

MAP 433 : Introduction aux méthodes statistiques. Cours 4

21 février 2014

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

- 1 M -estimation, rappel du Cours 3
 - Principe de maximum de vraisemblance
- 2 EMV, asymptotique des Z - et M - estimateurs
- 3 Méthode d'estimation dans le modèle de régression
 - Modèle de régression, notion de « design »
 - Régression à design déterministe
 - La droite des moindres carrés
 - Régression linéaire multiple

M -estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

M-estimation

- Situation : on observe X_1, \dots, X_n de loi \mathbb{P}_ϑ sur \mathbb{R} et $\vartheta \in \Theta$.
- Principe : Se donner une application $\psi : \Theta \times \mathbb{R} \rightarrow \mathbb{R}_+$ telle que, pour tout $\vartheta \in \Theta \subset \mathbb{R}^d$,

$$a \rightsquigarrow \mathbb{E}_\vartheta [\psi(a, X)] = \int \psi(a, x) \mathbb{P}_\vartheta(dx)$$

admet **un maximum en $a = \vartheta$** .

Définition

On appelle *M-estimateur* associé à ψ tout estimateur $\hat{\vartheta}_n$ satisfaisant

$$\sum_{i=1}^n \psi(\hat{\vartheta}_n, X_i) = \max_{a \in \Theta} \sum_{i=1}^n \psi(a, X_i).$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des Z- et M-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Un exemple classique : paramètre de localisation

- $\Theta = \mathbb{R}$, $\mathbb{P}_{\vartheta}(dx) = f(x - \vartheta)dx$, et $\int_{\mathbb{R}} xf(x)dx = 0$,
 $\int_{\mathbb{R}} x^2 \mathbb{P}_{\vartheta}(dx) < +\infty$ pour tout $\vartheta \in \mathbb{R}$. On pose

$$\psi(a, x) = -(a - x)^2$$

- La fonction

$$a \rightsquigarrow \mathbb{E}_{\vartheta} [\psi(a, X)] = - \int_{\mathbb{R}} (a - x)^2 f(x - \vartheta) dx$$

admet un **maximum** en $a = \mathbb{E}_{\vartheta} [X] = \int_{\mathbb{R}} xf(x - \vartheta)dx = \vartheta$.

- **M-estimateur associé :**

$$\sum_{i=1}^n (X_i - \hat{\vartheta}_n)^2 = \min_{a \in \mathbb{R}} \sum_{i=1}^n (X_i - a)^2.$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des Z- et M-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Paramètre de localisation

- C'est **aussi** un Z -estimateur associé à $\phi(a, x) = 2(x - a)$: on résout

$$\sum_{i=1}^n (a - X_i) = 0 \text{ d'où } \hat{\vartheta}_n = \bar{X}_n.$$

- Dans cet **exemple très simple**, tous les points de vue coïncident.
- Si, dans le même contexte, $\int_{\mathbb{R}} x^2 \mathbb{P}_{\vartheta}(dx) = +\infty$ et $f(x) = f(-x)$, on peut utiliser Z -estimateur avec $\phi(a, x) = \text{Arctg}(x - a)$. Méthode robuste, mais est-elle optimale? Peut-on faire mieux **si f est connue? A suivre...**

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Lien entre Z - et M - estimateurs

- **Pas d'inclusion** entre ces deux classes d'estimateurs **en général** :
 - Si ψ non-régulière, M -estimateur \nRightarrow Z -estimateur
 - Si une équation d'estimation admet plusieurs solutions distinctes, Z -estimateur \nRightarrow M -estimateur (cas d'un extremum local).
- Toutefois, si ψ **est régulière**, les M -estimateurs **sont** des Z -estimateurs : si $\Theta \subset \mathbb{R}$ ($d = 1$), en posant

$$\phi(a, x) = \partial_a \psi(a, x),$$

on a

$$\sum_{i=1}^n \partial_a \psi(\vartheta, X_i) \Big|_{a=\hat{\vartheta}_n} = \sum_{i=1}^n \phi(\hat{\vartheta}_n, X_i) = 0.$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M -estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Maximum de vraisemblance

- Principe **fondamental** et **incontournable** en statistique. Cas particuliers connus depuis le XVIIIème siècle. Définition générale : Fisher (1922).
- Fournit une première **méthode systématique** de construction d'un M -estimateur (souvent un Z -estimateur, souvent aussi *a posteriori* un estimateur par substitution simple).
- Procédure **optimale** (dans quel sens ?) sous des hypothèses de **régularité** de la famille $\{\mathbb{P}_\vartheta, \vartheta \in \Theta\}$ (Cours 6).
- Parfois difficile à mettre en oeuvre en pratique \rightarrow **méthodes numériques**, statistique computationnelle.

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M -estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Fonction de vraisemblance

- La famille $\{\mathbb{P}_\vartheta, \vartheta \in \Theta\}$ est dominée par une mesure σ -finie μ . On se donne, pour $\vartheta \in \Theta$

$$f(\vartheta, x) = \frac{d\mathbb{P}_\vartheta}{d\mu}(x), \quad x \in \mathbb{R}.$$

Définition

Fonction de vraisemblance du n -échantillon associée à la famille $\{f(\vartheta, \cdot), \vartheta \in \Theta\}$:

$$\vartheta \rightsquigarrow \mathcal{L}_n(\vartheta, X_1, \dots, X_n) = \prod_{i=1}^n f(\vartheta, X_i)$$

- C'est une fonction aléatoire (définie μ -presque partout).

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M -estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Exemples

- Exemple 1 : **Modèle de Poisson**. On observe

$$X_1, \dots, X_n \sim_{\text{i.i.d.}} \text{Poisson}(\vartheta),$$

$\vartheta \in \Theta = \mathbb{R}_+ \setminus \{0\}$ et prenons $\mu(dx) = \sum_{k \in \mathbb{N}} \delta_k(dx)$.

- La densité de \mathbb{P}_ϑ par rapport à μ est

$$f(\vartheta, x) = e^{-\vartheta} \frac{\vartheta^x}{x!}, \quad x = 0, 1, 2, \dots$$

- La **fonction de vraisemblance** associée s'écrit

$$\begin{aligned} \vartheta \rightsquigarrow \mathcal{L}_n(\vartheta, X_1, \dots, X_n) &= \prod_{i=1}^n e^{-\vartheta} \frac{\vartheta^{X_i}}{X_i!} \\ &= \frac{1}{\prod_{i=1}^n X_i!} e^{-n\vartheta} \vartheta^{\sum_{i=1}^n X_i}. \end{aligned}$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Exemples

- Exemple 2 **Modèle de Cauchy**. On observe

$$X_1, \dots, X_n \sim_{\text{i.i.d.}} \text{Cauchy},$$

$$\vartheta \in \Theta = \mathbb{R} \text{ et } \mu(dx) = dx \text{ (par exemple).}$$

- On a alors

$$\mathbb{P}_{\vartheta}(dx) = f(\vartheta, x)dx = \frac{1}{\pi(1 + (x - \vartheta)^2)}dx.$$

- La **fonction de vraisemblance** associée s'écrit

$$\vartheta \rightsquigarrow \mathcal{L}_n(\vartheta, X_1, \dots, X_n) = \frac{1}{\pi^n} \prod_{i=1}^n (1 + (X_i - \vartheta)^2)^{-1}.$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Principe de maximum de vraisemblance

- Cas d'une famille de lois **restreinte à deux points**

$$\Theta = \{\vartheta_1, \vartheta_2\} \subset \mathbb{R},$$

avec \mathbb{P}_{ϑ_i} discrète et $\mu(dx)$ la mesure de comptage.

- **A priori**, pour tout (x_1, \dots, x_n) , et pour $\vartheta \in \{\vartheta_1, \vartheta_2\}$,

$$\begin{aligned}\mathbb{P}_{\vartheta} [X_1 = x_1, \dots, X_n = x_n] &= \prod_{i=1}^n \mathbb{P}_{\vartheta} [X_i = x_i] \\ &= \prod_{i=1}^n f(\vartheta, x_i).\end{aligned}$$

La probabilité d'avoir la réalisation fixée (x_1, \dots, x_n) .

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Principe de maximum de vraisemblance

- A posteriori, on observe (X_1, \dots, X_n) . L'événement

$$\left\{ \prod_{i=1}^n f(\vartheta_1, X_i) > \prod_{i=1}^n f(\vartheta_2, X_i) \right\} \quad (\text{Cas 1})$$

ou bien l'événement

$$\left\{ \prod_{i=1}^n f(\vartheta_2, X_i) > \prod_{i=1}^n f(\vartheta_1, X_i) \right\} \quad (\text{Cas 2})$$

est réalisé. (On ignore le cas d'égalité.)

- Principe de maximum de vraisemblance :

$$\hat{\vartheta}_n^{\text{mv}} = \vartheta_1 1_{\{\text{Cas 1}\}} + \vartheta_2 1_{\{\text{Cas 2}\}}.$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Estimateur du maximum de vraisemblance

- On généralise le principe précédent pour une famille de lois et un ensemble de paramètres **quelconques**.
- Situation : $X_1, \dots, X_n \sim_{\text{i.i.d.}} \mathbb{P}_\vartheta$, $\{\mathbb{P}_\vartheta, \vartheta \in \Theta\}$ dominée, $\Theta \subset \mathbb{R}^d$, $\vartheta \mapsto \mathcal{L}_n(\vartheta, X_1, \dots, X_n)$ vraisemblance associée.

Définition

On appelle **estimateur du maximum de vraisemblance** tout estimateur $\hat{\vartheta}_n^{\text{mv}}$ satisfaisant

$$\mathcal{L}_n(\hat{\vartheta}_n^{\text{mv}}, X_1, \dots, X_n) = \max_{\vartheta \in \Theta} \mathcal{L}_n(\vartheta, X_1, \dots, X_n).$$

- Existence, unicité...

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des Z- et M-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Remarques

- Log-vraisemblance :

$$\begin{aligned}\vartheta \rightsquigarrow \ell_n(\vartheta, X_1, \dots, X_n) &= \log \mathcal{L}_n(\vartheta, X_1, \dots, X_n) \\ &= \sum_{i=1}^n \log f(\vartheta, X_i).\end{aligned}$$

Bien défini si $f(\vartheta, \cdot) > 0$ μ -pp.

Max. vraisemblance = max. log-vraisemblance.

- L'estimateur du maximum de vraisemblance **ne dépend pas** du choix de la mesure dominante μ .
- Notion de **racine de l'équation de vraisemblance** : tout estimateur $\hat{\vartheta}_n^{\text{rv}}$ vérifiant

$$\nabla_{\vartheta} \ell_n(\hat{\vartheta}_n^{\text{rv}}, X_1, \dots, X_n) = 0.$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Exemple : modèle normal

L'expérience statistique est engendrée par un n -échantillon de loi $\mathcal{N}(\mu, \sigma^2)$, le paramètre est $\vartheta = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times \mathbb{R}_+ \setminus \{0\}$.

■ Vraisemblance

$$\mathcal{L}_n((\mu, \sigma^2), X_1, \dots, X_n) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - \mu)^2\right).$$

■ Log-vraisemblance

$$\ell_n((\mu, \sigma^2), X_1, \dots, X_n) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - \mu)^2.$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Exemple : modèle normal

Equation(s) de vraisemblance

$$\begin{cases} \partial_{\mu} \ell_n((\mu, \sigma^2), X_1, \dots, X_n) &= \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu) \\ \partial_{\sigma^2} \ell_n((\mu, \sigma^2), X_1, \dots, X_n) &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (X_i - \mu)^2. \end{cases}$$

Solution de ces équations (pour $n \geq 2$) :

$$\hat{\vartheta}_n^{\text{rv}} = \left(\bar{X}_n, \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \right)$$

et on vérifie que $\hat{\vartheta}_n^{\text{rv}} = \hat{\vartheta}_n^{\text{mv}}$.

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des Z- et M-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Exemple : modèle de Poisson

■ Vraisemblance

$$\mathcal{L}_n(\vartheta, X_1, \dots, X_n) = \frac{1}{\prod_{i=1}^n X_i!} e^{-n\vartheta} \vartheta^{\sum_{i=1}^n X_i}.$$

■ Log-vraisemblance

$$\ell_n(\vartheta, X_1, \dots, X_n) = c(X_1, \dots, X_n) - n\vartheta + \sum_{i=1}^n X_i \log \vartheta.$$

■ Equation de vraisemblance

$$-n + \sum_{i=1}^n X_i \frac{1}{\vartheta} = 0, \text{ soit } \boxed{\hat{\vartheta}_n^{\text{rv}} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}_n}$$

et on vérifie que $\hat{\vartheta}_n^{\text{rv}} = \hat{\vartheta}_n^{\text{mv}}$.

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Exemple : modèle de Laplace

L'expérience statistique est engendrée par un n -échantillon de loi de Laplace de paramètre $\vartheta \in \Theta = \mathbb{R}$. La densité par rapport à la mesure de Lebesgue :

$$f(\vartheta, x) = \frac{1}{2\sigma} \exp\left(-\frac{|x - \vartheta|}{\sigma}\right),$$

où $\sigma > 0$ est **connu**.

■ Vraisemblance

$$\mathcal{L}_n(\vartheta, X_1, \dots, X_n) = (2\sigma)^{-n} \exp\left(-\frac{1}{\sigma} \sum_{i=1}^n |X_i - \vartheta|\right)$$

■ Log-vraisemblance

$$\ell_n(\vartheta, X_1, \dots, X_n) = -n \log(2\sigma) - \frac{1}{\sigma} \sum_{i=1}^n |X_i - \vartheta|.$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Exemple : modèle de Laplace

Maximiser $\mathcal{L}_n(\vartheta, X_1, \dots, X_n)$ revient à minimiser la fonction $\vartheta \rightsquigarrow \sum_{i=1}^n |X_i - \vartheta|$, dérivable presque partout de dérivée constante par morceaux. **Equation de vraisemblance :**

$$\sum_{i=1}^n \text{sign}(X_i - \vartheta) = 0.$$

Soit $X_{(1)} \leq \dots \leq X_{(n)}$ la statistique d'ordre.

- n pair : $\hat{\vartheta}_n^{\text{mv}}$ **n'est pas unique** ; tout point de l'intervalle $[X_{(\frac{n}{2})}, X_{(\frac{n}{2}+1)}]$ est un EMV.
- n impair : $\hat{\vartheta}_n^{\text{mv}} = X_{(\frac{n+1}{2})}$, l'EMV est unique. Mais $\hat{\vartheta}_n^{\text{rv}}$ n'existe pas.
- **pour tout** n , la médiane empirique est un EMV.

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Exemple : modèle de Cauchy

■ Vraisemblance

$$\mathcal{L}_n(\vartheta, X_1, \dots, X_n) = \pi^{-n} \prod_{i=1}^n \frac{1}{1 + (X_i - \vartheta)^2}.$$

■ Log-vraisemblance

$$\ell_n(\vartheta, X_1, \dots, X_n) = -n \log \pi - \sum_{i=1}^n \log (1 + (X_i - \vartheta)^2)$$

■ Equation de vraisemblance

$$\sum_{i=1}^n \frac{X_i - \vartheta}{1 + (X_i - \vartheta)^2} = 0$$

pas de solution explicite et admet en général plusieurs solutions.

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Maximum de vraisemblance = M -estimateur

- Une inégalité de convexité : μ mesure σ -finie sur \mathbb{R} ; f, g deux **densités de probabilités** par rapport à μ . Alors

$$\int_{\mathbb{R}} f(x) \log f(x) \mu(dx) \geq \int_{\mathbb{R}} f(x) \log g(x) \mu(dx)$$

(si les intégrales sont finies) avec égalité **ssi** $f = g$ μ -pp.

- Preuve : à montrer

$$\int_{\mathbb{R}} f(x) \log \frac{g(x)}{f(x)} \mu(dx) \leq 0.$$

(avec une convention de notation appropriée)

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M -estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Une inégalité de convexité

- On a $\log(1+x) \leq x$ pour $x \geq -1$ avec égalité ssi $x = 0$.
- Donc

$$\log \frac{g(x)}{f(x)} = \log \left(1 + \left(\frac{g(x)}{f(x)} - 1 \right) \right) \leq \frac{g(x)}{f(x)} - 1$$

(avec égalité ssi $f(x) = g(x)$).

- Finalement

$$\begin{aligned} \int_{\mathbb{R}} f(x) \log \frac{g(x)}{f(x)} \mu(dx) &\leq \int_{\mathbb{R}} f(x) \left(\frac{g(x)}{f(x)} - 1 \right) \mu(dx) \\ &= \int_{\mathbb{R}} g(x) \mu(dx) - \int_{\mathbb{R}} f(x) \mu(dx) \\ &= 0. \end{aligned}$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Conséquence pour l'EMV

- On pose

$$\psi(a, x) := \log f(a, x), \quad a \in \Theta, \quad x \in \mathbb{R}$$

(avec une convention pour le cas où on n'a pas $f(a, \cdot) > 0$.)

- La fonction

$$a \rightsquigarrow \mathbb{E}_{\vartheta} [\psi(a, X)] = \int_{\mathbb{R}} \log f(a, x) f(\vartheta, x) \mu(dx)$$

a un maximum en $a = \vartheta$ d'après **l'inégalité de convexité**.

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Choix de modèle statistique

- Le statisticien a le choix de la famille $\{\mathbb{P}_\vartheta, \vartheta \in \Theta\}$. L'EMV dépend de ce choix.
- Exemple : on a l'échantillon ($n = 10$) :

$$\underbrace{0.92, -0.20, -1.80, 0.02, 0.49, 1.41, -1.59, -1.29, 0.34, 100.}_{\text{tirage de } \mathcal{N}(0,1)}$$

- On prend $\mathbb{P}_\vartheta(dx) = f(x - \vartheta)dx$ pour deux f différents :
- f densité de la loi normale $\Rightarrow \hat{\vartheta}_n^{\text{mv}} = \bar{X}_n = 9.83$.
- f densité de loi de Laplace \Rightarrow tout point de l'intervalle $[0.02, 0.34]$ est un $\hat{\vartheta}_n^{\text{mv}}$, en particulier, la médiane :

$$\hat{\vartheta}_n^{\text{mv}} = M_n = (0.02 + 0.34)/2 = 0.18.$$

- Autre choix de modèle...

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Un M -estimateur qui n'est pas un Z -estimateur

- On observe $X_1, \dots, X_n \sim_{\text{i.i.d.}}$ uniformes sur $[0, \vartheta]$, $\vartheta \in \Theta = \mathbb{R}_+ \setminus \{0\}$.

- On a

$$\mathbb{P}_{\vartheta}(dx) = \vartheta^{-1} 1_{[0, \vartheta]}(x) dx$$

et

$$\begin{aligned} \mathcal{L}_n(\vartheta, X_1, \dots, X_n) &= \vartheta^{-n} \prod_{i=1}^n 1_{[0, \vartheta]}(X_i) \\ &= \vartheta^{-n} 1_{\{\max_{1 \leq i \leq n} X_i \leq \vartheta\}} \end{aligned}$$

- La fonction de vraisemblance **n'est pas régulière**.
- **L'estimateur du maximum de vraisemblance est**
 $\hat{\vartheta}_n^{\text{mv}} = \max_{1 \leq i \leq n} X_i$.

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

Principe de
maximum de
vraisemblance

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Asymptotique des Z - et M -estimateurs

- Problème général **délicat**. Dans ce cours : conditions suffisantes.
- **Convergence** : critère simple pour les M -estimateurs.
- **Vitesse de convergence** : technique simple pour les Z -estimateurs, à condition de savoir que l'estimateur est convergent.
- Sous des hypothèses de régularité, un M -estimateur est un Z -estimateur.

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M -estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Convergence des M -estimateurs

- Situation : on observe X_1, \dots, X_n i.i.d. de loi dans la famille $\{\mathbb{P}_\vartheta, \vartheta \in \Theta\}$.
- $\psi : \Theta \times \mathbb{R} \rightarrow \mathbb{R}$ **fonction de contraste**.
- **Loi des grands nombres** :

$$M_n(a) = \frac{1}{n} \sum_{i=1}^n \psi(a, X_i)$$

converge en \mathbb{P}_ϑ -probabilité vers

$$M(a, \vartheta) = \mathbb{E}_\vartheta [\psi(a, X)]$$

qui atteint son maximum en $a = \vartheta$

- « à montrer » :

$$\hat{\vartheta}_n = \arg \max_{a \in \Theta} M_n(a) \xrightarrow{\mathbb{P}_\vartheta} \arg \max_{a \in \Theta} \mathbb{E}_\vartheta [\psi(a, X)] = \vartheta.$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M -estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Convergence des M -estimateurs

Proposition

Si le M -estimateur $\hat{\vartheta}_n$ associé à la fonction de contraste est bien défini et si

- $\sup_{a \in \Theta} |M_n(a) - M(a, \vartheta)| \xrightarrow{\mathbb{P}_\vartheta} 0,$
- $\forall \varepsilon > 0, \sup_{|a - \vartheta| \geq \varepsilon} M(a, \vartheta) < M(\vartheta, \vartheta)$ (*condition de maximum*)

alors $\hat{\vartheta}_n \xrightarrow{\mathbb{P}_\vartheta} \vartheta.$

- La condition 1 (convergence uniforme) peut être délicate à montrer...

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M -estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Loi limite des Z -estimateurs

- Situation : on observe X_1, \dots, X_n i.i.d. de loi dans la famille $\{\mathbb{P}_\vartheta, \vartheta \in \Theta\}$, $\Theta \subset \mathbb{R}$.
- $\hat{\vartheta}_n$: Z -estimateur associé à $\phi : \Theta \times \mathbb{R} \rightarrow \mathbb{R}$ vérifie

$$\sum_{i=1}^n \phi(\hat{\vartheta}_n, X_i) = 0$$

- Si $\hat{\vartheta}_n$ est un M -estimateur associé à la fonction de contraste ψ régulière, alors c'est un Z -estimateur associé à la fonction $\phi(a, x) = \partial_a \psi(a, x)$.
- On suppose $\hat{\vartheta}_n$ convergent. Que dire de sa loi limite ?

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M -estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Loi limite des Z -estimateurs : principe

■ Loi des grands nombres

$$Z_n(a) = \frac{1}{n} \sum_{i=1}^n \phi(a, X_i) \xrightarrow{\mathbb{P}_\vartheta} Z(a, \vartheta) = \mathbb{E}_\vartheta [\phi(a, X)]$$

■ Principe. Développement de Taylor autour de ϑ :

$$0 = Z_n(\hat{\vartheta}_n) = Z_n(\vartheta) + (\hat{\vartheta}_n - \vartheta) Z'_n(\vartheta) + \frac{1}{2} (\hat{\vartheta}_n - \vartheta)^2 Z''(\tilde{\vartheta}_n).$$

■ On **néglige** le reste :

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta) \approx \frac{-\sqrt{n}Z_n(\vartheta)}{Z'_n(\vartheta)}$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Loi limite des Z-estimateurs : principe

- Convergence du **numérateur**

$$\sqrt{n}Z_n(\vartheta) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \phi(\vartheta, X_i) \xrightarrow{d} \mathcal{N}(0, \mathbb{E}_{\vartheta} [\phi(\vartheta, X)^2])$$

si $\mathbb{E}_{\vartheta} [\phi(\vartheta, X)] = 0$ et $\mathbb{E}_{\vartheta} [\phi(\vartheta, X)^2] < +\infty$.

- Convergence du **dénominateur**

$$Z'_n(\vartheta) = \frac{1}{n} \sum_{i=1}^n \partial_{\vartheta} \phi(\vartheta, X_i) \xrightarrow{\mathbb{P}_{\vartheta}} \mathbb{E}_{\vartheta} [\partial_{\vartheta} \phi(\vartheta, X)]$$

$\neq 0$ (à supposer).

- + hypothèses techniques pour **contrôler le reste** (besoin de la convergence de $\hat{\vartheta}_n$).

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z- et M-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Loi limite des Z-estimateurs

Proposition (Convergence des Z-estimateurs)

- Soit Θ un ouvert de \mathbb{R} . Pour tout $\vartheta \in \Theta$, $\hat{\vartheta}_n \xrightarrow{\mathbb{P}_\vartheta} \vartheta$,
 $\mathbb{E}_\vartheta [\phi(\vartheta, X)^2] < +\infty$ et

$$\mathbb{E}_\vartheta [\phi(\vartheta, X)] = 0, \quad \mathbb{E}_\vartheta [\partial_\vartheta \phi(\vartheta, X)] \neq 0.$$

- (*Contrôle reste*) pour tout $\vartheta \in \Theta$, pour tout a dans un voisinage de ϑ ,

$$|\partial_a^2 \phi(a, x)| \leq g(x), \quad \mathbb{E}_\vartheta [g(X)] < +\infty.$$

Alors

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta) \xrightarrow{d} \mathcal{N}\left(0, \frac{\mathbb{E}_\vartheta[\phi(\vartheta, X)^2]}{(\mathbb{E}_\vartheta[\partial_\vartheta \phi(\vartheta, X)])^2}\right).$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z- et M-
estimateurs

Méthode
d'estimation
dans le modèle
de régression



Influence d'une variable sur une autre

- Principe : on part de l'observation d'un n -échantillon

$$Y_1, \dots, Y_n \quad (Y_i \in \mathbb{R})$$

- A chaque observation Y_i est associée une observation auxiliaire $\mathbf{X}_i \in \mathbb{R}^k$.
- On suspecte l'échantillon

$$\mathbf{X}_1, \dots, \mathbf{X}_n \quad (\mathbf{X}_i \in \mathbb{R}^k)$$

de contenir la « majeure partie de la variabilité des Y_i ».

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
« design »

Régression à
design
déterministe

La droite des
moindres carrés

Régression
linéaire multiple

Modélisation de l'influence

- Si \mathbf{X}_i contient toute la variabilité de Y_i , alors Y_i est mesurable par rapport à \mathbf{X}_i : il existe $r : \mathbb{R}^k \rightarrow \mathbb{R}$ telle que

$$Y_i = r(\mathbf{X}_i),$$

mais peu réaliste (ou alors problème d'interpolation numérique).

- Alternative : représentation précédente avec erreur additive : on postule

$$Y_i = r(\mathbf{X}_i) + \xi_i,$$

ξ_i erreur aléatoire centrée (pour des raisons d'identifiabilité).

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
<< design >>

Régression à
design
déterministe

La droite des
moindres carrés

Régression
linéaire multiple

Motivation : meilleure approximation L^2

- Meilleure approximation L^2 . Si $\mathbb{E}[Y^2] < +\infty$, la meilleure approximation de Y par une variable aléatoire \mathbf{X} -mesurable est donnée par **l'espérance conditionnelle** $\mathbb{E}[Y|\mathbf{X}]$:

$$\mathbb{E}[(Y - r(\mathbf{X}))^2] = \min_h \mathbb{E}[(Y - h(\mathbf{X}))^2]$$

- où

$$r(\mathbf{x}) = \mathbb{E}[Y|\mathbf{X} = \mathbf{x}], \quad \mathbf{x} \in \mathbb{R}^k.$$

- On appelle $r(\cdot)$ **fonction de régression de Y sur \mathbf{X}** .

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
« design »

Régression à
design
déterministe

La droite des
moindres carrés

Régression
linéaire multiple

Régression

- On définit :

$$\xi = Y - \mathbb{E}[Y | \mathbf{X}] \implies \mathbb{E}[\xi] = 0.$$

- On a alors naturellement la représentation désirée

$$Y = r(\mathbf{X}) + \xi, \quad \mathbb{E}[\xi] = 0$$

si l'on pose

$$r(\mathbf{x}) = \mathbb{E}[Y | \mathbf{X} = \mathbf{x}], \quad \mathbf{x} \in \mathbb{R}^k$$

- On observe alors un n -échantillon

$$(\mathbf{X}_1, Y_1), \dots, (\mathbf{X}_n, Y_n)$$

où

$$Y_i = r(\mathbf{X}_i) + \xi_i, \quad \mathbb{E}[\xi_i] = 0$$

avec comme paramètre la fonction $r(\cdot)$ + un jeu d'hypothèses sur la loi des ξ_i .

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z - et M -
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
<< design >>

Régression à
design
déterministe

La droite des
moindres carrés

Régression
linéaire multiple

Modèle de régression à design aléatoire

Définition

Modèle de régression à **design aléatoire** = donnée de l'observation

$$(\mathbf{X}_1, Y_1), \dots, (\mathbf{X}_n, Y_n)$$

avec $(Y_i, \mathbf{X}_i) \in \mathbb{R} \times \mathbb{R}^k$ *i.i.d.*, et

$$Y_i = r(\vartheta, \mathbf{X}_i) + \xi_i, \quad \mathbb{E}[\xi_i | \mathbf{X}_i] = 0, \quad \vartheta \in \Theta \subset \mathbb{R}^d.$$

- $\mathbf{x} \rightsquigarrow r(\vartheta, \mathbf{x})$ fonction de *régression*, connue au paramètre ϑ près.
- \mathbf{X}_i = variables explicatives, co-variables, prédicteurs ;
 $(\mathbf{X}_1, \dots, \mathbf{X}_n) = \text{design}$.

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z- et M-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
« design »

Régression à
design
déterministe

La droite des
moindres carrés

Régression
linéaire multiple

Modèle alternatif : signal+bruit

- Principe : **sur un exemple**. On observe

$$Y_i = r(\vartheta, i/n) + \xi_i, \quad i = 1, \dots, n$$

où $r(\vartheta, \cdot) : [0, 1] \rightarrow \mathbb{R}$ est une fonction connue au paramètre $\vartheta \in \Theta \subset \mathbb{R}^d$ près, et les ξ_i sont i.i.d., $\mathbb{E}[\xi_i] = 0$.

- **But** : reconstruire $r(\vartheta, \cdot)$ c'est-à-dire **estimer ϑ** .
- Plus généralement, on observe

$$Y_i = r(\vartheta, \mathbf{x}_i) + \xi_i, \quad i = 1, \dots, n$$

où $\mathbf{x}_1, \dots, \mathbf{x}_n$ sont des points de \mathbb{R}^k **déterministes**.

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z- et M-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
<< design >>

Régression à
design
déterministe

La droite des
moindres carrés

Régression
linéaire multiple

Modèle de régression à design déterministe

Définition

Modèle de régression à design déterministe = donnée de l'observation

$$(\mathbf{x}_1, Y_1), \dots, (\mathbf{x}_n, Y_n)$$

avec $Y_i \in \mathbb{R}$, $\mathbf{x}_i \in \mathbb{R}^k$, et

$$Y_i = r(\vartheta, \mathbf{x}_i) + \xi_i, \quad \mathbb{E}[\xi_i] = 0, \quad \vartheta \in \Theta \subset \mathbb{R}^d.$$

- \mathbf{x}_i déterministes, donnés (ou choisis) : plan d'expérience, points du « design ».
- Hypothèses sur les ξ_i : à débattre. *Pour simplifier*, les ξ_i sont i.i.d. (*hypothèse restrictive*).
- \implies les Y_i ne sont *pas* i.i.d.

Question : Comment estimer θ dans ce modèle ?

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z- et M-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
« design »

Régression à
design
déterministe

La droite des
moindres carrés

Régression
linéaire multiple

Régression gaussienne

- Modèle de régression à design déterministe :

$$Y_i = r(\vartheta, \mathbf{x}_i) + \xi_i, \quad \vartheta \in \Theta \subset \mathbb{R}^d.$$

- Supposons : $\xi_i \sim \mathcal{N}(0, \sigma^2)$, i.i.d.
- On a alors le modèle de **régression gaussienne**. Comment estimer ϑ ? **On sait explicier la loi de l'observation**
 $Z = (Y_1, \dots, Y_n) \implies$ appliquer le principe du maximum de vraisemblance.
- La loi de Y_i :

$$\mathbb{P}^{Y_i}(dy) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y - r(\vartheta, \mathbf{x}_i))^2\right) dy \\ \ll dy.$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
<< design >>

Régression à
design
déterministe

La droite des
moindres carrés

Régression
linéaire multiple

EMV pour régression gaussienne

- Le modèle $\{\mathbb{P}_\vartheta^n = \text{loi de } (Y_1, \dots, Y_n), \vartheta \in \mathbb{R}^k\}$ est dominé par $\mu^n(dy_1 \dots dy_n) = dy_1 \dots dy_n$.

- D'où

$$\begin{aligned}\frac{d\mathbb{P}_\vartheta^n}{d\mu^n}(y_1, \dots, y_n) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_i - r(\vartheta, \mathbf{x}_i))^2\right) \\ &= \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - r(\vartheta, \mathbf{x}_i))^2\right).\end{aligned}$$

- La fonction de vraisemblance

$$\mathcal{L}_n(\vartheta, Y_1, \dots, Y_n) \propto \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - r(\vartheta, \mathbf{x}_i))^2\right)$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z- et M-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
<< design >>

Régression à
design
déterministe

La droite des
moindres carrés

Régression
linéaire multiple

Estimateur des moindres carrés

Maximiser la **vraisemblance** en régression gaussienne = minimiser la somme des carrés :

$$\sum_{i=1}^n (Y_i - r(\vartheta, \mathbf{x}_i))^2 \rightarrow \min_{\vartheta \in \Theta}.$$

Définition

*Estimateur des **moindres carrés** : tout estimateur $\hat{\vartheta}_n^{\text{mc}}$ t.q.*
$$\hat{\vartheta}_n^{\text{mc}} \in \arg \min_{\vartheta \in \Theta} \sum_{i=1}^n (Y_i - r(\vartheta, \mathbf{x}_i))^2.$$

- L'EMC est un M-estimateur. Pour le modèle de régression gaussienne : EMV = EMC.
- **Existence, unicité.**
- Propriétés remarquables si la régression est linéaire :
$$r(\vartheta, \mathbf{x}_i) = \vartheta^T \mathbf{x}_i.$$

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des Z- et M-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
<< design >>

Régression à
design
déterministe

La droite des
moindres carrés
Régression
linéaire multiple

Droite de régression

- Modèle le plus simple $r(\vartheta, x) = a + bx$

$$Y_i = a + bx_i + \xi_i, \quad i = 1, \dots, n$$

avec $\vartheta = (a, b)^T \in \Theta = \mathbb{R}^2$ et les (x_1, \dots, x_n) donnés.

- L'estimateur des moindres carrés :

$$\hat{\vartheta}_n^{\text{mc}} = (\hat{a}, \hat{b}) = \arg \min_{(a,b) \in \mathbb{R}^2} \sum_{i=1}^n (Y_i - a - bx_i)^2.$$

- **Solution explicite** existe toujours, sauf cas pathologique quand tous les x_i sont les mêmes (Poly, page 112).

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
<< design >>

Régression à
design
déterministe

La droite des
moindres carrés

Régression
linéaire multiple

Régression linéaire simple

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
<< design >>

Régression à
design
déterministe

**La droite des
moindres carrés**

Régression
linéaire multiple

Régression linéaire simple

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
<< design >>

Régression à
design
déterministe

**La droite des
moindres carrés**

Régression
linéaire multiple

Régression linéaire multiple (=Modèle linéaire)

- La fonction de régression est $r(\vartheta, \mathbf{x}_i) = \vartheta^T \mathbf{x}_i$. On observe

$$(\mathbf{x}_1, Y_1), \dots, (\mathbf{x}_n, Y_n)$$

avec

$$Y_i = \vartheta^T \mathbf{x}_i + \xi_i, \quad i = 1, \dots, n$$

où $\vartheta \in \Theta = \mathbb{R}^k$, $\mathbf{x}_i \in \mathbb{R}^k$.

- Matriciellement

$$\mathbf{Y} = \mathbf{M}\vartheta + \boldsymbol{\xi}$$

avec $\mathbf{Y} = (Y_1, \dots, Y_n)^T$, $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)^T$ et \mathbf{M} la matrice $(n \times k)$ dont les **lignes** sont les \mathbf{x}_i .

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
<< design >>

Régression à
design
déterministe

La droite des
moindres carrés

Régression
linéaire multiple

EMC en régression linéaire multiple

- Estimateur des **moindres carrés** en régression linéaire multiple : tout estimateur $\hat{\vartheta}_n^{\text{mc}}$ satisfaisant

$$\sum_{i=1}^n (Y_i - (\hat{\vartheta}_n^{\text{mc}})^T \mathbf{x}_i)^2 = \min_{\vartheta \in \mathbb{R}^k} \sum_{i=1}^n (Y_i - \vartheta^T \mathbf{x}_i)^2.$$

- En notations matricielles :

$$\begin{aligned} \|\mathbf{Y} - \mathbb{M} \hat{\vartheta}_n^{\text{mc}}\|^2 &= \min_{\vartheta \in \mathbb{R}^k} \|\mathbf{Y} - \mathbb{M} \vartheta\|^2 \\ &= \min_{v \in V} \|\mathbf{Y} - v\|^2 \end{aligned}$$

où $V = \text{Im}(\mathbb{M}) = \{v \in \mathbb{R}^n : v = \mathbb{M} \vartheta, \vartheta \in \mathbb{R}^k\}$.

Projection orthogonale sur V .

MAP 433 :
Introduction
aux méthodes
statistiques.
Cours 4

M-estimation,
rappel du
Cours 3

EMV,
asymptotique
des *Z*- et *M*-
estimateurs

Méthode
d'estimation
dans le modèle
de régression

Modèle de
régression,
notion de
<< design >>

Régression à
design
déterministe

La droite des
moindres carrés

Régression
linéaire multiple