x}_2')', \dots, (y_n | \mathbf{x}_n')'$ una muestra aleatoria simple de una población con distribución $\mathcal{N}(\mu, \Sigma)$ con μ y Σ particionadas como en un MRL. Los estimadores máximo verosímiles para μ y Σ son

$$\hat{\mu}_{MV} = \begin{pmatrix} \bar{y} \\ \bar{\mathbf{x}} \end{pmatrix} \quad y \quad \hat{\Sigma}_{MV} = \frac{n-1}{n} \begin{pmatrix} s_{yy} & s'_{\mathbf{x}y} \\ s_{\mathbf{x}y} & S_{\mathbf{x}\mathbf{x}} \end{pmatrix}.$$

¿Cómo estimamos los parámetros en este tipo de modelos?

Teorema de invarianza de estimadores máximo verosímiles

Si $\hat{\theta}$ es el estimador máximo verosímil del vector de parámetros θ entonces $g(\hat{\theta})$ es el estimador máximo verosímil para $g(\theta)$, donde g es una función continua.

Teorema

Si $(y_1 | \mathbf{x}_1')', (y_2 | \mathbf{x}_2')', \dots, (y_n | \mathbf{x}_n')'$ es una muestra aleatoria simple de una población con distribución $\mathcal{N}(\mu, \Sigma)$ con μ y Σ particionadas como antes. Los estimadores máximo verosímiles para β_0, β_1 y σ^2 son

$$\hat{\beta}_0 = \bar{y} - s'_{\mathbf{x}y} S_{\mathbf{xx}}^{-1} \bar{\mathbf{x}} \quad \text{y} \quad \hat{\beta}_1 = S_{\mathbf{xx}}^{-1} s_{\mathbf{x}y} \quad (32)$$

$$\hat{\sigma}^2 = \frac{n-1}{n} s^2 \quad \text{donde} \quad s^2 = s_{yy} - s'_{\mathbf{x}y} S_{\mathbf{xx}}^{-1} s_{\mathbf{x}y} \quad (33)$$

1 Modelo Lineal General

- Planteamiento del problema
- Planteamiento del modelo lineal general
- Estimación de parámetros: Método MCO
- Propiedades estadísticas del método MCO
- Medidas de Bondad de Ajuste del modelo
- Modelo Lineal Centrado

2 Modelo Lineal Clásico

- Planteamiento del modelo lineal clásico
- Estimación de parámetros: Método de Máxima Verosimilitud
- Propiedades estadísticas del modelo lineal clásico

3 Modelo de Regresión Lineal Múltiple

- Planteamiento del modelo de regresión lineal múltiple
- Estimación de parámetros
- Propiedades estadísticas del modelo de regresión lineal múltiple

¿Qué propiedades tienen los estimadores en un MRL?

- 1 Los estimadores $\hat{\beta}_0, \hat{\beta}_1$ tienen la misma forma algebraica que los estimadores MCO y MV cuando las variables x eran determinísticas.
- 2 $\hat{\beta} = (\hat{\beta}_0 | \hat{\beta}_1')'$ no posee distribución normal. Sin embargo, aún así es un estimador insesgado para β .
- 3 A pesar que $\hat{\beta}$ no posee distribución normal, los test F y t que se ven en pruebas de hipótesis para el caso de x determinístico son aún apropiados. Los cambios solo ocurren en el cálculo de las potencias de las pruebas.
- 4 Para medir el ajuste del modelo se usa el estimador

$$\hat{\rho}_{y|x}^2 = \frac{s'_{xy} S_{xx}^{-1} s_{xy}}{s_{yy}} = R^2 \quad (34)$$

¿Preguntas?

