

Quantification asymétrique optimale pour l'estimation d'un paramètre de centrage dans un bruit de loi symétrique

Rodrigo CABRAL FARIAS, Jean-Marc BROSSIER

GIPSA-lab, Grenoble INP, 11, rue des Mathématiques, BP 46, 38402 Saint Martin d'Hères, France
{rodrigo.cabral-farias,jean-marc.brossier}@gipsa-lab.grenoble-inp.fr

Résumé — Nous traitons de l'estimation d'un paramètre de centrage à partir d'observations bruitées quantifiées sur deux niveaux. Le comportement de la BCR (Borne de Cramér-Rao) est étudié en fonction du centrage du quantifieur pour différentes distributions symétriques de bruit. Nous montrons que, contrairement à ce qui est mentionné dans la littérature, l'emplacement optimal du centrage du quantifieur dépend radicalement de la loi de bruit et que son emplacement sur le paramètre de centrage, un choix intuitif vu la symétrie du problème, peut donner la performance d'estimation localement la plus mauvaise.

Abstract — In this paper, estimation of a location parameter based on noisy and binary quantized measurements is considered. The behavior of the Cramér-Rao bound as a function of the quantizer threshold is studied for different symmetric noise distributions. Differently from the literature, it is shown that the optimal threshold position depends strongly on the noise distribution and that placing the threshold exactly at the location parameter value, an intuitive choice given the symmetry of the problem, leads to locally worst estimation performance.

1 Introduction

La réduction des coûts des capteurs et des dispositifs de communication a permis l'émergence d'un nouveau domaine du traitement du signal, les réseaux de capteurs. Dans ce nouveau domaine, une grande quantité de capteurs est utilisée pour réaliser l'estimation d'une ou plusieurs quantités d'intérêt (ex. : température, humidité) ou la détection de certains événements (ex. : incendie, rupture d'un barrage). En raison de la grande quantité de capteurs, qui normalement communiquent sans fil, la bande de communication est fortement contrainte. Pour respecter cette contrainte, on peut considérer que les mesures sont quantifiées de façon très grossière, par exemple en utilisant seulement deux niveaux de quantification.

Motivé par ce contexte, dans ce travail, on étudie le problème d'estimation à partir de mesures quantifiées sur deux niveaux. Ce problème peut être détaillé de la façon suivante : pour $k = 1, \dots, N$, on réalise les mesures quantifiées $i_k = Q(Y_k)$ avec $Y_k = x + V_k$ où x est le paramètre réel constant à estimer, V_k le bruit (réel, centré, i.i.d.) et Q un quantifieur binaire tel que $Q(v) = \text{sign}(v - \tau_0) \in \{-1, +1\}$.

Le but principal de ce travail est la caractérisation de la performance d'estimation pour différentes distributions du bruit en fonction du centrage τ_0 du quantifieur. Dans la suite on note F la fonction de répartition du bruit et f sa densité de probabilité (d.d.p.). On suppose $f(v)$ strictement positive, paire et décroissante pour $v > 0$. Cette forme très générale englobe la loi normale, la loi de Laplace

et bien d'autres lois comme les gaussiennes généralisées.

2 Analyse de la BCR

Pour l'analyse de la performance optimale d'estimation, on utilisera une borne inférieure sur la variance des estimateurs non biaisés, la BCR. On sait que sous certaines contraintes de régularité, on peut atteindre la BCR asymptotiquement $N \rightarrow \infty$ avec l'estimateur du MV (Maximum de Vraisemblance).

L'indépendance entre les N observations binaires $i_{1:N} = \{i_j\}_{j=1 \dots N}$ donne la BCR suivante

$$\text{BCR} = N^{-1}B,$$

où B est l'inverse de l'information de Fisher I pour une unique mesure quantifiée. On note \mathbb{E} l'espérance mathématique et $\mathbb{P}(i = +1; x) = 1 - F(\tau_0 - x)$, $\mathbb{P}(i = -1; x) = F(\tau_0 - x)$ les probabilités des mesures quantifiées $i = +1$ et $i = -1$, B est donnée par

$$\begin{aligned} B &= \frac{1}{I} = \frac{1}{\mathbb{E} \left\{ \left[\frac{\partial \log \mathbb{P}(i; x)}{\partial x} \right]^2 \right\}} \\ &= \left\{ [1 - F(\tau_0 - x)] \left\{ \frac{\partial \log [1 - F(\tau_0 - x)]}{\partial x} \right\}^2 \right. \\ &\quad \left. + F(\tau_0 - x) \left[\frac{\partial \log F(\tau_0 - x)}{\partial x} \right]^2 \right\}^{-1} \\ &= \frac{F(\tau_0 - x) [1 - F(\tau_0 - x)]}{f^2(\tau_0 - x)}. \end{aligned}$$

On constate que la BCR est une fonction de l'écart entre le seuil du quantifieur et le paramètre à estimer $\varepsilon = \tau_0 - x$

$$\text{BCR}(\varepsilon) = N^{-1}B(\varepsilon) \text{ avec } B(\varepsilon) = \frac{F(\varepsilon)[1 - F(\varepsilon)]}{f^2(\varepsilon)}. \quad (1)$$

Dans ce qui suit, on étudie le comportement de la fonction $B(\varepsilon)$ pour différentes lois de bruit symétriques et unimodales.

2.1 La BCR est minimale en $\varepsilon = 0$ pour un bruit de Gauss ou de Laplace

L'étude du cas gaussien, $f(\varepsilon) = [\delta\sqrt{\pi}]^{-1} \exp[-\varepsilon^2/\delta^2]$, réalisée dans [2, 3], révèle que le choix $\varepsilon = 0$ (seuil sur le paramètre) minimise la BCR et que pour ce choix la quantification accroît la BCR d'un facteur $\pi/2$ seulement.

Pour un bruit distribué selon une loi de Laplace $f(\varepsilon) = \frac{1}{2\delta} \exp(-|\frac{\varepsilon}{\delta}|)$ on obtient la fonction suivante :

$$B(\varepsilon) = \delta^2 \left[2 \exp\left(\left|\frac{\varepsilon}{\delta}\right|\right) - 1 \right],$$

ce qui conduit à un résultat du même type : la position optimale du seuil du quantifieur est donnée par la valeur du paramètre x . De plus, dans ce cas, on peut montrer que la BCR pour les observations quantifiées avec $\tau_0 = x$ est égale à la BCR pour les observations à valeurs continues ; par conséquent, asymptotiquement, la quantification binaire optimale n'entraîne aucune perte de performance d'estimation.

Alors que l'intuition pourrait conduire à penser que ces résultats s'étendent aux lois symétriques, nous montrons que non seulement cela n'est pas le cas, mais qu'il arrive que ce choix soit localement le plus mauvais.

2.2 Une condition pour que la BCR soit minimale en $\varepsilon = 0$

Comment se comporte en général $B(\varepsilon)$ au voisinage de $\varepsilon = 0$?

Pour des lois unimodales symétriques $f(\varepsilon) = f(-\varepsilon)$ et $F(\varepsilon) = 1 - F(-\varepsilon)$, ainsi $B(\varepsilon)$ reste paire. On note $f^{(n)}$ la dérivée d'ordre n . On a :

$$\frac{dB}{d\varepsilon} = \frac{f^2(\varepsilon)[1 - 2F(\varepsilon)] - 2F(\varepsilon)[1 - F(\varepsilon)]f^{(1)}(\varepsilon)}{f^3(\varepsilon)},$$

$1 - 2F(0)$ étant nul, seul le second terme du numérateur peut être non nul en $\varepsilon = 0$. Si $f^{(1)}(0) = 0$, $\varepsilon = 0$ est localement un extremum, éventuellement un maximum. Pour le savoir, calculons la dérivée seconde :

$$\begin{aligned} \frac{d^2B}{d\varepsilon^2} = & -2 + \frac{1}{f^4(\varepsilon)} \times \left\{ -3f^2(\varepsilon)f^{(1)}(\varepsilon)[1 - 2F(\varepsilon)] + \right. \\ & \left. + F(\varepsilon)[1 - F(\varepsilon)] \left[6f^{(1)2}(\varepsilon) - 2f(\varepsilon)f^{(2)}(\varepsilon) \right] \right\}. \end{aligned}$$

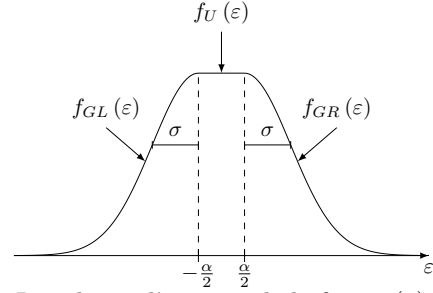


FIGURE 1 – Loi plate à l'origine, de la forme (3), qui viole la condition (2).

Ainsi $\frac{d^2B}{d\varepsilon^2} \Big|_{\varepsilon=0} = -\frac{1}{2} \frac{f^{(2)}(0)}{f^3(0)} - 2$ et une condition pour que $\varepsilon = 0$ soit un minimum local de $B(\varepsilon)$ est :

$$-f^{(2)}(0) > 4f^3(0). \quad (2)$$

On vérifie facilement que cette condition est vraie pour la loi de Gauss. Recherchons maintenant des lois pour lesquelles ce ne serait pas le cas.

2.3 Quelques lois qui violent cette condition.

2.3.1 Loi ad-hoc

La condition (2) n'est clairement pas vérifiée si la d.d.p. est plate au voisinage de l'origine, *i.e.* si $f^{(2)}(0) = 0$. Cette remarque incite à construire une d.d.p. ad-hoc telle que la suivante (cf. fig. 1) :

$$f(\varepsilon) = \begin{cases} f_{GL}(\varepsilon) = \frac{1}{C\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{\varepsilon+\frac{\alpha}{2}}{\sigma}\right)^2\right] & \text{pour } \varepsilon < -\frac{\alpha}{2}, \\ f_U(\varepsilon) = \frac{1}{C\sqrt{2\pi}\sigma} & \text{pour } -\frac{\alpha}{2} \leq \varepsilon \leq \frac{\alpha}{2}, \\ f_{GR}(\varepsilon) = \frac{1}{C\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{\varepsilon-\frac{\alpha}{2}}{\sigma}\right)^2\right] & \text{pour } \varepsilon > \frac{\alpha}{2}, \end{cases} \quad (3)$$

où $C = 1 + \frac{\alpha}{\sqrt{2\pi}\sigma}$ est une constant de normalisation.

La fonction $B(\varepsilon)$ pour cette loi est ¹

$$\begin{aligned} B(\varepsilon) &= \frac{F(\varepsilon)[1 - F(\varepsilon)]}{f^2(\varepsilon)} = \\ &= \begin{cases} 2\pi\sigma^2 \exp\left[\left(\frac{\varepsilon+\frac{\alpha}{2}}{\sigma}\right)^2\right] \Phi\left(\frac{\varepsilon+\frac{\alpha}{2}}{\sigma}\right) \left[C - \Phi\left(\frac{\varepsilon+\frac{\alpha}{2}}{\sigma}\right)\right] & \text{pour } \varepsilon < -\frac{\alpha}{2}, \\ 2\pi\sigma^2 \left[\left(\frac{1}{2} + \frac{\alpha}{2} \frac{1}{\sqrt{2\pi}\sigma}\right)^2 - \frac{\varepsilon^2}{2\pi\sigma^2}\right] & \text{pour } -\frac{\alpha}{2} \leq \varepsilon \leq \frac{\alpha}{2}, \\ 2\pi\sigma^2 \exp\left[\left(\frac{\varepsilon-\frac{\alpha}{2}}{\sigma}\right)^2\right] \left[\frac{\alpha}{\sqrt{2\pi}\sigma} + \Phi\left(\frac{\varepsilon-\frac{\alpha}{2}}{\sigma}\right)\right] \times \\ \quad \times \left[1 - \Phi\left(\frac{\varepsilon-\frac{\alpha}{2}}{\sigma}\right)\right] & \text{pour } \varepsilon > \frac{\alpha}{2} \end{cases} \end{aligned}$$

1. $\Phi(x)$ est la fonction de répartition de la loi de Gauss standard (moyenne nulle et $\sigma = 1$).

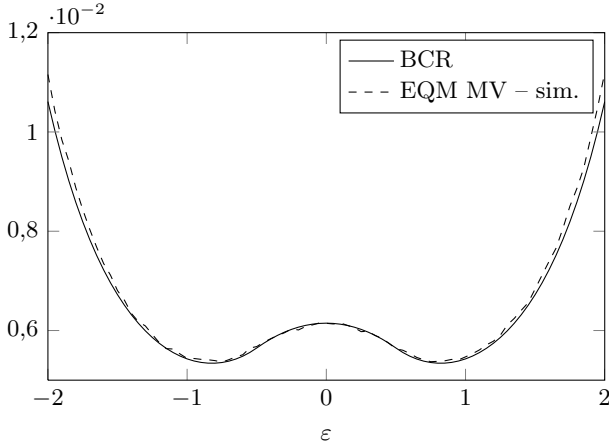


FIGURE 2 – BCR et EQM simulée pour l’estimateur du MV pour un bruit qui suit la distribution (3). La borne et l’EQM simulée ont été obtenues avec un nombre d’échantillons $N = 500$ et pour $\varepsilon \in [-2, 2]$. L’EQM pour l’estimateur du MV a été obtenue par simulation Monte-Carlo avec 10^5 réalisations, des blocs de 500 échantillons ont été utilisés. Les paramètres utilisés pour caractériser le bruit sont $\alpha = \sigma = 1$.

et la figure 2 illustre sa dépendance en ε . Pour vérifier que le comportement de la borne est proche de celui de l’estimation, nous simulons l’estimateur du MV 10^5 fois pour $N = 500$, les résultats de simulation ont été utilisés pour évaluer l’EQM (Erreur Quadratique Moyenne).

La borne est concave en $\varepsilon = 0$, placer le seuil du quantifieur sur le paramètre est localement le plus mauvais choix. Il existe deux minima symétriques autour de $\varepsilon = 0$. Pour ce bruit, la quantification binaire doit être asymétrique en décalant le seuil vers une région proche du bord de la zone uniforme.

2.3.2 Loi Gaussienne généralisée

Un exemple moins évident de lois qui ne respectent pas la condition (2) est celui des Gaussiennes généralisées² $f(\varepsilon) = \frac{\beta}{2\delta\Gamma(\frac{1}{\beta})} \exp\left(-|\frac{\varepsilon}{\delta}|^\beta\right)$. La dérivée première de f en 0 n’est pas définie pour $\beta \leq 1$, la dérivée seconde tend vers $-\infty$ pour $1 < \beta < 2$. En $\beta = 2$, $\varepsilon = 0$ est un minimum et l’on retrouve la loi de Gauss et le comportement vu en 2.1. En revanche, pour $\beta > 2$, $\varepsilon = 0$ devient un maximum, comme pour la loi ad-hoc du paragraphe 2.3.1 et :

$$B(\varepsilon) = \frac{\delta^2 \Gamma^2\left(\frac{1}{\beta}\right)}{\beta^2} \left[1 - \frac{\gamma^2\left(\frac{1}{\beta}, \left|\frac{\varepsilon}{\delta}\right|^\beta\right)}{\Gamma^2\left(\frac{1}{\beta}\right)} \right] \exp\left(2\left|\frac{\varepsilon}{\delta}\right|^\beta\right). \quad (4)$$

La figure 3 illustre la borne correspondante à cette fonction et les performances de l’estimateur.

2. On note $\Gamma(x) = \int_0^{+\infty} z^{x-1} \exp(-z) dz$ la fonction Gamma et $\gamma(x, w) = \int_0^w z^{x-1} \exp(-z) dz$ la Gamma incomplète.

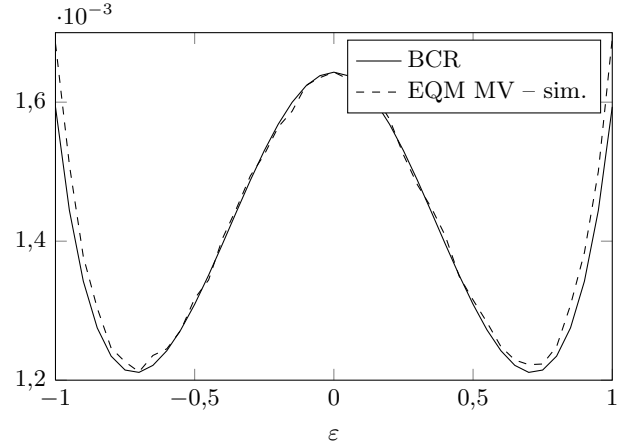


FIGURE 3 – BCR et EQM simulée pour l’estimateur du MV quand le bruit suit une distribution Gaussienne généralisée. La borne et l’EQM simulée ont été obtenues avec un nombre d’échantillons $N = 500$ et pour $\varepsilon \in [-1, 1]$. L’EQM pour l’estimateur du MV a été obtenue par simulation Monte-Carlo, 10^5 réalisations de blocs de 500 échantillons ont été utilisées. Les paramètres utilisés pour caractériser le bruit sont $\beta = 4$ et $\delta = 1$.

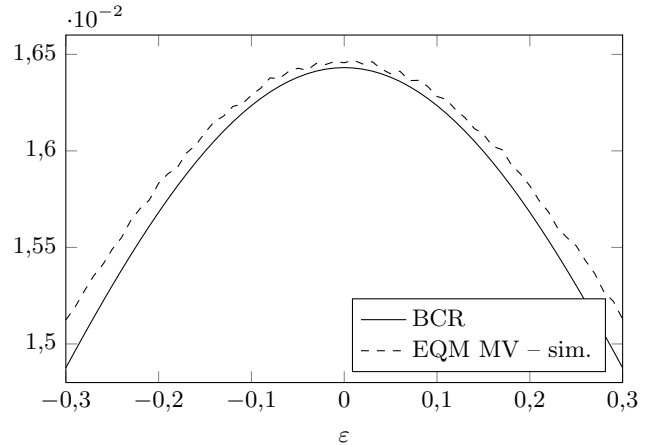


FIGURE 4 – BCR et EQM simulée pour l’estimateur du MV et un bruit distribué selon une loi Gaussienne généralisée. Les paramètres de simulation sont $\varepsilon \in [-0, 3; 0, 3]$, $N = 50$, nombre de réalisations 4×10^6 , $\beta = 4$ et $\delta = 1$.

La performance d’estimation est proche de la borne dans ce cas aussi. Pour quantifier de façon optimale, le seuil doit être décalé de la position du paramètre. La valeur optimale du décalage dépend de l’échelle δ du bruit, ce qui montre encore une fois la dépendance directe du quantifieur optimal par rapport aux caractéristiques du bruit.

Pour vérifier si la performance d’estimation est concave autour de $\varepsilon = 0$ même quand la BCR n’est pas valable, *i.e.* dans le cas non asymptotique, nous simulons l’estimateur du MV (aussi pour une loi de Gauss généralisée) avec $N = 50$. La BCR et l’EQM simulée sont montrées à la figure 4.

Les résultats sont donnés pour un intervalle plus petit que celui utilisé pour $N = 500$, $\varepsilon \in [-0,3; 0,3]$ dans ce cas. Ceci est dû au fait que pour un petit nombre d'échantillons, les grands écarts $|\varepsilon| > 0,3$ peuvent générer des réalisations de l'estimateur d'amplitude infinie avec une probabilité non négligeable (l'estimateur du MV diverge lorsque toutes les mesures quantifiées ont la même valeur).

On constate que, même si la performance d'estimation n'est pas très proche de la borne (en comparaison avec le cas $N = 500$), elle a la même forme concave autour de $\varepsilon = 0$.

3 Conclusions

Contrairement à ce qui est affirmé dans la littérature [4, page 265], le choix « seuil sur paramètre » n'est pas toujours optimal pour l'estimation d'un paramètre de centrage mesuré avec un bruit de d.d.p. symétrique. Même si pour des lois de bruit courantes (Gauss et Laplace) ce choix s'avère optimal, on peut trouver des cas (lois plates, lois de Gauss généralisées) pour lesquels, non seulement, ce choix n'est pas optimal, mais est, au contraire, localement le plus mauvais. Par conséquent, le choix du seuil optimal dépend radicalement de la loi du bruit. Quand la d.d.p. du bruit est uniforme autour de zéro, ce résultat pourrait être expliqué par le fait qu'asymptotiquement ($N \rightarrow \infty$) les bords de la distribution uniforme sont très informatifs (l'estimateur du MV d'un paramètre de centrage pour une distribution uniforme centrée est lié à la valeur maximale

ou minimale des échantillons observés). Cependant, dans le cas où la loi du bruit est une Gaussienne généralisée, on ne peut pas utiliser le même raisonnement, et le résultat paraît contre-intuitif.

Finalement, notons que cette asymétrie apparaît également pour la détection optimale basée sur des observations binaires de signaux à faible rapport signal sur bruit. Dans ce cas, on peut montrer que la performance de détection dépend de façon directe de $B(\varepsilon)$ [1].

Références

- [1] S. Kassam. Optimum quantization for signal detection. *IEEE Transactions on Communications*, 25(5) :479–484, 1977.
- [2] H.C. Papadopoulos, G.W. Wornell, and A.V. Oppenheim. Sequential signal encoding from noisy measurements using quantizers with dynamic bias control. *IEEE Transactions on Information Theory*, 47(3) :978–1002, 2001.
- [3] A. Ribeiro and G.B. Giannakis. Bandwidth-constrained distributed estimation for wireless sensor networks-part I : Gaussian case. *IEEE Transactions on Signal Processing*, 54(3) :1131–1143, 2006.
- [4] L.Y. Wang, G. Yin, J.F. Zhang, and Y. Zhao. *System identification with quantized observations*. Birkhauser, 2010.