Chapitre 3

Méthodes d'Estimation

1. INTRODUCTION

Dans le chapitre II, nous avons étudié la théorie de la détection où une version bruitée du signal est reçue par le récepteur. Dans le cas de la détection M-aires, une décision peut être prise parmi M-hypothèses possibles. Néanmoins, un récepteur binaire doit choisir entre l'hypothèse nulle H_0 ou l'hypothèse alternative H_1.

Dans ce chapitre, nous supposons que le récepteur a pris une décision en faveur de H_1. Cependant, certains paramètres du signal peuvent être inconnus. Auquel cas, le but est d'estimer ces paramètres de manière optimale et ce seulement à partir de la connaissance d'un nombre fini d'échantillons du signal.

Pour ce faire, soient $Y_1, Y_2, \ldots, Y_K$, *K* échantillons indépendants et identiquement distribués (Independent and Indentically Distributed , IID) issus d'une même variable aléatoire **Y**, dont la fonction de densité de probabilité dépend d'un paramètre inconnu θ. Soient $y_1, y_2, \ldots, y_K$ les valeurs correspondantes des échantillons $Y_1, Y_2, \ldots, Y_K$ et $g(Y_1, Y_2, \ldots, Y_K)$ une fonction (une statistique) des échantillons utilisés pour estimer le paramètre θ. Nous appelons:

$$\hat{\theta} = g(Y_1, Y_2, \ldots, Y_K) \tag{1}$$

l'*estimateur* de θ. La valeur que prend la statistique est appelée estimation de θ et est égale à $\hat{\theta} = g(y_1, y_2, \ldots, y_K)$. Pour éviter toute confusion entre une variable aléatoire et la valeur que celle-ci peut avoir, notons que $\hat{\theta}$, l'estimation de θ, est actuellement $g(Y_1, Y_2, \ldots, Y_K)$. Conséquemment, quand nous parlons de la moyenne de $\hat{\theta}$, $E[\hat{\theta}]$, nous nous référons à $E[g(Y_1, Y_2, \ldots, Y_K)]$.

Le paramètre à estimer peut ou ne pas être aléatoire. L'estimation des paramètres aléatoires est connue sous le nom de l'estimation de Bayes (Bayes Estimator), alors que celle des paramètres non aléatoires, est connue sous le nom de l'estimation du maximum de vraisemblance (Maximum Likelihood Estimator, MLE).

2. ESTIMATEUR DU MAXIMUM DE VRAISENBLANCE

Soient $Y_1, Y_2, \ldots, Y_K$, *K* observations de la variable aléatoire ***Y***, ayant pour valeurs des échantillons $y_1, y_2, \ldots, y_K$. Ces variables aléatoires sont IID. $f_{Y|\Theta}(y \mid \theta)$ dénote la densité de probabilité conditionnelle de la variable aléatoire ***Y***. Notons que la densité de probabilité de ***Y*** dépend du paramètre θ, $\theta \in \Theta$, lequel a besoin d'être estimé. La fonction de vraisemblance (likelihood function), $L(\theta)$, est donnée par:

$$L(\theta) = f_{Y|\Theta}(y\,|\,\theta) = f_{Y_1,\dots,Y_K|\Theta}(y_1, y_2, \dots, y_K\,|\,\theta) = \prod_{k=1}^{K} f_{Y_k|\Theta}(y_k\,|\,\theta) \tag{2}$$

La valeur de $\hat{\theta}$ qui maximise la fonction de vraisemblance est dite estimateur du maximum de vraisemblance de θ. Comme la fonction logarithme $\ln x$ est monotone et croissante de x, la maximisation $L(\theta)$ est équivalente à la maximisation de $\ln L(\theta)$. Par conséquent, il peut être montré qu'une condition nécessaire mais non suffisante pour obtenir le MLE $\hat{\theta}$ est la résolution de l'équation de vraisemblance (likelihood equation*)* suivante:

$$\frac{\partial}{\partial\theta} \ln f_{\mathbf{Y}|\Theta}(\mathbf{y}\,|\,\theta) = 0 \tag{3}$$

***Propriété d'invariance*.** Soient $L(\theta)$ la fonction de vraisemblance de θ et $g(\theta)$ une fonction bijective de θ ; Autrement dit, si $g(\theta_1) = g(\theta_2) \Leftrightarrow \theta_1 = \theta_2$. Si $\hat{\theta}$ est le MLE de θ, alors $g(\hat{\theta})$ est le MLE de $g(\theta)$.

<u>Exemple 1</u>:

Dans l'exemple 2 du Chapitre II, le signal reçu sous les hypothèses H_1 et H_0 était:

$$\begin{aligned} H_1 : Y_k &= m + N_k\,,\ \ k = 1, 2, \dots, K \\ H_0 : Y_k &= \phantom{m + {}} N_k\,,\ \ k = 1, 2, \dots, K \end{aligned}$$

(a) En supposant que la constante *m* ne soit pas connue, obtenir le MLE $\hat{m}_{ml}$ de la moyenne $\theta = m$.

(b) Supposer maintenant que la moyenne *m* soit connue, mais la variance σ^2 est inconnue. Obtenir le MLE $\hat{\theta}_{ml}$ de $\theta = \sigma^2$.

<u>Réponses</u>:

(a) Le paramètre $\hat{\theta}$ à déterminer est $\hat{m}_{ml}$, où $m \in M$. Comme les échantillons sont IID, la fonction de vraisemblance, (Cf. équation (2)), est:

$$f_{\mathbf{Y}|M}(\mathbf{y}\,|\,m) = \prod_{k=1}^{K} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y_k - m)^2}{2\sigma^2}\right] = \frac{1}{(2\pi)^{K/2}\sigma^K} \exp\left[-\sum_{k=1}^{K} \frac{(y_k - m)^2}{2\sigma^2}\right]$$

En prenant le logarithme des deux côtés, nous obtenons:

$$\ln f_{\mathbf{Y}|M}(\mathbf{y}\,|\,m) = \ln\left[\frac{1}{(2\pi)^{K/2}\sigma^K}\right] - \sum_{k=1}^{K} \frac{(y_k - m)^2}{2\sigma^2}$$

Le MLE est obtenu en égalant à zéro, la dérivée du logarithme de la fonction de vraisemblance comme cela est montré en (3). Ainsi,

$$\frac{\partial \ln f_{\mathbf{Y}|M}(\mathbf{y}\,|\,m)}{\partial m} = \sum_{k=1}^{K} \frac{y_k - m}{\sigma^2} = \sum_{k=1}^{K} \frac{y_k}{\sigma^2} - \frac{Km}{\sigma^2} = \frac{K}{\sigma^2}\left(\frac{1}{K}\sum_{k=1}^{K} y_k - m\right) = 0$$

ou $m=(1/k)\sum_{k=1}^{K} y_k$. Le MLE de m est $\hat{m}_{ml}=(1/k)\sum_{k=1}^{K} Y_k$.

(b) La fonction de vraisemblance est:

$$L(\sigma^2)=\frac{1}{(2\pi)^{\frac{K}{2}}\sigma^K}\exp\left[-\sum_{k=1}^{K}\frac{(y_k-m)^2}{2\sigma^2}\right]$$

En prenant le logarithme, nous obtenons:

$$\ln L(\sigma^2)=-\frac{K}{2}\ln 2\pi-K\ln\sigma-\sum_{k=1}^{K}\frac{(y_k-m)^2}{2\sigma^2}$$

Notons que la maximisation de $\ln L(\sigma^2)$ par rapport à σ^2 est équivalente à celle de:

$$g(\sigma^2)=K\ln\sigma+\sum_{k=1}^{K}\frac{(y_k-m)^2}{2\sigma^2}$$

En utilisant la propriété d'invariance, il est plus simple de différencier $g(\sigma^2)$ par rapport à σ pour obtenir $\hat{\sigma}_{ml}$ le MLE de σ, au lieu de $\hat{\sigma}^2_{ml}$ le MLE de σ^2. Par suite:

$$\frac{dg(\sigma^2)}{d\sigma}=\frac{K}{\sigma}-\sum_{k=1}^{K}\frac{(y_k-m)^2}{\sigma^3}=0 \quad \text{ou} \quad \hat{\sigma}=\sqrt{\frac{1}{K}\sum_{k=1}^{K}(y_k-m)^2}$$

Le MLE de σ^2 est donc $\hat{\sigma}^2_{ml}=(1/K)\sum_{k=1}^{K}(Y_k-m)^2$.

3. CRITERES D'UN BON ESTIMATEUR

Comme l'estimateur $\hat{\theta}$ est une variable aléatoire qui peut prendre plus d'une valeur, nous donnons ci-après quelques critères qui permettent de savoir si un estimateur est 'bon'.

Estimateur non biaisée: $\hat{\theta}$ est un estimateur non biaisé θ, si

$$E[\hat{\theta}]=\theta \tag{4}$$

Biais d'un estimateur: Le biais d'un estimateur $\hat{\theta}$ est donné par:

$$E[\hat{\theta}]=\theta+b(\theta) \tag{5}$$

1. Si $b(\theta)$ ne dépend pas de $\theta[b(\theta)=b]$, alors l'estimateur $\hat{\theta}$ a un biais connu. C'est à dire, $(\hat{\theta}-b)$ est un estimateur non biaisé.

2. Si $b(\theta)\neq b$, comme θ est inconnu, alors un estimateur non biaisé ne peut être obtenu. Dans ce cas, nous disons que l'estimateur a un biais inconnu.

Le fait qu'un estimateur soit non biaisé (sa valeur moyenne est proche de la valeur réelle), ne garantit pas nécessairement qu'il soit 'bon'. Ceci peut être facilement vérifié (Cf. Figure 4) à travers un exemple de densité de probabilité *a priori* de θ. En effet, malgré le fait que l'estimateur soit non biaisé, comme sa variance est grande, de grandes erreurs sont vraisemblables. Cependant, si la variance était petite, la variabilité de l'estimateur autour de sa moyenne serait petite aussi. Par conséquent, comme l'estimateur est non biaisé, le fait que sa variabilité soit proche de la vraie valeur constitue une caractéristique désirable. En définitive, nous disons que la seconde mesure de qualité d'un estimateur est celle d'avoir une petite variance.

Estimateur non biaisé à variance minimale: $\hat{\theta}$ est dit estimateur non biaisé à variance minimale de θ, si, pour tous les estimateurs θ' tel que $E[\theta'] = \theta$, nous avons $\mathrm{var}[\hat{\theta}] \leq \mathrm{var}[\theta'] \quad \forall \theta'$. Autrement dit, $\hat{\theta}$ a la plus petite variance parmi toutes celles des estimateurs non biaisés de θ.

Estimateur consistant **:** $\hat{\theta}$ est un estimateur consistant de θ, basé sur *K* échantillons observés, si

$$\lim_{K\to\infty} P\left(\left|\hat{\theta}-\theta\right|>\varepsilon\right)=0 \quad \forall\varepsilon>0 \tag{6}$$

où $P(\cdot)$ dénote la probabilité.

La vérification de la consistance d'un estimateur à travers l'application de la définition ci-dessus, n'est pas aussi simple que cela puisse paraitre. A cet effet, nous pouvons utiliser le théorème suivant:

Théorème: Soit $\hat{\theta}$ un estimateur de θ basé sur *K* échantillons observés. Si

$$\lim_{K\to\infty} E[\hat{\theta}] = \theta \tag{7}$$

et si

$$\lim_{K\to\infty} \mathrm{var}\, E[\hat{\theta}] = 0 \tag{8}$$

alors $\hat{\theta}$ est un estimateur consistant de θ.

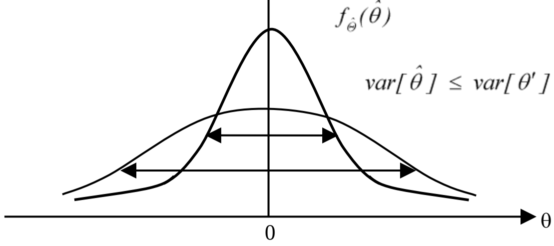

Figure 1. Fonction de densité de probabilité de l'estimateur $\hat{\theta}$.

Exemple 2:

(a) Vérifier si l'estimateur $\hat{m}_{ml}$ de l'exemple 1 est un estimateur non biaisé de *m* ?

(b) Vérifier si l'estimateur $\hat{\sigma}^2_{ml}$ de l'exemple 1 est un estimateur non biaisé de σ^2 ?

Réponses:

(a) L'estimateur $\hat{m}_{ml}$ est non biaisé si $E[\hat{m}_{ml}] = m$. Après substitution, nous obtenons:

$$E[\hat{m}_{ml}] = E\left[\frac{1}{K}\sum_{k=1}^{K} Y_K\right] = \frac{1}{K}E\left[\sum_{k=1}^{K} Y_K\right] = \frac{1}{K}Km = m$$

Donc, $\hat{m}_{ml}$ est non biaisé.

(b) L'estimateur $\hat{\sigma}^2_{ml}$ est non biaisé si $E[\hat{\sigma}^2_{ml}] = \sigma^2$. C'est à dire,

$$E\left[\frac{1}{K}\sum_{k=1}^{K} (Y_k - m)^2\right] = \frac{1}{K}E\left[Km^2 + \sum_{k=1}^{K} Y_k^2 - 2m\sum_{k=1}^{K} Y_k\right] = \sigma^2$$

Donc, $\hat{\sigma}^2_{ml}$ est non biaisé.

4. INEGALITE DE CRAMER-RAO

Pour vérifier si l'estimateur ML est 'bon', nous devons calculer son biais et la variance de l'erreur. Il est possible qu'il soit difficile d'obtenir une expression de la variance de l'erreur. Dans ce cas, pour savoir si l'estimateur est 'bon', nous devons l'étudier en termes de limite inférieure de la variance de l'erreur. Cette limite est connue sous le nom de *Cramer-Rao bound*. Le Cramer-Rao bound d'un paramètre constant est donné par le théorème suivant:

Théorème. Soit le vecteur $\boldsymbol{Y} = [Y_1, Y_2, ..., Y_K]^T$ représentant *K* observations et $\hat{\theta}$ un estimateur non biaisé de θ, alors:

$$\mathrm{var}[(\hat{\theta} - \theta) \mid \theta] \geq \frac{1}{E\left\{\left[\frac{\partial \ln f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta)}{\partial \theta}\right]^2\right\}} \tag{9}$$

où

$$E\left\{\left[\frac{\partial \ln f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta)}{\partial \theta}\right]^2\right\} = -E\left[\frac{\partial^2 \ln f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta)}{\partial \theta^2}\right] \tag{10}$$

Démonstration : Pour un estimateur non biaisé $\hat{\theta}$, nous avons:

$$E[\hat{\theta} \mid \theta] = \theta \tag{11}$$

Par conséquent:

$$E[(\hat{\theta} - \theta) \mid \theta] = \int_{-\infty}^{\infty} (\hat{\theta} - \theta) f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta) d\boldsymbol{y} = 0 \tag{12}$$

En différenciant (12) par rapport à θ, nous obtenons:

$$\int_{-\infty}^{\infty} (\hat{\theta} - \theta) \frac{\partial f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta)}{\partial \theta} d\boldsymbol{y} = \int_{-\infty}^{\infty} f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta) d\boldsymbol{y} \tag{13}$$

La deuxième intégrale est égale à l'unité. En utilisant le fait que:

$$\frac{\partial \ln g(x)}{\partial x} = \frac{1}{g(x)} \frac{\partial g(x)}{\partial x} \tag{14}$$

où $g(x)$ est une fonction de *x*, nous pouvons exprimer $\partial f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta) / \partial \theta$ comme suit:

$$\frac{\partial f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta)}{\partial \theta} = f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta) \frac{\partial \ln f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta)}{\partial \theta} \tag{15}$$

En substituant (15) dans (13), nous obtenons:

$$\int_{-\infty}^{\infty} (\hat{\theta} - \theta) f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta) \frac{\partial \ln f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta)}{\partial \theta} d\boldsymbol{y} = 1 \tag{16}$$

L'inégalité de Schwarz stipule que:

$$\left[\int_{-\infty}^{\infty} x^2(t) dt \right] \left[\int_{-\infty}^{\infty} y^2(t) dt \right] \geq \left[\int_{-\infty}^{\infty} x(t) y(t) dt \right]^2 \tag{17}$$

où $x(t)\ et\ y(t)$ sont deux fonctions de *t*. L'égalité est vérifiée si et seulement si $y(t) = cx(t)$, avec *c* une constante. Pour pouvoir utiliser l'inégalité de Schwarz inégalité, nous réécrivons (16):

$$\int_{-\infty}^{\infty} \left[\frac{\partial \ln f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta)}{\partial \theta} \sqrt{f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta)} \right] [\, (\hat{\theta} - \theta) \sqrt{f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y} \mid \theta)} \,] d\boldsymbol{y} = 1 \tag{18}$$

Ou alors:

$$\left[\int_{-\infty}^{\infty} (\hat{\theta} - \theta)^2 f_{\mathrm{Y}|\Theta}(\mathrm{y} \mid \theta) dy \right] \left\{ \int_{-\infty}^{\infty} \left[\frac{\partial \ln f_{\mathrm{Y}|\Theta}(\mathrm{y} \mid \theta)}{\partial \theta} \right]^2 f_{\mathrm{Y}|\Theta}(\mathrm{y} \mid \theta) d\mathrm{y} \right\} \geq 1 \tag{19}$$

La première intégrale entre crochets est actuellement $\mathrm{var}[(\hat{\theta} - \theta) \mid \theta]$. Par conséquent, l'inégalité devient:

$$\mathrm{var}[(\hat{\theta}-\theta)\,|\,\theta] \geq \frac{1}{E\left\{\left[\dfrac{\partial \ln f_{Y|\Theta}(y\,|\,\theta)}{\partial\theta}\right]^2\right\}} \tag{20}$$

Ceci prouve (9).

Maintenant, nous prouvons (10), laquelle exprime le Cramer-Rao Bound sous une forme différente. Nous savons que:

$$\int_{-\infty}^{\infty} f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y}\,|\,\theta)d\boldsymbol{y} = 1 \tag{21}$$

En différenciant les deux parties de l'équation par rapport à θ, donne lieu à:

$$\int_{-\infty}^{\infty} \frac{\partial f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y}\,|\,\theta)}{\partial\theta} d\boldsymbol{y} = 0 \tag{22}$$

En réécrivant (22) et en utilisant (14), nous avons:

$$\int_{-\infty}^{\infty} \frac{\partial \ln f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y}\,|\,\theta)}{\partial\theta} f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y}\,|\,\theta)d\boldsymbol{y} = 0 \tag{23}$$

En différenciant une nouvelle fois par rapport à θ, nous obtenons:

$$\int_{-\infty}^{\infty} \frac{\partial^2 \ln f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y}\,|\,\theta)}{\partial\theta^2} f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y}\,|\,\theta)d\boldsymbol{y} + \int_{-\infty}^{\infty} \frac{\partial \ln f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y}\,|\,\theta)}{\partial\theta} \frac{\partial f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y}\,|\,\theta)}{\partial\theta} = 0 \tag{24}$$

En substituant (15) pour le second terme de la seconde intégrale de (24), et en réarrangeant les termes, implique:

$$E\left[\frac{\partial^2 \ln f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y}\,|\,\theta)}{\partial\theta^2}\right] = -E\left\{\left[\frac{\partial \ln f_{\boldsymbol{Y}|\Theta}(\boldsymbol{y}\,|\,\theta)}{\partial\theta}\right]^2\right\} \tag{25}$$

laquelle est identique à (10).

L'égalité dans (9) est vérifiée si et seulement si:

$$\frac{\partial \ln f_{\boldsymbol{Y}|\boldsymbol{\Theta}}(\boldsymbol{y}\,|\,\theta)}{\partial\theta} = c(\theta)[\hat{\theta}-\theta] \tag{26}$$

N'importe quel estimateur satisfaisant l'égalité dans l'inégalité de Cramer-Rao de (9) est dit estimateur efficient.

Si un estimateur efficient existe, il est facile de montrer qu'il correspond à l'estimateur ML. L'équation ML est donnée par:

$$\left.\frac{\partial \ln f_{\mathbf{Y}|\Theta}(\boldsymbol{y}\,|\,\theta)}{\partial\theta}\right|_{\theta=\hat{\theta}_{ml}} = 0 \tag{27}$$

En utilisant (26), si un estimateur efficient existe, nous avons:

$$\left.\frac{\partial \ln f_{\mathrm{Y}|\Theta}(\mathrm{y}\,|\,\theta)}{\partial\theta}\right|_{\theta=\hat{\theta}_{ml}} = c(\theta)[\hat{\theta}-\theta]\Big|_{\theta=\hat{\theta}_{ml}} = 0 \tag{28}$$

lequel est égal à zéro pour $\hat{\theta}=\hat{\theta}_{ml}$.

<u>Exemple 3</u>:

Considérer *K* observations, tel que:

$Y_k = m + N_k, \quad k = 1, 2, \ldots, K$

où *m* est inconnu et les N_ks sont des variables aléatoires Gaussiennes statistiquement indépendantes de moyenne nulle et de variance σ^2.

(a) Trouver les estimateurs $\hat{m}$ *et* $\hat{\sigma}^2$ pour m *et* σ^2, respectivement.

(b) est-ce que $\hat{m}$ est un estimateur efficient?

(c) Trouver la variance conditionnelle du vecteur $\mathrm{var}[(\hat{m}-m)\,|\,m]$.

<u>Réponses</u>:

(a) En utilisant (2), nous pouvons déterminer simultanément $\hat{m}$ *et* $\hat{\sigma}^2$. La fonction densité conditionnelle de ***Y*** sachant m *et* σ^2 est:

$$f_{\mathbf{Y}}(\boldsymbol{y}\,|\,m,\sigma^2) = \prod_{k=1}^{K}\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(y_k-m)^2}{2\sigma^2}\right]$$

En prenant le logarithme, nous obtenons:

$$\ln f_{\mathbf{Y}}(\boldsymbol{y}\,|\,m,\sigma^2) = -\frac{K}{2}\ln\left(2\pi\sigma^2\right) - \sum_{k=1}^{K}\frac{(y_k-m)^2}{2\sigma^2}$$

Nous prenons la dérivée de l'équation ci-dessus par rapport à m *et* σ^2 pour obtenir deux équations à deux inconnus:

$$\frac{\partial \ln f_{\mathbf{Y}}(\boldsymbol{y}\,|\,m,\sigma^2)}{\partial m} = 2\sum_{k=1}^{K}\frac{y_k-m}{2\sigma^2} = 0 \quad \text{et} \quad \frac{\partial \ln f_{\mathbf{Y}}(\boldsymbol{y}\,|\,m,\sigma^2)}{\partial \sigma^2} = -\frac{K}{2\sigma^2} + \sum_{k=1}^{K}\frac{(y_k-m)^2}{2\sigma^4} = 0$$

Résolvant simultanément pour $\hat{m}_{ml}$ *et* $\hat{\sigma}^2_{ml}$, nous obtenons

$$\hat{m}_{ml} = \frac{1}{K}\sum_{k=1}^{K} y_k \text{ et } \hat{\sigma}_{ml}^2 = \frac{1}{K}\sum_{k=1}^{K}\left(y_k - \frac{1}{K}\sum_{k=1}^{K} y_k\right)^2 = \frac{1}{K}\sum_{k=1}^{K}(y_k - \hat{m}_{ml})^2$$

(b) $\hat{m}_{ml}$ est un estimateur non biaisé car:

$$E[\hat{m}_{ml}] = \frac{1}{K} E\left[\sum_{k=1}^{K} y_k\right] = m$$

Pour vérifier si l'estimateur est efficient, nous utilisons (50) pour obtenir:

$$\frac{\partial \ln f_{\mathbf{Y}}(\boldsymbol{y} \mid m, \sigma^2)}{\partial m} = \sum_{k=1}^{K} \frac{y_k - m}{\sigma^2} = \frac{K}{\sigma^2}\left(\frac{1}{K}\sum_{k=1}^{K} y_k - m\right)$$

où $c(m) = K/\sigma^2 \; et \; \hat{m} = (1/K)\sum_{k=1}^{K} y_k = \hat{m}_{ml}$. Donc, l'estimateur est efficient.

(c) Pour déterminer la variance conditionnelle de l'erreur, nous utilisons (9) et (10). En prenant la dérivée de l'équation de vraisemblance par rapport à *m*, nous obtenons:

$$\frac{\partial^2 \ln f_{\mathbf{Y}}(\boldsymbol{y} \mid m, \sigma^2)}{\partial m^2} = -\frac{K}{\sigma^2}$$

Donc,

$$\mathrm{var}[(\hat{m} - m) \mid m] = -\frac{1}{E\left[\dfrac{\partial^2 \ln f_{\mathrm{Y}}(\mathrm{y} \mid m, \sigma^2)}{\partial m^2}\right]} = \frac{\sigma^2}{K}$$

5. BRUIT BLANC GAUSSIEN

Soit Y une VA à valeurs réelles (X∈R). Soit $\{y_1, y_2, \dots, y_n, \dots, y_N\}$ un ensemble de N réalisations de la VA Y. Autrement dit, les y_n (n=1, 2, …, N) constituent N répliques indépendantes, générées successivement, de la VA Y. Nous pouvons y associer un temps discret, i.e., N instants $\{t_1, t_2, \dots, t_n, \dots, t_N\}$, et écrire $y_n = y(t_n)$. Nous supposerons pour simplifier que le pas de temps $\Delta t = t_{n+1} - t_n$ est constant.

La séquence $\{y(t_1), y(t_2), \dots, y(t_n), \dots, y(t_N)\}$ peut être vue comme une réalisation d'un processus temporel « purement aléatoire ». Nous nous intéressons ici de plus au cas où la loi de probabilité de la VA (génératrice du processus) est une loi gaussienne de moyenne nulle. Toute séquence finie constitue alors un vecteur gaussien.

Le processus aléatoire $Y(t_n)$ ainsi défini peut être considéré comme un exemple très particulier de fonction aléatoire (gaussienne et purement aléatoire) en temps discret.

Le bruit blanc gaussien proprement dit, noté Y(t)ou f(t), est un processus purement aléatoire gaussien en temps continu, que nous obtenons, intuitivement, comme limite du processus purement aléatoire gaussien défini plus haut lorsque le pas de temps $\Delta t \to 0$.

Le bruit blanc peut être interprété comme un mélange en égales proportions de bruits aléatoires de fréquences diverses, depuis les fréquences nulles jusqu''aux fréquences infinies (d'où le nom de bruit « blanc », comme la lumière « blanche ».

Définition et moments du bruit blanc stationnaire gaussien

Un bruit blanc gaussien f(t) est une fonction aléatoire en temps continu (t∈R ou R*) possédant les propriétés suivantes:

- ✓ f(t) est gaussienne,
- ✓ f(t) est de moyenne nulle,
- ✓ f(t) est stationnaire à tous les ordres, i.e., au sens strict,
- ✓ f(t) a une fonction d'auto-covariance proportionnelle à la « fonction delta » (Dirac).

Par conséquent,

$$f_T(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu_f)^2}{2\sigma^2}} \quad \text{où } \mu_f = \langle f(t) \rangle = 0 \tag{29}$$

et

$$C_{ff}(\tau) = \langle f(t+\tau) f(t) \rangle = C_0 \, \delta(\tau) \tag{30}$$

Comme la variance $Var\{f(t)\} = C_{ff}(0)$, ceci implique que la variance d'un bruit blanc diverge, i.e., tend vers l'infini. Le bruit blanc parfait n'est donc pas physiquement réalisable, mais il est réalisable par contre en version lissée, régularisée ou tronquée.

Nous appelons C_0 l'intensité du bruit blanc. Elle a l'unité de $[f^2(t)]$. Par exemple, si $f(t)$ représente un processus de déplacement $X(t)$, avec X en mètres et t en secondes, alors C_0 est en m^2s.

Il est intéressant d'exprimer l'intensité $C_0 = \sigma^2 T$.

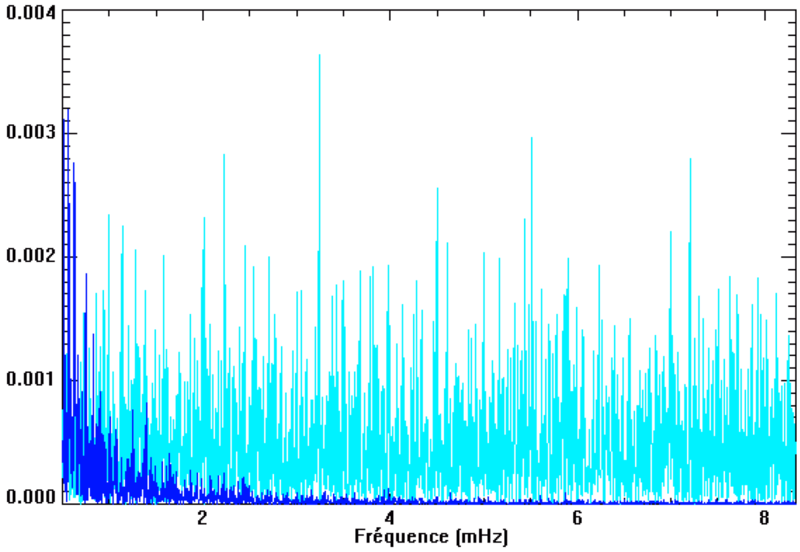

Figure 2. Les spectres de puissance de différents types de bruits : bruit instrumental en bleu foncé, bruit gaussien en bleu clair. Le spectre du bruit gaussien ne présente pas de fréquence privilégiée, contrairement au bruit instrumental.

Régularisation temporelle du bruit blanc (lissage)

Pour obtenir une version physiquement réalisable du bruit blanc, nous proposons de construire un bruit blanc « lissé » *f*(t)* à partir du bruit blanc théorique *f(t)* défini plus haut. Nous définissons *f*(t)* comme la moyenne glissante de *f(t)*sur une durée *T*, soit:

$$f^*(t) = \frac{1}{2T} \int_{t-T/2}^{t+T/2} f(s)ds \tag{31}$$

Le nouveau processus aléatoire f*(t) est aussi de moyenne nulle :

$$\langle f^*(t) \rangle = \frac{1}{T} \int_{t-T/2}^{t+T/2} \langle f(s) \rangle ds = 0 \tag{32}$$

puisque $\langle f(t) \rangle = 0$.

Sa fonction d'autocorrélation en (t, t+τ) est:

$$C_{f^*f^*}(\tau) = \langle f^*(t+\tau) f^*(t) \rangle = \frac{1}{T^2} \int_{t+\tau-T/2}^{t+\tau+T/2} ds' \int_{t-T/2}^{t+T/2} \langle f(s') f(s'') \rangle ds'' \tag{33}$$

En insérant $C_{ff}(s', s'') = C_0\, \delta(s' - s'')$ *dans l'intégrale, ceci donne* :

Donc, une auto-covariance stationnaire, de forme triangulaire:

$$C_{f^*f^*}(\tau) = \frac{C_0}{T}\, Max(0; 1 - \frac{|\tau|}{T}) \tag{34}$$

ou encore

$$C_{f^*f^*}(\tau) = \sigma^2 \; Max(0; 1 - \frac{|\tau|}{T}) \tag{35}$$

avec $\sigma^2 = Var\{f^*(t)\} = \frac{C_0}{T}$

En résumé, nous voyons bien que l'intensité C_0 du bruit blanc théorique peut s'interpréter comme le produit de la variance d'un bruit blanc « moyenné » (effectivement réalisable et observable) et du temps caractéristiaue de prise de moyenne T (assimilable à la résolution temporelle effectivement utilisée lors de l'observation).

Notons que la fonction de covariance triangulaire tend bien vers un Dirac quand $T \to 0$. Notons aussi que le produit $C_0 = \sigma^2 T$ ets un invariant, qui ne dépend pas de la résolution T choisie pour observer le bruit blanc.

Régularisation spectrale du bruit blanc (troncature en fréquence)

Nous obtenons un résultat à peu près équivalent au précédent en éliminant du spectre des fluctuations de *f(t)* celles des fréquences supérieures à un seuil ω_{Max} que nous pouvons ajuster (filtre passe-bas). Nous poserons donc $\omega_{Max=\Pi/T}$ pour comparer avec l'approche précédente.