MAP 553. Corrigé de l'examen. 13/12/2010

Exercice 1.

1). On utilise la Proposition 3.3 du polycopié du cours avec $G(u_1, \dots, u_n) = \sum_{i=1}^n (y_i - u_i)^2$ et $\Omega(t) = t^2$.

2). Soit $f(x) = \sum_{i=1}^n \theta_i \, K(x_i, x)$. Alors,

$$|f|_{\mathcal{F}}^2 \underset{\uparrow}{=} \langle f(\cdot), f(\cdot) \rangle_{\mathcal{F}} = \langle f(\cdot), \sum_{i=1}^n \theta_i K(x_i, \cdot) \rangle_{\mathcal{F}}$$

$\langle \cdot, \cdot \rangle_{\mathcal{F}}$ = produit scalaire associé à $|\cdot|_{\mathcal{F}}$

$$\underset{\uparrow}{=} \sum_{i=1}^n \theta_i f(x_i) = \sum_{i=1}^n \theta_i \sum_{j=1}^n \theta_j K(x_j, x_i) =$$

propriété de reproduction

$$= \theta^T \mathbb{K} \theta, \quad \text{où} \quad \mathbb{K} = (K(x_j, x_i))_{i,j=1,\dots,n}.$$

D'après la Question 1), $\hat{\theta}$ est solution de

$$\min_{\theta \in \mathbb{R}^n} \left(\sum_{i=1}^n \Big(y_i - \sum_{j=1}^n \theta_j K(x_j, x_i)\Big)^2 + \lambda \theta^T \mathbb{K} \theta \right)$$

$$= \min_{\theta \in \mathbb{R}^n} \left(|Y - \mathbb{K}\theta|_2^2 + \lambda \theta^T \mathbb{K} \theta \right).$$

La condition nécessaire et suffisante du minimum de la fonction quadratique

$$\theta \mapsto |Y - K\theta|_2^2 + \lambda \theta^T K \theta$$

est :

$$K^T Y = (K^T K + \lambda K)\theta$$

Comme K est inversible,

$$\hat{\theta} = (K^T K + \lambda K)^{-1} K^T Y.$$

Exercice 2.

1). $$P\left(\left|\tfrac{1}{n} X^T (y - X\theta^*)\right|_\infty \le \lambda \right) =$$

$$= P\left(\left|\tfrac{1}{n} X^T \xi\right|_\infty \le \lambda \right),$$

$$\left|\tfrac{1}{n} X^T \xi\right|_\infty = \max_{1 \le j \le M} |\mathcal{S}_j|,$$

où $\mathcal{S}_j = x_{(j)}^T \xi / n$ et $x_{(j)}$ est la jème colonne de X. Les variables aléatoires $\mathcal{S}_j$ sont gaussiennes de moyenne 0 et de variance

$$E(\mathcal{S}_j^2) = \frac{1}{n^2} E\left(x_{(j)}^T \xi \xi^T x_{(j)} \right) = \frac{1}{n^2} |x_{(j)}|_2^2 = \frac{1}{n},$$

car $|x_{(j)}|_2^2 / n$ est le jème élément diagonal de la matrice $\frac{1}{n}X^T X$.

$\Rightarrow \quad \zeta_j \sim \mathcal{N}(0, \frac{1}{n})$. Notons $\eta_j = \zeta_j \sqrt{n}$

$\Rightarrow \quad \eta_j \sim \mathcal{N}(0,1)$.

$$\mathbb{P}\left(\left| \frac{1}{n} X^T \xi \right|_\infty \le \lambda \right) = \mathbb{P}\left(\frac{1}{\sqrt{n}} \max_{1 \le j \le M} |\eta_j| \le \lambda \right).$$

D'après le cours,

$$\mathbb{P}\left(\max_{1\le j \le M} |\eta_j| > \sqrt{2 \log M} \right) \le \frac{1}{\sqrt{\pi \log M}}.$$

$$\Rightarrow \mathbb{P}\left(\left| \frac{1}{n} X^T \xi \right|_\infty \le \sqrt{\frac{2 \log M}{n}} \right) > 1 - \frac{1}{\sqrt{\pi \log M}}$$

$$\to 1, \quad M \to \infty.$$

2). Si $\frac{1}{n} X^T X = I_M$, $\hat{\theta}^D$ est solution de

$$\min_{\theta : |y_j - \theta_j| \le \lambda, \, j = 1, \dots, M} \sum_{j=1}^M |\theta_j|$$

, où y_j est le $j$ème élément du vecteur $\frac{1}{n} X^T y$

Il suffit de résoudre le problème

$$\min_{x \in \mathbb{R} : |y_j - x| \le \lambda} |x|$$

Ce problème a comme solution $x^* = y_j\left(1 - \frac{\lambda}{|y_j|}\right)_+$.

$$\Rightarrow \quad \hat{\theta}_j^{\mathcal{D}} = y_j \left(1 - \frac{\lambda}{|y_j|}\right)_+ , \quad j = 1, \dots, M.$$

C'est l'estimateur de seuillage doux avec le seuil $\tau = \lambda$.

Exercice 3

1)
$$\hat{p}_{n,s}(x) = \frac{1}{2\pi} \int \mathcal{F}[\hat{p}_{n,s}](\omega)\, e^{-i\omega x}\, d\omega$$

$$= \frac{1}{2\pi} \int (-i\omega)^s \varphi_n(\omega)\, \lambda(\omega)\, e^{-i\omega x}\, d\omega$$

$$= \frac{d^s}{dx^s} \left(\frac{1}{2\pi} \int \varphi_n(\omega)\, \lambda(\omega)\, e^{-i\omega x}\, d\omega \right)$$

$$= \frac{d^s}{dx^s} \left[\frac{1}{n} \sum_{i=1}^n \Lambda(x - X_i) \right] =$$

$$\Lambda(x) = \mathcal{F}^{-1}[\lambda](x), \qquad \lambda = \mathcal{F}[\Lambda]$$

$$= \frac{1}{n} \sum_{i=1}^n \Lambda^{(s)}(x - X_i)$$

(Condition: $\Lambda^{(s)}$ existe si $\int |\omega|^s |\lambda(\omega)|\, d\omega < \infty$)

2) $\text{MISE} = E \int \left(\hat{p}_{n,s} - p^{(s)}\right)^2 \overset{\text{Plancherel}}{=}$

$$= \frac{1}{2\pi} E \int |\omega|^{2s} \left| \varphi_n(\omega)\lambda(\omega) - \varphi(\omega) \right|^2 d\omega$$

$$= \frac{1}{2\pi} E\Big[\int |\omega|^{2s} E|\varphi_n(\omega)-\varphi(\omega)|^2 |\lambda(\omega)|^2 d\omega$$ (5.)

$$+ \int |\omega|^{2s} |1-\lambda(\omega)|^2 |\varphi(\omega)|^2 d\omega \Big]$$

$$= \frac{1}{2\pi} \Big[\int \frac{1}{n}\left(1-|\varphi(\omega)|^2\right)|\omega|^{2s} |\lambda(\omega)|^2 d\omega$$

$$+ \int |1-\lambda(\omega)|^2 |\omega|^{2s} |\varphi(\omega)|^2 d\omega \Big]$$

3). Posons $\varepsilon^2(\omega) \overset{\text{déf}}{=} \frac{1}{n}\left(1-|\varphi(\omega)|^2\right)$. Il suffit de minimiser l'expression sous l'intégrale pour tout ω fixé.

$$|\omega|^{2s}\left[\varepsilon^2(\omega)|\lambda(\omega)|^2 + |1-\lambda(\omega)|^2|\varphi(\omega)|^2\right] \to \min_{\lambda(\omega)}$$

$$\Rightarrow \boxed{\lambda^*(\omega) = \frac{|\varphi(\omega)|^2}{\varepsilon^2(\omega)+|\varphi(\omega)|^2}} \text{ solution.}$$

$$MISE^* = \frac{1}{2\pi}\Big[\int \varepsilon^2(\omega)|\omega|^{2s} |\lambda^*(\omega)|^2 d\omega$$

$$+ \int |1-\lambda^*(\omega)|^2 |\omega|^{2s} |\varphi(\omega)|^2 d\omega \Big]$$

$$= \frac{1}{2\pi} \int |\omega|^{2s} \frac{\varepsilon^2(\omega)|\varphi(\omega)|^2}{\varepsilon^2(\omega)+|\varphi(\omega)|^2} d\omega .$$

4). $\lambda(\omega) = \hat{K}(h\omega)$
$\quad = I(|h\omega| \le 1)$ $\quad\Big|\quad K(u) = \frac{\sin u}{\pi u}$

$$\Lambda(x) = \frac{1}{h} K\left(\frac{x}{h}\right) = \frac{1}{h}\left(\frac{\sin(x/h)}{(\pi x/h)}\right) = \frac{\sin(x/h)}{\pi x}$$

4.1). L'estimateur est de la forme (cf. Question 1)):

$$\hat{p}_{n,s}(x) = \frac{d^s}{dx^s}\left(\frac{1}{n}\sum_{i=1}^{n}\Lambda(x-X_i)\right)$$

$$= \frac{d^s}{dx^s}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\sin((x-X_i)/h)}{\pi(x-X_i)}\right) =$$

$$= \frac{d^s}{dx^s}\left(\frac{1}{nh}\sum_{i=1}^{n}\frac{\sin((x-X_i)/h)}{\pi(x-X_i)/h}\right) =$$

$$= \frac{1}{nh}\sum_{i=1}^{n}\frac{d^s}{dx^s}K\left(\frac{x-X_i}{h}\right) =$$

$$= \frac{1}{nh^{s+1}}\sum_{i=1}^{n}K^{(s)}\left(\frac{x-X_i}{h}\right).$$

4.2). Le terme de biais.

$$\int |1-\hat{K}(h\omega)|^2 |\omega|^{2s} |\varphi(\omega)|^2 d\omega$$

$$= \int_{|\omega h|>1} |\varphi(\omega)|^2 |\omega|^{2s} d\omega$$

$$= \int_{|\omega|>1/h} |\omega|^{2\beta} |\varphi(\omega)|^2 |\omega|^{2s-2\beta} d\omega$$

$$\leq h^{2(\beta-s)} \int |\omega|^{2\beta} |\varphi(\omega)|^2 d\omega \leq \boxed{2\pi L^2 h^{2(\beta-s)}}$$

Le terme de variance.

$$\int \frac{1}{n}\left(1-|\varphi(\omega)|^2\right)|\omega|^{2s}\, I\left(|h\omega|\le 1\right)d\omega$$

$$\le \frac{1}{n}\int |\omega|^{2s}\, I\left(|h\omega|\le 1\right)d\omega =$$

$$= \frac{1}{n}\int_{|\omega|\le 1/h} |\omega|^{2s}\, d\omega = \boxed{\frac{2}{(2s+1)\,n\,h^{2s+1}}}.$$

$$\Rightarrow \quad MISE \le \frac{1}{2\pi}\left(2\pi L^2 h^{2(\beta-s)} + \frac{2}{(2s+1)\,n\,h^{2s+1}}\right).$$

$$h_{opt} = \underset{h>0}{\operatorname{argmin}}\left(L^2 h^{2(\beta-s)} + \frac{1}{(2s+1)\,\pi\, n\, h^{2s+1}}\right)$$

$$\Rightarrow \quad h_{opt} = C(\beta, s, L)\, n^{-\frac{1}{2\beta+1}}$$

où $C(\beta, s, L)$ ne dépend que de β, s, L.

$$MISE_{opt} = O\left(h_{opt}^{2(\beta-s)}\right) = O\left(n^{-\frac{2(\beta-s)}{2\beta+1}}\right), \quad n\to\infty.$$

5).

$$E\,|\varphi_n(\omega)|^2 = \left(1-\frac{1}{n}\right)|\varphi(\omega)|^2 + \frac{1}{n}.$$

$$|\varphi(\omega)|^2 = \left(E\,|\varphi_n(\omega)|^2 - 1/n\right)\frac{1}{\left(1-\frac{1}{n}\right)}$$

$$1-|\varphi(\omega)|^2 = 1 - \frac{E\,|\varphi_n(\omega)|^2 - 1/n}{1-1/n} =$$

$$= \frac{1-E\,|\varphi_n(\omega)|^2}{1-1/n}.$$

$$\Rightarrow \quad E(\tilde{J}(\lambda)) = \left(1-\frac{1}{n}\right) \cdot 2\pi \cdot \text{MISE}.$$

Définissons $\hat{\lambda}$, l'estimateur adaptatif de λ, par:

$$\hat{\lambda} = \arg\min_{\lambda} \tilde{J}(\lambda).$$

$$\tilde{J}(\lambda) = \int \varepsilon_n^2(\omega)\, |\omega|^{2s} |\lambda(\omega)|^2 d\omega +$$

$$+ \int |1-\lambda(\omega)|^2 |\omega|^{2s} g_n(\omega)\, d\omega$$

avec $\varepsilon_n^2(\omega) = \frac{1}{n}\left(1-|\varphi_n(\omega)|^2\right)$, $g_n(\omega) = |\varphi_n(\omega)|^2 - \frac{1}{n}$.

Si $g_n(\omega) > 0$, le minimum est atteint par minimisation de l'expression sous l'intégrale pour tout ω fixé:

$$\min_{\lambda \in \mathbb{R}} \left(\varepsilon_n^2(\omega)\, \lambda^2 + (1-\lambda)^2 g_n(\omega)\right)$$

$$\Rightarrow \quad \lambda^*(\omega) = \frac{g_n(\omega)}{\varepsilon_n^2(\omega) + g_n(\omega)} \qquad \text{(solution)}$$

$$= \frac{n}{n-1}\left(1 - \frac{1}{n|\varphi_n(\omega)|^2}\right).$$

Si $g_n(\omega) \leq 0$, la meilleure solution est $\lambda^*(\omega) = 0$.

$\Rightarrow$ dans tous les cas, $\lambda^*(\omega) = \frac{n}{n-1}\left(1 - \frac{1}{n|\varphi_n(\omega)|^2}\right)_+$

Pour $\lambda(\omega) = \hat{K}(h\omega)$, on cherche

$$\hat{h} = \arg\min_{h>0} \tilde{J}(\hat{K}(h\cdot)).$$