Signaux Aléatoires

Estimation Bayésienne

Jean-Yves Tourneret[1]

(1) University of Toulouse, ENSEEIHT-IRIT-TéSA
jyt@n7.fr

Année $2023 - 2024$

Bibliographie

Quelques références

- A. Papoulis and S. U. Pillai, Probability, Random Variable and Stochastic Processes, McGraw Hill Higher Education, 4th edition, 2002.
- S. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice Hall, 1997.

Rappels sur l'estimation statistique

Notations

- Observations

$$x_1, ..., x_n$$

- Échantillon

$$X_1, ..., X_n$$

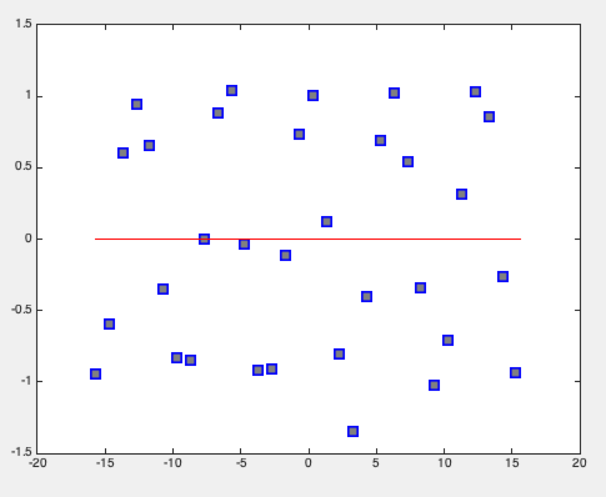
n va iid associées aux observations

- Estimateur

$$\widehat{\theta}(X_1, ..., X_n) \text{ ou } \widehat{\theta}_n \text{ ou } \widehat{\theta}$$

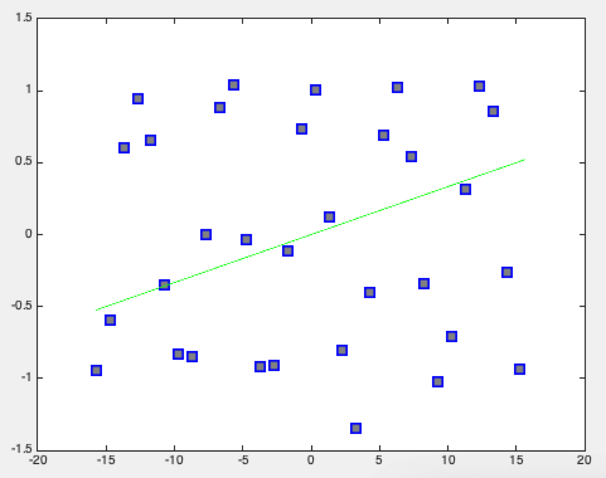
Exemple

Modèle 1 : $x_i = a + e_i$ avec $e_i \sim \mathcal{N}(0, \sigma^2)$



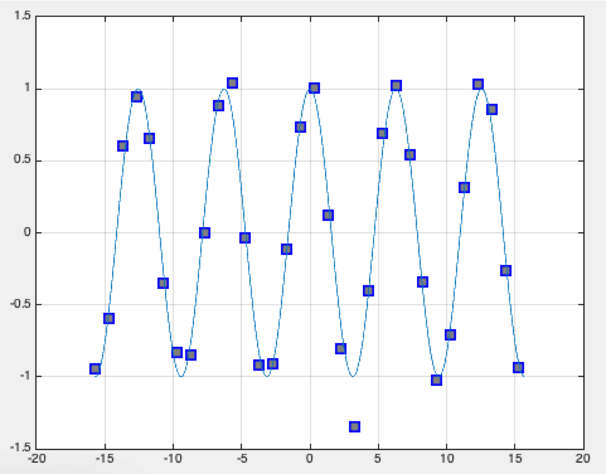
Exemple

Modèle 2 : $x_i = ai + b + e_i$ avec $e_i \sim \mathcal{N}(0, \sigma^2)$



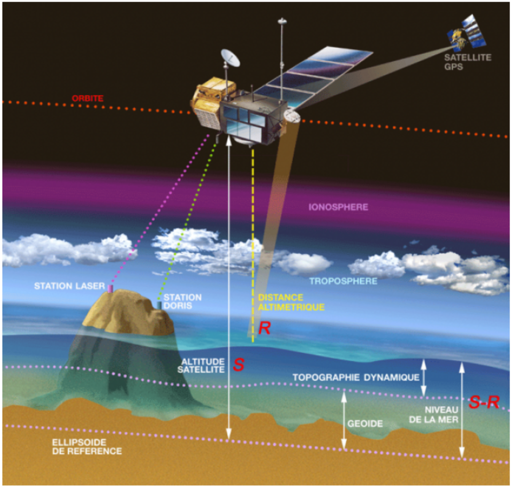
Exemple

Modèle 3 : $x_i = a\cos(i\phi) + e_i$ avec $e_i \sim \mathcal{N}(0, \sigma^2)$



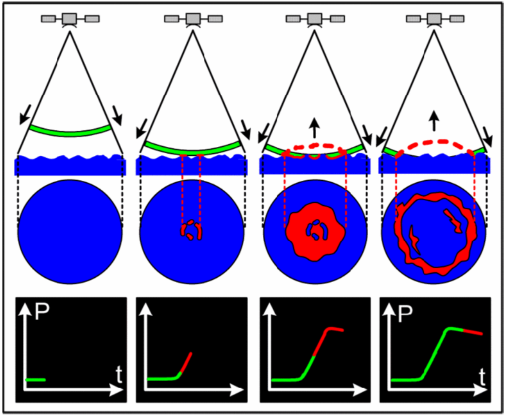
Application réelle

Altimétrie



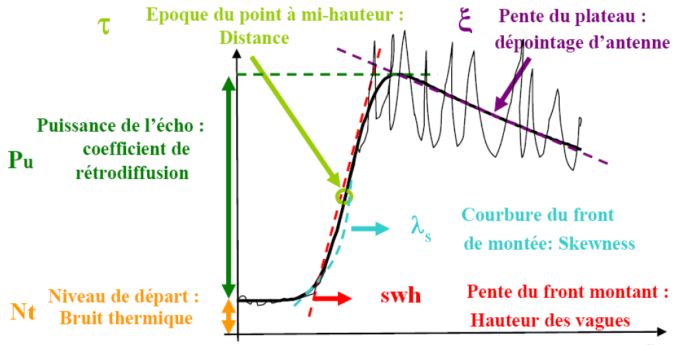
Application réelle

Formation de l'écho altimétrique



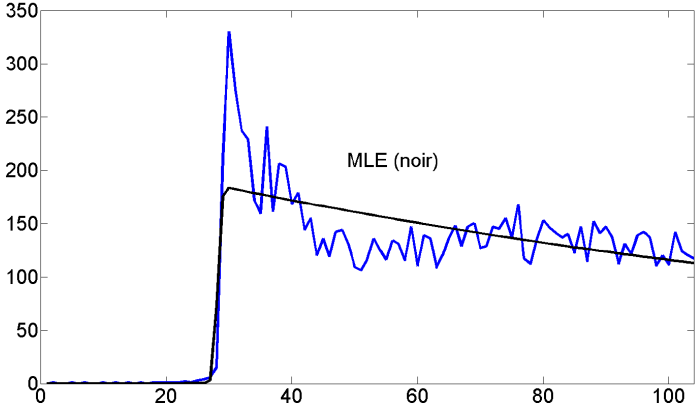
Application réelle

Modèle de Brown



Application réelle

Modèle de Brown



Estimation Bayésienne

Hypothèses

Soit $(X_1, ..., X_n)$ un échantillon (n variables aléatoires indépendantes et de même loi) dont la loi dépend d'un vecteur paramètre inconnu $\boldsymbol{\theta} \in \mathbb{R}^p$. On cherche à estimer $\boldsymbol{\theta}$ à partir d'observations $(x_1, ..., x_n)$ et à partir d'une information a priori sur le paramètre $\boldsymbol{\theta}$. L'approche Bayésienne suppose que l'on connaît les quantités suivantes :

- la densité a priori du paramètre $\boldsymbol{\theta}$ notée $p(\boldsymbol{\theta})$ qui résume l'information dont on dispose sur le paramètre $\boldsymbol{\theta}$,
- la loi de la variable aléatoire X_i conditionnelle au vecteur paramètre $\boldsymbol{\theta}$, i.e., la densité de X_i sachant $\boldsymbol{\theta}$ notée $p(x_i|\boldsymbol{\theta})$ (cas continu) ou la probabilité d'avoir x_i sachant $\boldsymbol{\theta}$ notée $P[X_i = x_i|\boldsymbol{\theta}]$ (cas discret).

Problème

Comment estimer le vecteur paramètre $\boldsymbol{\theta}$ à l'aide des observations $(x_1, ..., x_n)$?

Estimateur du maximum de vraisemblance

- Vraisemblance

$$L(x_1, ..., x_n; \theta) = \begin{cases} X_i \text{ va discrète : } P[X_1 = x_1, ..., X_n = x_n; \theta] \\ X_i \text{ va continue : } p(x_1, ..., x_n; \theta) \end{cases}$$

- Estimateur du maximum de vraisemblance

$$\widehat{\boldsymbol{\theta}}_{\text{MV}} = \arg\max_{\boldsymbol{\theta}} L(X_1, ..., X_n; \boldsymbol{\theta})$$

- Exemple : $X_i \sim \mathcal{N}(m, \sigma^2)$ avec $\boldsymbol{\theta} = (m, \sigma^2)^T$.

Estimateur du maximum a posteriori (MAP)

- Définition

$$\widehat{\boldsymbol{\theta}}_{\text{MAP}} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|X_1, ..., X_n)$$

avec

$$p(\boldsymbol{\theta}|x_1, ..., x_n) = \frac{p(x_1, ..., x_n|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(x_1, ..., x_n)},$$

i.e.,

$$\widehat{\boldsymbol{\theta}}_{\text{MAP}} = \arg\max_{\boldsymbol{\theta}} \{p(x_1, ..., x_n|\boldsymbol{\theta})p(\boldsymbol{\theta})\}$$

- Exemple
  - Vraisemblance

$$X_i \sim \mathcal{N}(\theta, \sigma^2)$$

  - Loi a priori

$$\theta \sim \mathcal{N}(\mu, \nu^2)$$

  - Estimateur MAP

$$\widehat{\theta}_{\text{MAP}} = \overline{X}\left(\frac{n\nu^2}{n\nu^2 + \sigma^2}\right) + \mu\left(\frac{\sigma^2}{\sigma^2 + n\nu^2}\right)$$

Estimateur MMSE (minimum mean square error)

- Définition

$$\widehat{\boldsymbol{\theta}}_{\mathsf{MMSE}} = \arg\max_{\boldsymbol{\theta}} E\left[\left(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\right)^T \left(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\right)\right] = E[\boldsymbol{\theta}|X_1, ..., X_n]$$

- Exemple
  - Vraisemblance

$$X_i \sim \mathcal{N}(\theta, \sigma^2)$$

  - Loi a priori

$$\theta \sim \mathcal{N}(\mu, \nu^2)$$

  - Estimateur MMSE

$$\widehat{\theta}_{\mathrm{MMSE}} = \overline{X}\left(\frac{n\nu^2}{n\nu^2 + \sigma^2}\right) + \mu\left(\frac{\sigma^2}{\sigma^2 + n\nu^2}\right)$$

Estimateur MMSE

Preuve de $\widehat{\boldsymbol{\theta}}_{\text{MMSE}} = E[\boldsymbol{\theta}|X_1, ..., X_n]$ dans le cas où $\boldsymbol{\theta} \in \mathbb{R}$

$$\begin{aligned} E[(\widehat{\theta} - \theta)^2] &= E[E[(\widehat{\theta} - \theta)^2|\boldsymbol{x}]] \\ &= \int E[(\widehat{\theta} - \theta)^2|\boldsymbol{x}]p(\boldsymbol{x})d\boldsymbol{x}, \ \boldsymbol{x} = (x_1, ..., x_n)^T \end{aligned}$$

Puisque $p(\boldsymbol{x}) \geq 0$, l'estimateur $\widehat{\theta}(\boldsymbol{x})$ qui minimise $E[(\widehat{\theta} - \theta)^2]$ minimise $E[(\widehat{\theta} - \theta)^2|\boldsymbol{x}]$ pour tout $\boldsymbol{x}$. Mais

$$\begin{aligned} E[(\widehat{\theta} - \theta)^2|\boldsymbol{x}] &= E[\theta^2|\boldsymbol{x}] - 2\widehat{\theta}(\boldsymbol{x})E[\boldsymbol{\theta}|\boldsymbol{x}] + \widehat{\theta}^{\,2}(\boldsymbol{x}) \\ &= \mathsf{var}[\theta|\boldsymbol{x}] + \left\{\widehat{\theta}(\boldsymbol{x}) - E[\boldsymbol{\theta}|\boldsymbol{x}]\right\}^2 \\ &\geq \mathsf{var}[\theta|\boldsymbol{x}] \end{aligned}$$

avec égalité lorsque

$$\widehat{\theta}(\boldsymbol{x}) = E[\boldsymbol{\theta}|\boldsymbol{x}]$$

Que faut-il savoir ?

- Signification de l'estimateur du maximum a posteriori d'un vecteur paramètre inconnu $\boldsymbol{\theta}$
- Calculer l'estimateur du maximum a posteriori d'un vecteur paramètre inconnu $\boldsymbol{\theta}$
- Signification de l'estimateur MMSE d'un vecteur paramètre inconnu $\boldsymbol{\theta}$
- Calculer l'estimateur MMSE d'un vecteur paramètre inconnu $\boldsymbol{\theta}$