Prevision in the simple linear model

Model: $Y = \beta_0 + \beta_1 X + \varepsilon$

observations: $(x_1, y_1), \dots, (x_n, y_n)$

x_i is associated to X_i (fixed)

y_i ———————— Y_i (random)

random variables $\underbrace{\varepsilon_i}_{\text{noise}}$ (non observable)

$$U = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \vdots \\ \varepsilon_n \end{pmatrix} \sim \underbrace{\mathcal{N}\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \sigma^2 I_n \right)}_{\text{gaussian vector of size } n}$$

Rk: since $\sigma^2 I_n$ is diagonal, because the random vector is gaussian

$\Rightarrow$ we have the independance of the components of the vector let consider a new observation x_{n+1}.

We would like to get information about y_{n+1} which is an observation of Y_{n+1}.

we assume that

$$Y_{n+1} = \beta_0 + \beta_1 X_{n+1} + \varepsilon_{n+1}$$

with

$$\mathbb{E}[\varepsilon_{n+1}] = 0$$

$$V(\varepsilon_{n+1}] = \sigma^2$$

$$\operatorname{cov}(\varepsilon_{n+1}, \varepsilon_i) = 0 \quad \text{for } i \in \{1, \dots, n\}$$

$$\varepsilon_{n+1} \sim \mathcal{N}(0, \sigma^2)$$

We predict Y_{n+1} by

$$\hat{Y}^{(T)}_{n+1} = \hat{\beta}_0 + \hat{\beta}_1 X_{n+1}$$

Rk: $\hat{\beta}_0$ and $\hat{\beta}_1$ have been constructed thanks to the learning sample, thanks to $(X_1, Y_1), \dots, (X_n, Y_n)$.

Prevision error: $\hat{\varepsilon}^{(T)}_{n+1} = Y_{n+1} - \hat{Y}^{(T)}_{n+1}$

<u>Resultats:</u>

- $$E\left[\hat{\varepsilon}^{(T)}_{n+1}\right] = 0$$
- $$V\left[\hat{Y}^{(T)}_{n+1}\right] = \sigma^2 \left(\frac{1}{n} + \frac{(X_{n+1} - \bar{X}_n)^2}{\sum_{i=1}^{n} (X_i - \bar{X}_n)^2} \right)$$

$$V\left[\hat{\varepsilon}_{n+1}^{(P)}\right] = \sigma^2\left(1 + \frac{1}{n} + \frac{(X_{n+1} - \bar{X}_n)^2}{\sum (X_i - \bar{X}_n)^2}\right)$$

<u>Rk</u>: The 2 variances are function of

$$(X_{n+1} - \bar{X}_n)^2 .$$

More and more the distance between X_{n+1} and

$\bar{X}_n$ increases, more and more the variances are big.

↳ the quality of the prevision becomes bad.

Proof:

$$E\left[\hat{\varepsilon}^{(T)}_{n+1}\right] = E\left[Y_{n+1} - \hat{Y}^{(T)}_{n+1}\right]$$

$$= E\left[\left(\beta_0 + \beta_1 X_{n+1} + \varepsilon_{n+1}\right) - \left(\hat{\beta}_0 + \hat{\beta}_1 X_{n+1}\right)\right]$$

$$= \beta_0 + \beta_1 X_{n+1} + \overbrace{E\left[\varepsilon_{n+1}\right]}^{=0} - \underbrace{E\left[\hat{\beta}_0\right]}_{=\beta_0} - \underbrace{E\left[\hat{\beta}_1\right]}_{\beta_1} X_{n+1}$$

$$= 0$$

$$V\left[\hat{y}^{(r)}_{n+1}\right] = \mathbb{E}\left[\left(\hat{y}^{(r)}_{n+1} - \mathbb{E}\left[\hat{y}^{(r)}_{n+1}\right]\right)^2\right]$$

or $\hat{y}^{(r)}_{n+1} = \begin{pmatrix} 1 & x_{n+1} \end{pmatrix} \hat{\beta}$

$$\hookrightarrow \mathbb{E}\left[\hat{y}^{(r)}_{n+1}\right] = \mathbb{E}\left[\underbrace{\begin{pmatrix} 1 & x_{n+1} \end{pmatrix}}_{\text{not random}} \underbrace{\hat{\beta}}_{\text{random vector}}\right]$$

$$= \begin{pmatrix} 1 & x_{n+1} \end{pmatrix} \mathbb{E}\left[\hat{\beta}\right] = \begin{pmatrix} 1 & x_{n+1} \end{pmatrix} \beta$$

$$\hat{Y}_{n+1}^{(P)} - \mathbb{E}\left[\hat{Y}_{n+1}^{(P)}\right] = \begin{pmatrix} 1 & X_{n+1} \end{pmatrix} \underbrace{\begin{pmatrix} \hat{\beta} - \beta \end{pmatrix}}$$

$$\hat{\beta} = \left({}^t\mathbb{X}\mathbb{X} \right)^{-1} {}^t\mathbb{X}\,\underline{Y}$$

$$\underline{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad \mathbb{X} = \begin{pmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_n \end{pmatrix}$$

$$\underline{Y} = \mathbb{X}\beta + u \qquad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

$$\hat{\beta} = \left(\mathbb{X}^T\mathbb{X}\right)^{-1}\mathbb{X}^T\left(\mathbb{X}\beta + u\right)$$

$$= \underbrace{\left(\mathbb{X}^T\mathbb{X}\right)^{-1}\left(\mathbb{X}^T\mathbb{X}\right)}_{I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}\beta + \left(\mathbb{X}^T\mathbb{X}\right)^{-1}\mathbb{X}^T u$$

$$\hat{\beta} = \beta + ({}^t\mathbb{X}\mathbb{X})^{-1}\,{}^t\mathbb{X}\,U$$

$$\Rightarrow \hat{\beta} - \beta = ({}^t\mathbb{X}\mathbb{X})^{-1}\,{}^t\mathbb{X}\,U$$

let denote $A = \begin{pmatrix} 1 & X_{n+1} \end{pmatrix}$

$$V\left[\hat{Y}^{(P)}_{n+1}\right] = V\left[A(\hat{\beta} - \beta)\right]$$

$$= V\left[\underbrace{A\left({}^t\mathbb{X}\mathbb{X}\right)^{-1}{}^t\mathbb{X}}_{B\,:\ \text{deterministic}}\overbrace{U}^{\text{random vector}}\right]$$

$$V[BU] = E\left[(BU - E(BU))^t \cdot (BU - E(BU))\right]$$

But:

$$E[BU] = B\,E[U] = 0$$

$$V[BU] = E\left[BU \cdot {}^t(BU)\right]$$

$$= E\left[B\underbrace{U\,{}^tU}_{\text{random matrix}}\,{}^tB\right] = B\,E\left(U \cdot {}^tU\right)\,{}^tB$$

$$E\left[U \cdot {}^tU\right] = E\left[\left(U - E[U]\right) \cdot {}^t\left(U - E[U]\right)\right]$$

$$= V[U]$$

$$= \sigma^2 I_n$$

$$\Rightarrow \quad V[BU] = \sigma^2 B\, {}^tB$$

$$\Rightarrow \quad V\left[\hat{Y}^{(T)}_{n+1}\right] = \sigma^2 \underbrace{A\left(X^T X\right)^{-1} X^T}_{B} \, \underbrace{X \left(X^T X\right)^{-1} A^T}_{B^T}$$

$$= \sigma^2 A \left(X^T X\right)^{-1} A^T$$

$$\mathbb{X} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

$${}^t\mathbb{X}\mathbb{X} = \begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix}$$

$$(\mathbb{X}^T \mathbb{X})^{-1} = \frac{1}{n\sum x_i^2 - \left(\sum x_i\right)^2} \begin{pmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & n \end{pmatrix}$$

$$A(\mathbb{X}^T \mathbb{X})^{-1} = \frac{1}{n\sum x_i^2 - \left(\sum x_i\right)^2} \left(\sum x_i^2 - x_{n+1}\sum x_i, \; -\sum x_i + n x_{n+1}\right)$$

$$A\left({}^tXX\right)^{-1}{}^tA = \frac{1}{n\sum X_i^2 - \left(\sum X_i\right)^2}\left(\sum X_i^2 - X_{n+1}\underbrace{\sum X_i}_{=n\bar{X}_n} - X_{n+1}\sum X_i + nX_{n+1}^2\right)$$

$$= \frac{1}{n\sum X_i^2 - n^2\bar{X}_n^2}\left(\sum X_i^2 - 2nX_{n+1}\bar{X}_n + nX_{n+1}^2\right)$$

$$\mathbf{a}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{a} = \frac{n(X_{n+1}-\bar{X}_n)^2 - n\bar{X}_n^2 + \sum X_i^2}{n\left(\sum X_i^2 - n\bar{X}_n^2\right)}$$

$$= \sum (X_i - \bar{X}_n)^2$$

$$= \frac{1}{n} + \frac{(X_{n+1}-\bar{X}_n)^2}{\sum (X_i - \bar{X}_n)^2}$$

$$V\left[\hat{Y}^{(r)}_{n+1}\right] = \sigma^2\left(\frac{1}{n} + \frac{(x_{n+1} - \bar{x}_n)^2}{\sum (x_i - \bar{x}_n)^2}\right)$$