Multiple linear regression

Jean-Michel Marin

University of Montpellier Faculty of Sciences

HAX912X - 2025/2026

- Homoscedastic models
 - Hypotheses
 - Estimations of β and σ^2
 - Geometric properties

Homoscedastic models

Data

- ightharpoonup n individuals : i = 1, ..., n
- $ightharpoonup y_i \in \mathbb{R}$ variable to explain (response)
- ▶ $\mathbf{x}_i = (\mathbf{x}_{i1}, \dots, \mathbf{x}_{ip}) \in \mathbb{R}^p$ values of p explanatory variables including the constant term constant $\mathbf{x}_{i1} = 1, \forall i = 1, \dots, n$

Matrix notation

$$\mathbf{y} = (y_1, \dots, y_n) \quad ; \quad \mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} 1 & x_{12} & \dots & x_{1p} \\ 1 & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n2} & \dots & x_{np} \end{bmatrix}$$

Homoscedastic models

The multiple linear regression model is such that

$$y = X\beta + u$$

$$\beta=(\beta_1,\dots,\beta_p)$$
 is an unknown parameter $u=(u_1,\dots,u_n)$ vector of residuals

We can write, $\forall i = 1, ..., n$

$$y_i = \beta_1 + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} + u_i$$

 y_i is a random variable for which we have a realization \mathbf{x}_i is a fixed vector



Homoscedastic models Hypotheses

- the constant regressor is present in all models
- n > p more observations than variables
- ➤ X^TX invertible, the rank of X is equal to p, the model is identifiable
- $ightharpoonup \mathbb{E}[\mathbf{u}] = \mathbf{0}_{\mathfrak{n}}$ residuals are centered
- ▶ $\mathbb{E}[\mathbf{u}\mathbf{u}^T] = \mathbb{V}(\mathbf{u}) = \sigma^2 I_n$ residuals are decorrelated and have the same variance σ^2

Homoscedastic models Estimations of β and σ^2

Ordinary least squares (OLS) estimator of β

$$\boldsymbol{\hat{\beta}} \in \text{arg} \min_{\boldsymbol{\beta} \in \mathbb{R}^p} \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 \left(= \text{arg} \min_{\boldsymbol{\beta} \in \mathbb{R}^p} \sum_{i=1}^n \left(\boldsymbol{y}_i - \boldsymbol{x}_i^\mathsf{T} \boldsymbol{\beta} \right)^2 \right)$$

$$\boldsymbol{\hat{\beta}} = \left(\boldsymbol{X}^T \boldsymbol{X} \right)^{-1} \boldsymbol{X}^T \boldsymbol{y}$$

Unbiased estimator of σ^2

$$\hat{\sigma}^2 = \frac{\|y - X\hat{\beta}\|^2}{n - p} = \frac{RSS}{n - p}$$

Estimator of the standard error of $\hat{\beta}_j \mid \hat{\sigma}(\hat{\beta}_j) = \sqrt{\hat{\sigma}^2 (\mathbf{X}^T \mathbf{X})_{(j,j)}^{-1}}$

$$\hat{\boldsymbol{\sigma}}(\boldsymbol{\hat{\beta}}_{j}) = \sqrt{\hat{\boldsymbol{\sigma}}^{2} \left(\boldsymbol{X}^{T} \boldsymbol{X}\right)_{(j,j)}^{-1}}$$

Homoscedastic models

Geometric properties

 χ vector subspace of \mathbb{R}^n generated by the columns of X $H = X \left(X^T X \right)^{-1} X^T$ orthogonal projector on χ $K = I_n - H$ projector on χ^\perp

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X} \left(\mathbf{X}^{\mathsf{T}} \mathbf{X} \right)^{-1} \mathbf{X}^{\mathsf{T}} \mathbf{y} = \mathsf{H} \mathbf{y}$$

 $\hat{\mathbf{y}}$ orthogonal projection of \mathbf{y} on χ

$$\hat{\mathbf{u}} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathsf{H}\mathbf{y} = (\mathsf{I}_{\mathsf{n}} - \mathsf{H})\mathbf{y} = \mathsf{K}\mathbf{y}$$

 $\hat{\mathbf{u}}$ orthogonal projection of \mathbf{y} on χ^{\perp}

$$\hat{\mathbf{y}} \perp \hat{\mathbf{u}}$$



Homoscedastic models Geometric properties

According to the Pythagorean theorem

$$\|\mathbf{y} - \bar{\mathbf{y}}\mathbf{1}_n\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \|\hat{\mathbf{y}} - \bar{\mathbf{y}}\mathbf{1}_n\|^2$$

$$\boxed{\mathsf{TSS} = \mathsf{RSS} + \mathsf{ESS}}$$

Coefficient of determination

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

$$y = X\beta + u$$

n > p, X of full rank, identifiable model

$$\mathbb{E}(\textbf{u})=\textbf{0}_n$$

$$\mathbb{V}(\mathbf{u}) = \mathbb{E}\left[\mathbf{u}\mathbf{u}^\mathsf{T}\right] = \sigma^2\Omega \text{ and } \Omega \neq \mathrm{I}_{\mathfrak{n}}$$

 $\mathbb{C}(u_i,u_j)=\sigma^2\Omega_{ij}\;$ heteroccedasticity and error correlation

It is assumed that Ω is known

Let R be such that $\Omega^{-1} = R^T R$

Set $y^* = Ry$, $X^* = RX$ and $u^* = Ru$.

$$\begin{split} \boldsymbol{y}^* &= R\boldsymbol{y} = R(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}) = R\boldsymbol{X}\boldsymbol{\beta} + R\boldsymbol{u} = \boldsymbol{X}^*\boldsymbol{\beta} + \boldsymbol{u}^* \\ \mathbb{E}(\boldsymbol{u}^*) &= \boldsymbol{0}_n \quad ; \quad \mathbb{V}(\boldsymbol{u}^*) = R\mathbb{V}(\boldsymbol{u})R^\mathsf{T} = \sigma^2\boldsymbol{I}_n \end{split}$$

We have decorrelated the residuals and returned to the case of classical linear regression (homoscedastic)

We can calculate the Ordinary Least Squares (OLS) estimator for the homoscedastic model

$$\boxed{y^* = X^*\beta + u^*}$$

$$\hat{\beta} \in \text{arg} \min_{\beta \in \mathbb{R}^p} \|y^* - X^*\beta\|^2$$

$$\hat{\boldsymbol{\beta}} = \left((\boldsymbol{X}^*)^T \boldsymbol{X}^* \right)^{-1} (\boldsymbol{X}^*)^T \boldsymbol{y}^* = \left(\boldsymbol{X}^T \boldsymbol{\Omega}^{-1} \boldsymbol{X} \right)^{-1} \boldsymbol{X}^T \boldsymbol{\Omega}^{-1} \boldsymbol{y}$$

This is the Generalized Least Squares (GLS) estimator of β

It's difficult to estimate Ω without additional constraints. Ω of particular parametric forms

General estimation method

If we have an estimator $\hat{\Omega}(\beta)$ of Ω , we can iterate until stabilization, we can iterate to stabilize

estimate β with

$$\hat{\boldsymbol{\beta}}^{(t)} = \left(\boldsymbol{X}^T \left[\hat{\boldsymbol{\Omega}}\left(\hat{\boldsymbol{\beta}}^{(t-1)}\right)\right]^{-1} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \left[\hat{\boldsymbol{\Omega}}\left(\hat{\boldsymbol{\beta}}^{(t-1)}\right)\right]^{-1} \boldsymbol{y}$$

estimate Ω with

$$\hat{\Omega}\left(\hat{\beta}^{(t)}\right)$$

