

Bornes de performances

Philippe Ciblat

École Nationale Supérieure des Télécommunications, Paris, France

Plan

1. Bornes de Cramer-Rao classiques
 - avec paramètres déterministes
 - avec paramètres aléatoires (approche bayésienne)
2. Bornes de Cramer-Rao avec paramètres de nuisance
3. Autres bornes
 - Barankin
 - Ziv-Zakai

Exemple-conducteur : estimation de fréquence sans séquence d'apprentissage

Partie 1 : Bornes de Cramer-Rao classiques

Borne de Cramer-Rao déterministe (I)

- Soit θ un paramètre (réel) déterministe à estimer
- Soit $\hat{\theta}$ un estimateur sans biais de θ
- Soit \mathbf{y}_N le vecteur d'observations de taille N

Borne de Cramer-Rao

$$\mathbb{E}_Y \left[\left(\hat{\theta} - \theta \right) \left(\hat{\theta} - \theta \right)^T \right] \geq F(\theta)^{-1} = \text{CRB}(\theta)$$

Remarque :

- La matrice $F(\theta)$ se nomme matrice d'information de Fisher (FIM)

$$F(\theta) = \mathbb{E}_Y \left[\left(\frac{\partial \ln p_Y(\mathbf{y}_N | \theta)}{\partial \theta} \right) \left(\frac{\partial \ln p_Y(\mathbf{y}_N | \theta)}{\partial \theta} \right)^T \right]$$

- Un estimateur sans biais atteignant la CRB est dit *efficace*

Borne de Cramer-Rao déterministe (II)

Si la fonction $\ln p_Y(\mathbf{y}_N|\theta)$ admet une dérivée seconde, alors

$$F(\theta) = -\mathbb{E}_Y \left[\frac{\partial^2 \ln p_Y(\mathbf{y}_N|\theta)}{(\partial\theta)^2} \right]$$

Remarque :

- $p_Y(\mathbf{y}_N|\theta)$ est la vraisemblance
- $\ln p_Y(\mathbf{y}_N|\theta)$ est la log-vraisemblance

Borne de Cramer-Rao bayésienne

- Soit θ un paramètre (réel) aléatoire à estimer admettant une distribution *a priori* $p_\theta(\theta)$
- Soit $\hat{\theta}$ un estimateur (réel) sans biais de θ

$$\mathbb{E}_{Y,\theta} \left[\left(\hat{\theta} - \theta \right) \left(\hat{\theta} - \theta \right)^T \right] \geq F^{-1} = \text{BCRB}$$

avec

$$F = \mathbb{E}_{Y,\theta} \left[\left(\frac{\partial \ln p_{Y,\theta}(\mathbf{y}_N, \theta)}{\partial \theta} \right) \left(\frac{\partial \ln p_{Y,\theta}(\mathbf{y}_N, \theta)}{\partial \theta} \right)^T \right]$$

Remarque :

- On a aussi

$$F = -\mathbb{E}_{Y,\theta} \left[\frac{\partial^2 \ln p_{Y,\theta}(\mathbf{y}_N, \theta)}{(\partial \theta)^2} \right]$$

- Pas de liens triviaux entre BCRB et $\mathbb{E}_\theta[\text{CRB}(\theta)]$

Partie 2 : Bornes de Cramer-Rao avec paramètres de nuisance

Exemple : estimation autodidacte de la fréquence

$$y(n) = a(n)e^{2i\pi f_0 n} + b(n), \quad n = 0, \dots, N-1 \Leftrightarrow \mathbf{y}_N = \mathbf{D}_N(f_0)\mathbf{a}_N + \mathbf{b}_N$$

avec

- $\mathbf{y}_N = [y(0), \dots, y(N-1)]^T$
- $\mathbf{D}_N(f_0) = \text{diag}([1, \dots, e^{2i\pi f_0(N-1)}])$
- $\mathbf{a}_N = [a(0), \dots, a(N-1)]^T$
- \mathbf{b}_N bruit blanc gaussien circulaire centré et de variance $2N_0$ supposée connue par souci de simplicité

Problème

- f_0 : paramètre d'intérêt (supposé déterministe)
 - \mathbf{a}_N : paramètres de nuisance
- ⇒ *A chaque hypothèse formulée sur les paramètres de nuisance correspondra UNE borne de Cramer-Rao*

CRB non-conditionnelle (I)

- $a(n)$ est vu comme une vraie nuisance \rightsquigarrow *stochastique*
- Vraisemblance non-conditionnée ou vraie vraisemblance

$$p_u(\mathbf{y}_N|f) = \mathbb{E}_{\mathbf{a}}[p(\mathbf{y}_N|f, \mathbf{a}_N)] = \int p(\mathbf{y}_N|f, \mathbf{a}_N)p(\mathbf{a}_N)d\mathbf{a}_N$$

Définition

$$\text{UCRB}(f) = \frac{1}{\mathbb{E}_{\mathbf{y}} \left[\left| \frac{\partial}{\partial f} \ln p_u(\mathbf{y}_N|f) \right|^2 \right]} = \frac{1}{\mathbb{E}_{\mathbf{y}} \left[\left| \frac{\partial}{\partial f} \ln \mathbb{E}_{\mathbf{a}}[p(\mathbf{y}_N|f, \mathbf{a}_N)] \right|^2 \right]}$$

CRB non-conditionnelle (II)

Dénomination :

- UCRB : *unconditional CRB*
- TCRB : *true CRB*
- SCRb : *stochastic CRB*

Inconvénient :

- Cette borne est souvent incalculable analytiquement car

$$p_u(\mathbf{y}_N|f) \propto \int e^{\frac{-\|\mathbf{y}_N - \mathbf{D}_N(f)\mathbf{a}_N\|^2}{2N_0}} p(\mathbf{a}_N) d\mathbf{a}_N$$

est difficile à exprimer simplement

- Approximation à bas RSB ($e^x = 1 + x + x^2/2 + o(x^2)$)

CRB conditionnelle

- $a(n)$ est vu comme un paramètre d'intérêt \rightsquigarrow *déterministe*
- Vraisemblance conditionnelle ou vraisemblance déterministe

$$p_c(\mathbf{y}_N|f) = p(\mathbf{y}_N|f, \hat{\mathbf{a}}_f) \quad \text{avec} \quad \left. \frac{\partial p(\mathbf{y}_N|f, \mathbf{a}_N)}{\partial \mathbf{a}_N} \right|_{\hat{\mathbf{a}}_f} = 0$$

Définition

$$\text{CCRB}(f) = \frac{1}{\mathbb{E}_{\mathbf{y}, \mathbf{a}} \left[\left| \frac{\partial}{\partial f} \ln p_c(\mathbf{y}_N|f) \right|^2 \right]}$$

Dénomination :

- CCRB : *conditional CRB*
- DCRB : *deterministic CRB*

Remarque :

- La CCRB est peu utilisée alors que le CML l'est beaucoup

CRB modifiée

- $a(n)$ est vu comme un paramètre *connu*

Définition

$$\text{MCRB}(f) = \frac{1}{\mathbb{E}_{\mathbf{y}, \mathbf{a}} \left[\left| \frac{\partial}{\partial f} \ln p(\mathbf{y}_N | f, \mathbf{a}_N) \right|^2 \right]}$$

Remarque

- Expressions analytiques *calculables*
- MCRB souvent utilisée pas toujours précise

CRB gaussienne

- $a(n)$ est vu comme un processus *gaussien*
- Vraisemblance gaussienne

$$p_g(\mathbf{y}_N|f) = \mathbb{E}_{\mathbf{a}}[p(\mathbf{y}_N|f, \mathbf{a}_N)]$$

où \mathbf{a}_N est un vecteur gaussien

Définition

$$\text{GCRB}(f) = \frac{1}{\mathbb{E}_{\mathbf{y}} \left[\left| \frac{\partial}{\partial f} \ln p_g(\mathbf{y}_N|f) \right|^2 \right]}$$

- Expressions analytiques *calculables*
- Approche *a priori* pas valide en coms. nums
- C'est une borne pour tous les estimateurs sans biais basés sur la matrice de covariance des observations

Liens entre xCRB

Toutes bornes (exceptée la GCRB) sont des bornes de l'EQM !

Résultat

$$\text{UCRB} \geq \text{MCRB} (*) \quad \text{et} \quad \text{CCRB} \geq \text{MCRB}$$

Deux cas particuliers :

- $\text{UCRB} = \text{MCRB}$ (si fort RSB et si les valeurs de paramètres de nuisance appartiennent à un ensemble discret)
- $\text{UCRB} = \text{GCRB}$ (si les paramètres de nuisance sont vraiment gaussiens)

↪ **MCRB souvent trop optimiste**

↪ **GCRB incapable de prendre en compte les HOS**

Preuve de (*)

$$\begin{aligned}
\mathbb{E}_{\mathbf{y}} \left[\left(\frac{\partial \ln p(\mathbf{y}|f)}{\partial f} \right)^2 \right] &= \int \left(\frac{\partial p(\mathbf{y}|f)}{\partial f} \right)^2 \frac{1}{p(\mathbf{y}|f)} d\mathbf{y} \\
&= \int \left(\int \frac{\partial \ln p(\mathbf{y}|f, \mathbf{a})}{\partial f} p(\mathbf{y}, \mathbf{a}|f) d\mathbf{a} \right)^2 \frac{1}{p(\mathbf{y}|f)} d\mathbf{y} \\
&\leq \int \left(\int \left(\frac{\partial \ln p(\mathbf{y}|f, \mathbf{a})}{\partial f} \right)^2 p(\mathbf{y}, \mathbf{a}|f) d\mathbf{a} \right) \\
&\quad \times \left(\int p(\mathbf{y}, \mathbf{a}|f) d\mathbf{a} \right) \frac{1}{p(\mathbf{y}|f)} d\mathbf{y} \\
&\leq \int \int \left(\frac{\partial \ln p(\mathbf{y}|f, \mathbf{a})}{\partial f} \right)^2 p(\mathbf{y}, \mathbf{a}|f) d\mathbf{y} d\mathbf{a} \\
&\leq \mathbb{E}_{\mathbf{y}, \mathbf{a}} \left[\left(\frac{\partial \ln p(\mathbf{y}|f, \mathbf{a})}{\partial f} \right)^2 \right]
\end{aligned}$$

Exemple : estimation autodidacte de la fréquence (I)

a(n) blanc, centré, à valeurs discrètes et non-circulaire à l'ordre deux

Résultat

$$\text{MCRB} = \frac{3\sigma^2}{2\pi^2 N^3} \quad \text{et} \quad \text{GCRB} = \frac{3[(1 - |u|^2) + 2\sigma^2 + \sigma^4]}{4\pi^2 |u|^2 N^3}$$

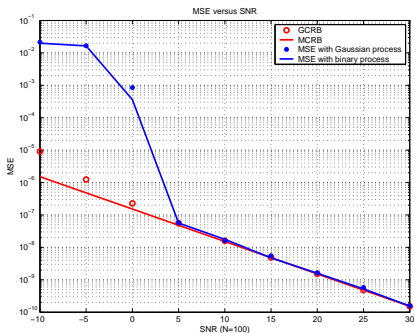
$$\text{UCRB}_{|\text{fort RSB}} = \frac{3\sigma^2}{2\pi^2 N^3} \quad \text{et} \quad \text{UCRB}_{|\text{faible RSB}} = \frac{3\sigma^4}{4\pi^2 |u|^2 N^3}$$

avec variance $\mathbb{E}[|a(n)|^2] = 1$ et pseudo-variance $\mathbb{E}[a(n)^2] = u$

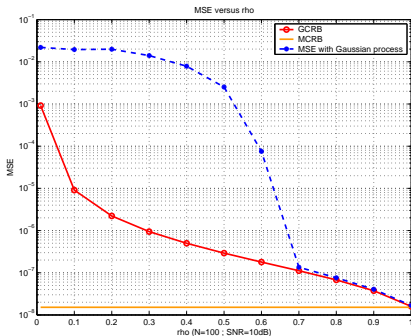
Remarque :

- MCRB utile à fort RSB mais pas à faible RSB
- GCRB utile à faible RSB (si non-circularité au second ordre)
- GCRB utile à fort et faible RSB ssi $a(n) \in \text{MDA}$ (i.e., $u = 1$)
- GCRB inutile pour les MDP et MAQ

Exemple : estimation autodidacte de la fréquence (II)



EQM en fct. du RSB

EQM en fct. de $|u|$

Partie 3 : Autres bornes

Autres bornes

Remarque

xCRB incapable de prédire et d'analyser l'effet de décrochement

Solutions

Introduire des bornes plus précises

- Approche déterministe
 - ↪ Borne de Barankin
- Approche aléatoire
 - ↪ Borne de Ziv-Zakai
- Présentation pour un scalaire réel et sans paramètres de nuisance

Borne de Barankin

- On considère des points-tests $\{\theta^{(1)}, \dots, \theta^{(p)}\}$
- On forme $\mathbf{B}_p = (B_{k,l})_{1 \leq k, l \leq p}$ de la manière suivante

$$B_{k,l} = \mathbb{E}_{\mathbf{y}} \left[\frac{\rho(\mathbf{y}|\theta^{(k)})\rho(\mathbf{y}|\theta^{(l)})}{\rho(\mathbf{y}|\theta)^2} \right]$$

Définition

$$BB_p(\theta) = \sup_{\mathcal{E}_p} \underbrace{\mathcal{E}_p^T (\mathbf{B}_p - \mathbf{1}_p \mathbf{1}_p^T)^{-1} \mathcal{E}_p}_{S_p(\mathcal{E}_p)}$$

avec $\mathcal{E}_p = [\theta^{(1)} - \theta, \dots, \theta^{(p)} - \theta]^T$ et $\mathbf{1}_p = \text{ones}(p, 1)$

- *EQM de tout estimateur sans biais est plus grand que BB_p*
- *Quand $p \rightarrow \infty$, BB_∞ devient la plus fine des bornes*

Borne de Ziv-Zakai

Définition

$$\text{ZZB} = \int \varepsilon \left[\int \min(p(\theta), p(\theta + \varepsilon)) P_e(\theta, \theta + \varepsilon) d\theta \right] d\varepsilon$$

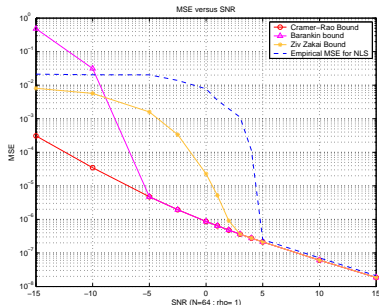
où

- $p(\cdot)$ densité *a priori* de θ
- $P_e(\theta, \theta + \varepsilon)$ est la probabilité d'erreur du détecteur optimal pour le problème suivant

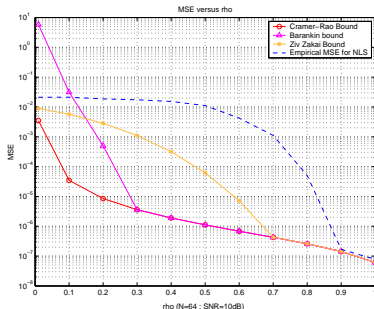
$$\begin{cases} H_0 : y(n) = f_n(\theta) \\ H_1 : y(n) = f_n(\theta + \varepsilon) \end{cases}$$

Simulations

$a(n)$ blanc centré gaussien à valeurs complexes de variance 1 et de pseudo-variance $\mathbb{E}[a(n)^2] = u$



EQM en fct. du RSB



EQM en fct. de u

- Ecart important entre BB et l'estimateur de l'élevation au carré
- Petit écart entre ZZB et l'estimateur de l'élevation au carré

Bibliographie

- H.L. Van Trees, "Detection, Estimation, and Modulation Theory", Partie 1, 1968.
- J. Ziv, "Some lower bounds on Signal Processing, IEEE Trans. on Information Theory, Mai 1969.
- B. Porat, "Digital Processing of Random Signals : Theory and Methods", 1994.
- M. Moeneclaey, "On the True and the Modified CRB for the estimation of a scalar paramter in the presence of nuisance parameter", IEEE Trans. on Communications, Nov. 1998.
- G. Vazquez, "Non-Data-Aided Digital Synchronization" dans le livre "Signal Processing Advances in Communications" dirigé par G. Giannakis et al., 2000.
- P. Forster et P. Larzabal, "On lower bounds for deterministic paramter estimation", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Mai 2002.
- P. Ciblat et al., "Harmonic retrieval in the presence of non-circular Gaussian multiplicative noise : Performance bounds", EURASIP Signal Processing, Avril 2005.
- P. Ciblat et M. Ghogho, "Ziv-Zakai bound for harmonic retrieval in multiplicative and additive Gaussian noise", IEEE Workshop on Statistical Signal Processing (SSP), Juillet 2005.
- A. Renaux, "Contribution à l'analyse des performances d'estimation en traitement statistique du signal ", Thèse ENS Cachan, Juillet 2006.