

Un nouvel indice de performance et sa borne inférieure

Abdelhamid ZAIDI

Laboratoire de Modélisation et Calcul
BP 53, 38041 Grenoble Cedex 9, France
Abdelhamid.Zaidi@imag.fr

Résumé – Nous proposons, dans ce papier, un nouvel indice, invariant, pour mesurer les performances des méthodes de séparation d'un mélange instantané de sources colorées. Nous en donnons une interprétation statistique et nous déterminons l'expression de sa borne inférieure en fonction des densités spectrales des différentes sources. Puis nous comparons, par simulation, cet indice à un critère usuellement utilisé dans l'analyse des performances des méthodes de séparation.

Abstract – We propose, in this paper, a new invariant index to measure the performances of the separating method of an instantaneous mixture of colored sources. We give a statistical interpretation to this index and we determine the expression of its lower bound in terms of the spectral densities of the different sources. Then we compare, by computer simulation, this index to the one usually used in the analysis of the performances of the separating methods.

1 Indice de performance I_{EPF}

Nous introduisons ici un nouvel indice de performance pour mesurer la qualité des méthodes de séparation d'un mélange linéaire instantané de sources colorées décrit par le modèle :

$$X(t) = AS(t), \quad t = 1, \dots, T,$$

où $X(\cdot)$ est le processus observable, A est une matrice régulière de taille (K, K) , appelée matrice de mélange, et $S(\cdot)$ est le processus source. Les sources $S_k(t)$, $k = 1, \dots, K$, composantes de $S(t)$, sont gaussiennes autorégressives d'ordre p_k , mutuellement indépendantes. Chaque modèle $AR(p_k)$ est paramétré par les coefficients du polynôme autorégressif associé : $a_k(j)$, $j = 1, \dots, p_k$. On note a le paramètre de la structure des sources constitué par l'ensemble $\{a_k(j), k = 1, \dots, K, j = 1, \dots, p_k\}$. Ainsi $\theta = (B, a)$ constitue le paramètre global du problème, où $B = A^{-1}$ est la matrice de séparation. En fait a est un paramètre fantôme, car seule la matrice de séparation est nécessaire pour la reconstitution. Enfin notons $\theta_0 = (B_0, a^0)$ la vraie valeur du paramètre θ .

Nous suivons ici la même démarche qu'Akaike [1] dans l'élaboration du critère, basé sur l'Erreur de Prédiction Finale, dans l'estimation de l'ordre d'un modèle autorégressif. Cet indice, noté I_{EPF} , est défini en fonction de la matrice du système global $\hat{C}_{A_0} = \hat{B}A_0$ par :

$$I_{EPF} = \sum_{i,j=1}^K \text{EQM}([\hat{C}_{A_0}]_{ij}),$$

où \hat{B} est l'estimation de B et $\text{EQM}([\hat{C}_{A_0}]_{ij})$ est l'erreur quadratique moyenne relative au coefficient d'indices (i, j) de la matrice \hat{C}_{A_0} . Il se justifie de la façon suivante. On suppose que B a été estimé à partir des observations

$$X(t) = A_0S(t), \quad t = 1, \dots, T.$$

On considère une nouvelle observation $X(t_0) = A_0S(t_0)$, indépendante des précédentes, pour laquelle on calcule l'erreur quadratique moyenne commise en estimant $S(t_0)$ par

$$\hat{S}(t_0) = \hat{B}X(t_0) = \hat{C}_{A_0}S(t_0).$$

On a :

$$\begin{aligned} \mathbb{E} \left(\|\hat{S}(t_0) - S(t_0)\|^2 \right) &= \mathbb{E} \left(\|(\hat{C}_{A_0} - I)S(t_0)\|^2 \right) \\ &= \text{trace} \left\{ \mathbb{E} \left((\hat{C}_{A_0} - I)^t (\hat{C}_{A_0} - I) S(t_0) S(t_0)^t \right) \right\} \\ &= \text{trace} \left\{ \mathbb{E} \left((\hat{C}_{A_0} - I)^t (\hat{C}_{A_0} - I) \text{cov}(S(t_0)) \right) \right\}. \end{aligned}$$

En supposant de plus que les sources ont la même puissance, $\text{cov}(S) = I_K$, on obtient :

$$\mathbb{E} \left(\|\hat{S}(t_0) - S(t_0)\|^2 \right) = \sum_{k,j} \text{EQM}([\hat{C}_{A_0}]_{kj}) = I_{EPF}.$$

Il est clair que pour un estimateur invariant \hat{B} , l'indice I_{EPF} ne dépend pas de la matrice de mélange A_0 mais uniquement du processus des sources $S(\cdot)$.

Le terme $\text{EQM}([\hat{C}_{A_0}]_{kj})$, $k \neq j$, représente le *taux de contamination* de \hat{S}_k par S_j . L'indice I_{EPF} que nous proposons diffère de l'indice de performance I_{perf} , proposé par Cardoso et al [2], par l'injection dans I_{perf} des termes diagonaux $\text{EQM}([\hat{C}_{A_0}]_{kk})$, $k = 1, \dots, K$:

$$I_{EPF} = I_{perf} + \sum_{k=1}^K \text{EQM}([\hat{C}_{A_0}]_{kk}).$$

2 Borne inférieure de I_{EPF}

Nous déterminons, dans ce paragraphe, la borne inférieure de l'indice de performance I_{EPF} d'un estimateur sans biais de la matrice de séparation en fonction des paramètres

des modèles autorégressifs associés aux différentes sources. Asymptotiquement, cette borne inférieure est inversement proportionnelle à T et s'exprime uniquement à l'aide des densités spectrales des sources.

Si l'estimateur de la matrice de séparation est sans biais, ce qui est asymptotiquement le cas pour celui des méthodes MVE [4] et SOBI [2] utilisées en simulation ci-dessous, alors

$$I_{EPF} = \sum_{i,j=1}^K \text{cov}((\hat{C}_{A_0})_{ij}). \quad (1)$$

Soit C une matrice carrée de taille K . On définit le vecteur $[C]$ de \mathbb{R}^{K^2} par les lignes de la matrice C :

$$[C] = (C_1^t, \dots, C_K^t)^t.$$

Soit $\langle C \rangle_K$ la matrice bloc-diagonale de \mathbb{R}^{K^2} dont les blocs sont réduits à la matrice C . Alors

$$[\hat{C}_{A_0}] = \langle A_0^t \rangle_K [\hat{B}].$$

Sous l'hypothèse de sources de puissance unité, l'expression de la log-vraisemblance de la séquence observée,

$$X_T = [X(1)^t, \dots, X(T)^t],$$

est donnée par :

$$\begin{aligned} l_T(X_T; B, a) = & -\frac{1}{2} \sum_{k=1}^K \frac{1}{\sigma_{\varepsilon_k}^2} e_k^t B X_T (R_k^T(a_k))^{-1} X_T^t B^t e_k \\ & -\frac{1}{2} \sum_{k=1}^K \{ \log(\det(R_k^T(a_k))) + K \log \sigma_{\varepsilon_k}^2 \} \\ & + T \log |\det B| - \frac{KT}{2} \log(2\pi). \end{aligned}$$

La matrice d'information de Fisher relative au paramètre θ est définie par :

$$I_T(\theta_0) = \mathbb{E} \left[\left(\frac{\partial l_T}{\partial \theta} \right)_{\theta_0} \left(\frac{\partial l_T}{\partial \theta} \right)_{\theta_0}^t \right] = -\mathbb{E} \left[\left(\frac{\partial^2 l_T}{\partial \theta^2} \right)_{\theta_0} \right].$$

Le bloc supérieur gauche d'ordre K^2 de l'inverse de la matrice $I_T(\theta_0)$, noté $I_T(B_0)$, constitue la borne inférieure de la matrice de covariance du vecteur $[\hat{B}]$. Ce bloc est "minoré" par l'inverse du coin supérieur gauche d'ordre K^2 de la matrice I_T , noté $[I_T(B_0)]^{-1}$:

$$\text{cov}([\hat{B}]) \geq I_T^{-1}(B_0) \geq [I_T(B_0)]^{-1}.$$

L'expression de la matrice d'information de Fisher a été calculée dans [4]. On montre que la matrice $I_T(\theta_0)/T$ est asymptotiquement bloc-diagonale. Par conséquent les matrices $I_T^{-1}(B_0)$ et $[I_T(B_0)]^{-1}$ sont asymptotiquement égales. Ainsi la matrice de covariance du vecteur $[\hat{C}_{A_0}]$ est "minorée" par :

$$J_T = \langle A_0^t \rangle_K [I_T(B_0)]^{-1} \langle A_0 \rangle_K.$$

On en déduit l'expression de la borne inférieure de l'indice I_{EPF} :

$$I_{EPF} = \sum_{i,j=1}^K (J_T)_{ij,ij}.$$

Proposition 2.1 L'expression de la matrice $I_T(B_0)$ est donnée par :

$$\begin{aligned} -\mathbb{E} \left[\left(\frac{\partial^2 l_T}{\partial b_{ij} \partial b_{ki}} \right)_{\theta_0} \right] &= T a_{jk}^0 a_{ki}^0 + \\ \frac{\delta_{ki}}{\sigma_{\varepsilon_k}^2} \sum_{m,n=0}^{p_i} (T-m-n) a_i^0(m) a_i^0(n) \sum_{u=1}^K a_{ju}^0 a_{iu}^0 \gamma_u^0(m-n), \end{aligned}$$

où $\gamma_k^0(m)$ est l'autocovariance d'ordre m du processus $AR(p_k)$ de puissance unité.

Proposition 2.2 La matrice J_T^{-1} se réécrit sous la forme $T\Pi + D$, où Π est la matrice de permutation qui échange la ligne kj avec la ligne jk et conserve la ligne kk et où D est la matrice diagonale définie par :

$$d(kj, kj) = \frac{1}{\sigma_{\varepsilon_k}^2} \sum_{m,n=0}^{p_k} (T-m-n) a_k^0(m) a_k^0(n) \gamma_j^0(m-n).$$

Nous donnons maintenant l'expression des coefficients de la matrice J_T^{-1} à l'aide de la densité spectrale f_k^0 associée à la $k^{\text{ème}}$ source et de la fonction g_k^0 définie par :

$$g_k^0(\lambda) = \phi_k(e^{-i\lambda}),$$

où $\phi_k^0(\cdot)$ est le filtre linéaire du modèle $AR(p_k)$:

$$\phi_k^0(z) = \sum_{i=0}^{p_k} a_k^0(i) z^i.$$

La densité spectrale $f_k^0(\cdot)$ est liée à la fonction $g_k^0(\cdot)$ par la relation :

$$f_k^0(\lambda) = \frac{\sigma_{\varepsilon_k}^2}{2\pi} |g_k^0(\lambda)|^{-2}.$$

De plus, la fonction d'autocovariance $\gamma_k^0(\cdot)$ est la transformée de Fourier de la densité spectrale $f_k^0(\cdot)$:

$$\gamma_k^0(j) = \int_{-\pi}^{\pi} e^{i\lambda j} f_k^0(\lambda) d\lambda.$$

Le $kj^{\text{ème}}$ terme diagonal de la matrice D , introduite dans la Proposition 2.2, s'exprime en fonction de f_k^0 et g_k^0 . En effet,

$$\begin{aligned} d(kj, kj) &= \frac{1}{\sigma_{\varepsilon_k}^2} \sum_{m,n=0}^{p_k} (T-m-n) a_k^0(m) a_k^0(n) \gamma_j^0(m-n) \\ &= \frac{T}{\sigma_{\varepsilon_k}^2} \int_{-\pi}^{\pi} \sum_{m,n=0}^{p_k} a_k^0(m) a_k^0(n) e^{i\lambda(m-n)} f_j^0(\lambda) d\lambda \\ &\quad - \frac{1}{\sigma_{\varepsilon_k}^2} \int_{-\pi}^{\pi} \sum_{m,n=0}^{p_k} (m+n) a_k^0(m) a_k^0(n) e^{i\lambda(m-n)} f_j^0(\lambda) d\lambda \\ &= \frac{T}{2\pi} \int_{-\pi}^{\pi} \frac{f_j^0(\lambda)}{f_k^0(\lambda)} d\lambda - \frac{1}{\sigma_{\varepsilon_k}^2} \int_{-\pi}^{\pi} |g_k^0(\lambda)|^2 f_j^0(\lambda) d\lambda. \end{aligned}$$

D'après la Proposition 2.2, la matrice J_T^{-1} est bloc-diagonale dont les blocs sont réduits soit à l'élément $T + d(kk, kk)$, $k = 1, \dots, K$, soit à la matrice

$$\begin{pmatrix} d(kj, kj) & T \\ T & d(jk, jk) \end{pmatrix}, \quad k \neq j.$$

On en déduit l'expression de I_{EPF} en fonction de f_k^0 et g_k^0 :

$$I_{EPF} = \sum_{k < j} \frac{d(kj, kj) + d(jk, jk)}{d(kj, kj)d(jk, jk) - T^2} + \sum_k \frac{1}{T + d(kk, kk)}.$$

Notons que lorsque T tend vers $+\infty$, la matrice J_T^{-1}/T tend vers une matrice bloc-diagonale dont les blocs sont réduits soit à la valeur 2 soit à la matrice

$$\begin{pmatrix} \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{f_j^0(\lambda)}{f_k^0(\lambda)} d\lambda & 1 \\ 1 & \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{f_k^0(\lambda)}{f_j^0(\lambda)} d\lambda \end{pmatrix}, \quad k \neq j.$$

La limite de la matrice J_T^{-1}/T coïncide avec la limite de la borne supérieure de l'inverse de la matrice de covariance de l'élément $\tilde{C} = \tilde{B}A_0$, où \tilde{B} est l'estimateur de la matrice B obtenu par la méthode QMV utilisant des filtres séparateurs optimaux [3]. Asymptotiquement I_{EPF} est inversement proportionnel à T . En effet :

$$\lim_{T \rightarrow +\infty} T I_{EPF} = 2\pi \sum_{k < j} \frac{\int_{-\pi}^{\pi} \frac{f_j^0(\lambda)}{f_k^0(\lambda)} d\lambda + \int_{-\pi}^{\pi} \frac{f_k^0(\lambda)}{f_j^0(\lambda)} d\lambda}{\int_{-\pi}^{\pi} \frac{f_j^0(\lambda)}{f_k^0(\lambda)} d\lambda \int_{-\pi}^{\pi} \frac{f_k^0(\lambda)}{f_j^0(\lambda)} d\lambda - 4\pi^2} + \frac{K}{2}.$$

En faisant tendre vers l'infini les ordres p_k des modèles autorégressifs, cette expression est valable plus généralement pour des sources suffisamment régulières (par exemple ARMA inversibles).

2.1 Indice d'efficacité

On appelle indice d'efficacité le rapport, à valeurs dans $[0, 1]$, de la borne inférieure de l'indice de performance I_{EPF} sur l'indice I_{EPF} :

$$e = \frac{I_{EPF}}{I_{EPF}}. \quad (2)$$

Cet indice représente une mesure intrinsèque des performances d'une méthode de séparation. Ainsi, une méthode est d'autant plus efficace que son indice d'efficacité est proche de un. Une estimation de cet indice est obtenue à partir d'une évaluation empirique de I_{EPF} et de la valeur exacte de I_{EPF} .

3 Simulation

Nous avons choisi de comparer les deux indices de performance à travers deux méthodes de séparation : SOBI [2], qui est d'un usage très fréquent, et MVE [4] qui constitue en quelque sorte une référence pour l'efficacité de telles méthodes.

3.1 Illustration

Nous considérons un mélange instantané de trois sources autorégressives, gaussiennes, centrées, de puissance unité et d'ordres respectifs 2, 3 et 4. Les modèles autorégressifs sont paramétrés par les autocorrélations partielles :

$$\beta = \{0, -0.8; 0.5, -0.2, 0.7; 0.2, -0.7, -0.4, 0.9\}.$$

La matrice de mélange A et la matrice de séparation B sont données par :

$$A = \begin{bmatrix} 1 & 4 & 7 \\ 3 & 2 & 1 \\ 9 & 6 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} -0.2 & -1.1 & -0.5 \\ 0.3 & 2.9 & -1 \\ 0 & -1.5 & 0.5 \end{bmatrix}.$$

La taille T des séquences simulées ainsi que le nombre r de répétitions utilisées sont indiqués dans le Tableau 1. L'ordre des modèles autorégressifs estimés, dans le cas du maximum de vraisemblance, est fixé *a priori* à 6 et prend ainsi en compte les autocovariances empiriques correspondant aux retards $0, 1, \dots, 6$. La méthode SOBI est également basée sur ces mêmes retards. Pour chaque répétition, on détermine la matrice $\hat{C}_i = \hat{B}_i A$, par permutation et changement de signe éventuel de ses lignes, pour laquelle la somme des termes diagonaux (alors positifs) est maximale. On lui associe aussi la version normalisée $\tilde{C}_i = [\text{Diag}\{\hat{C}_i\}]^{-1} \hat{C}_i$. On estime alors I_{EPF} en utilisant la version empirique de l'EQM :

$$EQM\{\hat{C}(l, k)\} \simeq \frac{1}{r} \sum_{i=1}^r [\hat{C}_i(l, k) - \delta_{lk}]^2,$$

où δ_{lk} est le symbole de Kronecker. L'indice I_{perf} ne prend pas en compte les termes diagonaux : il élimine l'erreur d'estimation des puissances relatives des sources. C'est pourquoi on utilise \tilde{C} dans son estimation,

$$EQM\{\tilde{C}(l, k)\} \simeq \frac{1}{r} \sum_{i=1}^r \tilde{C}_i^2(l, k), \quad k \neq l$$

T	50	250	500	1000	10000	100000
r	10000	2000	1000	500	50	10

TAB. 1: r en fonction de T

3.2 Commentaires

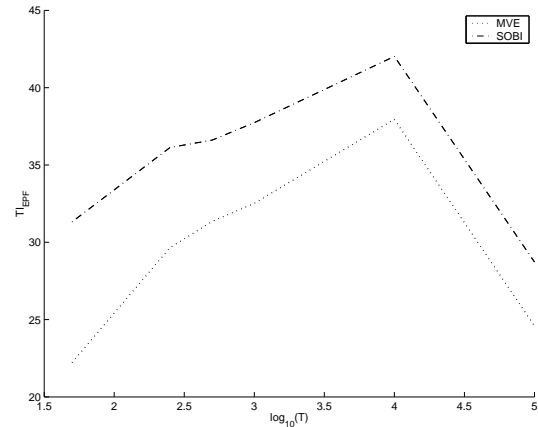


FIG. 1: $T \times I_{EPF}$ en fonction de $\log_{10}(T)$

Les Figures 1 et 2 représentent les variations de $T \times I_{EPF}$ et $T \times I_{perf}$ en fonction de $\log_{10}(T)$.

On constate que les deux critères évoluent en sens inverse l'un de l'autre. Par ailleurs, les valeurs prises par

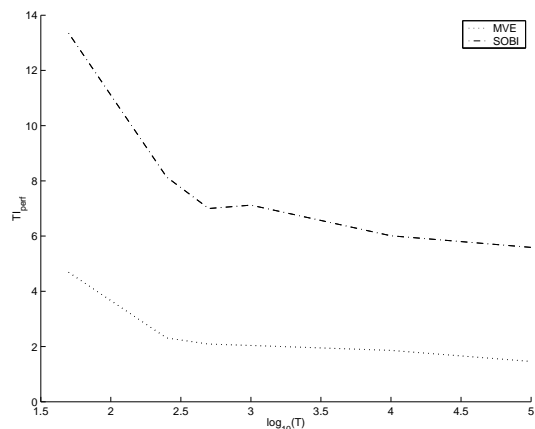


FIG. 2: $T \times I_{perf}$ en fonction de $\log_{10}(T)$

le premier indice ne sont pas du même ordre que celles prises par le second. Cependant, la comparaison des deux méthodes par les deux indices donne le même classement.

Références

- [1] Akaike, H., *Statistical predictor identification*, Ann. Inst. Math, 22, pp.203-217, 1970.
- [2] Belouchrani, A., K. Abed-Meraim, J.-F. Cardoso and E. Moulines, *A blind Source Separation Technique Using Second-Order Statistics*, IEEE, Trans. on Signal Processing, Vol. SP-45, N° 2, pp. 434-444, Feb. 1997.
- [3] Pham, D-T. and P. Garat, *Blind separation of mixture of independent sources through a quasi-maximum likelihood approach*, IEEE, Trans. on Signal Processing, vol. 45, N° 7, pp. 1712-1725, july 1997.
- [4] Zaidi, A., *Séparation aveugle d'un mélange instantané de sources autorégressives gaussiennes par la méthode du maximum de vraisemblance exact*, Thèse de Doctorat, Univ. J. Fourier, Grenoble, 2000.