

Caractérisations asymptotiques pour les composantes principales des M -estimateurs

Gordana DRAŠKOVIĆ¹, Arnaud BRELOY², Frédéric PASCAL¹

¹ L2S, CentraleSupélec, ² LEME, University Paris-Nanterre.

{gordana.draskovic, frederic.pascal}@l2s.centralesupelec.fr, abreloy@parisnanterre.fr

Résumé – Les paramètres de décomposition en valeurs propres (EVD) des statistiques de second ordre sont omniprésents en analyse statistique et en traitement du signal. Dans le but de caractériser le comportement de ces quantités, cet article propose de nouvelles caractérisations asymptotiques pour les paramètres de la EVD d'un M -estimateur pour des données suivant une distribution symétrique elliptique complexe (CES). Tout d'abord, leur distribution asymptotique gaussienne est obtenue en étendant les résultats standards pour la matrice de covariance empirique dans un contexte gaussien. Ensuite, leur vitesse de convergence vers les paramètres d'EVD d'un équivalent Wishart est dérivé. Ce deuxième résultat représente la contribution principale, en ce sens qu'il quantifie le moment où il est acceptable de utiliser directement des résultats établis pour l'EVD d'une matrice Wishart afin de caractériser les performances des M -estimateurs.

Abstract – The eigenvalue decomposition (EVD) parameters of the second order statistics are ubiquitous in statistical analysis and signal processing. Notably, the EVD of M -estimators of the scatter is a popular choice to perform robust probabilistic PCA or other dimension reduction related applications. Towards the goal of characterizing the behavior of these quantities, this paper proposes new asymptotics for the EVD parameters of the scatter M -estimator in the context of complex elliptically symmetric (CES) distributions. First, their Gaussian asymptotic distribution is obtained by extending standard results on the sample covariance matrix in a Gaussian context. Second, their convergence rate towards the EVD parameters of a Gaussian-Core Wishart Equivalent is derived. This second result represents the main contribution in the sense that it quantifies when it is acceptable to directly plug-in well-established results on the EVD of Wishart distributed matrix for characterizing the EVD of M -estimators.

1 Introduction

La statistique de second ordre joue un rôle important dans les applications de traitement du signal et d'apprentissage. En général, ce paramètre est inconnu et doit être estimé pour appliquer un processus dit adaptatif. Dans ce cadre, les M -estimateurs de matrice de dispersion [1, 2] ont suscité beaucoup d'intérêt pour leur robustesse aux échantillons aberrants et aux distributions à queue lourdes, i.e., là où la traditionnelle matrice de covariance empirique (SCM) ne permet pas une estimation précise.

La caractérisation statistique des M -estimateurs est un problème complexe car ceux-ci sont définis par des équations à point fixe. Alors que la SCM dans un cadre gaussien suit une distribution de Wishart, [3], la distribution exacte des M -estimateurs reste inconnue. Néanmoins, leur distribution asymptotique gaussienne a été dérivée dans [4] et étendue au cas complexe dans [2, 5]. Il a aussi été récemment montré que leur distribution peut être précisément approchée par un modèle d'équivalent Wishart (EW) [6, 7].

Par ailleurs, les valeurs/vecteurs propres de la matrice de dispersion sont des paramètres omniprésents en analyse statistique. La décomposition en éléments propres (EVD) des M -estimateurs est utilisée dans de nombreuses applications telles que l'analyse en composantes principales (PCA) [8, 9], ou la sélection d'ordre de modèle [10]. Par conséquent, caractériser avec précision la distribution de la EVD des M -estimateurs

représente un intérêt pour l'analyse et la conception de traitements robustes.

Dans cet article, nous dérivons de nouvelles propriétés asymptotiques pour l'EVD des M -estimateurs pour des données ayant une distribution complexe elliptique symétrique (CES). Nous dérivons ces résultats dans le régime asymptotique standard et puis dans le régime de EW. Finalement, nous illustrons comment les résultats proposés peuvent être exploités dans la question complexe de la caractérisation du biais intrinsèque [11] des M -estimateurs.

Notations : \mathbf{K} est la matrice de commutation qui transforme $\text{vec}(\mathbf{A})$ en $\text{vec}(\mathbf{A}^T)$. \mathcal{H}_M^{++} est l'ensemble des matrices définies positives hermitiennes. L'ensemble des matrices semi-unitaires est noté $\mathcal{U}_r^p = \{\mathbf{U} \in \mathbb{C}^{p \times r}, \mathbf{U}^H \mathbf{U} = \mathbf{I}_r\}$. Enfin, $\mathcal{GCN}(\mathbf{0}, \mathbf{X}, \mathbf{Y})$ indique la distribution normale complexe non circulaire de moyenne nulle avec la matrice de covariance \mathbf{X} et la matrice de pseudo-covariance \mathbf{Y} [2].

2 Contexte

2.1 Distributions CES

Soit \mathbf{z} un vecteur aléatoire complexe circulaire de taille p . \mathbf{z} suit une distribution CES, notée $\mathcal{CES}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_{\mathbf{z}})$, si sa densité de probabilité (DDP) peut s'écrire

$$f_{\mathbf{z}}(\mathbf{z}) = |\boldsymbol{\Sigma}|^{-1} g_{\mathbf{z}}((\mathbf{z} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \boldsymbol{\mu})) \quad (1)$$

où $g_z : [0, \infty) \rightarrow [0, \infty)$ est une fonction telle que (1) définit une DDP, $\boldsymbol{\mu}$ est l'espérance de \mathbf{z} et $\boldsymbol{\Sigma}$ sa matrice de dispersion. $\boldsymbol{\Sigma}$ contient la structure de la matrice de covariance de \mathbf{z} , i.e., cette matrice de covariance (si elle existe) est proportionnelle à $\boldsymbol{\Sigma}$. La loi gaussienne, notée $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, est un cas particulier de distribution CES obtenu pour $g_{\mathbf{x}}(x) \propto e^{-x}$. Dans ce travail, on suppose $\boldsymbol{\mu} = \mathbf{0}$.

Représentation par noyaux gaussiens : Dans [7], le modèle de "noyau gaussien" a été utilisé comme alternative à la représentation stochastique classique [12] des vecteurs distribués CES. Un vecteur aléatoire $\mathbf{z} \sim \mathcal{CES}(\mathbf{0}, \boldsymbol{\Sigma}, g_z)$ peut être représenté par

$$\mathbf{z} \stackrel{d}{=} \frac{\sqrt{\mathcal{Q}}}{\|\mathbf{g}\|} \mathbf{A}\mathbf{g} \quad (2)$$

où $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^H$ est une factorisation de $\boldsymbol{\Sigma}$ et $\mathbf{g} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. \mathcal{Q} est une variable aléatoire réelle non négative, indépendante de \mathbf{g} avec une DDP dépendant uniquement de g_z . Nous appelons $\mathbf{x} = \mathbf{A}\mathbf{g}$ le noyau gaussien de \mathbf{z} .

2.2 Distribution de Wishart

Soit $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ un n -échantillon de vecteurs complexes indépendants de taille p avec $\mathbf{x}_i \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$. Alors, la SCM donnée par

$$\widehat{\boldsymbol{\Sigma}}_{SCM} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^H \quad (3)$$

suit au facteur $1/n$ près une distribution de Wishart complexe, notée $\mathcal{CW}_p(n, \boldsymbol{\Sigma})$. La DDP de $\mathbf{W} \sim \mathcal{CW}_p(n, \boldsymbol{\Sigma})$ est donnée par

$$f(\mathbf{W}) = \frac{|\mathbf{W}|^{n-p}}{|\boldsymbol{\Sigma}|^n \Gamma_p(n)} \exp[-\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{W})].$$

2.3 M -estimateurs

Soit $(\mathbf{z}_1, \dots, \mathbf{z}_n)$ un n -échantillon de vecteurs complexes indépendants de taille p avec $\mathbf{z}_i \sim \mathcal{CES}(\mathbf{0}, \boldsymbol{\Sigma}, g_z)$. Un M -estimateur de $\boldsymbol{\Sigma}$, noté $\widehat{\boldsymbol{\Sigma}}$, est défini comme la solution de l'équation suivante

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^n u(\mathbf{z}_i^H \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{z}_i) \mathbf{z}_i \mathbf{z}_i^H \quad (4)$$

où u est une fonction réelle de pondération définie sur $[0, \infty)$ qui n'est pas nécessairement reliée à la DDP de la distribution CES. La matrice de dispersion théorique de l'échantillon est définie comme la solution de

$$\mathbb{E}[\varphi(\mathbf{z}^H \boldsymbol{\Sigma}_\sigma^{-1} \mathbf{z}) \mathbf{z} \mathbf{z}^H] = \boldsymbol{\Sigma}_\sigma = \sigma^{-1} \boldsymbol{\Sigma}, \quad (5)$$

où $\sigma > 0$ est solution de $\mathbb{E}[\psi(\sigma\tau)] = p$ avec $\psi(\sigma\tau) = u(\sigma\tau)\sigma\tau$ et $\tau = \mathbf{z}^H \boldsymbol{\Sigma}^{-1} \mathbf{z}$.

2.4 Propriétés asymptotiques

Théorème 2.1 (Régime standard) Si $\widehat{\boldsymbol{\Sigma}}$ est un M -estimateur respectant les conditions de [1], alors

$$\sqrt{n} \text{vec} \left(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_\sigma \right) \stackrel{d}{\rightarrow} \mathcal{GCN}(\mathbf{0}, \mathbf{C}, \mathbf{P})$$

où les matrices de covariance et de pseudo-covariance asymptotiques sont

$$\begin{cases} \mathbf{C} = \vartheta_1 \boldsymbol{\Sigma}_\sigma^T \otimes \boldsymbol{\Sigma}_\sigma + \vartheta_2 \text{vec}(\boldsymbol{\Sigma}_\sigma) \text{vec}(\boldsymbol{\Sigma}_\sigma)^H, \\ \mathbf{P} = \vartheta_1 \left(\boldsymbol{\Sigma}_\sigma^T \otimes \boldsymbol{\Sigma}_\sigma \right) \mathbf{K} + \vartheta_2 \text{vec}(\boldsymbol{\Sigma}_\sigma) \text{vec}(\boldsymbol{\Sigma}_\sigma)^T. \end{cases} \quad (6)$$

Les constantes $\vartheta_1 > 0$ et $\vartheta_2 > -\vartheta_1/p$ sont données dans [5].

Modèle équivalent Wishart : Soit $(\mathbf{z}_1, \dots, \mathbf{z}_n)$ un n -échantillon où $\mathbf{z}_i = \sqrt{\mathcal{Q}_i}/\|\mathbf{g}_i\| \mathbf{A}\mathbf{g}_i$ suivant $\mathcal{CES}(\mathbf{0}, \boldsymbol{\Sigma}, g_z)$, $i = 1, \dots, n$, et

- $\widehat{\boldsymbol{\Sigma}}$ est un M -estimateur construit avec $(\mathbf{z}_1, \dots, \mathbf{z}_n)$ en utilisant Eq. (4),
- $\widehat{\boldsymbol{\Sigma}}_{EW} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^H$ est le SCM construit avec des noyaux gaussiens $\mathbf{x}_i = \mathbf{A}\mathbf{g}_i \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$ de \mathbf{z}_i , $i = 1, \dots, n$, donnés par Eq. (2) qui représentent des données fictives (car non accessibles) utilisées pour l'analyse théorique. Notons que $\widehat{\boldsymbol{\Sigma}}_{EW}$ suit donc, par construction, une loi de Wishart.

Théorème 2.2 (EW) Soit σ donné par Eq. (5). Ensuite, la distribution asymptotique de $\sigma \widehat{\boldsymbol{\Sigma}} - \widehat{\boldsymbol{\Sigma}}_{EW}$ est donnée par [7]

$$\sqrt{n} \text{vec} \left(\sigma \widehat{\boldsymbol{\Sigma}} - \widehat{\boldsymbol{\Sigma}}_{EW} \right) \stackrel{d}{\rightarrow} \mathcal{GCN}(\mathbf{0}, \widetilde{\mathbf{C}}, \widetilde{\mathbf{P}}) \quad (7)$$

où $\widetilde{\mathbf{C}}$ et $\widetilde{\mathbf{P}}$ sont définis par

$$\begin{cases} \widetilde{\mathbf{C}} = \sigma_1 \boldsymbol{\Sigma}^T \otimes \boldsymbol{\Sigma} + \sigma_2 \text{vec}(\boldsymbol{\Sigma}) \text{vec}(\boldsymbol{\Sigma})^H, \\ \widetilde{\mathbf{P}} = \sigma_1 \left(\boldsymbol{\Sigma}^T \otimes \boldsymbol{\Sigma} \right) \mathbf{K} + \sigma_2 \text{vec}(\boldsymbol{\Sigma}) \text{vec}(\boldsymbol{\Sigma})^T \end{cases} \quad (8)$$

avec σ_1 et σ_2 donnés dans [7].

Une remarque importante est que les facteurs σ_1 et σ_2 sont beaucoup plus petits que ceux du régime asymptotique standard (Eq. (6)), ce qui signifie que le comportement d'un M -estimateur peut être mieux approché par le comportement de la matrice de Wishart correspondante que par son régime asymptotique classique.

3 Contribution théorique

L'EVD d'une matrice (de dispersion) $\boldsymbol{\Sigma}$ est définie comme

$$\boldsymbol{\Sigma} \stackrel{\text{EVD}}{=} \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H \quad (9)$$

avec $\mathbf{U}^H \mathbf{U} = \mathbf{I}$ où $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_p] \in \mathcal{U}_p^p$ et $\boldsymbol{\Lambda} = \text{diag}(\boldsymbol{\lambda})$, $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_p] \in \mathbb{R}^p$. Dans la suite, nous supposons que les valeurs propres sont ordonnées $\lambda_1 > \dots > \lambda_p > 0$ et qu'un élément de chaque \mathbf{u}_j est réel positif (pour l'unicité).

Nous définissons les opérateurs $\boldsymbol{\theta}_j$ et ϕ_j retournant respectivement le $j^{\text{ème}}$ vecteur propre et valeur propre comme

$$\begin{cases} \mathbf{u}_j = \boldsymbol{\theta}_j(\boldsymbol{\Sigma}), \\ \lambda_j = \phi_j(\boldsymbol{\Sigma}). \end{cases} \quad (10)$$

En considérant à nouveau le modèle équivalent Wishart proposé dans la section 2.4, on note :

$$\begin{cases} \widehat{\mathbf{u}}_j^M = \boldsymbol{\theta}_j(\widehat{\boldsymbol{\Sigma}}) & \widehat{\mathbf{u}}_j^{EW} = \boldsymbol{\theta}_j(\widehat{\boldsymbol{\Sigma}}_{EW}), \\ \widehat{\lambda}_j^M = \phi_j(\widehat{\boldsymbol{\Sigma}}) & \widehat{\lambda}_j^{EW} = \phi_j(\widehat{\boldsymbol{\Sigma}}_{EW}). \end{cases} \quad (11)$$

avec $\boldsymbol{\phi} = [\phi_1, \dots, \phi_p]$.

Théorème 3.1 (Régime standard) La distribution asymptotique des valeurs propres et des vecteurs propres d'un M -estimateur (Eqs. (11)) est caractérisée par

$$\begin{cases} \sqrt{n}(\sigma\widehat{\boldsymbol{\lambda}}^M - \boldsymbol{\lambda}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \vartheta_1\boldsymbol{\Lambda}^2 + \vartheta_2\boldsymbol{\lambda}\boldsymbol{\lambda}^T), \\ \sqrt{n}\boldsymbol{\Pi}_j^\perp \widehat{\mathbf{u}}_j^M \xrightarrow{d} \mathcal{CN}(\mathbf{0}, \boldsymbol{\Xi}_j). \end{cases} \quad (12)$$

avec

$$\boldsymbol{\Xi}_j = \vartheta_1\lambda_j\mathbf{U}\boldsymbol{\Lambda}(\lambda_j\mathbf{I} - \boldsymbol{\Lambda})^+{}^2\mathbf{U}^H \quad (13)$$

où $\boldsymbol{\Pi}_j^\perp = \mathbf{I} - \mathbf{u}_j\mathbf{u}_j^H$ et ϑ_1, ϑ_2 données par Eq. (6).

Théorème 3.2 (EW) La distribution asymptotique de la différence entre les valeurs propres et les vecteurs propres d'un M -estimateur et de son EW est donnée par

$$\begin{cases} \sqrt{n}(\sigma\widehat{\boldsymbol{\lambda}}^M - \widehat{\boldsymbol{\lambda}}^{\text{EW}}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \sigma_1\boldsymbol{\Lambda}^2 + \sigma_2\boldsymbol{\lambda}\boldsymbol{\lambda}^T), \\ \sqrt{n}\boldsymbol{\Pi}_j^\perp (\widehat{\mathbf{u}}_j^M - \widehat{\mathbf{u}}_j^{\text{EW}}) \xrightarrow{d} \mathcal{CN}(\mathbf{0}, \sigma_1/\vartheta_1\boldsymbol{\Xi}_j). \end{cases} \quad (14)$$

avec $\boldsymbol{\Xi}_j$ et σ_1, σ_2 donnés par Eqs. (13) and (8), respectivement.

Preuve 3.1 Pour prouver ces deux théorèmes, il faut caractériser les dérivées des opérateurs θ_j et ϕ_j en étendant les résultats de [13] au cas complexe. La preuve se conclue en appliquant la méthode Delta aux résultats des théorèmes 2.1 et 2.2.

4 Validation et application des résultats

4.1 Validation

Afin de valider les résultats théoriques, nous générons des données de dimension $p = 20$ selon une loi de Student- t avec $d = 3$ degrés de libertés. La matrice de dispersion est Toeplitz, avec $\rho = 0.9(1 + \sqrt{-1})/\sqrt{2}$. Pour estimer la matrice de dispersion, nous utiliserons un M -estimateur de Student. Considérons $\widehat{\boldsymbol{\lambda}}^t = \phi(\widehat{\boldsymbol{\Sigma}}_t)$ où $\widehat{\boldsymbol{\Sigma}}_t$ est un M -estimateur de Student et $\widehat{\boldsymbol{\lambda}}^{\text{EW}} = \phi(\widehat{\boldsymbol{\Sigma}}_{\text{EW}})$ où $\widehat{\boldsymbol{\Sigma}}_{\text{EW}}$ est la SCM construite avec les noyaux gaussiens (non observables) des données observées (EW). Sur la figure 1, on observe que les résultats empiriques (courbes rouge $\text{MSE}(\widehat{\boldsymbol{\lambda}}^t - \boldsymbol{\lambda})$ et bleue $\text{MSE}(\widehat{\boldsymbol{\lambda}}^t - \widehat{\boldsymbol{\lambda}}^{\text{EW}})$) tendent vers les résultats théoriques correspondants (courbes noires $\mathcal{T}(\widehat{\boldsymbol{\lambda}}^t - \boldsymbol{\lambda}) = \text{Tr}(\vartheta_1\boldsymbol{\Lambda}^2 + \vartheta_2\boldsymbol{\lambda}\boldsymbol{\lambda}^T)$ et $\mathcal{T}(\widehat{\boldsymbol{\lambda}}^t - \widehat{\boldsymbol{\lambda}}^{\text{EW}}) = \text{Tr}(\sigma_1\boldsymbol{\Lambda}^2 + \sigma_2\boldsymbol{\lambda}\boldsymbol{\lambda}^T)$) lorsque n augmente. Une autre remarque est que l'erreur de $\widehat{\boldsymbol{\lambda}}^t$ est beaucoup plus petite en comparant à $\widehat{\boldsymbol{\lambda}}^{\text{EW}}$ qu'à $\boldsymbol{\lambda}$. Ceci confirme l'idée que la distribution des valeurs propres d'un M -estimateur peut être mieux approchée avec celle des valeurs propres de son EW qu'avec la distribution normale asymptotique autour des valeurs propres de la vraie matrice de dispersion.

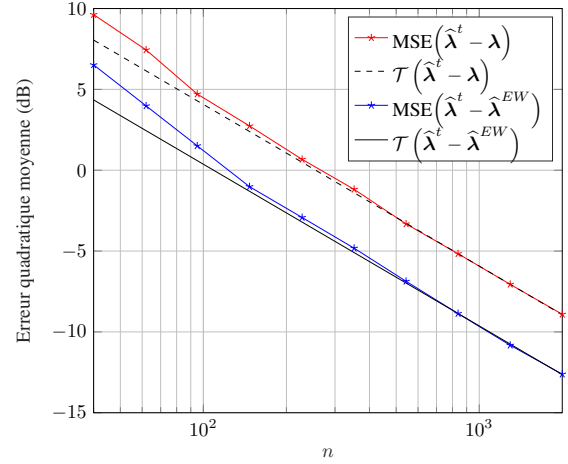


FIGURE 1 – Erreurs quadratiques moyennes (MSE) empiriques et théoriques (\mathcal{T}) sur les valeurs propres : résultats pour le régime asymptotique standard ($\widehat{\boldsymbol{\lambda}}^t - \boldsymbol{\lambda}$) et pour l'EW ($\widehat{\boldsymbol{\lambda}}^t - \widehat{\boldsymbol{\lambda}}^{\text{EW}}$).

4.2 Application : biais intrinsèque

Dans [11] a été établi un équivalent intrinsèque (i.e., pour des paramètres vivants dans une variété) de l'inégalité de Cramér-Rao (CRB). Dans le contexte de l'estimation de la matrice de covariance, ces résultats permettent notamment de borner la distance riemannienne naturelle (plutôt que la norme de Frobenius), définie par :

$$d_{\text{nat}}^2(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2) = \left\| \ln \left(\boldsymbol{\Sigma}_1^{-1/2} \boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1^{-1/2} \right) \right\|_F^2 = \sum_{j=1}^p \ln^2 \lambda_j^d$$

où λ_j^d est la $j^{\text{ème}}$ valeur propre de $\boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_2$. Cette analyse révèle également des propriétés inattendues, telles que le biais de la SCM quant à la métrique naturelle sur \mathcal{H}_M^{++} . La CRB intrinsèque est établie pour la SCM dans un contexte gaussien dans [11, Corollaire 5], sous la forme :

$$\mathbb{E} \left[d_{\text{nat}}^2 \left(\widehat{\boldsymbol{\Sigma}}_{\text{GCWE}}, \boldsymbol{\Sigma} \right) \right] \geq \frac{p^2}{n} + p\eta(p, n)^2. \quad (15)$$

Le terme $\eta(p, n)$ dans (15) est lié au biais intrinsèque (IB) de la SCM donné dans [11, Théorème 7] par $\mathbb{E} \left[\exp_{\boldsymbol{\Sigma}}^{-1} \widehat{\boldsymbol{\Sigma}}_{\text{GCWE}} \right] = -\eta(p, n)\boldsymbol{\Sigma}$, avec

$$\begin{aligned} \eta(p, n) &= \frac{1}{p} \{ p \ln n + p - \psi(n - p + 1) \\ &\quad + (n - p + 1)\psi(n - p + 2) \\ &\quad + \psi(n + 1) - (n + 1)\psi(n + 2) \} \end{aligned} \quad (16)$$

et où $\psi(x) = \Gamma'(x)/\Gamma(x)$ est la fonction digamma.

Pour les échantillons suivant une distribution CES, la CRB sur d_{nat}^2 est donnée dans [14] pour tout estimateur non biaisé $\widehat{\boldsymbol{\Sigma}}$ comme

$$\mathbb{E} \left[d_{\text{nat}}^2 \left(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma} \right) \right] \geq \frac{p^2 - 1}{n\alpha} + (n(\alpha + p\beta))^{-1}. \quad (17)$$

avec $\alpha = \left(1 - \frac{\mathbb{E}[\mathcal{Q}^2 \phi'(\mathcal{Q})]}{p(p+1)}\right)$ and $\beta = \alpha - 1$. Étendre le corollaire 5 de [11] dans ce contexte nécessite de dériver le biais intrinsèque d'un M -estimateur obtenu avec des échantillons ayant une distribution CES. Ce problème semble intractable, puisque ce résultat repose principalement sur la distribution des valeurs propres d'une matrice Wishart. Toutefois, l'équivalent EW du théorème 3.2 (ainsi que les résultats de simulations précédents) donne un motif théorique raisonnable pour l'approximation $\mathbb{E}[\exp_{\Sigma}^{-1} \hat{\Sigma}_M] \simeq -\eta(p, n)\Sigma$ pour tout M -estimateur consistant (en termes d'échelle) dans le cadre des CES. Par conséquent, nous pouvons proposer d'incorporer un terme de biais équivalent dans (17) pour obtenir une approximation précise de la valeur CRB intrinsèque pour les M -estimateurs obtenus avec des échantillons CES. La figure 2 illustre ce point pour des données suivant une loi de Student- t : la première courbe montre que le biais intrinsèque empirique obtenu avec le M -estimateur de Student coïncide avec le biais intrinsèque empirique de la SCM construite avec des noyaux gaussiens correspondants, ainsi qu'avec le résultat théorique attendu en (Eq. (16)). La seconde montre que la moyenne empirique de d_{nat} (notée $\epsilon^N(\hat{\Sigma}_t)$) peut être mieux caractérisée par la CRB biaisée approchée.

5 Conclusion

Nous avons proposé deux caractérisations asymptotiques pour l'EVD des M -estimateurs dans un cadre CES : un régime gaussien asymptotique standard et un équivalent Wishart [7]. Ce dernier nous a permis de dériver un biais intrinsèque approché et d'étendre les résultats de [11] aux M -estimateurs en ce qui concerne la borne sur l'estimation des matrices de dispersion établie pour la distance naturelle riemannienne.

Références

- [1] R. A. Maronna, "Robust M -estimators of multivariate location and scatter," *Annals of Statistics*, vol. 4, no. 1, pp. 51–67, January 1976.
- [2] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex elliptically symmetric distributions : Survey, new results and applications," *Signal Processing, IEEE Transactions on*, vol. 60, no. 11, pp. 5597–5625, November 2012.
- [3] R. J. Muirhead, "Aspects of multivariate statistical analysis." *JOHN WILEY & SONS, INC., 605 THIRD AVE., NEW YORK, NY 10158, USA, 1982, 656*, 1982.
- [4] D. E. Tyler, "Radial estimates and the test for sphericity," *Biometrika*, vol. 69, no. 2, p. 429, 1982.
- [5] M. Mahot, F. Pascal, P. Forster, and J.-P. Ovarlez, "Asymptotic properties of robust complex covariance matrix estimates," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3348–3356, July 2013.
- [6] G. Drašković and F. Pascal, "New properties for tyler's covariance matrix estimator," in *2016 50th Asilomar Conference on Signals, Systems and Computers*, Nov 2016, pp. 820–824.
- [7] —, "New insights into the statistical properties of m -estimators," *IEEE Transactions on Signal Processing*, vol. 66, no. 16, pp. 4253–4263, Aug 2018.
- [8] C. Croux and G. Haesbroeck, "Principal component analysis based on robust estimators of the covariance or correlation matrix : Influence functions and efficiencies," *BIOMETRIKA*, vol. 87, pp. 603–618, 2000.
- [9] J. Zhao and Q. Jiang, "Probabilistic pca for t distributions," *Neurocomputing*, vol. 69, no. 16–18, pp. 2217–2226, 2006.
- [10] E. Terreaux, J.-P. Ovarlez, and F. Pascal, "Robust model order selection in large dimensional elliptically symmetric noise," *arXiv :1710.06735*.
- [11] S. T. Smith, "Covariance, subspace, and intrinsic Cramér-Rao bounds," *IEEE Transactions on Signal Processing*, vol. 53, no. 5, pp. 1610–1630, 2005.
- [12] K. Yao, "A representation theorem and its applications to spherically invariant random processes," *Information Theory, IEEE Transactions on*, vol. 19, no. 5, pp. 600–608, September 1973.
- [13] T. Kollo and H. Neudecker, "Asymptotics of eigenvalues and unit-length eigenvectors of sample variance and correlation matrices," *Journal of Multivariate Analysis*, vol. 47, no. 2, pp. 283 – 300, 1993.
- [14] A. Breloy, G. Ginolhac, A. Renaux, and F. Bouchard, "Intrinsic cramer-rao bounds for scatter and shape matrices estimation in CES distributions," *submitted to IEEE Signal Processing Letters*.

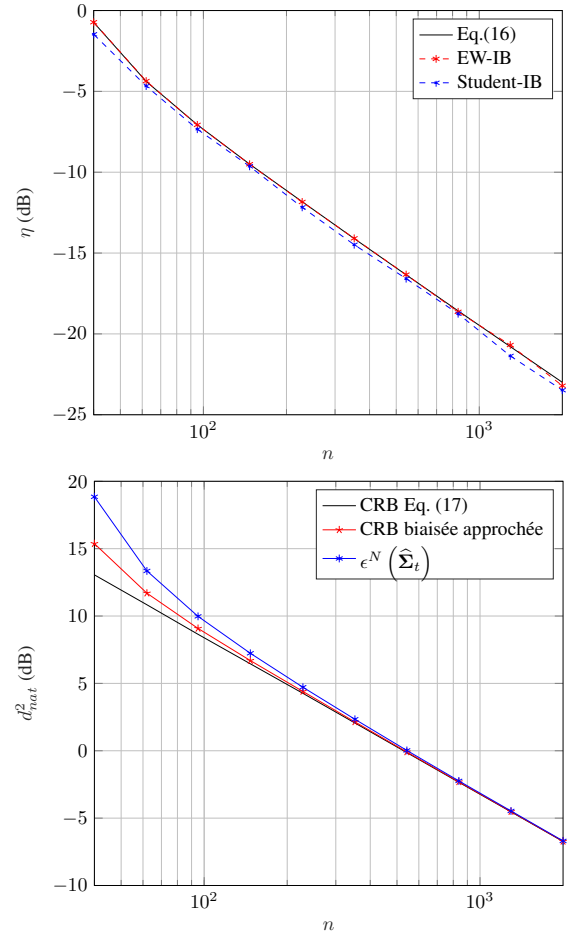


FIGURE 2 – (Haut) Biais intrinsèque empirique pour le M -estimateur de Student (Student-IB) et la SCM du modèle EW (EW-IB) comparé à (Eq. (16)). (Bas) Moyenne empirique de $\epsilon^N(\hat{\Sigma}_t) = d_{\text{nat}}^2(\hat{\Sigma}_t, \Sigma)$ comparée à la CRB intrinsèque et à la CRB intrinsèque biaisée approchée.