

Bornes de Cramér-Rao intrinsèques et matrices de covariance

F. Bouchard, A. Breloy, G. Ginolhac, A. Renaux

I Problématique

II Bornes de Cramér-Rao « euclidiennes »

III Borne de Cramér-Rao intrinsèque – Généralités
(estimation intrinsèque)

IV Borne de Cramér-Rao intrinsèque – Estimation d'une
matrice de covariance

V Perspectives

✓ I Problématique

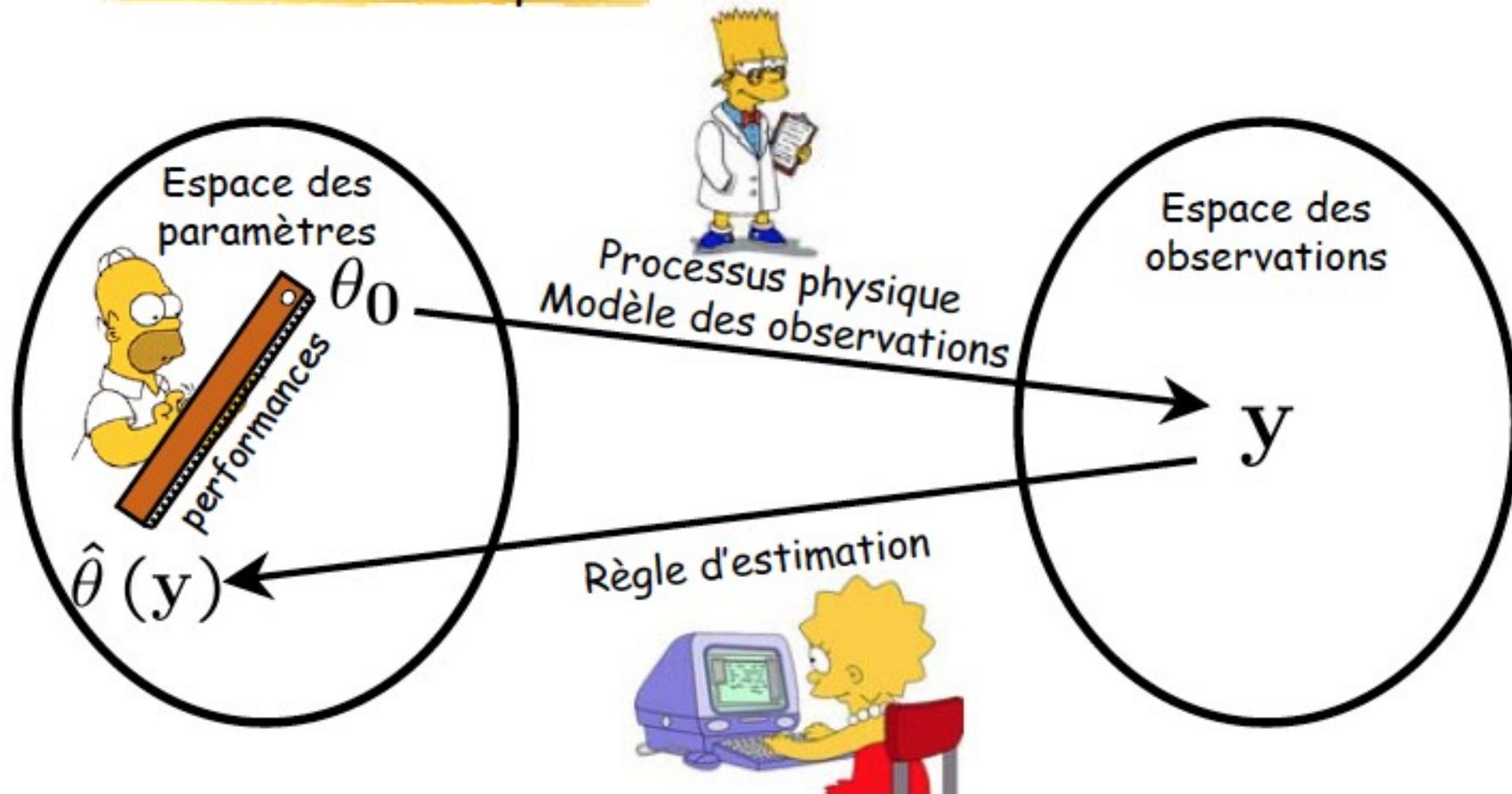
II Bornes de Cramér-Rao « euclidiennes »

III Borne de Cramér-Rao intrinsèque – Généralités

IV Borne de Cramér-Rao intrinsèque – Estimation d'une matrice de covariance

V Perspectives

Cadre statistique



Exemples : modèles \longrightarrow estimateurs

$$y(t) = \theta + b(t), t = 1, \dots, T \quad \longrightarrow \quad \hat{\theta} = \frac{1}{T} \sum_{t=1}^T y(t)$$

$$y(t) = \cos(2\pi\theta t) + b(t), t = 1, \dots, T \quad \longrightarrow \quad \hat{\theta} = \arg \max_{\theta} TF \{y(1), \dots, y(T)\}$$

$$\mathbf{y}(t) = \mathbf{b}(t) \sim \mathcal{N}(0, \Sigma), t = 1, \dots, T \quad \longrightarrow \quad \hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \mathbf{y}(t) \mathbf{y}^T(t)$$

Estimateurs : moindres carrés, méthode MUSIC, méthodes des moments, maximum de vraisemblance, etc.

Remarques : notations, hypothèses statistiques

Performances



L'estimateur est une variable (vecteur, matrice) **aléatoire**

$$y(t) = \theta + b(t), \quad t = 1, \dots, T$$

$$b(t) \sim \mathcal{N}(0, \sigma^2)$$

$$b(t) \perp b(t') \text{ pour } t \neq t'$$

$$\longrightarrow \hat{\theta} = \frac{1}{T} \sum_{t=1}^T y(t) \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{T}\right)$$

$$y(t) = \cos(2\pi\theta t) + b(t), \quad t = 1, \dots, T \longrightarrow \hat{\theta} = \arg \max_{\theta} TF\{y(1), \dots, y(T)\}$$

$$y(t) = \mathbf{b}(t) \sim \mathcal{N}(0, \Sigma), \quad t = 1, \dots, T \longrightarrow \hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \mathbf{y}(t) \mathbf{y}^T(t) \sim \text{Wishart}$$

Indicateurs de performances : consistance, **biais**, **variance (bornes)**, etc.

Performances

- Biais $\mathbb{E}(\hat{\theta}) - \theta$ Distance moyenne



On aimerait bien que le biais soit nul (ou asymptotiquement nul)

- Variance / covariance Distance quadratique moyenne

$$\mathbb{E} \left(\left(\hat{\theta} - \mathbb{E}(\hat{\theta}) \right)^2 \right) \stackrel{\text{biais}=0}{=} \mathbb{E} \left(\left(\hat{\theta} - \theta \right)^2 \right)$$

$$\mathbb{E} \left(\left(\hat{\theta} - \mathbb{E}(\hat{\theta}) \right) \left(\hat{\theta} - \mathbb{E}(\hat{\theta}) \right)^T \right) \stackrel{\text{biais}=0}{=} \mathbb{E} \left(\left(\hat{\theta} - \theta \right) \left(\hat{\theta} - \theta \right)^T \right)$$



On aimerait bien que la variance soit petite

Remarque : $\hat{\theta} = \frac{1}{T} \sum_{t=1}^T y(t) \sim \mathcal{N} \left(\theta, \frac{\sigma^2}{T} \right)$ Biais nul et variance asymptotiquement nulle

Performances

- Le biais et la variance dépendent de la stratégie d'estimation (existe t'il un meilleurs estimateur?)
- Pour des problèmes modernes de traitement du signal il est difficile d'obtenir des expressions analytiques du biais et de la variance

Exemple : $\hat{\theta} = \arg \max_{\theta} TF \{y(1), \dots, y(T)\} \longrightarrow \mathbb{E}(\hat{\theta}) = ?$

Simulation de Monte-Carlo



Trouver une « limite » indépendante de l'estimateur (seulement du modèle des observations), « proche » de la variance de l'estimateur et calculable facilement.

I Problématique

✓ II Bornes de Cramér-Rao « euclidiennes »

III Borne de Cramér-Rao intrinsèque – Généralités

IV Borne de Cramér-Rao intrinsèque – Estimation d'une matrice de covariance

V Perspectives

Bornes de Cramér-Rao « euclidiennes »

Historique et bibliographie

- M. Frechet. Sur l'extension de certaines évaluations statistiques au cas de petit échantillons. Rev. Inst. Int. Statist., 11 :182–205, 1943.
- G. Darmais. Sur les lois limites de la dispersion de certaines estimations. Rev. Inst. Int. Statist., 13 :9–15, 1945.
- H. Cramér. Mathematical Methods of Statistics. Princeton University, Press, New York, 1946.
- C. R. Rao (99 ans en septembre !!!!). Information and accuracy attainable in the estimation of statistical parameters. Bull. Calcutta Math. Soc., 37 :81–91, 1945.
- S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice Hall, 1993
- E. L. Lehmann. Theory of Point Estimation. Wiley, New York, 1983.

Bornes de Cramér-Rao « euclidiennes »

Cadre et hypothèses

- Modèle des observations avec sa densité de probabilité $f_{\theta}(\mathbf{y})$ continue en \mathbf{y}

Identifiable $f_{\theta_1}(\mathbf{y}) = f_{\theta_2}(\mathbf{y}) \Rightarrow \theta_1 = \theta_2$

- L'espace des paramètres Θ est un ouvert de \mathbb{R}^P **Euclidien !!!**

- L'ensemble $\mathcal{A} = \{\mathbf{y} : f_{\theta}(\mathbf{y}) > 0\}$ est indépendant de θ

- $\forall \mathbf{y} \in \mathcal{A}$, $f_{\theta}(\mathbf{y})$ est différentiable (deux fois) par rapport à θ

- $$\frac{\partial}{\partial \theta} \int (\cdot) d\mathbf{y} = \int \frac{\partial}{\partial \theta} (\cdot) d\mathbf{y}$$



Paramètres discrets, densités non dérivable (uniforme), le paramètre ne peut pas être sur les « bords » de la loi, etc.

Bornes de Cramér-Rao « euclidiennes »

Inégalité de Cramér-Rao

Pour un estimateur non biaisé

● Cas scalaire
$$\mathbb{E} \left((\hat{\theta} - \theta)^2 \right) \geq \frac{1}{\mathbb{E} \left(\left(\frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta} \right)^2 \right)} = \frac{1}{-\mathbb{E} \left(\frac{\partial^2 \ln f_{\theta}(\mathbf{y})}{\partial \theta^2} \right)}$$

Remarques (valables pour le cas vectoriel)

- Il s'agit bien d'une « limite » indépendante de l'estimateur (seulement du modèle des observations)

- Le terme $F(\theta) = \mathbb{E} \left(\left(\frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta} \right)^2 \right)$ s'appelle information de Fisher et est

supposé inversible. L'information de Fisher (et donc la borne de Cramér-Rao) dépend de θ

- Le terme $\frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}$ s'appelle le score (à moyenne nulle). L'information de

Fisher est la variance du score.

Bornes de Cramér-Rao « euclidiennes »

Preuve $\mathbf{v} = \begin{pmatrix} \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta} \\ \hat{\theta} - \theta \end{pmatrix}$

$$\mathbb{E}(\mathbf{v}\mathbf{v}^T) = \begin{pmatrix} F(\theta) & \mathbb{E}\left(\left(\hat{\theta} - \theta\right) \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right) \\ \mathbb{E}\left(\left(\hat{\theta} - \theta\right) \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right) & \text{Var}(\hat{\theta}) \end{pmatrix}$$

$$\mathbb{E}\left(\left(\hat{\theta} - \theta\right) \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right) = \mathbb{E}\left(\hat{\theta} \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right) - \theta \mathbb{E}\left(\frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right)$$

$$\mathbb{E}\left(\frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right) = \int_{\Omega} \frac{\partial f_{\theta}(\mathbf{y})}{\partial \theta} d\mathbf{y} = \frac{\partial}{\partial \theta} \int_{\Omega} f_{\theta}(\mathbf{y}) d\mathbf{y} = 0$$

Bornes de Cramér-Rao « euclidiennes »

Preuve

$$\mathbf{v} = \begin{pmatrix} \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta} \\ \hat{\theta} - \theta \end{pmatrix}$$

$$\mathbb{E}(\mathbf{v}\mathbf{v}^T) = \begin{pmatrix} F(\theta) & \mathbb{E}\left(\left(\hat{\theta} - \theta\right) \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right) \\ \mathbb{E}\left(\left(\hat{\theta} - \theta\right) \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right) & \text{Var}(\hat{\theta}) \end{pmatrix}$$

$$\mathbb{E}\left(\left(\hat{\theta} - \theta\right) \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right) = \mathbb{E}\left(\hat{\theta} \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right) - \theta \mathbb{E}\left(\frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right)$$

$$\mathbb{E}\left(\hat{\theta} \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right) = \int_{\Omega} \hat{\theta} \frac{\partial f_{\theta}(\mathbf{y})}{\partial \theta} d\mathbf{y} = \frac{\partial}{\partial \theta} \int_{\Omega} \hat{\theta} f_{\theta}(\mathbf{y}) d\mathbf{y} = \frac{\partial}{\partial \theta} (-\theta) = -1$$

Bornes de Cramér-Rao « euclidiennes »

Preuve

$$\mathbf{v} = \begin{pmatrix} \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta} \\ \hat{\theta} - \theta \end{pmatrix}$$

$$\mathbb{E}(\mathbf{v}\mathbf{v}^T) = \begin{pmatrix} F(\theta) & \mathbb{E}\left(\left(\hat{\theta} - \theta\right) \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right) \\ \mathbb{E}\left(\left(\hat{\theta} - \theta\right) \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta}\right) & \text{Var}(\hat{\theta}) \end{pmatrix}$$

$$= \begin{pmatrix} F(\theta) & -1 \\ -1 & \text{Var}(\hat{\theta}) \end{pmatrix}$$

$$|\mathbb{E}(\mathbf{v}\mathbf{v}^T)| \geq 0 \Rightarrow F(\theta) \text{Var}(\hat{\theta}) - 1 \geq 0 \Rightarrow \text{Var}(\hat{\theta}) \geq \frac{1}{F(\theta)}$$

Bornes de Cramér-Rao « euclidiennes »

Exemple de calcul

$$y(t) = \theta + b(t), \quad t = 1, \dots, T \quad \longrightarrow \quad \hat{\theta} = \frac{1}{T} \sum_{t=1}^T y(t) \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{T}\right)$$
$$b(t) \sim \mathcal{N}(0, \sigma^2)$$
$$b(t) \perp b(t') \text{ pour } t \neq t'$$

$$f_{\theta}(\mathbf{y}) = \prod_{t=1}^T \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y(t)-\theta)^2}{2\sigma^2}} = \frac{1}{(\sqrt{2\pi}\sigma)^T} e^{-\frac{1}{2\sigma^2} \sum_{t=1}^T (y(t)-\theta)^2}$$

$$\ln f_{\theta}(\mathbf{y}) = -T \ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2} \sum_{t=1}^T (y(t) - \theta)^2$$

$$\frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta} = \frac{1}{\sigma^2} \sum_{t=1}^T (y(t) - \theta)$$

$$\frac{\partial^2 \ln f_{\theta}(\mathbf{y})}{\partial \theta^2} = \frac{1}{\sigma^2} \sum_{t=1}^T -1 = -\frac{T}{\sigma^2}$$

Bornes de Cramér-Rao « euclidiennes »

Exemple de calcul

$$y(t) = \theta + b(t), \quad t = 1, \dots, T \quad \longrightarrow \quad \hat{\theta} = \frac{1}{T} \sum_{t=1}^T y(t) \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{T}\right)$$
$$b(t) \sim \mathcal{N}(0, \sigma^2)$$
$$b(t) \perp b(t') \text{ pour } t \neq t'$$

$$f_{\theta}(\mathbf{y}) = \prod_{t=1}^T \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y(t)-\theta)^2}{2\sigma^2}} = \frac{1}{(\sqrt{2\pi}\sigma)^T} e^{-\frac{1}{2\sigma^2} \sum_{t=1}^T (y(t)-\theta)^2}$$

$$\ln f_{\theta}(\mathbf{y}) = -T \ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2} \sum_{t=1}^T (y(t) - \theta)^2$$

$$\frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta} = \frac{1}{\sigma^2} \sum_{t=1}^T (y(t) - \theta)$$

$$\frac{\partial^2 \ln f_{\theta}(\mathbf{y})}{\partial \theta^2} = \frac{1}{\sigma^2} \sum_{t=1}^T -1 = -\frac{T}{\sigma^2}$$

$$\mathbb{E}\left(\left(\hat{\theta} - \theta\right)^2\right) \geq \frac{\sigma^2}{T}$$

Estimateur **efficace**

(...et calcul facile)



Bornes de Cramér-Rao « euclidiennes »

Remarques (valables pour le cas vectoriel mais plus facile à expliquer en scalaire...)

- Score $\frac{\partial \ln f_{\theta}(y)}{\partial \theta}$ sensibilité du modèle à un changement infinitésimal de la valeur du paramètre

→ Information de Fisher est la « force » (variance) de cette sensibilité

Plus l'information de Fisher est grande, plus il sera donc « facile » d'estimer le paramètre (cad BCR petite)

- Divergence de Kullback-Leibler (entropie relative)

$$D_{KL}(f_{\theta}, f_{\theta'}) = \int_{\Omega} f_{\theta}(y) \ln \frac{f_{\theta}(y)}{f_{\theta'}(y)} dy$$

$$D_{KL}(f_{\theta}, f_{\theta'}) \underset{\theta' \rightarrow \theta}{\approx} 0 + \frac{1}{2} (\theta - \theta')^2 \frac{\partial^2 D_{KL}(f_{\theta}, f_{\theta'})}{\partial \theta^2} \Big|_{\theta=\theta'} + o((\theta - \theta')^3)$$

$$\frac{\partial^2 D_{KL}(f_{\theta}, f_{\theta'})}{\partial \theta'^2} = - \int_{\Omega} f_{\theta}(y) \frac{\frac{\partial^2 f_{\theta'}(y)}{\partial \theta'^2} f_{\theta'}(y) - \left(\frac{\partial f_{\theta'}(y)}{\partial \theta'}\right)^2}{f_{\theta'}^2(y)} dy \underset{\theta=\theta'}{=} \int_{\Omega} f_{\theta}(y) \frac{\left(\frac{\partial f_{\theta}(y)}{\partial \theta}\right)^2}{f_{\theta}^2(y)} dy = \mathbb{E} \left(\left(\frac{\partial \ln f_{\theta}(y)}{\partial \theta} \right)^2 \right)$$

→ Localement, la divergence de Kullback-Leibler se comporte comme l'information de Fisher

Bornes de Cramér-Rao « euclidiennes »

Inégalité de Cramér-Rao

Pour un estimateur non biaisé (+ inversibilité)

● Cas vectoriel

$$\mathbb{E} \left((\hat{\theta} - \theta) (\hat{\theta} - \theta)^T \right) \succeq \left(\mathbb{E} \left(\frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta} \frac{\partial \ln f_{\theta}(\mathbf{y})}{\partial \theta^T} \right) \right)^{-1} = - \left(\mathbb{E} \left(\frac{\partial^2 \ln f_{\theta}(\mathbf{y})}{\partial \theta \partial \theta^T} \right) \right)^{-1} = \mathbf{F}^{-1}(\theta)$$

$$\text{avec } \frac{\partial}{\partial \theta} = \begin{pmatrix} \frac{\partial}{\partial \theta_1} \\ \frac{\partial}{\partial \theta_2} \\ \vdots \\ \frac{\partial}{\partial \theta_P} \end{pmatrix} \quad \text{et} \quad \frac{\partial^2}{\partial \theta \partial \theta^T} = \begin{pmatrix} \frac{\partial^2}{\partial \theta_1^2} & \frac{\partial^2}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2}{\partial \theta_1 \partial \theta_P} \\ \frac{\partial^2}{\partial \theta_1 \partial \theta_2} & \frac{\partial^2}{\partial \theta_2^2} & & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2}{\partial \theta_1 \partial \theta_P} & \cdots & \cdots & \frac{\partial^2}{\partial \theta_P^2} \end{pmatrix}$$

$\mathbf{A} \succeq \mathbf{B} \iff \mathbf{A} - \mathbf{B}$ matrice (symétrique) semi-définie positive

On a toujours $Var(\theta_i) \geq \left\{ (\mathbf{F}(\theta))^{-1} \right\}_{i,i} \quad i = 1, \dots, P$

Mais aussi $tr \left(\mathbb{E} \left((\hat{\theta} - \theta) (\hat{\theta} - \theta)^T \right) \right) \geq tr \left((\mathbf{F}(\theta))^{-1} \right)$

Bornes de Cramér-Rao « euclidiennes »

Exemple de résultat de calcul

$$y(t) = \theta_1 + b(t), \quad t = 1, \dots, T$$

$$b(t) \sim \mathcal{N}(0, \theta_2)$$

$$b(t) \perp b(t') \text{ pour } t \neq t'$$

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$$

$$\mathbf{F}(\theta) = \begin{pmatrix} -\mathbb{E} \left(\frac{\partial^2 \ln f_{\theta}(\mathbf{y})}{\partial \theta_1^2} \right) & -\mathbb{E} \left(\frac{\partial^2 \ln f_{\theta}(\mathbf{y})}{\partial \theta_1 \partial \theta_2} \right) \\ -\mathbb{E} \left(\frac{\partial^2 \ln f_{\theta}(\mathbf{y})}{\partial \theta_1 \partial \theta_2} \right) & -\mathbb{E} \left(\frac{\partial^2 \ln f_{\theta}(\mathbf{y})}{\partial \theta_2^2} \right) \end{pmatrix} = \begin{pmatrix} \frac{T}{\sigma^2} & 0 \\ 0 & \frac{T}{2\sigma^4} \end{pmatrix}$$

Influence de l'estimation d'un paramètre sur un autre (nuisance)

$$\mathbf{F}(\theta) = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \quad \left\{ (\mathbf{F}(\theta))^{-1} \right\}_{1,1} = \frac{c}{ac - b^2} = \frac{1}{a} = \left(\{ \mathbf{F}(\theta) \}_{1,1} \right)^{-1}$$


Si et seulement si découplage ($c=0$)

On peut aussi invoquer Jensen...

Bornes de Cramér-Rao « euclidiennes »

Il reste 2 questions importantes !!! Rappelons nous que



Trouver une « limite » indépendante de l'estimateur (seulement du modèle des observations), « proche » de la variance de l'estimateur et calculable facilement.

Bornes de Cramér-Rao « euclidiennes »

Réponse 1 : avec les conditions de régularités du début + une autre, l'estimateur du **maximum de vraisemblance** vérifie

$$\sqrt{T} (\hat{\theta} - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{F}^{-1}(\theta))$$

Réponse 2 (formule de Slepian-Bang) : si les observations sont gaussiennes à moyenne et covariance paramétrées $y \sim \mathcal{N}(\mu(\theta), \Gamma(\theta))$

$$\{F(\theta)\}_{i,j} = \frac{\partial \mu^T(\theta)}{\partial \theta_i} \Gamma^{-1}(\theta) \frac{\partial \mu(\theta)}{\partial \theta_j} + \frac{1}{2} \text{Tr} \left\{ \Gamma^{-1}(\theta) \frac{\partial \Gamma(\theta)}{\partial \theta_i} \Gamma^{-1}(\theta) \frac{\partial \Gamma(\theta)}{\partial \theta_j} \right\}$$

Bornes de Cramér-Rao « euclidiennes »

Quelques remarques sur **les** bornes de Cramér-Rao

- Borne de Cramér-Rao bayésienne et hybride
- Prise en compte du biais ou d'une transformation des paramètres
- Borne de Cramér-Rao pour des modèles mal posés
- Borne de Cramér-Rao semi-paramétrique
- Formule de Slepian-Bang autre que pour des données gaussiennes et application à une pléthore de problèmes
- Planification d'expérience et matrice d'information de Fisher
- Borne de Cramér-Rao pour l'estimation d'état
- Borne de Cramér-Rao quantique
- Bornes autres que Cramér-Rao
- Bornes « périodiques »
- Borne de Cramér-Rao-Leibnitz
- **Borne de Cramér-Rao sous contraintes**
- **Borne de Cramér-Rao intrinsèque**

I Problématique

II Bornes de Cramér-Rao « euclidiennes »

✓ III Borne de Cramér-Rao intrinsèque – Généralités

IV Borne de Cramér-Rao intrinsèque – Estimation d'une matrice de covariance

V Perspectives

Borne de Cramér-Rao : prise en compte des contraintes entre les paramètres



Si les éléments du vecteur de paramètre $\boldsymbol{\theta}$ sont liés par des équations de la forme

exemple
$$\begin{cases} \theta_1 + \theta_2 = 5 \\ \theta_1 + \frac{1}{\theta_3} = 12 \\ \dots \end{cases} \quad \begin{aligned} h_k(\theta_1, \theta_2, \dots, \theta_P) &= 0 \\ k &\in 1, \dots, M \end{aligned}$$

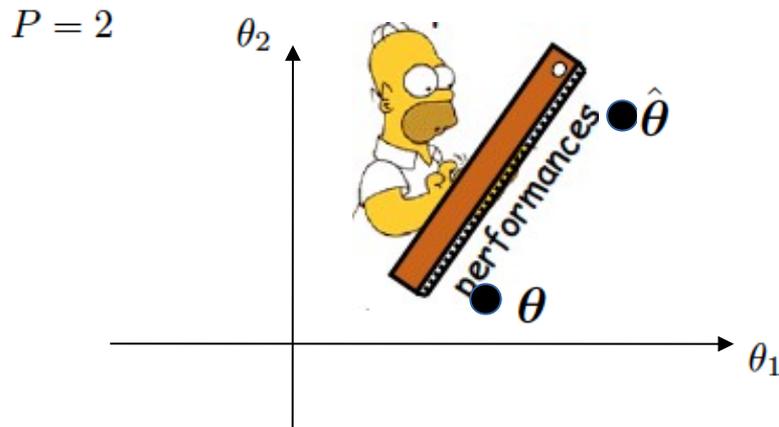
Problème d'identifiabilité, matrice d'information de Fisher non inversible

Soient $\mathbf{h}(\boldsymbol{\theta}) = \mathbf{0} \in \mathbb{R}^M$ et $\mathbf{H}(\boldsymbol{\theta}) = \frac{\partial \mathbf{h}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \in \mathbb{R}^{M \times P}$ alors il existe une matrice $\mathbf{U}(\boldsymbol{\theta}) \in \mathbb{R}^{P \times M}$ tel que $\mathbf{H}(\boldsymbol{\theta}) \mathbf{U}(\boldsymbol{\theta}) = \mathbf{0}$ et $\mathbf{U}^T(\boldsymbol{\theta}) \mathbf{U}(\boldsymbol{\theta}) = \mathbf{I}_M$ et on peut prouver l'inégalité matricielle :

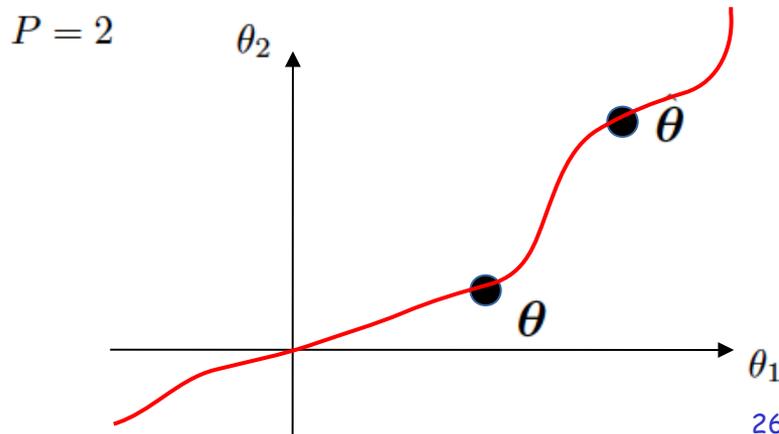
$$\mathbb{E} \left((\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T \right) \succeq \mathbf{U}(\boldsymbol{\theta}) (\mathbf{U}^T(\boldsymbol{\theta}) \mathbf{F}(\boldsymbol{\theta}) \mathbf{U}(\boldsymbol{\theta}))^{-1} \mathbf{U}^T(\boldsymbol{\theta})$$

iCRB – Généralités

Jusqu'ici, le vecteur de paramètre appartenait à \mathbb{R}^P (ou un ouvert de \mathbb{R}^P) et la distance entre $\hat{\theta}$ et θ était la ligne droite classique dans le plan.



Si on contraint les éléments de θ (et donc de $\hat{\theta}$) à rester sur une courbe (variété)



Redéfinir :

- Distance
- Biais
- Variance
- Inégalité de Cramér-Rao

Bibliographie

S.T. Smith *IEEE TSP* 2005 « Covariance, subspace, and instrinsic Cramér-Rao bounds »

N. Boumal PhD Thesis 2014 « Optimization and estimation on manifolds »

et

- Harrie Hendriks, “A Cramér-Rao type lower bound for estimators with values in a manifold,” *Journal of Multivariate Analysis*, vol. 38, no. 2, pp. 245–261, 1991.
- Joao Xavier and Victor Barroso, “Intrinsic distance lower bound for unbiased estimators on riemannian manifolds,” in *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*. IEEE, 2002, vol. 2, pp. II–1141.
- Joao Xavier and Victor Barroso, “Intrinsic variance lower bound (IVLB): an extension of the Cramér-Rao bound to riemannian manifolds,” in *Acoustics, Speech, and Signal Processing (ICASSP), 2005, IEEE International Conference on*. IEEE, 2005, vol. 5, pp.v–1033.
- Nicolas Boumal, “On intrinsic Cramér-Rao bounds for riemannian submanifolds and quotient manifolds,” *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1809–1821, 2013.
- Axel Barrau and Silvere Bonnabel, “A note on the intrinsic Cramér-Rao bound,” in *Geometric Science of Information*, pp. 377–386. Springer, 2013.
- Shun-ichi Amari and Hiroshi Nagaoka, *Methods of information geometry*, vol. 191, American Mathematical Soc., 2007.
- Silvère Bonnabel, Axel Barrau An intrinsic Cramér-Rao bound on Lie groups, *Conference Geometric Sciences of Information GSI15*
- J. M. Oller and J. M. Corcuera Intrinsic Analysis of Statistical Estimation *The Annals of Statistics* Vol. 23, No. 5 (Oct., 1995), pp. 1562-1581
- P.E.Jupp A van Trees inequality for estimators on manifolds , *Journal of Multivariate Analysis* Volume 101, Issue 8, September 2010, Pages 1814-1825

Variétés (différentielles) : \mathcal{M}

espace **topologique** localement **euclidien** (calcul différentiel et intégral)

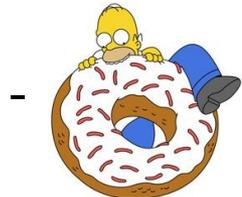
→ Limite
Continuité

→ On a pas tout perdu...

Riemannienne : munie d'une structure supplémentaire (métrique) permettant de calculer le produit scalaire de deux vecteurs tangents à la variété en un même point.

Exemples de variétés

- la sphère
- l'espace des matrices (symétriques) définies positives (matrices de covariances)
- les sous-espaces



- (pendule double)

- Groupe spécial orthogonal (calibration de réseaux de caméra)

$$\text{SO}(n) = \{R \in \mathbb{R}^{n \times n} : R^T R = I \text{ and } \det R = 1\}$$

- etc.

Remarque : difficile (impossible) à traiter avec la borne de Cramér-Rao sous contraintes et, en plus ce n'est pas *naturel*

Notations et définitions

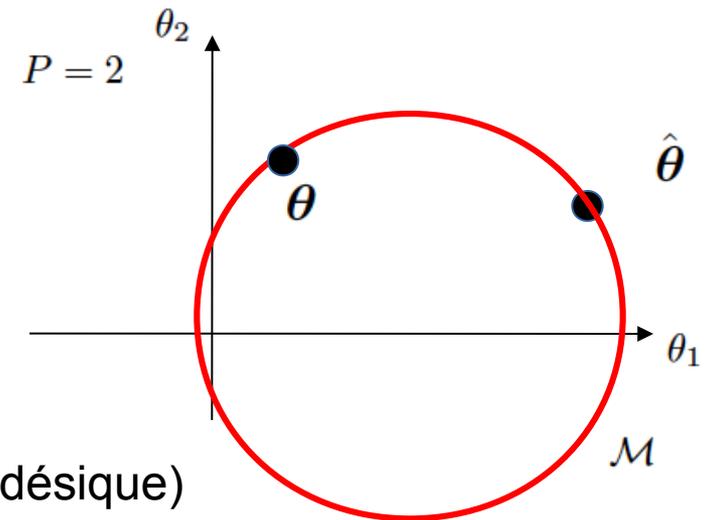
Désormais $\theta \in \mathcal{M}$ et $\hat{\theta} \in \mathcal{M}$

On a toujours $y \sim f_{\theta}(y) \longrightarrow \hat{\theta}$

● Erreur d'estimation

On veut aller de θ à $\hat{\theta}$ en se promenant
« intelligemment » sur \mathcal{M}

Intelligemment = suivre le chemin le plus court (géodésique)



Notations et définitions

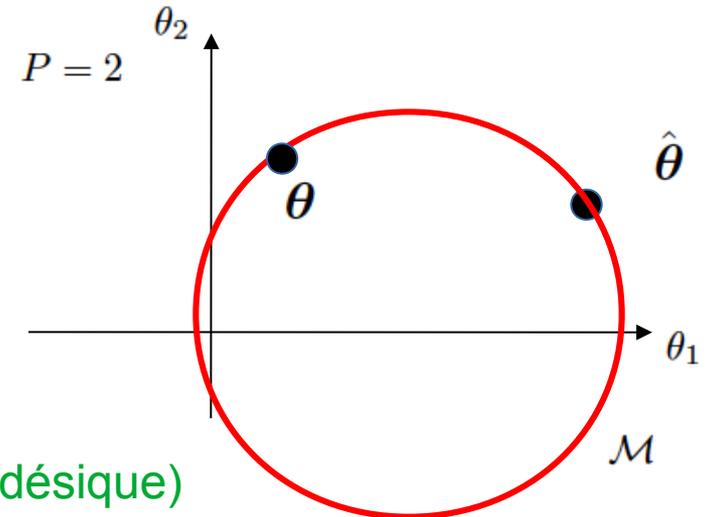
Désormais $\theta \in \mathcal{M}$ et $\hat{\theta} \in \mathcal{M}$

On a toujours $y \sim f_{\theta}(y) \longrightarrow \hat{\theta}$

● Erreur d'estimation

On veut aller de θ à $\hat{\theta}$ en se promenant
« intelligemment » sur \mathcal{M}

Intelligemment = suivre le chemin le plus court (géodésique)



Unicité, etc.

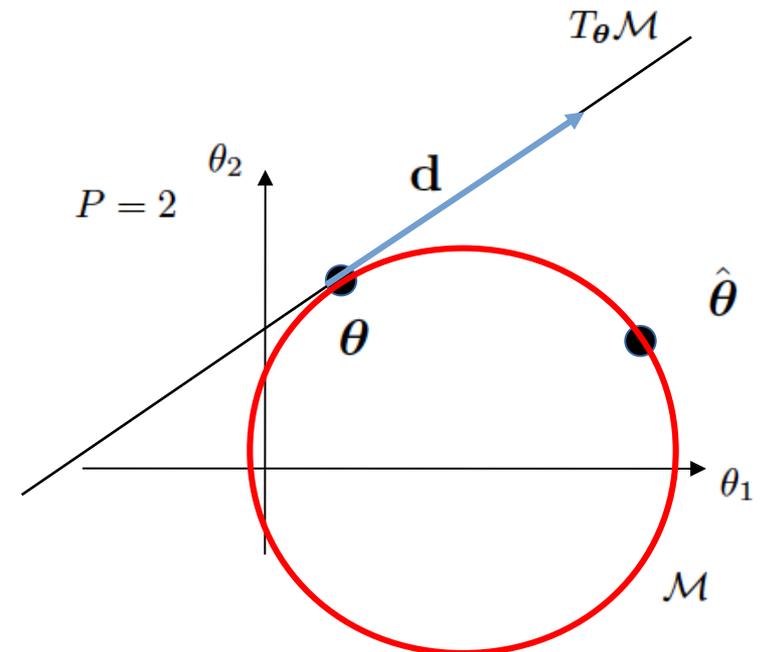
Notations et définitions

Désormais $\theta \in \mathcal{M}$ et $\hat{\theta} \in \mathcal{M}$

On a toujours $y \sim f_{\theta}(y) \longrightarrow \hat{\theta}$

● Erreur d'estimation

On veut aller de θ à $\hat{\theta}$ en se promenant
« intelligemment » sur \mathcal{M}



d vecteur de l'espace tangent (espace vectoriel facile pour travailler (linéarité))
 $T_{\theta}\mathcal{M}$ qui va nous indiquer l'« intensité » qu'il faut utiliser (vitesse)

Notations et définitions

Désormais $\theta \in \mathcal{M}$ et $\hat{\theta} \in \mathcal{M}$

On a toujours $y \sim f_{\theta}(y) \xrightarrow{\text{red arrow}} \hat{\theta}$

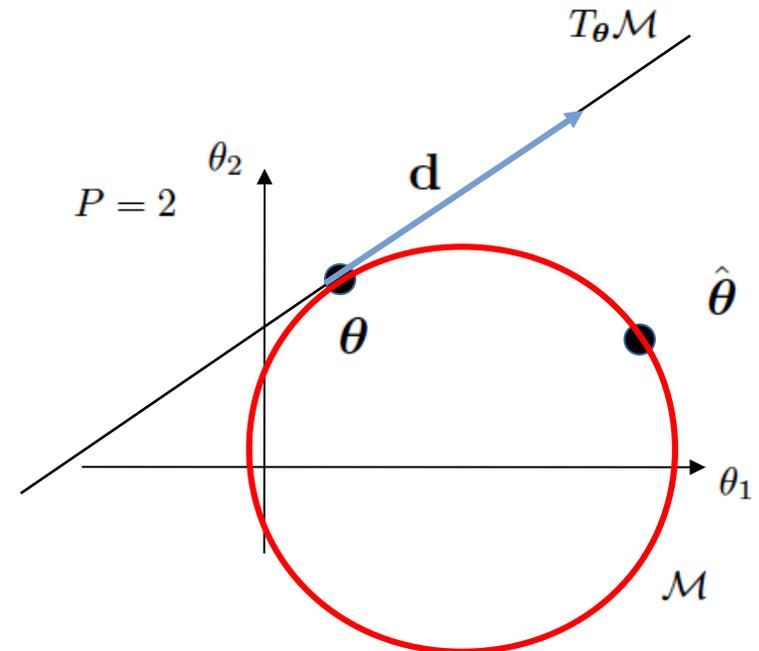
● Erreur d'estimation

On veut aller de θ à $\hat{\theta}$ en se promenant
« intelligemment » sur \mathcal{M}

$e_{\theta} = \exp_{\theta}^{-1}(\hat{\theta})$ application exponentielle inverse (ou application logarithmique) ou comment comparer 2 vecteurs sur une variété

$$\exp_{\theta}^{-1}(\cdot) : \mathcal{M} \rightarrow T_{\theta}\mathcal{M}$$

$$\exp_{\theta}^{-1}(\hat{\theta}) \rightarrow \mathbf{d} \quad \text{tel que} \quad \|\mathbf{d}\|_{\theta} = \text{dist}(\theta, \hat{\theta})$$



Notations et définitions

Variété \mathcal{M}

Espace tangent $T_{\theta}\mathcal{M}$ au point θ

Métrique = produit scalaire sur $T_{\theta}\mathcal{M}$ $\forall \theta \in \mathcal{M}$ $g_{\theta}(\cdot, \cdot) = \langle \cdot, \cdot \rangle_{\theta}$

→ géodésique, distance

Forme bilinéaire, symétrique et définie positive

$\{\Omega_i\}$ Base de $T_{\theta}\mathcal{M}$ orthonormé par rapport à g_{θ}

- Erreur d'estimation (avec coordonnées)

$$e_{\theta} = \exp_{\theta}^{-1}(\hat{\theta})$$

Le vecteur d'erreur $e(\theta)$ a pour éléments $[e(\theta)]_i = g_{\theta}(e_{\theta}, \Omega_i)$

$$\text{et } e^T(\theta) e(\theta) = \text{dist}^2(\theta, \hat{\theta})$$

Notations et définitions

● Biais intrinsèque $\mathbf{b}(\boldsymbol{\theta}) = \mathbb{E}(\mathbf{e}_{\boldsymbol{\theta}}) = \mathbb{E}\left(\exp_{\boldsymbol{\theta}}^{-1}(\hat{\boldsymbol{\theta}})\right) \in T_{\boldsymbol{\theta}}\mathcal{M}$

Avec le système de coordonnées $\{\boldsymbol{\Omega}_i\}$ et par rapport à $g_{\boldsymbol{\theta}}(\cdot, \cdot)$

$$[\mathbf{b}(\boldsymbol{\theta})]_i = \mathbb{E}([e(\boldsymbol{\theta})]_i) = \mathbb{E}(g_{\boldsymbol{\theta}}(\mathbf{e}_{\boldsymbol{\theta}}, \boldsymbol{\Omega}_i))$$

L'estimateur est non biaisé si

$$\forall \boldsymbol{\theta} \in \mathcal{M}, \mathbf{b}(\boldsymbol{\theta}) \equiv \mathbf{0}$$

Dans S.T. Smith *IEEE TSP* 2005 « Covariance, subspace, and intrinsic Cramér-Rao bounds » tout les calculs sont menés pour un biais quelconque

Notations et définitions

● Covariance intrinsèque

Si $\hat{\theta}$ non biaisé

$$\mathbf{C}(\boldsymbol{\theta}) = \mathbb{E}(\mathbf{e}_{\boldsymbol{\theta}} \mathbf{e}_{\boldsymbol{\theta}}^T) = \mathbb{E} \left(\exp_{\boldsymbol{\theta}}^{-1}(\hat{\boldsymbol{\theta}}) \left(\exp_{\boldsymbol{\theta}}^{-1}(\hat{\boldsymbol{\theta}}) \right)^T \right)$$

Avec le système de coordonnées $\{\Omega_i\}$ et par rapport à $g_{\boldsymbol{\theta}}(\cdot, \cdot)$

$$[\mathbf{C}(\boldsymbol{\theta})]_{i,j} = \mathbb{E}(g_{\boldsymbol{\theta}}(\mathbf{e}_{\boldsymbol{\theta}}, \Omega_i) g_{\boldsymbol{\theta}}(\mathbf{e}_{\boldsymbol{\theta}}, \Omega_j))$$

Remarque

$$\text{tr}(\mathbf{C}(\boldsymbol{\theta})) = \sum_i \mathbb{E}(g_{\boldsymbol{\theta}}^2(\mathbf{e}_{\boldsymbol{\theta}}, \Omega_i)) = \mathbb{E}(e^T(\boldsymbol{\theta}) e(\boldsymbol{\theta})) = \mathbb{E}(\text{dist}^2(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}))$$

C'est ce qu'on va chercher à borner

Mais il nous manque un ingrédient !!!

Notations et définitions

● Métrique d'information de Fisher

$$\forall \Omega \in T_{\theta} \mathcal{M} \quad g_{\theta}^{fisher}(\Omega, \Omega) = -\mathbb{E} \left(\left. \frac{d^2}{dt^2} \ln f_{\theta+t\Omega}(\mathbf{y}) \right|_{t=0} \right)$$

Identité de polarisation (norme \longrightarrow produit scalaire)

$$g_{\theta}^{fisher}(\Omega_i, \Omega_j) = \frac{1}{4} \left(g_{\theta}^{fisher}(\Omega_i + \Omega_j, \Omega_i + \Omega_j) - g_{\theta}^{fisher}(\Omega_i - \Omega_j, \Omega_i - \Omega_j) \right)$$

● Matrice d'information de Fisher

$\{\Omega_i\}$ Base de $T_{\theta} \mathcal{M}$ orthonormé **par rapport à g_{θ} !!!**

$$[\mathbf{F}(\theta)]_{i,j} = g_{\theta}^{fisher}(\Omega_i, \Omega_j)$$

Remarques

- Il s'agit bien d'une métrique (forme bilinéaire définie positive)
- L'erreur est mesurée par rapport à la métrique g_{θ} et pas nécessairement g_{θ}^{fisher}

Bornes de Cramér-Rao intrinsèques

Smith 05

Courbure Riemannienne

$$\mathbf{C} \geq \mathbf{M}_b \mathbf{G}^{-1} \mathbf{M}_b^T - \frac{1}{3} (\mathbf{R}_m(\mathbf{M}_b \mathbf{G}^{-1} \mathbf{M}_b^T) \mathbf{G}^{-1} \mathbf{M}_b^T + \mathbf{M}_b \mathbf{G}^{-1} \mathbf{R}_m(\mathbf{M}_b \mathbf{G}^{-1} \mathbf{M}_b^T)^T)$$

avec $\mathbf{M}_b = \mathbf{I} - \frac{1}{3} \|b\|^2 \mathbf{K}(b) + \nabla b$

Dérivé covariante

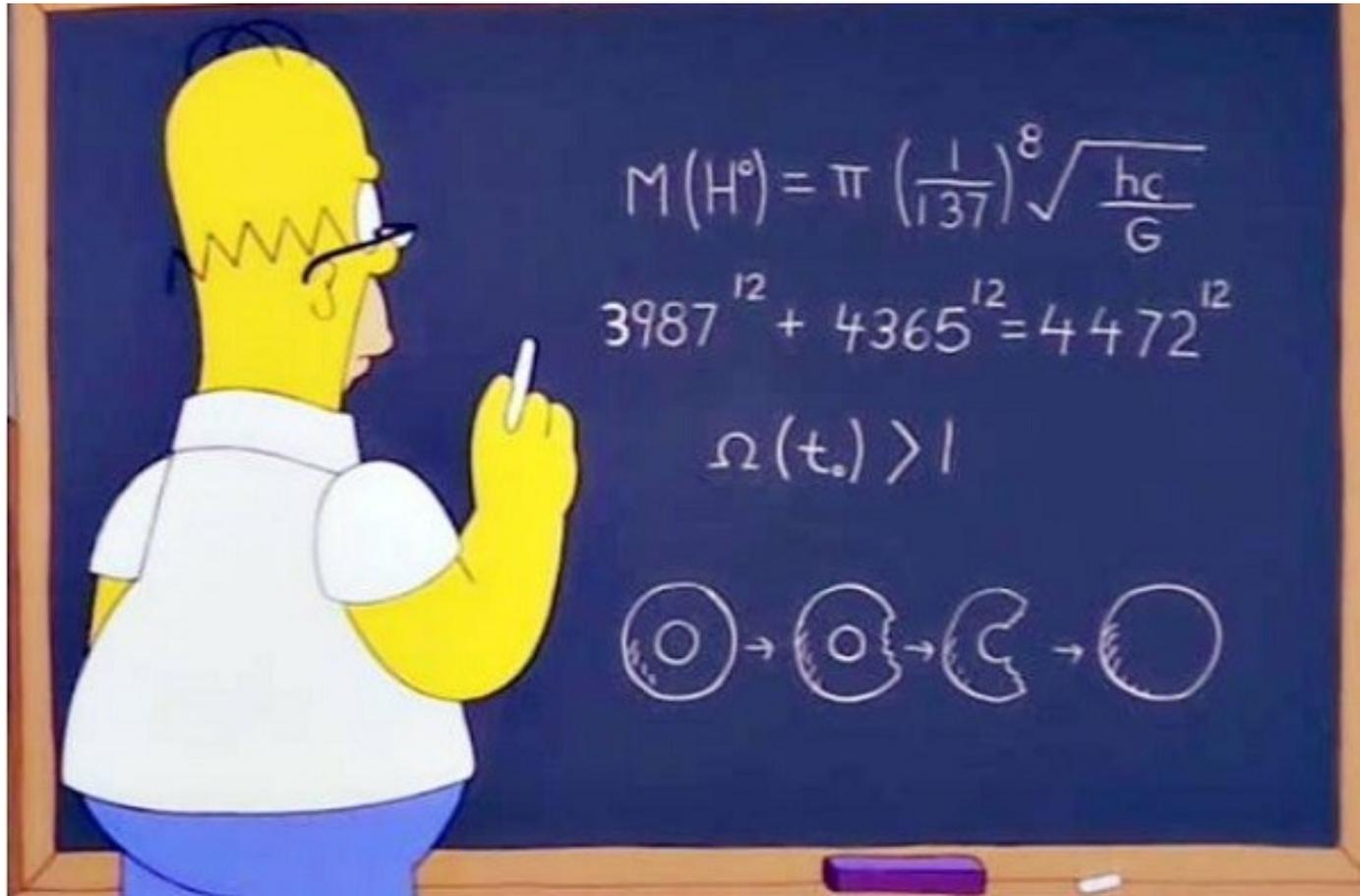
Courbure sectionnelle

Boumal 14
$$\mathbf{C} \succeq \mathbf{F}^{-1} - \frac{1}{3} (\mathbf{F}^{-1} \mathbf{R}_m(\mathbf{F}^{-1}) + \mathbf{R}_m(\mathbf{F}^{-1}) \mathbf{F}^{-1})$$

Si on néglige tout
$$\mathbf{C}(\boldsymbol{\theta}) \succeq \mathbf{F}^{-1}(\boldsymbol{\theta})$$

Et en particulier
$$tr(\mathbf{C}(\boldsymbol{\theta})) = \mathbb{E} \left(dist^2(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) \right) \geq tr(\mathbf{F}^{-1}(\boldsymbol{\theta}))$$

Preuve



Recette de cuisine pour calculer une borne de Cramér-Rao intrinsèque

1. Calcul analytique de la métrique de Fisher puis polarisation de la métrique

$$g_{\boldsymbol{\theta}}^{fisher}(\boldsymbol{\Omega}, \boldsymbol{\Omega}) = -\mathbb{E} \left(\left. \frac{d^2}{dt^2} \ln f_{\boldsymbol{\theta}+t\boldsymbol{\Omega}}(\mathbf{y}) \right|_{t=0} \right)$$

$$g_{\boldsymbol{\theta}}^{fisher}(\boldsymbol{\Omega}_i, \boldsymbol{\Omega}_j) = \frac{1}{4} \left(g_{\boldsymbol{\theta}}^{fisher}(\boldsymbol{\Omega}_i + \boldsymbol{\Omega}_j, \boldsymbol{\Omega}_i + \boldsymbol{\Omega}_j) - g_{\boldsymbol{\theta}}^{fisher}(\boldsymbol{\Omega}_i - \boldsymbol{\Omega}_j, \boldsymbol{\Omega}_i - \boldsymbol{\Omega}_j) \right)$$

2. Passage à la matrice d'information de Fisher (choix d'une base de $T_{\boldsymbol{\theta}}\mathcal{M} : \{\boldsymbol{\Omega}_1, \dots, \boldsymbol{\Omega}_n\}$)

$$[\mathbf{F}(\boldsymbol{\theta})]_{i,j} = g_{\boldsymbol{\theta}}^{fisher}(\boldsymbol{\Omega}_i, \boldsymbol{\Omega}_j)$$

2 bis. Passage de la métrique à la distance naturelle (géodésique) sur la variété \mathcal{M}

$$\gamma : [a, b] \rightarrow \mathcal{M}$$

$$\boldsymbol{\theta}_a, \boldsymbol{\theta}_b \in \mathcal{M}$$

$$dist(\boldsymbol{\theta}_a, \boldsymbol{\theta}_b) = \inf_{\gamma} \int_a^b \sqrt{g_{fim}(\gamma'(t), \gamma'(t))} dt$$

Calcul direct, équation des géodésiques (symboles de Christoffel), connexion de Levi-Civita.

3. Inversion et calcul de la trace (si possible) $\mathbb{E} \left(dist^2(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) \right) \geq tr(\mathbf{F}^{-1}(\boldsymbol{\theta}))$

I Problématique

II Bornes de Cramér-Rao « euclidiennes »

III Borne de Cramér-Rao intrinsèque – Généralités

✓ IV Borne de Cramér-Rao intrinsèque – Estimation d'une matrice de covariance

V Perspectives

iCRB – Matrice de covariance

● Modèle des observations $\mathbf{y}(t) = \mathbf{b}(t) \sim \mathcal{N}(0, \Sigma), t = 1, \dots, T$

$$b(t) \perp b(t') \text{ pour } t \neq t'$$

$$T > N$$

$\hat{\Sigma} = ?$ Problème central en traitement du signal

Détection Radar

Méthode MUSIC (traitement d'antenne)

Séparation de sources

Imagerie SAR

Calibration Radioastronomie

Etc.

Paramètres $\Sigma \in \mathcal{M} = \mathcal{S}_N^{++}$ Espace des matrices symétriques $N \times N$ définies positives

iCRB – Matrice de covariance

- Modèle des observations $\mathbf{y}(t) = \mathbf{b}(t) \sim \mathcal{N}(0, \Sigma)$, $t = 1, \dots, T$
 $b(t) \perp b(t')$ pour $t \neq t'$
- Estimateur (maximum de vraisemblance) $\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \mathbf{y}(t) \mathbf{y}^T(t)$
- Biais (classique) $\mathbb{E}(\hat{\Sigma}) = \frac{1}{T} \sum_{t=1}^T \mathbb{E}(\mathbf{y}(t) \mathbf{y}^T(t)) = \Sigma$
- Covariance (classique) $tr(\mathbf{C}(\theta)) = \frac{1}{T} (tr(\Sigma^2) + tr^2(\Sigma))$
 $\theta = vec(\Sigma)$
- Borne de Cramér-Rao $tr(\mathbf{F}^{-1}(\theta)) = \frac{1}{T} (tr(\Sigma^2) + tr^2(\Sigma))$



L'estimateur du maximum de vraisemblance est sans biais et efficace **à distance finie**

iCRB – Matrice de covariance

Démonstrations

Estimateur

$$f_{\Sigma}(\mathbf{y}(t)) = \frac{1}{(2\pi)^{N/2} |\Sigma|^{1/2}} e^{-\frac{1}{2} \mathbf{y}^T(t) \Sigma^{-1} \mathbf{y}(t)}$$

$$f_{\Sigma}(\mathbf{y}(1), \dots, \mathbf{y}(T)) = \prod_{t=1}^T f_{\Sigma}(\mathbf{y}(t)) = \frac{1}{(2\pi)^{NT/2} |\Sigma|^{T/2}} e^{-\frac{1}{2} \sum_{t=1}^T \mathbf{y}^T(t) \Sigma^{-1} \mathbf{y}(t)}$$

$$\ln f_{\Sigma}(\mathbf{y}(1), \dots, \mathbf{y}(T)) = Cst - \frac{T}{2} \ln |\Sigma| - \frac{1}{2} \sum_{t=1}^T \mathbf{y}^T(t) \Sigma^{-1} \mathbf{y}(t)$$

Equation de vraisemblance $\left. \frac{\partial \ln f_{\Sigma}(\mathbf{y}(1), \dots, \mathbf{y}(T))}{\partial \text{vec}(\Sigma)} \right|_{\hat{\Sigma}} = 0$

$$\begin{aligned} \partial \ln f_{\Sigma}(\mathbf{y}(1), \dots, \mathbf{y}(T)) &= -\frac{T}{2} \partial \ln |\Sigma| - \frac{1}{2} \sum_{t=1}^T \mathbf{y}^T(t) \partial (\Sigma^{-1}) \mathbf{y}(t) \\ &= -\frac{T}{2} \text{vec}^T(\Sigma^{-1}) \partial \text{vec}(\Sigma) + \frac{1}{2} \sum_{t=1}^T \text{vec}^T(\Sigma^{-1} \mathbf{y}(t) \mathbf{y}^T(t) \Sigma^{-1}) \partial \text{vec}(\Sigma) \end{aligned}$$

$$\begin{aligned} \sum_{t=1}^T \text{vec}(\hat{\Sigma}^{-1} \mathbf{y}(t) \mathbf{y}^T(t) \hat{\Sigma}^{-1}) &= T \text{vec}(\hat{\Sigma}^{-1}) \iff \sum_{t=1}^T (\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1}) \text{vec}(\mathbf{y}(t) \mathbf{y}^T(t)) = T \text{vec}(\hat{\Sigma}^{-1}) \\ &\iff \frac{1}{T} \sum_{t=1}^T \text{vec}(\mathbf{y}(t) \mathbf{y}^T(t)) = (\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1})^{-1} \text{vec}(\hat{\Sigma}^{-1}) = \text{vec}(\hat{\Sigma} \hat{\Sigma}^{-1} \hat{\Sigma}) \end{aligned}$$

Formules utiles

$$\partial \text{Tr}(\mathbf{X}) = \text{Tr}(\partial \mathbf{X})$$

$$\partial \text{vec}(\mathbf{X}) = \text{vec}(\partial \mathbf{X})$$

$$\partial \mathbf{A}^{-1} = -\mathbf{A}^{-1} \partial \mathbf{A} \mathbf{A}^{-1}$$

$$\partial \ln(|\mathbf{M}|) = \text{Tr}(\mathbf{M}^{-1} \partial \mathbf{M})$$

$$\text{Tr}(\mathbf{A}\mathbf{B}) = \text{Tr}(\mathbf{B}\mathbf{A})$$

$$\mathbf{M}^{-1} \otimes \mathbf{M}^{-1} = (\mathbf{M} \otimes \mathbf{M})^{-1}$$

$$\text{Tr}(\mathbf{A}^H \mathbf{B}) = \text{vec}^H(\mathbf{A}) \text{vec}(\mathbf{B})$$

$$\text{vec}(\mathbf{A}\mathbf{B}\mathbf{C}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B})$$

iCRB – Matrice de covariance

● Covariance (attention prise en compte de la redondance)

$$\begin{aligned} \text{tr}(\mathbf{C}(\boldsymbol{\theta})) &= \text{tr}\left(\mathbb{E}\left(\text{vec}\left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right)\text{vec}^T\left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right)\right)\right) = \mathbb{E}\left(\text{tr}\left(\text{vec}^T\left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right)\text{vec}\left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right)\right)\right) \\ &= \mathbb{E}\left(\text{vec}^T\left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right)\text{vec}\left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right)\right) = \mathbb{E}\left(\text{tr}\left(\left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right)^T\left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right)\right)\right) = \text{tr}\left(\mathbb{E}\left(\hat{\boldsymbol{\Sigma}}^2\right) - \boldsymbol{\Sigma}^2\right) \end{aligned}$$

$$\mathbb{E}\left(\hat{\boldsymbol{\Sigma}}^2\right) = \frac{1}{T^2} \sum_{t=1}^T \sum_{t'=1}^T \mathbb{E}\left(\mathbf{y}(t)\mathbf{y}^T(t)\mathbf{y}(t')\mathbf{y}^T(t')\right)$$

$$\mathbb{E}\left(\mathbf{y}(t)\mathbf{y}^T(t)\mathbf{y}(t')\mathbf{y}^T(t')\right) = \boldsymbol{\Sigma}^2 \text{ si } t \neq t'$$

$$\mathbb{E}\left(\mathbf{y}(t)\mathbf{y}^T(t)\mathbf{y}(t)\mathbf{y}^T(t)\right) = 2\boldsymbol{\Sigma}^2 + \text{tr}(\boldsymbol{\Sigma})\boldsymbol{\Sigma} \text{ si } t = t'$$

$$\text{tr}(\mathbf{C}(\boldsymbol{\theta})) = \frac{1}{T^2} \text{tr}\left(\left(T\left(2\boldsymbol{\Sigma}^2 + \text{tr}(\boldsymbol{\Sigma})\boldsymbol{\Sigma}\right) + T(T-1)\boldsymbol{\Sigma}^2\right) - \boldsymbol{\Sigma}^2\right) = \frac{1}{T} \left(\text{tr}(\boldsymbol{\Sigma}^2) + \text{tr}^2(\boldsymbol{\Sigma})\right)$$

● Borne de Cramér-Rao :

on s'en fiche va la retrouver tout à l'heure
(ou Slepian-Bang ou différentielle ordre 2)

Borne de Cramér-Rao intrinsèque

Métrique de Fisher $g_{\boldsymbol{\theta}}^{fisher}(\boldsymbol{\Omega}, \boldsymbol{\Omega}) = -\mathbb{E} \left(\left. \frac{d^2}{dt^2} \ln f_{\boldsymbol{\theta}+t\boldsymbol{\Omega}}(\mathbf{y}) \right|_{t=0} \right) \quad \boldsymbol{\theta} = \boldsymbol{\Sigma}$

$$\ln f_{\boldsymbol{\theta}+t\boldsymbol{\Omega}}(\mathbf{y}) = -T \left(\ln |\boldsymbol{\theta} + t\boldsymbol{\Omega}| + \text{tr} \left((\boldsymbol{\theta} + t\boldsymbol{\Omega})^{-1} \mathbf{Y} \right) \right)$$

$$f(\boldsymbol{\Sigma} + \partial\boldsymbol{\Sigma}) = f(\boldsymbol{\Sigma}) + \partial f(\boldsymbol{\Sigma}) + \frac{1}{2} \partial^2 f(\boldsymbol{\Sigma}) \quad f(\cdot) = \ln |\cdot| \quad \text{et} \quad f(\cdot) = \text{tr}(\cdot)$$

$$\partial \ln |\boldsymbol{\Sigma}| = \text{tr}(\boldsymbol{\Sigma}^{-1} \partial \boldsymbol{\Sigma})$$

$$\partial^2 \ln |\boldsymbol{\Sigma}| = \partial \text{tr}(\boldsymbol{\Sigma}^{-1} \partial \boldsymbol{\Sigma}) = \text{tr}(\partial \boldsymbol{\Sigma}^{-1} \partial \boldsymbol{\Sigma} + \boldsymbol{\Sigma}^{-1} \partial^2 \boldsymbol{\Sigma}) = -\text{tr} \left((\boldsymbol{\Sigma}^{-1} \partial \boldsymbol{\Sigma})^2 \right)$$

$$\partial \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Y}) = -\text{tr}(\partial \boldsymbol{\Sigma} \boldsymbol{\Sigma}^{-1} \mathbf{Y} \boldsymbol{\Sigma}^{-1})$$

$$\partial^2 \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Y}) = -\partial \text{tr}(\partial \boldsymbol{\Sigma} \boldsymbol{\Sigma}^{-1} \mathbf{Y} \boldsymbol{\Sigma}^{-1}) = \dots = 2 \text{tr} \left((\boldsymbol{\Sigma}^{-1} \partial \boldsymbol{\Sigma})^2 \mathbf{Y} \boldsymbol{\Sigma}^{-1} \right)$$

Borne de Cramér-Rao intrinsèque

Métrique de Fisher

$$\begin{aligned} \ln f_{\boldsymbol{\theta}+t\boldsymbol{\Omega}}(\mathbf{y}) &= -T \left(\text{tr}(\mathbf{Y}\boldsymbol{\Sigma}^{-1}) - \text{tr}(t\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}\mathbf{Y}\boldsymbol{\Sigma}^{-1}) + \text{tr}\left((t\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1})^2\mathbf{Y}\boldsymbol{\Sigma}^{-1}\right) \right) \\ &\quad + \ln|\boldsymbol{\Sigma}| + \text{tr}(t\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}) - \frac{1}{2}\text{tr}\left((t\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1})^2\right) \end{aligned}$$

$$\left. \frac{d^2}{dt^2} \ln f_{\boldsymbol{\theta}+t\boldsymbol{\Omega}}(\mathbf{y}) \right|_{t=0} = T \left(\text{tr}\left((\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1})^2\right) - 2\text{tr}\left((\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1})^2\mathbf{Y}\boldsymbol{\Sigma}^{-1}\right) \right)$$

$$g_{\boldsymbol{\theta}}^{\text{fisher}}(\boldsymbol{\Omega}, \boldsymbol{\Omega}) = T \text{tr}\left((\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1})^2\right)$$

iCRB – Matrice de covariance

Borne de Cramér-Rao intrinsèque

Identité de polarisation

$$\begin{aligned} g_{\theta}^{fisher}(\Omega_i, \Omega_j) &= \frac{1}{4} \left(g_{\theta}^{fisher}(\Omega_i + \Omega_j, \Omega_i + \Omega_j) - g_{\theta}^{fisher}(\Omega_i - \Omega_j, \Omega_i - \Omega_j) \right) \\ &= \frac{T}{4} \left(\text{tr} \left(((\Omega_i + \Omega_j) \Sigma^{-1})^2 \right) - \text{tr} \left(((\Omega_i - \Omega_j) \Sigma^{-1})^2 \right) \right) = T \text{tr}(\Omega_i \Sigma^{-1} \Omega_j \Sigma^{-1}) \end{aligned}$$

Passage à la matrice d'information de Fisher (choix d'une base de $T_{\theta}\mathcal{M}$
par rapport à une métrique qcq)

L'espace tangent des matrices symétriques définies positives est l'ensemble des matrices symétriques

iCRB – Matrice de covariance

Borne de Cramér-Rao intrinsèque

Passage à la matrice d'information de Fisher (choix d'une base de $T_{\theta}\mathcal{M}$)

$\{\Omega_1, \dots, \Omega_n\}$ avec $n = \frac{N(N+1)}{2}$

// Métrique euclidienne

$$g_{\theta}^{euclid}(\Omega_1, \Omega_2) = Tr(\Omega_1 \Omega_2)$$

D_{ll}^{diag} Matrice symétrique $N \times N$ dont le $l^{i\grave{e}me}$ élément diagonal égal 1 (zéro ailleurs)

$D_{lk}^{réel}$ Matrice symétrique $N \times N$ dont les éléments (l, k) et (k, l) sont égaux à 1

(zéro ailleurs) $l < k$

Exemple : $N = 2$ $n = 3$

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \sqrt{2}c \begin{pmatrix} 0 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 0 \end{pmatrix}$$

D_{11}^{diag}
 Ω_1

D_{22}^{diag}
 Ω_2

$D_{12}^{réel}$
 Ω_3

Borne de Cramér-Rao intrinsèque

Base orthonormal en Σ // métrique Riemannienne

$$g_{\theta}^{fisher}(\Omega_i, \Omega_j) = Ttr(\Omega_i \Sigma^{-1} \Omega_j \Sigma^{-1})$$

$$\tilde{\mathbf{D}}_{i,j}^? = \Sigma^{1/2} \mathbf{D}_{i,j}^? \Sigma^{1/2}$$

Borne de Cramér-Rao intrinsèque

Distance naturelle $dist_{nat}(\Sigma_1, \Sigma_2) = \sqrt{\sum_k \ln^2 \lambda_k}$

λ_k valeurs propres obtenue par $|\lambda \Sigma_1 - \Sigma_2| = 0$

Distance Euclidienne

$$dist_{euclid}(\Sigma_1, \Sigma_2) = \sqrt{tr\left((\Sigma_1 - \Sigma_2)^2\right)} = \|\Sigma_1 - \Sigma_2\|_{Fro}$$

Borne de Cramér-Rao intrinsèque

Passage à la matrice d'information de Fisher $[\mathbf{F}(\boldsymbol{\theta})]_{i,j} = g_{\boldsymbol{\theta}}^{fisher}(\boldsymbol{\Omega}_i, \boldsymbol{\Omega}_j)$

$$g_{\boldsymbol{\theta}}^{fisher}(\boldsymbol{\Omega}_i, \boldsymbol{\Omega}_j) = T \text{tr}(\boldsymbol{\Omega}_i \boldsymbol{\Sigma}^{-1} \boldsymbol{\Omega}_j \boldsymbol{\Sigma}^{-1})$$

$$\longrightarrow \mathbf{F}(\boldsymbol{\theta}) = \frac{T}{2} \mathbf{I}_{\frac{N(N+1)}{2}}$$

$$\mathbb{E}(\text{dist}_{nat}(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)) \geq \text{tr}\left((\mathbf{F}(\boldsymbol{\theta}))^{-1}\right) = \frac{N(N+1)}{T}$$

Borne de Cramér-Rao intrinsèque

Passage à la matrice d'information de Fisher $[\mathbf{F}(\boldsymbol{\theta})]_{i,j} = g_{\boldsymbol{\theta}}^{fisher}(\boldsymbol{\Omega}_i, \boldsymbol{\Omega}_j)$

$$g_{\boldsymbol{\theta}}^{fisher}(\boldsymbol{\Omega}_i, \boldsymbol{\Omega}_j) = T \text{tr}(\boldsymbol{\Omega}_i \boldsymbol{\Sigma}^{-1} \boldsymbol{\Omega}_j \boldsymbol{\Sigma}^{-1})$$

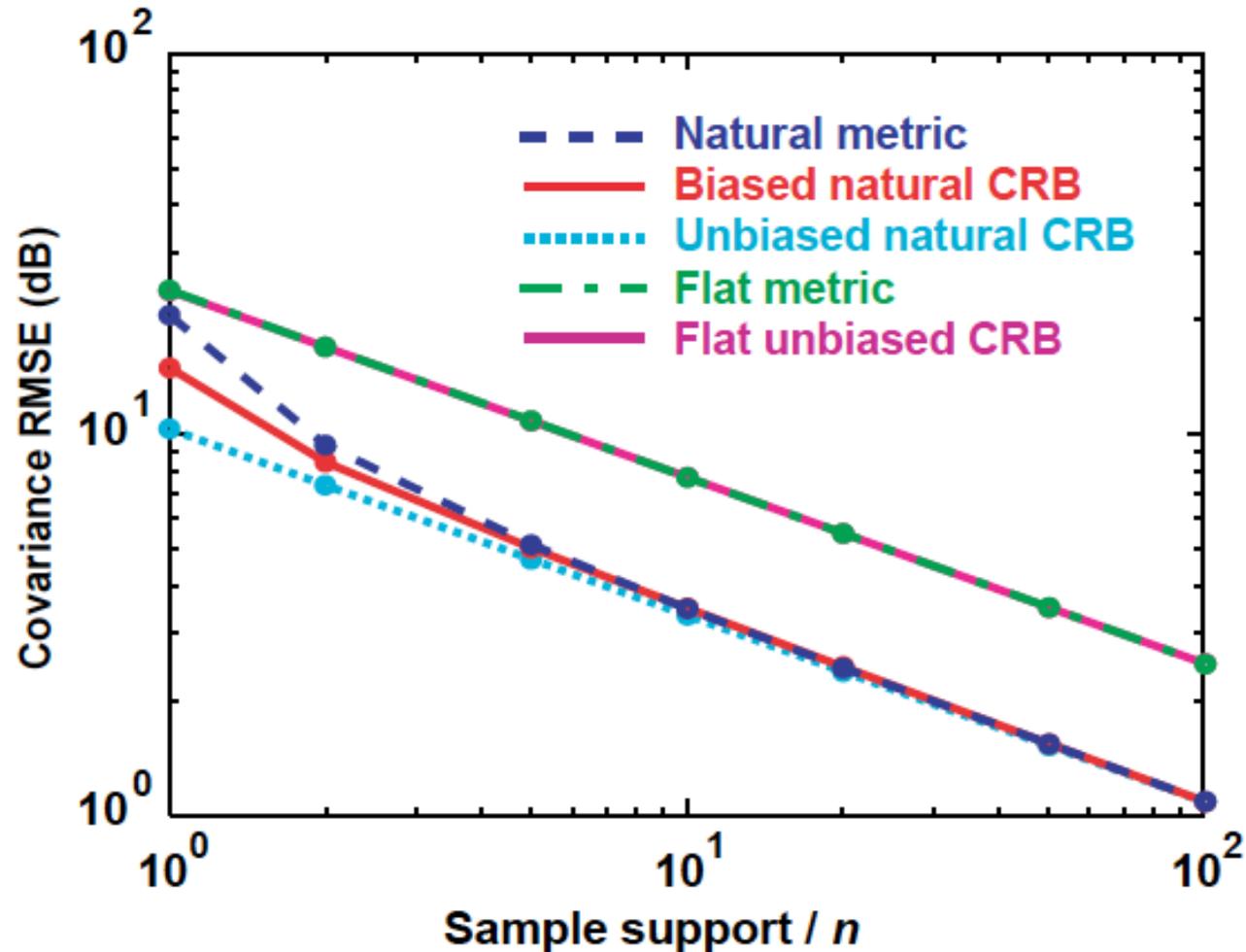
$$g_{\boldsymbol{\theta}}^{euclid}(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2) = T \text{tr}(\boldsymbol{\Omega}_1 \boldsymbol{\Omega}_2)$$

$$\mathbb{E}(\text{dist}_{euclid}^2(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)) = \mathbb{E}\left(\|\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2\|_{Fro}^2\right) \geq \frac{1}{T} (\text{tr}(\boldsymbol{\Sigma}^2) + \text{tr}^2(\boldsymbol{\Sigma}))$$

Les deux bornes coïncident lorsque $\boldsymbol{\Sigma} = \mathbf{I}_N$

iCRB : Estimation d'une matrice de covariance

Résultats de simulation (Smith 05)



I Problématique

II Bornes de Cramér-Rao « euclidiennes »

III Borne de Cramér-Rao intrinsèque – Généralités

IV Borne de Cramér-Rao intrinsèque – Estimation d'une matrice de covariance

✓ V Perspectives

- Application a des distributions plus générales (CES) et prise en compte de contraintes supplémentaires
- Estimation de sous espaces dans le contexte gaussien et CES
- Séparation aveugle de sources

- Borne de Cramér-Rao bayésienne ?



Merci !!!!