

iXWaves
Sophia Antipolis
laurent.kopp@ixwaves.com

TRAITEMENT D'ANTENNE

- Laurent KOPP -

Bornes de CRAMER-RAO
en traitement d'antenne

Calcul et utilisation

Bornes de CRAMER-RAO en traitement d'antenne / Calcul et utilisation

Résumé :	3
1 Inégalité de CRAMER-RAO	4
1.1 Dérivation des moyennes	4
1.2 Inégalité de CAUCHY-SCHWARZ matricielle	5
1.3 Démonstration de l'inégalité de CRAMER-RAO	6
2 Calcul de matrice de FISHER dans le cas Gaussien	6
2.1 Formule générale	6
2.2 Passif monosource	8
3 Efficacité du traitement classique	11
3.1 Maximum de vraisemblance	11
3.2 Covariance d'estimation du traitement classique	12
3.3 Modèle « retard pur »	14
3.4 Champ proche	15
3.5 Répartition optimale des capteurs d'une antenne linéaire	17
4 Approximation parabolique du lobe principal et largeur à 3 dB	19
ANNEXES	21
Annexe 1 : Formule de LIOUVILLE (Dérivée d'un déterminant)	21
Annexe 2 : Densité Complexe Circulaire	22

Résumé :

1. Inégalité de CRAMER-RAO

1.1 Formule de dérivation des moyennes $\underline{\partial} E\{\underline{Z}^T\} = E\{\underline{\partial} L \cdot \underline{Z}^T\}$

1.2 Inégalité de CAUCHY-SCHWARZ matricielle $C_{11} - C_{12}C_{22}^{-1}C_{21} \geq 0$

1.3 Démonstration de l'inégalité de CRAMER-RAO $V - (I + \underline{\partial} \underline{b}^T)^T F^{-1} (I + \underline{\partial} \underline{b}^T) \geq 0$

2. Matrice de FISHER dans le cas Gaussien

2.1 Formule générale $F_{kl} = -Tr\{\partial_k \Gamma^{-1} \partial_l \Gamma\} + 2Re\{\partial_k \underline{\mu}^+ \Gamma^{-1} \partial_l \underline{\mu}\}$

2.2 Passif une source $F = -\frac{(K_S)^2}{1 + K_S} (\partial^2 D)_0$

3. Efficacité du traitement Classique

3.1 Maximum de vraisemblance $C(\underline{\theta}) = \underline{d}(\underline{\theta})^+ \hat{\Gamma} \underline{d}(\underline{\theta})$

3.2 Covariance d'estimation du traitement classique $V \cong \frac{1}{N} \frac{1 + K_S}{(K_S)^2} (-\partial^2 D)_0^{-1} = F^{-1}$

3.3 Modèle « retard pur » $(\partial^2 D)_0 = \frac{2(2\pi f)^2}{K} T^T P T$

3.4 Champ proche $T^T P T = \frac{1}{c^2} \begin{pmatrix} \left(S_2 - \frac{S_1^2}{K} \right) & \frac{1 - \alpha^2}{2} \left(S_3 - \frac{S_1 S_2}{K} \right) \\ \frac{1 - \alpha^2}{2} \left(S_3 - \frac{S_1 S_2}{K} \right) & \left(\frac{1 - \alpha^2}{2} \right)^2 \left(S_4 - \frac{S_2^2}{K} \right) \end{pmatrix}$

3.5 Répartition optimale des capteurs d'une antenne linéaire

4. Approximation parabolique du lobe principal et largeur à 3 dB en $\cos \theta$ et R^{-1}

$$2\theta_3 = \frac{\sqrt{6}}{\pi} \frac{\lambda}{L} \left(\frac{K-1}{K+1} \right)^{1/2} \quad \text{et} \quad 2\alpha_3 = \frac{6\sqrt{10}}{\pi} \frac{\lambda}{L^2 \sin^2 \theta} \left(\frac{(K-1)^3}{(K+1)(K^2-4)} \right)^{1/2}$$

1 Inégalité de CRAMER-RAO

La démonstration de cette inégalité est fondée sur deux résultats préalables : La formule de dérivation des moyennes et l'inégalité de CAUCHY-SCHWARZ matricielle.

1.1 Dérivation des moyennes

La dérivée de la moyenne d'une fonction n'est pas égale à la moyenne de la dérivée.

Soit $\underline{f}(X, \underline{\theta})$ une fonction de l'observation X et du paramètre $\underline{\theta}$ vectoriel.

Par définition la moyenne de cette fonction est définie par :

$$E\{\underline{f}(X, \underline{\theta})\} = \int_{\Omega} \underline{f}(X, \underline{\theta}) p(X / \underline{\theta}) dX$$

$p(X / \underline{\theta})$ désignant la densité de probabilité de l'observation.

Dérivant par rapport à $(\underline{\theta})_k$, ce qu'on note $\partial_k(\bullet) = \frac{\partial(\bullet)}{\partial(\underline{\theta})_k}$, on obtient (Ω ne dépend pas de $\underline{\theta}$)

$$\partial_k E\{\underline{f}\} = \int_{\Omega} \partial_k \underline{f} \cdot p \cdot dX + \int_{\Omega} \underline{f} \cdot \partial_k L \cdot p \cdot dX$$

avec $L = Ln(p(X / \underline{\theta}))$ la vraisemblance logarithmique.

Ce qu'on peut écrire aussi :

$$\partial_k E\{\underline{f}\} = E\{\partial_k \underline{f}\} + E\{\partial_k L \cdot \underline{f}\}$$

ou encore, matriciellement :

$$\underline{\partial} E\{\underline{f}^T\} = E\{\underline{\partial} \underline{f}^T\} + E\{\underline{\partial} L \cdot \underline{f}^T\}$$

De cette relation générale on tire trois conséquences :

a) Si $\underline{f} = 1$ $E\{\underline{\partial} L\} = \underline{0}$

b) Si $\underline{f} = \underline{\partial} L$ $E\{\underline{\partial} L \cdot \underline{\partial} L^T\} = -E\{\underline{\partial}^2 L\} = F$, la matrice de FISHER, par définition.

c) Si $\underline{f} = \underline{Z}(X)$ est un estimateur quelconque (fonction de X uniquement), on trouve :

$$\underline{\partial} E\{\underline{Z}^T\} = E\{\underline{\partial} L \cdot \underline{Z}^T\}$$

C'est le principal résultat utile pour la suite.

Si on note $\underline{Z} = \hat{\underline{\theta}}(X)$ un estimateur de $\underline{\theta}$, on aura

$$E\left\{\underline{\partial}L \cdot \left(\hat{\underline{\theta}}(X) - E\{\hat{\underline{\theta}}(X)\}^T\right)\right\} = \underline{\partial}E\left\{\hat{\underline{\theta}}^T(X)\right\} = (\underline{\partial}\underline{b}^T + I)$$

Sachant que $\underline{\partial}L$ est une variable aléatoire centrée et en définissant le biais d'estimation :

$$\underline{b} = E\{\hat{\underline{\theta}}(X)\} - \underline{\theta}$$

1.2 Inégalité de CAUCHY-SCHWARZ matricielle

Cette inégalité bien connue en analyse et en algèbre vectorielle se généralise aussi à l'algèbre matricielle.

Soit \underline{X}_1 et \underline{X}_2 deux vecteurs aléatoires de C^N et $\underline{\alpha}, \underline{\beta}$ deux vecteurs fixes de C^N .

Introduisant la variable aléatoire scalaire complexe :

$$z = \underline{\alpha}^+ \underline{X}_1 + \underline{\beta}^+ \underline{X}_2$$

On aura, bien sur :

$$E\{|z|^2\} \geq 0 \quad \forall \underline{\alpha}, \underline{\beta}$$

développant $|z|^2$, on trouve $\forall \underline{\alpha}, \underline{\beta}$

$$\underline{\alpha}^+ C_{11} \underline{\alpha} + \underline{\beta}^+ C_{22} \underline{\beta} + \underline{\alpha}^+ C_{12} \underline{\beta} + \underline{\beta}^+ C_{21} \underline{\alpha} \geq 0$$

avec $C_{ij} = E\{\underline{X}_i \underline{X}_j^+\}$ et donc $C_{12} = C_{21}^+$

Si on précise que $\underline{\beta} = -C_{22}^{-1} C_{21} \underline{\alpha}$ (en supposant que C_{22} est une matrice inversible), on trouvera que :

$$\underline{\alpha}^+ [C_{11} - C_{12} C_{22}^{-1} C_{21}] \underline{\alpha} \geq 0 \quad \forall \underline{\alpha}$$

ce qu'on écrit :

$$A = C_{11} - C_{12} C_{22}^{-1} C_{21} \geq 0$$

La matrice A est non négative.

1.3 Démonstration de l'inégalité de CRAMER-RAO

Appliquant l'inégalité de CAUCHY-SCHWARZ matricielle au cas où :

$$\begin{aligned}\underline{X}_1 &= \hat{\underline{\theta}}(X) - E\{\hat{\underline{\theta}}(X)\} \\ \underline{X}_2 &= \underline{\partial}L\end{aligned}$$

on obtient :

$$\begin{aligned}C_{11} &= V = E\left\{\left(\hat{\underline{\theta}}(X) - E\{\hat{\underline{\theta}}(X)\}\right)\left(\hat{\underline{\theta}}(X) - E\{\hat{\underline{\theta}}(X)\}\right)^T\right\} \\ C_{22} &= F = E\{\underline{\partial}L\underline{\partial}L^T\} \\ C_{21} &= E\left\{\underline{\partial}L\left(\hat{\underline{\theta}}(X) - E\{\hat{\underline{\theta}}(X)\}\right)^T\right\}\end{aligned}$$

Dans ces expressions V représente la matrice de covariance de l'erreur d'estimation et F la matrice de FISHER.

La quantité C_{21} se déduit de la formule de dérivation des moyennes :

$$C_{21} = \left(I + \underline{\partial}\underline{b}^T\right)$$

On obtient finalement le résultat :

$$V - \left(I + \underline{\partial}\underline{b}^T\right)^T F^{-1} \left(I + \underline{\partial}\underline{b}^T\right) \geq 0$$

2 Calcul de matrice de FISHER dans le cas Gaussien

Le cas Gaussien permet le calcul explicite de la matrice de FISHER et montre le lien étroit qui la relie à la fonction d'ambiguïté : La partie « géométrique » de la matrice de FISHER est proportionnelle au Hessien de la fonction d'ambiguïté (Courbure ou Largeur à 3 dB au sens large comme on le verra plus loin).

2.1 Formule générale

Considérons tout d'abord le cas d'un seul vecteur \underline{X} complexe circulaire, de moyenne $\underline{\mu}$, de covariance Γ .

Sa densité de probabilité s'écrit :

$$p(\underline{X} / \underline{\theta}) = |\pi\Gamma|^{-1} \exp\left[-(\underline{X} - \underline{\mu})^\dagger \Gamma^{-1} (\underline{X} - \underline{\mu})\right]$$

avec $\underline{\mu} = E(\underline{X})$, $\hat{\Gamma} = (\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})^\dagger$ et $\Gamma = E(\hat{\Gamma})$.

Les paramètres $\underline{\theta}$ sont cachés dans Γ et $\underline{\mu}$

La vraisemblance logarithmique s'écrit :

$$\begin{aligned} L &= Ln[p(\underline{X} / \underline{\theta})] \\ &= -Ln|\pi\Gamma| - Tr(\hat{\Gamma}\Gamma^{-1}) \end{aligned}$$

Sa dérivée par rapport à $(\underline{\theta})_k$ s'écrit :

$$\partial_k L = -\partial_k (Ln|\pi\Gamma|) - Tr(\partial_k \hat{\Gamma}\Gamma^{-1}) - Tr(\hat{\Gamma}\partial_k \Gamma^{-1})$$

Mais on a :

$$\begin{aligned} \partial_k Ln|\Gamma| &= Tr(\Gamma^{-1}\partial_k \Gamma) \\ \partial_k \Gamma^{-1} &= -\Gamma^{-1}(\partial_k \Gamma)\Gamma^{-1} \\ \partial_k \hat{\Gamma} &= -\partial_k \underline{\mu}(\underline{X} - \underline{\mu})^+ - (\underline{X} - \underline{\mu})\partial_k \underline{\mu}^+ \end{aligned}$$

donc :

$$\partial_k L = -Tr\{\Gamma^{-1}(\Gamma - \hat{\Gamma})\Gamma^{-1}\partial_k \Gamma\} - Tr\{\Gamma^{-1}\partial_k \hat{\Gamma}\}$$

On remarquera que $E(\Gamma - \hat{\Gamma}) = 0$ ainsi que $E(\partial_k L) = 0$

Pour calculer la matrice de FISHER on utilisera la forme suivante de son terme général :

$$F_{kl} = -E\{\partial_{kl}^2 L\}$$

Lorsque l'on va effectuer la dérivation de $\partial_k L$ par rapport à $(\underline{\theta})_l$, différents termes vont apparaître (somme de produits) dont il faudra ensuite prendre la moyenne.

Il est inutile de calculer les termes dans lesquels resteront $(\Gamma - \hat{\Gamma})$ et $\partial_k L$ dont les moyennes sont nulles. Au bout du compte, il ne restera que deux termes non nuls :

$$E\{\partial_{kl}^2 L\} = -Tr\{\Gamma^{-1}\partial_l \Gamma \Gamma^{-1}\partial_k \Gamma\} - Tr\{\Gamma^{-1}E(\partial_{kl}^2 \hat{\Gamma})\}$$

et on a $E\{\partial_{kl}^2 \hat{\Gamma}\} = \partial_k \underline{\mu} \partial_l \underline{\mu}^+ + \partial_l \underline{\mu} \partial_k \underline{\mu}^+$, d'où finalement :

$$F_{kl} = -Tr\{\partial_k \Gamma^{-1} \partial_l \Gamma\} + 2 \operatorname{Re}\{\partial_k \underline{\mu}^+ \Gamma^{-1} \partial_l \underline{\mu}\}$$

Dans la suite on s'intéressera uniquement au cas où $\underline{\mu} = \underline{0}$, correspondant au problème de l'écoute passive en acoustique sous-marine. La matrice de FISHER est donc réduite à :

$$F_{kl} = -Tr\{\partial_k \Gamma^{-1} \partial_l \Gamma\}$$

2.2 Passif monosource

On se limite ici au problème de l'estimation des paramètres d'une source unique dans du bruit incohérent.

On s'intéressera donc ici au cas où :

$$\Gamma = \gamma \underline{d} \underline{d}^+ + \sigma I$$

On supposera aussi que le vecteur source remplit la condition $\underline{d}^+(\underline{\theta}) \underline{d}(\underline{\theta}) = K$, ce qui correspond au modèle fréquent où la propagation se réduit à « un retard pur », pour lequel :

$$(\underline{d})_k = \exp(2i\pi f \tau_k) \quad \forall k = 1, \dots, N$$

K désigne le nombre de capteurs.

Les paramètres à estimer sont $\{\sigma, \gamma, \underline{\theta}\}$ où $\underline{\theta}$ désigne un vecteur de paramètres géométriques cachés dans \underline{d} . Il est intéressant de remarquer que les paramètres $\{\sigma, \gamma\}$ d'une part, et $\underline{\theta}$ d'autre part, sont découplés.

En effet :

$$F(\sigma, (\underline{\theta})_k) = -\text{Tr}\{\partial_\sigma \Gamma \partial_k \Gamma^{-1}\}, \quad \partial_\sigma \Gamma = I \quad \text{et} \quad \partial_k \Gamma^{-1} = -\frac{s}{\sigma} \frac{1}{1 + Ks} (\underline{u}_k \underline{d}^+ + \underline{d} \underline{u}_k^+)$$

On a utilisé la notation $\underline{u}_k = \partial_k \underline{d}$ et le lemme d'inversion matricielle (WOODBURY) :

$$\Gamma^{-1} = \frac{1}{\sigma} I - \frac{s}{\sigma} \frac{\underline{d} \underline{d}^+}{1 + Ks}$$

avec $s = \frac{\gamma}{\sigma}$

$$F(\sigma, (\underline{\theta})_k) = \frac{s}{\sigma} \frac{1}{1 + Ks} (\underline{u}_k^+ \underline{d} + \underline{d}^+ \underline{u}_k) = 0$$

car

$$\underline{u}_k^+ \underline{d} + \underline{d}^+ \underline{u}_k = \partial_k (\underline{d}^+ \underline{d}) = \partial_k (K) = 0$$

De même

$$\begin{aligned} F(\gamma, (\underline{\theta})_k) &= \frac{s}{\sigma} \frac{1}{1 + Ks} \text{Tr}[\underline{d} \underline{d}^+ (\underline{u}_k \underline{d}^+ + \underline{d} \underline{u}_k^+)] \\ &= \frac{s}{\sigma} \frac{1}{1 + Ks} K (\underline{u}_k^+ \underline{d} + \underline{d}^+ \underline{u}_k) = 0 \end{aligned}$$

La matrice de FISHER est donc bloc-diagonale. Il n'y a pas de mélanges dans l'estimation de $\{\sigma, \gamma, \underline{\theta}\}$.

Si on se limite maintenant au calcul de la seule partie géométrique de cette matrice, que l'on notera $F(\underline{\theta}, \underline{\theta})$, on obtient son terme général :

$$F_{kl} = F[(\underline{\theta})_k, (\underline{\theta})_l] = -Tr(\partial_k \Gamma^{-1} \partial_l \Gamma)$$

$$\text{avec } \partial_k \Gamma = \gamma(\underline{u}_k \underline{d}^+ + \underline{d} \underline{u}_k^+) \text{ et } \partial_l \Gamma^{-1} = -\frac{s}{\sigma} \frac{1}{1 + Ks} (\underline{u}_l \underline{d}^+ + \underline{d} \underline{u}_l^+)$$

il vient :

$$F_{kl} = \frac{s^2}{1 + Ks} \left\{ (\underline{d}^+ \underline{u}_l) (\underline{d}^+ \underline{u}_k) + (\underline{u}_k^+ \underline{d}) (\underline{u}_l^+ \underline{d}) + K(\underline{u}_k^+ \underline{u}_l + \underline{u}_l^+ \underline{u}_k) \right\}$$

En utilisant la propriété $\underline{d}^+(\underline{\theta}) \underline{d}(\underline{\theta}) = K$, on trouve, par dérivation:

$$\underline{u}_{kl}^+ \underline{d} + \underline{d}^+ \underline{u}_{kl} + \underline{u}_k^+ \underline{u}_l + \underline{u}_l^+ \underline{u}_k = 0$$

avec la notation $\underline{u}_{kl} = \partial_k \underline{u}_l$

Et de même, si l'on considère la fonction de directivité du traitement classique

$$D(\underline{\theta}, \underline{\theta}_0) = \frac{1}{K^2} \left| \underline{d}^+(\underline{\theta}) \underline{d}(\underline{\theta}_0) \right|^2$$

On notera $\underline{d} = \underline{d}(\underline{\theta})$ et $\underline{d}_0 = \underline{d}(\underline{\theta}_0)$

On a d'abord :

$$\begin{aligned} \partial_k D &= \frac{1}{K^2} \left\{ \underline{u}_k^+ \underline{d}_0 \underline{d}_0 \underline{d} + \underline{d}^+ \underline{d}_0 \underline{d}_0^+ \underline{u}_k \right\} \\ \partial_{kl}^2 D &= \frac{1}{K^2} \left\{ \underline{u}_{kl}^+ \underline{d}_0 \underline{d}_0 \underline{d} + \underline{d}^+ \underline{d}_0 \underline{d}_0^+ \underline{u}_{kl} + \underline{u}_k^+ \underline{d}_0 \underline{d}_0^+ \underline{u}_l + \underline{u}_l^+ \underline{d}_0 \underline{d}_0^+ \underline{u}_k \right\} \end{aligned}$$

En $\underline{\theta} = \underline{\theta}_0$ on obtient :

$$\left(\partial_{kl}^2 D \right)_0 = \frac{1}{K^2} \left\{ -(\underline{u}_k^+ \underline{d}_0) (\underline{u}_l^+ \underline{d}_0) - K(\underline{u}_k^+ \underline{u}_l + \underline{u}_l^+ \underline{u}_k) \right\}$$

d'où l'on déduit :

$$F_{kl} = \frac{(Ks)^2}{1 + Ks} \left(-\partial_{kl}^2 D \right)_0$$

et, sous forme matricielle,

$$F(\underline{\theta}, \underline{\theta}) = \frac{(Ks)^2}{1 + Ks} \left(-\partial^2 D \right)_0$$

On a obtenu au passage un résultat utile

$$\left(\partial_{kl}^2 D \right)_0 = -\frac{1}{K} \left\{ \underline{u}_k^+ P_0 \underline{u}_l + \underline{u}_l^+ P_0 \underline{u}_k \right\}$$

avec $P_0 = I - \frac{\underline{d}_0 \underline{d}_0^+}{K}$, projecteur orthogonal à \underline{d}_0

Si maintenant on s'intéresse au cas où l'observation complète est constitué de plusieurs observations indépendantes $X = \{\underline{X}_n\}_{n=1,N}$, la vraisemblance logarithmique s'écrit :

$$L(X, \underline{\theta}) = \sum_{n=1,N} L(\underline{X}_n, \underline{\theta})$$

$$\partial^2 L(X, \underline{\theta}) = \sum_{n=1,N} \partial^2 L(\underline{X}_n, \underline{\theta}) \quad \text{et} \quad F = \sum_{n=1,N} E\{-\partial^2 L(\underline{X}_n, \underline{\theta})\}$$

D'où on déduit :

$$F(\underline{\theta}, \underline{\theta}) = \frac{N(K_S)^2}{1 + K_S} (-\partial^2 D)_0$$

Si l'on considère enfin le cas de la représentation spectrale complète, on aura:

$$L(X, \underline{\theta}) = \sum_{f \in F} L(X(f), \underline{\theta})$$

$$F = \sum_{f \in F} F(f)$$

Ce que l'on écrit encore :

$$F = T \sum_{f \in F} \delta F(f) \cong T \int_0^B F(f) df$$

que l'on note encore $F = BT \langle F(f) \rangle$, - où la notation $\langle X(f) \rangle$ désigne la moyenne en fréquence de la fonction $X(f)$ selon la fréquence f dans la bande B -

$$\langle X(f) \rangle = \frac{1}{B} \int_0^B X(f) df$$

T est la durée de l'analyse spectrale. On a $\delta f = T^{-1}$
On aura alors :

$$F(\underline{\theta}, \underline{\theta}) = BT \left\langle \frac{(K_S)^2}{1 + K_S} (-\partial^2 D)_0 \right\rangle$$

En développant au deuxième ordre la fonction de directivité autour de $\underline{\theta}_0$, on obtient :

$$D(\underline{\theta}, \underline{\theta}_0) \cong 1 + \frac{1}{2} (\underline{\theta} - \underline{\theta}_0)^T (\partial^2 D)_0 (\underline{\theta} - \underline{\theta}_0)$$

$$\cong 1 - \frac{1}{2} (\underline{\theta} - \underline{\theta}_0)^T \Theta_{-3}^{-2} (\underline{\theta} - \underline{\theta}_0)$$

$(\partial^2 D)_0 = -(\Theta_{-3})^{-2}$ définit l'angle de CRAMER-RAO, ce qui fait référence à la demi-largeur d'ambiguïté « à -3 dB » (ou à mi-hauteur) dans le cas des mesures angulaires.

On remarquera ici le rôle primordial joué par la fonction de directivité, et le lien qu'elle présente naturellement avec la matrice de *FISHER*.

3 Efficacité du traitement classique

Le but de cette section est de démontrer que le traitement classique atteint la borne de CRAMER-RAO, au moins en première approximation.

3.1 Maximum de vraisemblance

Dans le cas d'une observation Gaussienne de densité : $p(\underline{X} / \underline{\theta}) = |\pi\Gamma|^{-1} \exp\left[-\underline{X}^+ \Gamma^{-1} \underline{X}\right]$

avec $\Gamma = \gamma \underline{d}(\underline{\theta}_0) \underline{d}(\underline{\theta}_0)^+ + \sigma I$ et on notera $\underline{d}_0 = \underline{d}(\underline{\theta}_0)$. On en déduit :

$$\Gamma^{-1} = \frac{1}{\sigma} I - \frac{s}{\sigma} \frac{\underline{d}_0 \underline{d}_0^+}{1 + Ks} \quad \text{avec} \quad s = \frac{\gamma}{\sigma}$$

Il s'agit de maximiser par rapport à $\underline{\theta}$, la quantité : $-Ln(|\Gamma|) + \frac{s}{1 + Ks} \left| \underline{X}^+ \underline{d}(\underline{\theta}) \right|^2$

On a $|\Gamma| = \sigma^K (1 + Ks)$ indépendant de $\underline{\theta}$.

Reste donc à maximiser $\underline{d}(\underline{\theta})^+ \hat{\Gamma} \underline{d}(\underline{\theta}) = \underline{d}(\underline{\theta})^+ \underline{X} \underline{X}^+ \underline{d}(\underline{\theta})$.

Ce résultat reste valable lorsque l'on dispose de plusieurs observations $X = \{\underline{X}_n\}_{n=1,N}$ indépendantes. Dans ce cas en effet on aura :

$$L = -NLn(|\Gamma|) - NTr\left\{\hat{\Gamma}\Gamma^{-1}\right\} \quad \text{avec} \quad \hat{\Gamma} = \frac{1}{N} \sum_{n=1,N} \underline{X}_n \underline{X}_n^+$$

$$L = -NLn(|\Gamma|) - NTr\left\{\hat{\Gamma}\left(\Gamma^{-1} = \frac{1}{\sigma} I - \frac{s}{\sigma} \frac{\underline{d}(\underline{\theta}) \underline{d}(\underline{\theta})^+}{1 + Ks}\right)\right\}$$

Il faut donc maximiser par rapport à $\underline{\theta}$ la quantité $Tr\left\{\hat{\Gamma} \underline{d}(\underline{\theta}) \underline{d}(\underline{\theta})^+\right\} = \underline{d}(\underline{\theta})^+ \hat{\Gamma} \underline{d}(\underline{\theta})$.

L'estimateur de $\underline{\theta}$, optimal au sens du maximum de vraisemblance, maximise donc :

$$C(\underline{\theta}) = \underline{d}(\underline{\theta})^+ \hat{\Gamma} \underline{d}(\underline{\theta})$$

C'est le traitement « Classique » (en fait il faudrait normaliser par K^2 , une constante qui ne change rien à la position du maximum).

3.2 Covariance d'estimation du traitement classique

La quantité obtenue précédemment est une fonction de $\underline{\theta}$ et de $\hat{\Gamma}$, ce que l'on fait apparaître de façon explicite par :

$$C(\underline{\theta}, \hat{\Gamma}) = \underline{d}(\underline{\theta})^+ \hat{\Gamma} \underline{d}(\underline{\theta})$$

Cette quantité possède un maximum en $\underline{\theta}$ autour de $\underline{\theta}_0$ en un point $\hat{\underline{\theta}}$ pour lequel

$$\underline{\partial} C(\hat{\underline{\theta}}, \hat{\Gamma}) = \underline{0}$$

On peut faire un développement limité de la fonction $\underline{\partial} C(\underline{\theta}, \hat{\Gamma})$ autour de $\underline{\theta}_0, \Gamma$ en la considérant comme une fonction de $\underline{\theta}$ et $\hat{\Gamma}$, ce qui donne :

$$\underline{\partial} C(\underline{\theta}, \hat{\Gamma}) = \underline{\partial} C(\underline{\theta}_0, \Gamma) + \left(\underline{\partial}^2 C \right)_{\underline{\theta}_0, \Gamma} (\underline{\theta} - \underline{\theta}_0) + \underline{\partial} C(\underline{\theta}_0, \hat{\Gamma} - \Gamma)$$

Le dernier terme du second membre de cette expression résulte de la linéarité de la fonction vis à vis de la variable $\hat{\Gamma}$.

On a alors :

$$\begin{aligned} \partial_k C(\underline{\theta}_0, \Gamma) &= \underline{d}_0^+ \Gamma \underline{u}_k(\underline{\theta}_0) + \underline{u}_k^+(\underline{\theta}_0) \Gamma \underline{d}_0 \\ &= (\sigma + K\gamma) (\underline{d}_0^+ \underline{u}_k(\underline{\theta}_0) + \underline{u}_k^+(\underline{\theta}_0) \underline{d}_0) = 0 \end{aligned}$$

D'où $\underline{\partial} C(\underline{\theta}_0, \Gamma) = \underline{0}$

Au point $\hat{\underline{\theta}}, \hat{\Gamma}$ on aura $\underline{\partial} C(\hat{\underline{\theta}}, \hat{\Gamma}) = \underline{0}$, ce qui permet d'en déduire :

$$\hat{\underline{\theta}} - \underline{\theta}_0 = \left(\underline{\partial}^2 C \right)_0^{-1} \underline{\partial} C(\underline{\theta}_0, \hat{\Gamma})$$

On en déduit donc : $E\{\hat{\underline{\theta}}\} = \underline{\theta}_0$ (au premier ordre, l'estimation est sans biais).

Pour calculer $\left(\underline{\partial}^2 C \right)_0 = \left(\underline{\partial}^2 C \right)_{\underline{\theta}_0, \Gamma}$ on peut remarquer que

$$\left(\partial_{kl}^2 C(\underline{\theta}, \hat{\Gamma}) \right)_{\underline{\theta}_0, \Gamma} = \left(\partial_{kl}^2 C(\underline{\theta}, \Gamma) \right)_{\underline{\theta}_0, \Gamma}$$

On en déduit que :

$$\begin{aligned} \left(\partial_{kl}^2 C(\underline{\theta}, \hat{\Gamma}) \right)_{\underline{\theta}_0, \Gamma} &= \left(\partial_{kl}^2 \left[\underline{d}^+ \Gamma \underline{d} \right] \right)_{\underline{\theta}_0, \Gamma} \\ &= \left(\partial_{kl}^2 \left[K\sigma + K^2 \gamma D \right] \right)_{\underline{\theta} = \underline{\theta}_0} \\ &= K^2 \gamma \left(\partial_{kl}^2 D \right)_0 \end{aligned}$$

Finalement on pourra écrire : $\hat{\underline{\theta}}_0 - \underline{\theta}_0 = \frac{1}{K^2 \gamma} (\partial_{kl}^2 D)_0^{-1} \hat{\underline{C}}(\underline{\theta}_0, \hat{\Gamma})$

La variance d'estimation s'écrira alors :

$$V = \frac{1}{K^4 \gamma^2} (\partial^2 D)_0^{-1} \Phi (\partial^2 D)_0^{-1}$$

avec $\Phi = E \left\{ \hat{\underline{C}}(\underline{\theta}_0, \hat{\Gamma}) \hat{\underline{C}}^+(\underline{\theta}_0, \hat{\Gamma}) \right\}$

et on a donc : $\Phi_{kl} = E \left\{ \left(\underline{d}_0^+ \hat{\Gamma} \underline{u}_k(\underline{\theta}_0) + \underline{u}_k^+(\underline{\theta}_0) \hat{\Gamma} \underline{d}_0 \right) \left(\underline{d}_0^+ \hat{\Gamma} \underline{u}_l(\underline{\theta}_0) + \underline{u}_l^+(\underline{\theta}_0) \hat{\Gamma} \underline{d}_0 \right)^+ \right\}$

Il faut donc calculer des expressions du type :

$$E \left\{ \underline{u}^+ \hat{\Gamma} A \hat{\Gamma} \underline{v} \right\} \text{ avec } \hat{\Gamma} = \frac{1}{N} \sum_{n=1, N} \underline{X}_n \underline{X}_n^+$$

On utilise pour cela une généralisation des formules classiques sur les moments d'ordre 4 des variables Gaussiennes :

$$E \left\{ \underline{u}^+ \hat{\Gamma} A \hat{\Gamma} \underline{v} \right\} = \underline{u}^+ \Gamma A \Gamma \underline{v} + \frac{1}{N} \underline{u}^+ \Gamma \underline{v} \text{Tr}(A \Gamma)$$

On trouve alors que :

$$\Phi_{kl} = \frac{1}{N} \left[\left(\underline{d}_0^+ \Gamma \underline{d}_0 \right) \left(\underline{u}_{k0}^+ \Gamma \underline{u}_{l0} + \underline{u}_{l0}^+ \Gamma \underline{u}_{k0} \right) + \left(\underline{u}_{k0}^+ \Gamma \underline{d}_0 \right) \left(\underline{u}_{l0}^+ \Gamma \underline{d}_0 \right) + \left(\underline{d}_0^+ \Gamma \underline{u}_{k0} \right) \left(\underline{d}_0^+ \Gamma \underline{u}_{l0} \right) \right]$$

Sachant que $\Gamma \underline{d}_0 = \sigma(1 + Ks) \underline{d}_0$, on obtient :

$$\begin{aligned} \Phi_{kl} &= \frac{1}{N} \sigma^2 (1 + Ks)^2 \left[K \left(\underline{u}_{k0}^+ \underline{u}_{l0} + \underline{u}_{l0}^+ \underline{u}_{k0} \right) - \left(\underline{u}_{k0}^+ \underline{d}_0 \underline{d}_0^+ \underline{u}_{l0} + \underline{u}_{l0}^+ \underline{d}_0 \underline{d}_0^+ \underline{u}_{k0} \right) \right] \\ &= \frac{K \sigma^2 (1 + Ks)}{N} \left\{ \underline{u}_{k0}^+ P_0 \underline{u}_{l0} + \underline{u}_{l0}^+ P_0 \underline{u}_{k0} \right\} \end{aligned}$$

Soit encore $\Phi_{kl} = -\frac{K^2 \sigma^2 (1 + Ks)}{N} (\partial_{kl}^2 D)_0$, compte tenu de l'expression obtenue dans la section 2.2 ci dessus.

On en déduit l'expression de la matrice de covariance de l'erreur d'estimation :

$$V = \frac{1}{K^4 \gamma^2} (\partial^2 D)_0^{-1} \left[-\frac{K^2 \sigma^2 (1 + Ks)}{N} (\partial^2 D)_0 \right] (\partial^2 D)_0^{-1}$$

$$\text{Soit, enfin : } V = -\frac{(1 + Ks)}{N(Ks)^2} (\partial^2 D)_0^{-1} = F^{-1}$$

Au second ordre près le traitement classique atteint la borne de CRAMER-RAO.

3.3 Modèle « retard pur »

On peut préciser l'expression de la matrice de FISHER lorsque le vecteur source s'écrit :

$$(\underline{d})_k = \exp\{2i\pi f \tau_k(\underline{\theta})\}$$

c'est à dire dans le cas , très fréquent, d'un modèle « retard pur » (ondes planes et sphériques par exemple).

$$\begin{aligned} \text{On a alors : } \underline{u}_k &= \partial_k \underline{d} = \text{col}\{2i\pi f \partial_k \tau_j \exp(2i\pi f \tau_j)\}_{j=1,K} \\ &= 2i\pi f \Delta \partial_k \underline{\tau} \end{aligned}$$

$$\text{Où : } \Delta = \text{diag}\{\underline{d}\}, \text{ c'est à dire } \Delta_{kl} = (\underline{d})_k \delta_{kl}$$

On peut alors calculer $(\partial_{kl}^2 D)_0$:

$$(\partial_{kl}^2 D)_0 = -\frac{1}{K} \{ \underline{u}_k^+ P_0 \underline{u}_l + \underline{u}_l^+ P_0 \underline{u}_k \}$$

avec $P_0 = I - \frac{\underline{d}_0 \underline{d}_0^+}{K}$, projecteur orthogonal à \underline{d}_0 .

$$(\partial_{kl}^2 D)_0 = -\frac{1}{K} \{ (2\pi f)^2 (\partial_k \underline{\tau}^T \Delta^+ P_0 \Delta \partial_l \underline{\tau} + \partial_l \underline{\tau}^T \Delta^+ P_0 \Delta \partial_k \underline{\tau}) \}$$

$$\text{Mais on a : } \Delta^+ P_0 \Delta = I - \frac{\Delta^+ \underline{d} \underline{d}^+ \Delta}{K} = I - \frac{\underline{1} \cdot \underline{1}^T}{K} = P(\underline{1}) \text{ où } \underline{1} = \text{col}\{1\}_{j=1,K}$$

$$\text{Finalement : } (\partial_{kl}^2 D)_0 = -\frac{(2\pi f)^2}{K} 2 \partial_k \underline{\tau}^T P(\underline{1}) \partial_l \underline{\tau}$$

En introduisant la matrice : $T = [\partial_1 \underline{\tau}, \partial_2 \underline{\tau}, \dots, \partial_N \underline{\tau}]$, on peut encore écrire :

$$(\partial^2 D)_0 = \frac{2(2\pi f)^2}{K} T^T P T$$

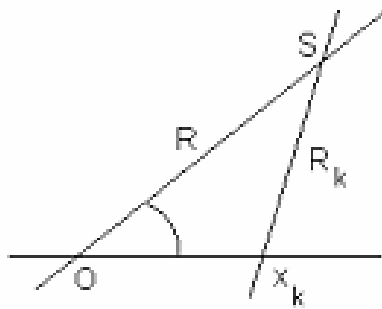
3.4 Champ proche

On peut encore préciser les choses sur l'exemple de l'antenne linéaire uniforme en champ proche. Dans ce cas on précise l'expression des retards :

$$(\underline{\tau})_k = \frac{R_k - R}{c}$$

R_k désigne ici la distance entre le capteur courant k et la source située à la distance R du point de référence. On suppose ici un milieu de propagation isocélère dans lequel les ondes sont sphériques et se propagent avec une célérité c . On doit aussi préciser que le retard dépend de deux paramètres géométriques (la distance R et le gisement θ), ce que l'on peut écrire :

$$\tau_k(\theta, R) = \frac{R_k - R}{c}$$



Si l'on désigne par x_k l'abscisse du capteur k par rapport au point de référence, l'expression de la distance entre le capteur et la source s'écrit :

$$\begin{aligned} R_k &= \left\{ R^2 + x_k^2 - 2Rx_k \cos \theta \right\}^{1/2} \\ &= R \left[1 + \left(\frac{x_k}{R} \right)^2 - 2 \left(\frac{x_k}{R} \right) \cos \theta \right]^{1/2} \end{aligned}$$

On posera $\alpha = \cos \theta$ et $\beta = R^{-1}$, ce qui conduit à la nouvelle expression :

$$R_k = \beta^{-1} \left(1 + x_k^2 \beta^2 - 2x_k \beta \alpha \right)^{1/2}$$

Les paramètres géométriques choisis pour décrire le problème sont α et β et non pas θ et R . On a alors besoin de calculer :

$$\partial_{\alpha} R_k = \beta^{-1} (1 + x_k^2 \beta^2 - 2x_k \beta \alpha)^{-1/2} \left(\frac{1}{2} (-2x_k \beta) \right) = -x_k \left(\frac{R}{R_k} \right)$$

$$\text{d'où } (\partial_{\alpha} \underline{\tau})_k \cong -\frac{x_k}{c}$$

De même :

$$\begin{aligned} \partial_{\beta} R_k &= -\beta^{-2} \cdot \beta R_k + \beta^{-1} \frac{1}{2} (\beta R_k)^{-1} (2\beta x_k^2 - 2x_k \alpha) \\ &= -\beta^{-3} R_k^{-1} \{ \beta^2 R_k^2 - \beta^2 x_k^2 + x_k \alpha \beta \} \end{aligned}$$

On en déduit :

$$\begin{aligned} \partial_{\beta} R_k &= -\beta^{-3} R_k^{-1} (1 - x_k \alpha \beta) \\ \partial_{\beta} (\underline{\tau})_k &= \frac{1}{c} (-\beta^{-3} R_k^{-1} (1 - x_k \alpha \beta) + \beta^{-2}) \\ &= \frac{\beta^{-2}}{c} (1 - (\beta R_k)^{-1} (1 - x_k \alpha \beta)) \end{aligned}$$

Un développement limité à l'ordre 2 de $(\beta R_k)^{-1}$ selon $x_k \beta$ donne finalement :

$$\begin{aligned} \partial_{\beta} (\underline{\tau})_k &\cong \frac{\beta^{-2}}{c} \left(1 - \left(1 + x_k \alpha \beta - \frac{1}{2} x_k^2 \beta^2 (1 - 3\alpha^2) \right) (1 - x_k \alpha \beta) \right) \\ &\cong \frac{\beta^{-2}}{c} \left(-\frac{1}{2} x_k^2 \beta^2 (1 - \alpha^2) \right) \end{aligned}$$

$$\text{Soit finalement } \partial_{\beta} (\underline{\tau})_k \cong -\frac{x_k^2}{2c} (1 - \alpha^2)$$

En posant : $\underline{\xi}_1 = \text{col}\{x_k\}_{k=1,K}$ et $\underline{\xi}_2 = \text{col}\{x_k^2\}_{k=1,K}$ on pourra écrire :

$$\partial_{\alpha} \underline{\tau} = -\frac{1}{c} \underline{\xi}_1 \text{ et } \partial_{\beta} \underline{\tau} = -\frac{1 - \alpha^2}{2c} \underline{\xi}_2$$

d'où l'expression de la matrice T précédente :

$$T = \begin{bmatrix} -\frac{1}{c} \underline{\xi}_1 & -\frac{1 - \alpha^2}{2c} \underline{\xi}_2 \end{bmatrix}$$

On peut alors calculer :

$$T^T P T = T^T T - \frac{1}{K} (T^T \underline{1})(T^T \underline{1})^T$$

On a tout d'abord : $T^T \underline{1} = -\frac{1}{c} \left(\underline{\xi}_1 \underline{1} \quad \frac{1-\alpha^2}{2} \underline{\xi}_2 \underline{1} \right)^T = -\frac{1}{c} \left(S_1 \quad \frac{1-\alpha^2}{2} S_2 \right)^T$

En utilisant la notation $S_m = \sum_{k=1, K} x_k^m$

On calculera de même : $T^T T = \frac{1}{c^2} \begin{pmatrix} S_2 & \frac{1-\alpha^2}{2} S_3 \\ \frac{1-\alpha^2}{2} S_3 & \left(\frac{1-\alpha^2}{2} \right)^2 S_4 \end{pmatrix}$

D'où vient finalement :

$$T^T P T = \frac{1}{c^2} \begin{pmatrix} \left(S_2 - \frac{S_1^2}{K} \right) & \frac{1-\alpha^2}{2} \left(S_3 - \frac{S_1 S_2}{K} \right) \\ \frac{1-\alpha^2}{2} \left(S_3 - \frac{S_1 S_2}{K} \right) & \left(\frac{1-\alpha^2}{2} \right)^2 \left(S_4 - \frac{S_2^2}{K} \right) \end{pmatrix}$$

Si on choisit la référence au centre de l'antenne $S_1 = 0$.

Si on suppose de plus que l'antenne est symétrique $S_3 = 0$.

Dans ce cas la matrice $T^T P T$ est diagonale.

3.5 Répartition optimale des capteurs d'une antenne linéaire

Jusqu'ici l'hypothèse de l'uniformité de l'antenne linéaire n'a pas été exploitée.

Pour l'instant supposons que la densité de capteurs est quelconque, décrite par une distribution .

On peut alors écrire :

$$S_2 = \sum_{k=1, K} x_k^2 = \frac{1}{\delta x} \int_{-L/2}^{+L/2} x^2 \mu(x) dx$$

On pourra évaluer δx de la manière suivante :

$$S_0 = \sum_{k=1, K} 1 = K = \frac{1}{\delta x} \int_{-L/2}^{+L/2} \mu(x) dx \quad \text{d'où} \quad \delta x = \frac{\int_{-L/2}^{+L/2} \mu(x) dx}{K}$$

Finalement on écrira $S_2 = K \langle x^2 \rangle$ en notant

$$\langle F \rangle = \frac{\int_{-L/2}^{L/2} F(x)\mu(x)dx}{\int_{-L/2}^{L/2} \mu(x)dx} \text{ la « valeur moyenne » de } F.$$

On aura de même $S_4 = K\langle x^4 \rangle$ et $S_3 = K\langle x^3 \rangle$.

La matrice de FISHER peut alors s'écrire :

$$F = \frac{2(Ks)^2}{1 + Ks} \left(\frac{2\pi}{\lambda} \right)^2 \begin{pmatrix} \langle x^2 \rangle & \frac{1-\alpha^2}{2} \langle x^3 \rangle \\ \frac{1-\alpha^2}{2} \langle x^3 \rangle & \left(\frac{1-\alpha^2}{2} \right)^2 (\langle x^4 \rangle - \langle x^2 \rangle^2) \end{pmatrix}$$

L'origine étant choisie de manière que $\langle x \rangle = 0$.

Sous cette forme on peut s'intéresser à la question suivante :

« Comment disposer les capteurs d'une antenne linéaire pour optimiser l'estimation de la distance, de l'angle ou de la position ? ».

Cette question peut être résolue en recherchant la distribution $\mu(x)$ telle que :

$\langle x^2 \rangle$ soit maximum pour l'estimation optimale de l'angle

$\langle x^4 \rangle - \langle x^2 \rangle^2$ soit maximum pour estimer au mieux la distance

$\langle x^2 \rangle (\langle x^4 \rangle - \langle x^2 \rangle^2)$ soit maximum pour estimer au mieux la position

Dans le cas de l'estimation de l'angle la meilleure répartition consiste à répartir la moitié des capteurs à chaque extrémité de l'antenne. Dans ce cas :

$$\langle x^2 \rangle = \left(\frac{L}{2} \right)^2 = \frac{L^2}{4}$$

Dans le cas de la distance il faut répartir les capteurs en trois fractions p, q, r avec $p + q + r = 1$ et $p = r$ (symétrie). Il faut maximiser par rapport à p :

$$2 \times p \times \left(\frac{L}{2} \right)^4 - 4 \times p^2 \times \left(\frac{L}{2} \right)^{2 \times 2} = \frac{L^4}{16} (2p - 4p^2)$$

On obtient $p=1/4$ et la meilleure répartition est donc $\left[\frac{1}{4} \quad \frac{1}{2} \quad \frac{1}{4} \right]$.

Dans le cas de la position il faut de même maximiser par rapport à p :

$$\frac{L^4}{16} (2p - 4p^2) 2p \frac{L^2}{4} = \frac{L^6}{16} (p^2 - 2p^3)$$

On trouve $2p - 6p^2 = 0$, soit $p = 1/3$.

La meilleure répartition pour mesurer la position sera $\left[\frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3} \right]$.

4 Approximation parabolique du lobe principal et largeur à 3 dB

Les résultats obtenus dans les sections précédentes permettent d'exprimer la largeur à 3 dB du lobe principal du traitement classique dans le cas général. Pour cela on écrira d'abord l'expression de la fonction de directivité du traitement classique sous la forme :

$$D(\underline{\theta}, \underline{\theta}_0) = \frac{|\underline{d}^+(\underline{\theta}) \underline{d}(\underline{\theta}_0)|^2}{\|\underline{d}^+(\underline{\theta})\|^2 \|\underline{d}^+(\underline{\theta}_0)\|^2}$$

Une approximation parabolique de la forme du lobe principal (autour de $\underline{\theta}_0$) est obtenue en écrivant le développement limité au second ordre de $D(\underline{\theta}, \underline{\theta}_0)$:

$$D(\underline{\theta}, \underline{\theta}_0) \cong 1 + \frac{1}{2} (\underline{\theta} - \underline{\theta}_0)^T (\partial^2 D)_0 (\underline{\theta} - \underline{\theta}_0)$$

On en déduit une expression de la largeur à 3 dB dans le cas multidimensionnel :

$$(\partial^2 D)_0 = (\Theta_3)^{-2}, \text{ c'est à dire } (2\Theta_3) = 2(-\partial^2 D)^{-1/2}$$

Si la matrice Hessienne $(\partial^2 D)_0$ est diagonale, on peut écrire $(\partial^2 D)_0 = \Lambda = \text{diag}\{\lambda_k\}_{k=1,N}$

Et donc $(2\theta_3)_k = 2(-\lambda_k)^{-1/2}$.

On peut appliquer les résultats obtenus précédemment :

$$(\partial_{kl}^2 D)_0 = -\frac{1}{K} \{ \underline{u}_k^+ P_0 \underline{u}_l + \underline{u}_l^+ P_0 \underline{u}_k \}$$

avec $P_0 = I - \frac{\underline{d}_0 \underline{d}_0^+}{K}$, projecteur orthogonal à \underline{d}_0 , $\underline{u}_k = \partial_k \underline{d}(\underline{\theta})$ et $\underline{d}_0 = \underline{d}[\underline{\theta}_0]$

$$(\partial^2 D)_0 = -\frac{2}{K} (\partial \underline{d}^+ P_0 \partial \underline{d})$$

On note ici : $\partial \underline{d} = [\partial_1 \underline{d} \quad \dots \quad \partial_N \underline{d}]$

Dans le cas d'un modèle « retard pur » et, plus précisément, pour des ondes sphériques, on obtient :

$$(\partial^2 D)_0 = \frac{2(2\pi f)^2}{K} T^T P T$$

avec

$$T^T P T = \frac{1}{c^2} \begin{pmatrix} \left(S_2 - \frac{S_1^2}{K} \right) & \frac{1-\alpha^2}{2} \left(S_3 - \frac{S_1 S_2}{K} \right) \\ \frac{1-\alpha^2}{2} \left(S_3 - \frac{S_1 S_2}{K} \right) & \left(\frac{1-\alpha^2}{2} \right)^2 \left(S_4 - \frac{S_2^2}{K} \right) \end{pmatrix}$$

et, plus précisément, pour une antenne linéaire uniforme avec référence au centre de l'antenne:

$$(\partial^2 D)_0 = -\frac{2\omega^2}{4c^2} \begin{pmatrix} \|\underline{\xi}_1\|^2 & 0 \\ 0 & \left(\|\underline{\xi}_2\|^2 - \frac{|\underline{\xi}_2^T \underline{\xi}_0|^2}{K} \right) \frac{\sin^4(\theta)}{4} \end{pmatrix}$$

et on peut préciser ici :

$$\|\underline{\xi}_1\|^2 = \frac{K(K^2 - 1)}{12} d^2 ; \quad \|\underline{\xi}_2\|^2 - \frac{|\underline{\xi}_2^T \underline{\xi}_0|^2}{K} = \frac{K(K^2 - 1)(K^2 - 4)}{180} d^4$$

D'où l'expression des largeurs d'ambiguïté en $\cos \theta$ et R^{-1} :

$$2\theta_3 = \frac{\sqrt{6}}{\pi} \frac{\lambda}{L} \left(\frac{K-1}{K+1} \right)^{1/2} \quad \text{et} \quad 2\alpha_3 = \frac{6\sqrt{10}}{\pi} \frac{\lambda}{L^2 \sin^2 \theta} \left(\frac{(K-1)^3}{(K+1)(K^2-4)} \right)^{1/2}$$

Formules que l'on peut approcher, si K est grand, par :

$$2\theta_3 = \frac{\sqrt{6}}{\pi} \frac{\lambda}{L} \quad \text{et} \quad 2\alpha_3 = \frac{6\sqrt{10}}{\pi} \frac{\lambda}{L^2 \sin^2 \theta}$$

ANNEXES

Annexe 1 : Formule de LIOUVILLE (Dérivée d'un déterminant).

- Calcul de $\partial|A|$

Ici, A est une matrice hermitienne, le symbole $|A|$ désigne le déterminant de A , et le symbole ∂X la dérivée de X par rapport à un paramètre (peu importe lequel, il n'y a pas d'ambiguïté).

A sera hermitienne, elle est donc diagonalisable par transformation unitaire U , ce que l'on peut écrire :

$$A = U\Lambda U^+$$

d'où l'on déduit :

$$|A| = |\Lambda| = \prod_{i=1,K} \lambda_i$$

On supposera que les valeurs propres $\{\lambda_k\}_{k=1,K}$ sont des fonctions dérivables des paramètres, alors :

$$\begin{aligned} \partial|A| &= \partial|\Lambda| = |\Lambda| \partial \ln |\Lambda| \\ &= |\Lambda| \sum_{k=1,K} \frac{\partial \lambda_k}{\lambda_k} \\ &= |A| \text{Tr}[\Lambda^{-1} \partial \Lambda] \end{aligned}$$

On obtient donc le résultat : $\partial \ln |A| = \text{Tr}[\Lambda^{-1} \partial \Lambda]$

mais on a :

$$\begin{aligned} \Lambda &= U^+ A U \\ \Lambda^{-1} &= U^+ A^{-1} U \\ \partial \Lambda &= (\partial U)^+ A U + U^+ (\partial A) U + U^+ A (\partial U) \\ \text{Tr}[\Lambda^{-1} \partial \Lambda] &= \text{Tr}[U^+ A^{-1} U (\partial U)^+ A U + U^+ A^{-1} (\partial A) U + U^+ (\partial U)] \end{aligned}$$

En utilisant la propriété d'invariance de la trace d'un produit par permutation circulaire de ses termes, on peut simplifier :

$$\text{Tr}[\Lambda^{-1} \partial \Lambda] = \text{Tr}[U (\partial U)^+ + A^{-1} (\partial A) + U^+ (\partial U)]$$

Cette dernière résultant de l'unitarité de la matrice U .

Finalement : $\partial \ln |A| = \text{Tr}[A^{-1} \partial A]$

- Calcul de ∂A^{-1}

De l'égalité $AA^{-1} = I$ on tire $(\partial A)A^{-1} + A\partial A^{-1} = 0$ soit : $\partial A^{-1} = -A^{-1}(\partial A)A^{-1}$

Annexe 2 : Densité Complexe Circulaire

On s'intéresse ici aux distributions statistiques de vecteurs à K composantes complexes.

$$\underline{X} = \underline{X}_R + i\underline{X}_I$$

En toute rigueur, il s'agit dans ce cas de la distribution du vecteur réel à $2K$ composantes réelles :

$$\underline{\xi} = \begin{pmatrix} \underline{X}_R \\ \underline{X}_I \end{pmatrix}$$

Par abus de langage, on parlera cependant de la densité de probabilité de \underline{X} plutôt que de celle de $\underline{\xi}$.

Lorsque $\underline{\xi}$ est gaussien, de moyenne $\underline{\mu}$ et de matrice de covariance C , sa densité de probabilité peut s'écrire :

$$p(\underline{\xi}) = |2\pi C|^{-1/2} \exp\left[-\frac{1}{2}(\underline{\xi} - \underline{\mu})^T C^{-1}(\underline{\xi} - \underline{\mu})\right]$$

Si l'on introduit dans cette expression la partition en partie réelle et imaginaire ($\underline{X}_R, \underline{X}_I$) on aura :

$$\underline{\mu} = \begin{pmatrix} \underline{m}_R \\ \underline{m}_I \end{pmatrix} \quad \text{et} \quad C = \begin{pmatrix} C_{RR} & C_{RI} \\ C_{IR} & C_{II} \end{pmatrix}$$

$$\begin{aligned} \underline{m}_R &= E\{\underline{X}_R\} \quad \text{et} \quad \underline{m}_I = E\{\underline{X}_I\} \\ C_{RR} &= E\{(\underline{X}_R - \underline{m}_R)(\underline{X}_R - \underline{m}_R)^T\} \\ C_{RI} &= E\{(\underline{X}_R - \underline{m}_R)(\underline{X}_I - \underline{m}_I)^T\} = C_{IR}^T \\ C_{II} &= E\{(\underline{X}_I - \underline{m}_I)(\underline{X}_I - \underline{m}_I)^T\} \end{aligned}$$

On introduit alors la propriété de circularité qui est une condition sur la structure de la matrice de covariance C .

$$C_{RR} = C_{II} = A \quad \text{et} \quad C_{IR} = -C_{RI} = B$$

Autrement dit, la matrice C se met sous la forme :

$$C = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

On rencontre ce type de symétrie dans les applications qui font intervenir la transformée de FOURIER (représentation spectrale).

L'intérêt de cette symétrie provient du fait que l'on peut exprimer la densité de probabilité de $\underline{\xi}$ en fonction des quantités :

$$\begin{aligned}\underline{X} &= \underline{X}_R + i\underline{X}_I \\ \underline{m} &= \underline{m}_R + i\underline{m}_I \\ \Gamma &= 2(A + iB) = E \left\{ (\underline{X} - \underline{m})(\underline{X} - \underline{m})^+ \right\}\end{aligned}$$

Le résultat se met sous la forme compacte suivante :

$$p(\underline{X}) = |\pi\Gamma|^{-1} \exp \left[-(\underline{X} - \underline{m})^+ \Gamma^{-1} (\underline{X} - \underline{m}) \right]$$

ce que l'on note simplement $p(\underline{X})$.

Cette formule n'est pas difficile à mémoriser puisqu'on l'obtient à partir de l'expression de $p(\underline{\xi})$ initiale en y remplaçant tous les facteurs 2 par 1 .

On va tout d'abord calculer $|C|$ en fonction de $|\Gamma|$.

Si Γ (Hermitienne) est diagonalisée sous la forme : $\Gamma = U\Lambda U^+$ on aura $|\Gamma| = |\Lambda| = \prod_{i=1,K} \lambda_i$

Si l'on pose $U = U_R + iU_I$ (matrice unitaire)

$$\begin{aligned}\Gamma &= 2(A + iB) = (U_R + iU_I)\Lambda(U_R - iU_I)^T \\ &= (U_R\Lambda U_R^T + U_I\Lambda U_I^T) + i(U_I\Lambda U_R^T - U_R\Lambda U_I^T)\end{aligned}$$

on en déduit A et B en fonction de U , Λ et, de même, on pourra écrire :

$$C = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = \begin{pmatrix} U_R & -U_I \\ U_I & U_R \end{pmatrix} \begin{pmatrix} \frac{\Lambda}{2} & 0 \\ 0 & \frac{\Lambda}{2} \end{pmatrix} \begin{pmatrix} U_R & -U_I \\ U_I & U_R \end{pmatrix}^T$$

La matrice $V = \begin{pmatrix} U_R & -U_I \\ U_I & U_R \end{pmatrix}$ est orthogonale, et donc $VV^T = V^T V = I_{2K}$

On en déduit $|C| = \left| \frac{\Lambda}{2} \right|^2 = \left| \frac{\Gamma}{2} \right|^2$ d'où $|2\pi C|^{-1/2} = |\pi\Gamma|^{-1}$

On pourra de même calculer $\frac{1}{2}(\underline{\xi} - \underline{\mu})^T C^{-1}(\underline{\xi} - \underline{\mu})$

On utilise pour cela le calcul de matrice C^{-1}

La matrice C^{-1} aura la même structure circulaire que C .

En effet, posons : $C^{-1} = \begin{pmatrix} U & -V \\ V & U \end{pmatrix}$, on aura $CC^{-1} = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \begin{pmatrix} U & -V \\ V & U \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$

Soit $AU - BV = I$ et $AV + BU = 0$

On en déduit que, $\Gamma = 2(A + iB)$ alors, $\Gamma^{-1} = \frac{1}{2}(U + iV)$

Puisque $\Gamma\Gamma^{-1} = (A + iB)(U + iV) = (AU - BV) + i(AV + BU)$

Finalement :

$$\begin{aligned} \frac{1}{2}(\underline{\xi} - \underline{\mu})^T C^{-1}(\underline{\xi} - \underline{\mu}) &= \frac{1}{2}[(\underline{X}_R - \underline{m}_R)^T, (\underline{X}_I - \underline{m}_I)^T] \begin{bmatrix} U & -V \\ V & U \end{bmatrix} \begin{bmatrix} (\underline{X}_R - \underline{m}_R) \\ (\underline{X}_I - \underline{m}_I) \end{bmatrix} \\ &= \frac{1}{2}(\underline{X}_R - \underline{m}_R)^T \{U(\underline{X}_R - \underline{m}_R) - V(\underline{X}_I - \underline{m}_I)\} \\ &\quad + \frac{1}{2}(\underline{X}_I - \underline{m}_I)^T \{V(\underline{X}_R - \underline{m}_R) + U(\underline{X}_I - \underline{m}_I)\} \\ &= (\underline{X} - \underline{m})^+ \Gamma^{-1}(\underline{X} - \underline{m}) \end{aligned}$$

En bref, on a démontré que :

$$|2\pi C|^{-1/2} = |\pi\Gamma|^{-1} \text{ et } \frac{1}{2}(\underline{\xi} - \underline{\mu})^T C^{-1}(\underline{\xi} - \underline{\mu}) = (\underline{X} - \underline{m})^+ \Gamma^{-1}(\underline{X} - \underline{m})$$

En conclusion,

$$p(\underline{\xi}) = |\pi\Gamma|^{-1} \exp[-(\underline{X} - \underline{m})^+ \Gamma^{-1}(\underline{X} - \underline{m})]$$

et, plutôt que de parler de $\underline{\xi}$, on écrira $p(\underline{X})$ au lieu de $p(\underline{\xi})$

Les vecteurs complexes circulaires ont quelques propriétés sympathiques, curieuses au premier abord.

On supposera $\underline{m} = \underline{0}$, donc $E\{\underline{X}\underline{X}^+\} = \Gamma$ et $E\{\underline{X}\underline{X}^T\} = 0$

En effet $E\{\underline{X}\underline{X}^T\} = E\{(\underline{X}_R + i\underline{X}_I)(\underline{X}_R + i\underline{X}_I)^T\} = (C_{RR} - C_{II}) + i(C_{RI} + C_{RI}) = 0$
d'après la condition de circularité.

Un cas particulier intéressant est celui où $K=I$, ($\underline{X} \in C$). La densité de probabilité de $z=x+iy$ est :

$$p(z) = (\pi\sigma^2)^{-1} \exp\left(-\frac{|z|^2}{\sigma^2}\right) \text{ avec } E\{z^2\} = 0 \text{ et } E\{|z|^2\} = \sigma^2$$

Les conditions de circularité s'écrivent :

$$\text{Var}(x) = \text{Var}(y) = \frac{\sigma^2}{2} \text{ et } E\{xy\} = 0$$