Devoir maison I

Exercice 1.

Soit le modèle $y = f + \xi$ dans $\mathbb{R}^p$ où ξ variable aléatoire telle que $\mathrm{E}\left(\xi_i \cdot \xi_j\right) = 0$ pour $i \neq j$ et $\mathrm{E}\xi_i^2 = \sigma^2$ pour $i = 1, \ldots, n$.

Soit $\widehat{\theta}^{\mathrm{MC}}$ l'estimateur des moindres carrés sur $\mathbb{R}^p$, *ie* $\widehat{\theta}^{\mathrm{MC}}$ solution du programme $\min_{\theta \in \mathbb{R}^p} \|y - X\theta\|^2$.

Montrer que l'on a la relation suivante :

$$\mathbb{E}\|\tilde{f} - f\|^2 = \min_{\theta \in \mathbb{R}^p} \|X\theta - f\|^2 + \frac{\sigma^2 R}{n}$$

Solution 1.

Notons $g(\theta) = \|y - X\theta\|^2$ à valeur dans $\mathbb{R}^p$. On a clairement g différentiable sur $\mathbb{R}^p$ ensemble convexe donc en vertu du lemme 1.3, on a

$$\forall \theta \in \mathbb{R}^p, \qquad < \nabla g(\widehat{\theta}^{\mathrm{MC}}), \theta - \widehat{\theta}^{\mathrm{MC}} > \geq 0$$

En particulier, en appliquant cette inégalité avec $\theta = \widehat{\theta}^{\mathrm{MC}} - \nabla g(\widehat{\theta}^{\mathrm{MC}})$ on a $\nabla g(\widehat{\theta}^{\mathrm{MC}}) = 0$. L'inégalité précedente devient une égalité et on a

$$\forall \theta \in \mathbb{R}^p, \qquad < 2X^t \cdot \left(X\widehat{\theta}^{\mathrm{MC}} - y\right), \theta - \widehat{\theta}^{\mathrm{MC}} > = 0 \quad \Leftrightarrow \quad 2 < X\widehat{\theta}^{\mathrm{MC}} - y, X\left(\theta - \widehat{\theta}^{\mathrm{MC}}\right) >= 0$$

En notant $\tilde{f} = X\widehat{\theta}^{\mathrm{MC}}$ et en utilisant la relation $y = f + \xi$, on a

$$\forall \theta \in \mathbb{R}^p, \qquad 2 < \tilde{f} - f - \xi, \quad X\theta - \tilde{f} >= 0$$

Par le théorème d'Al-Kashi, on a la relation suivante :

$$2 < \tilde{f} - f, X\left(\theta - \widehat{\theta}^{\mathrm{MC}}\right) >= \|X\theta - f\|^2 - \left\|\tilde{f} - f\right\|^2 - \left\|X\theta - \tilde{f}\right\|^2$$

D'où l'égalité suivante valable pour tout θ appartenant à $\mathbb{R}^p$

$$\|X\theta - f\|^2 - \left\|\tilde{f} - f\right\|^2 - \left\|X\theta - \tilde{f}\right\|^2 - 2 < \xi, \ X\theta - \tilde{f} >= 0$$

Soit $A = \left(a_{i,j}\right)$ la matrice de projection orthogonale sur $\operatorname{Im} X$. On a

$$2 < \xi, \ X\theta - \tilde{f} >= 2 < A\xi, \ X\theta - \tilde{f} >= \left\|A\xi + X\theta - \tilde{f}\right\|^2 - \left\|X\theta - \tilde{f}\right\|^2 - \|A\xi\|^2$$

Par ailleurs, on a par hypothèses sur ξ et en utilisant $A^t \cdot A = A$ (A est une projection orthogonale)

$$\mathrm{E}\,\|A\xi\|^2 = \mathrm{E}\frac{1}{n}\sum_{i,j} a_{i,j}\xi_i \cdot \xi_j = \frac{\sigma^2}{n}\sum_i a_{i,i} = \frac{\sigma^2}{n}\operatorname{tr} A = \frac{\sigma^2}{n}\operatorname{rg} A$$

D'où la relation

$$\|X\theta - f\|^2 - \left\|\tilde{f} - f\right\|^2 - \left\|A\xi + X\theta - \tilde{f}\right\|^2 + \frac{\sigma^2}{n}\operatorname{rg} A = 0$$

En remplaçant θ par $\theta_0 := \min_{\theta\in\mathbb{R}^p} \|f - X\theta\|^2$ et en utilisant $Ay = \tilde{f}$ et $Af = X\theta_0$, le troisième terme s'annule :

$$\left\|A\xi + X\theta_0 - \tilde{f}\right\|^2 = \left\|Ay - Af + X\theta_0 - \tilde{f}\right\|^2 = \left\|\tilde{f} - X\theta_0 + X\theta_0 - \tilde{f}\right\|^2 = 0$$

donnant bien la relation demandée.

Exercice 2.

Montrer le *lemme de Hoeffding (1963)*:

Si η est une variable aléatoire telle que $\mathbb{E}(\eta) = 0$ et $\eta \in [a,b]$ avec $a < b$ réels, alors

$$\mathbb{E}\exp(t\eta) \leq \exp\left(\frac{t^2(b-a)^2}{8}\right), \quad \forall t \in \mathbb{R}$$

Solution 2. Commençons par procéder à un changement de variable :

$$\eta = \theta \cdot a + (1-\theta) \cdot b$$

avec $\theta \in [0,1]$ variable aléatoire.

On a par convexité

$$\exp(t\cdot\eta) \leq \theta\exp(t\cdot a) + (1-\theta)\exp(t\cdot b)$$

.

Par la définition de θ et en utilisant le fait que η est centrée, on a que

$$\mathrm{E}(\theta) = \frac{b}{b-a}$$

En passant à l'espérance sur l'inégalité plus haut, on obtient :

$$\mathrm{E}\exp(t\cdot\eta) \leq \frac{1}{b-a}\{b\exp(t\cdot a) - a\exp(t\cdot b)\}$$

Etudions le membre de gauche. Posons

$$\varphi(t) = \log\left(\frac{1}{b-a}\{b\exp(t\cdot a) - a\exp(t\cdot b)\}\right)$$

pour t appartenant à $\mathbb{R}$. On a par calcul

$$\varphi'(t) = \frac{ab\{\exp(t\cdot a) - \exp(t\cdot b)\}}{b\exp(t\cdot a) - a\exp(t\cdot b)}$$

$$\varphi''(t) = \frac{-ab(a-b)^2\exp t(a+b)}{(b\exp(t\cdot a) - a\exp(t\cdot b))^2}$$

En utilisant le fait que $(b\exp(t\cdot a) - a\exp(t\cdot b))^2 + 4ab\exp(t\cdot(a+b)) \geq 0$, on a que $\varphi''(t) \leq \frac{(a-b)^2}{4}$.

On conclut la preuve en utilisant sur φ la formule de Taylor-Lagrange à l'ordre 2 entre 0 et t

$$\forall t \in \mathbb{R}, \exists s \in]0,t[\qquad \text{tel que } \varphi(t) = \varphi(0) + t\varphi'(0) + \frac{1}{2}t^2\varphi''(s) \leq \frac{(a-b)^2\cdot t^2}{8}.$$

Exercice 3.

Montrer que si η est une variable aléatoire σ sous-gaussienne alors $\mathbb{E}(\eta) = 0$.

Solution 3.

En appliquant l'inégalité de Jensen

$$\varphi(\mathrm{E}[X]) \leq \mathrm{E}[\varphi(X)]$$

avec $X = t \cdot \eta$ où t réel et $\varphi = \exp$ fonction convexe, on a l'inégalité suivante :

$$\forall t \in \mathbb{R}, \qquad \exp^{t \cdot \mathrm{E}(\eta)} \leq \mathrm{E}\left(\exp^{t \cdot \eta}\right) \leq exp^{\frac{\sigma^2 t^2}{2}}$$

la dernière inégalité provenant du caractère σ sous-gaussien de η.

L'exponentielle étant strictement croissante, il vient que l'on a

$$t \cdot \mathrm{E}(\eta) \leq \frac{\sigma^2 t^2}{2}$$

pour tout réel t.

Le polynôme du second degré

$$P(X) = \frac{\sigma^2}{2} \cdot X^2 - \mathrm{E}(\eta) \cdot X$$

possède donc au plus une racine réelle et en notant Δ son discriminant, on a donc

$$\Delta = \mathrm{E}(\eta)^2 \leq 0$$

ce qui permet de conclure.