

Apprentissage statistique

TD1 : Modèle statistique

Lucie Le Briquer

9 janvier 2018

Exercice 1

Modèle :

$$Y = x^1 \beta_1 + \dots + x^d \beta_d + \varepsilon \in \mathbb{R}$$

On suppose ε centrée et de variance bornée. n réalisations de ce modèle indépendantes à x fixé.

1. Modèle statistique : $\Omega = \mathbb{R}^n$, $\mathcal{F} = \mathcal{B}(\mathbb{R}^n)$, $Y = (Y_1, \dots, Y_n)$,

$$\mathcal{P} = \left\{ \mathbb{P} \text{ proba sur } (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)) \mid \mathbb{E}[Y_i] = \sum_{j=1}^d x_i^j \beta_j, \mathbb{E}[Y_i^2] < +\infty, (Y_i)_{1 \leq i \leq n} \text{ idp sous } \mathbb{P} \right\}$$

- 2.

$$\mathcal{P} = \{\mathbb{P}_{x, \beta, \sigma^2} = \mathcal{N}(x\beta, \sigma^2 I_n) \text{ où } x \text{ est la matrice } (x_i^j)_{i,j}\}$$

\mathcal{P} est dominé par la mesure de Lebesgue et les densités associées sont :

$$\alpha \exp\left(-\frac{\|y - x\beta\|^2}{2\sigma^2}\right)$$

3. Pour tout $(y_1, x_i^1, \dots, x_i^d)_{1 \leq i \leq n}$ montrons que :

$$\inf_{\beta \in \mathbb{R}^d} \sum (y_i - (x\beta)_i)^2 = \min(\dots)$$

Soit $\rho(\beta) = \|y - x\beta\|^2$, ρ est convexe. On considère $V \subset \mathbb{R}^n$ défini par :

$$V = \{z \mid z = x\beta, \beta \in \mathbb{R}^d\}$$

On a :

$$\inf_{\beta \in \mathbb{R}^d} \rho(\beta) = \inf_{z \in V} \|y - z\|^2 = \min_{z \in V} \|y - z\|^2$$

atteint en $z = p_V(y)$ (la projection orthogonale de y sur V) car V est un sev de dimension finie.

Remarque. $\hat{\beta}_n(x_1 \dots x_n, y_1 \dots y_n)$ est l'EMV de β dans le modèle gaussien.

4. On suppose maintenant x injective, $d \leq n$. Montrons que $\forall (x, y) \in \mathbb{R}^{(d+1)n}$, l'estimateur des moindres carrés est unique et donné par :

$$\hat{\beta}_n(x, y) = (x^T x)^{-1} x^T y$$

1ère étape. Montrons que $x^T x$ est inversible. Soit $\beta \in \text{Ker}(x^T x)$. $x^T x \beta = 0$ alors $\beta^T x^T x \beta = 0$ i.e. $\|x\beta\|^2 = 0$ donc $x\beta = 0 \Rightarrow \beta = 0$ car x injective.

2nde étape. On a plusieurs solutions :

- (a) Si x est injective alors ρ est fortement convexe i.e. $\nabla^2 \rho(\beta) > 0 \forall \beta \in \mathbb{R}^d$. Donc ρ admet un unique minimum tel que $\nabla f(\beta) = 0$ i.e. $-2x^T y + 2x^T x \beta = 0$ ainsi forcément :

$$\beta = (x^T x)^{-1} x^T y$$

- (b) Montrons que le minimum est unique et que $\hat{\beta}_n$ atteint le minimum.

- (c) Caractérisation de la projection orthogonale : $z = p_V(y)$ ssi $\forall \tilde{z} \in V$, $\langle y - z, \tilde{z} \rangle = 0$

5. On se place maintenant dans le cas gaussien et $x^T x$ inversible, $d \leq n$. On définit le vecteur des résidus comme :

$$\hat{\varepsilon}_n(Y) = Y - x\hat{\beta}_n = (I_n - x(x^T x)^{-1} x^T) Y$$

$Y \sim \mathcal{N}(x\beta, I_n \sigma^2)$, $\hat{\beta}_n \sim \mathcal{N}((x^T x)^{-1} x^T x \beta, \sigma^2 (x^T x)^{-1})$ car $x^T x$ est symétrique donc $(x^T x)^{-1}$ aussi on donc on a la loi de $\hat{\beta}_n$.

$$\hat{\varepsilon}_n(Y) \sim \mathcal{N}\left(0, \sigma^2 (I_n - x(x^T x)^{-1} x^T)^2\right)$$

B projection ? Soit $z \in \mathbb{R}^n$. $Bz \in V$, β projecteur sur F sev de V . Il suffit de montrer que $\forall z \in V$, $Bz = z$. $z \in V$, $\exists \beta \in \mathbb{R}^d$ tel que $z = x\beta$

$$\beta x \beta = \underbrace{x(x^T x)^{-1} x^T x \beta}_{=\text{id}} = x\beta = z$$

Donc $F = V$. Conclusion, β projecteur et $\text{id} - \beta$ aussi.

6. Un estimateur de β est $\hat{\beta}_n$. Un estimateur de σ^2 ?

$$\mathbb{E}[\|\hat{\varepsilon}_n\|^2] = \sigma^2 \text{Tr}(I - B) = \sigma^2(n - d)$$

Un estimateur de σ^2 est $\frac{\|\hat{\varepsilon}_n\|^2}{n-d}$.

7. $\bar{\beta}$ est dit linéaire si $\bar{\beta}(Y) = AY$ pour $A \in \mathcal{M}_{d,n}(\mathbb{R})$, $\mathbb{E}[\bar{\beta}] = \beta$. Montrons qu'il existe R définie positive telle que :

$$\text{Cov}(\bar{\beta}_A) = \text{Cov}(\hat{\beta}_n) + R$$

Solution : comme $\bar{\beta}_A$ est non biaisé, $\mathbb{E}[\bar{\beta}] = \beta$, or $\mathbb{E}[\bar{\beta}] = \mathbb{E}[AY] = A\mathbb{E}[Y] = Ax\beta$. Ainsi, $Ax\beta = \beta$.

$$\begin{aligned} \text{Cov}(\bar{\beta}_A) &= \mathbb{E}[(AY - \beta)(AY - \beta)] \\ &= A\text{Cov}(Y)A^T \end{aligned}$$