Estimation de la matrice d’échelle des distributions à symétrie elliptique

Mohamed Anis HADDOUCHE[1]

7 Mai 2020

[1] INSA de Rouen Normandie, France

Plan

Introduction

Distributions à symétrie elliptique

C’est une classe de distributions qui contient entre autres :

- Normale
- Student
- Kotz
- Laplace
- Cauchy

Définition

Différentes manières de définir cette classe :

- Fonction caractéristique
- Représentation stochastique

  [3] *K.T. Fang and Y.T. Zhang, Generalized multivariate analysis.* Science Press, Springer-Verlag, Beijing*, 1990.*

- Invariance par transformation orthogonales

  [5] *D. Fourdrinier, W.E. Strawderman and M.T. Wells, Shrinkage Estimation.* Springer*, 2018.*

Exemple de domaine d'application

- Finance, gestion de portefeuille

  [6] *M. Hamada and Valdez E. A. , CAPM and Option Pricing With Elliptically Contoured Distributions.* Journal of Risk and Insurance *, 75(2) :387-409, 2008.*

Densité

- Si le vecteur aléatoire $Y \in \mathbb{R}^p$

Matrice
d'échelle

$$Y \sim E(\theta, \boxed{\Sigma})$$

admet une densité $f(\cdot)$ par rapport à la mesure de Lebesgue, alors

$$f(y) = |\Sigma|^{-1/2} \, g\left[(y-\theta)^{\mathrm{T}} \Sigma^{-1} (y-\theta)\right]$$

- où

  la fonction $g(\cdot)$ est appelée fonction générative,

$$E_{\theta,\Sigma}[Y] = \theta \qquad \text{et} \qquad Cov[Y] = \frac{E_{\theta,\Sigma}\left[(y-\theta)^{\mathrm{T}} \Sigma^{-1} (y-\theta)\right]}{p} \, \Sigma \, .$$

$E_{\theta,\Sigma}$ désigne l'espérance par rapport à la densité $f(\cdot)$.

Exemples de fonctions génératives

Normale

$$g(u) \propto exp(-u/2)\,.$$

Student

$$g(u) \propto \left[1 + \frac{u}{m}\right]^{-(p+m)/2}\,, \quad m > 0\,.$$

Laplace

$$g(u) \propto exp(-\left|u\right|)\,.$$

Cauchy

$$g(u) \propto (1 + u)^{-(p+1)/2}\,.$$

Cas gaussien (matrice d'échelle = matrice de covariance)

- Si

$$Y \sim \mathcal{N}(\theta, \Sigma)\,,$$

- où

$$Cov[Y] = \frac{E_{\theta,\Sigma}\big((y-\theta)^{\mathrm{T}}\,\Sigma^{-1}(y-\theta)\big)}{p}\,\Sigma\,.$$

- Alors

$$Cov[Y] = \Sigma\,,$$

- car

$$(y-\theta)^{\mathrm{T}}\,\Sigma^{-1}(y-\theta) \sim \chi^2_p\,.$$

Les estimateurs usuels

- Pour estimer Σ, on a besoin de plusieurs copies du vecteur elliptique, soit

$$Y_{(1)}, \ldots, Y_{(m)} \,,$$

dans $\mathbb{R}^p$. En pratique, on forme la matrice $m \times p$

$$Y = (\,Y_{(1)}, \ldots, Y_{(m)})^{\mathrm{T}} \,.$$

- Les estimateurs usuels de Σ sont de la forme

$$\hat{\Sigma}_a = a\, Y^{\mathrm{T}}\; Y \,,$$

où a est une constante positive.

Matrice d'échelle empirique

Dans le cas gaussien, la matrice de d'échelle (covariance) empirique $\hat{\Sigma}_{1/m-1}$ est un estimateur de référence.

- Le contexte en petite dimension ($p \leq m$) :
    - sans biais
    - convergent
    - inversible
    - ses valeurs propres sont de bons estimateurs de ceux de Σ

- Le contexte en grande dimension ($p > m$) :
    - n'est pas inversible
    - ses valeurs propres sont des estimateurs biaisés pour ceux de Σ
    - Il est inadmissible

Comment les améliorer (Théorie de la décision)

- Un estimateur $\hat{\Sigma}$ est évalué au travers d'une fonction de coût

$$L(\Sigma, \hat{\Sigma})$$

  et son risque associé

$$R(\Sigma, \hat{\Sigma}) = E_{\theta,\Sigma}\left[L(\Sigma, \hat{\Sigma})\right] .$$

- Un estimateur $\hat{\Sigma}_1$ est dit meilleur que $\hat{\Sigma}_2$ si

$$R(\Sigma, \hat{\Sigma}_1) \leq R(\Sigma, \hat{\Sigma}_2) \qquad \forall\Sigma ,$$

  avec inégalité stricte pour au moins un Σ.
- L' estimateur $\hat{\Sigma}_2$ est dit inadmissible.

Phénomène Stein (cas gaussien)

- Dans le cas gaussien, James et Stein

  [7] *W. James and C. Stein, Estimation with Quadratic Loss.* Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, *1961.*

  montrent que les estimateurs usuels

  $$\hat{\Sigma}_a = a\, Y^{\mathrm{T}}\; Y \qquad \text{où} \qquad a > 0\,,$$

  sont inadmissibles dans

  - le contexte en grande dimension ($p > m$)
  - le contexte en petite dimension ($p \leq m$) lorsque $p \approx m$

- Ce phénomène persiste dans le cadre des distributions à symétrie elliptique.

Le modèle considéré

Considérons le modèle additif

$$Y = M + \mathcal{E}, \qquad \mathcal{E} \sim E\left(0_{mp}, I_m \otimes \Sigma\right) \tag{1.1}$$

où

- $Y \in \mathbb{R}^{m \times p}$ est une matrice d'observation
- M est une matrice de paramètres inconnus avec

$$rang\,(M) = q < m \wedge p. \tag{1.2}$$

- $\mathcal{E}$ est un bruit elliptique de matrice d'échelle Σ inversible.

La forme canonique du modèle additif

D'après (1.2), il existe une matrice semi-orthogonale $Q_2 \in \mathbb{R}^{m\times(m-q)}$ (i.e $Q_2^{\mathrm{T}} Q_2 = I_{m-q} \neq Q_2 Q_2^{\mathrm{T}}$) telle que

$$Q_2^{\mathrm{T}} M = 0\,.$$

On complète Q_2 par $Q_1 \in \mathbb{R}^{m\times q}$ pour former une matrice orthogonale $Q = (Q_1 Q_2)$

$$Q^{\mathrm{T}} Y = \begin{pmatrix} Q_1^{\mathrm{T}} \\ Q_2^{\mathrm{T}} \end{pmatrix} Y = \begin{pmatrix} Z \\ U \end{pmatrix} = \begin{pmatrix} Q_1^{\mathrm{T}} \\ Q_2^{\mathrm{T}} \end{pmatrix} M + Q^{\mathrm{T}} \mathcal{E} = \begin{pmatrix} \theta \\ 0 \end{pmatrix} + Q^{\mathrm{T}} \mathcal{E}.$$

- La densité de $Q^{\mathrm{T}} \mathcal{E}$ est la même que celle de $\mathcal{E}$.
- Il s'ensuit que la densité de $(Z^{\mathrm{T}} U^{\mathrm{T}})^{\mathrm{T}} = Q^{\mathrm{T}} Y$ est

$$(z, u) \mapsto |\Sigma|^{-m/2} f\left[\operatorname{tr}\{(z-\theta)\Sigma^{-1}(z-\theta)^{\mathrm{T}}\} + \operatorname{tr}\{\Sigma^{-1} u^{\mathrm{T}} u\}\right]. \tag{1.3}$$

Intérêt de la forme canonique

Rappelons

$$\begin{pmatrix} Q_1^{\mathrm{T}} \\ Q_2^{\mathrm{T}} \end{pmatrix} Y = \begin{pmatrix} Z \\ U \end{pmatrix} = \begin{pmatrix} \theta \\ 0 \end{pmatrix} + Q^{\mathrm{T}} \mathcal{E}.$$

- Exploiter uniquement l'information contenue dans la matrice résiduelle $U \in \mathbb{R}^{(m-q)\times p}$, qui est de plus petite dimension que $Y \in \mathbb{R}^{m\times p}$, pour estimer Σ.
- Dans ce cas, les estimateurs usuels sont de la forme

$$\hat{\Sigma}_a = a\, U^{\mathrm{T}}\, U = a\, S \qquad \text{où} \qquad a > 0\,,$$

Notons dans la suite, $n = m - q$.

Espérances liées

Soit

$$F^*(t) = \frac{1}{2}\int_t^\infty f(\nu)\,d\nu \quad \text{et} \quad F^{**}(t) = \frac{1}{2}\int_t^\infty F^*(\nu)\,d\nu.$$

Notons $E_{\theta,\Sigma}$ l'espérance par rapport à la densité (1.3), $E^*_{\theta,\Sigma}$ l'espérance par rapport à

$$(z,u) \mapsto \frac{1}{K^*}|\Sigma|^{-m/2}\,F^*\left[\operatorname{tr}\{(z-\theta)\Sigma^{-1}(z-\theta)^\top\} + \operatorname{tr}\{\Sigma^{-1}\,u^{\mathrm{T}}\,u\}\right]$$

et $E^{**}_{\theta,\Sigma}$ l'espérance par rapport à

$$(z,u) \mapsto \frac{1}{K^{**}}|\Sigma|^{-m/2}\,F^{**}\left[\operatorname{tr}\{(z-\theta)\Sigma^{-1}(z-\theta)^\top\} + \operatorname{tr}\{\Sigma^{-1}\,u^{\mathrm{T}}\,u\}\right]$$

où K^* et K^{**} sont des constantes de normalisation.

Sous classe de densités de type Berger

Nous considérons la sous classe de densités telles que

$$0 < c \leq \frac{F^*(t)}{f(t)} \leq b\,,$$

Exemples :

- la loi de type logistique

$$f(t) \propto \frac{\exp\left(-\beta t - \gamma\right)}{\left(1 + \exp\left(-\beta\, t - \gamma\right)\right)^2}$$

où $\beta > 0$ et $\gamma > 0$. Dans ce cas

$$c = \frac{1}{2\beta} \qquad b = \frac{(1 + e^{-\gamma})}{2\beta}.$$

- la loi normale : $c = b = 1$ puisque $F^* = f$.

[4] *D. Fourdrinier, F. Mezoued, and W. E. Strawderman. Bayes minimax estimators of a location vector for densities in the Berger class.* Electronic Journal of Statistics. *6 :783–809, 2012.*

Notre objectif

- Dans une approche unifiée ($p > n$ et $p \leq n$), on veut améliorer

$$\hat{\Sigma}_a = a\,S \qquad \text{où} \qquad a > 0\,,$$

par des estimateurs alternatifs de la forme

$$\hat{\Sigma}_{a,G} = a\,\left(S + SS^+ G(Z,S)\right)\,,$$

où S^+ est l'inverse de Moore-Penrose S et $SS^+G(Z,S)$ une matrice de correction.

- Sous le coût quadratique

$$L(\Sigma,\hat{\Sigma}) = \mathrm{tr}(\Sigma^{-1}\hat{\Sigma} - I_p)^2\,.$$

- Sous le coût basé sur les données

$$L_S(\Sigma,\hat{\Sigma}) = \mathrm{tr}(S^+\,\Sigma\,(\Sigma^{-1}\,\hat{\Sigma} - I_p)^2)\,.$$

Estimateurs de forme générale

Estimateurs alternatifs

- Considérons le risque quadratique

$$R(\hat{\Sigma}, \Sigma) = E_{\theta,\Sigma}[\mathrm{tr}(\Sigma^{-1}\hat{\Sigma} - I_p)^2].$$

- Pour $\hat{\Sigma} = \hat{\Sigma}_a = a\,S$, la meilleur constante a est

$$a_o = \frac{1}{K^{**}(n+p+1)}\,.$$

- Considérons des estimateurs alternatifs de la forme

$$\hat{\Sigma}_{a_o,G} = a_o\,\left(S + SS^{+}G(Z,S)\right),$$

  où $SS^{+}G(Z,S)$ est une matrice de correction symétrique.

- Les estimateurs $\hat{\Sigma}_{a_o,G}$ améliorent $\hat{\Sigma}_{a_o}$ si

$$\Delta(G) = R(\hat{\Sigma}_{a_o,G}, \Sigma) - R(\hat{\Sigma}_{a_o}, \Sigma) \leq 0.$$

Différence des risques

- La proposition suivante assure la finitude de la diff. des risques $\Delta(G)$.

Proposition 1

Supposons $E_{\theta,\Sigma}\left[\mathrm{tr}\left(\Sigma^{-1}\,S\right)^2\right] < \infty$ *et* $E_{\theta,\Sigma}\left[\mathrm{tr}(\Sigma^{-1}SS^{+}G)^2\right] < \infty$*.*

Alors

$$\Delta(G) = a_o^2\,E_{\theta,\Sigma}\left[\mathrm{tr}\left(\Sigma^{-1}SS^{+}(2S+G)\Sigma^{-1}SS^{+}G\right)\right] - 2\,a_o\,E_{\theta,\Sigma}\left[\mathrm{tr}\left(\Sigma^{-1}SS^{+}G\right)\right] < \infty.$$

- Sous quelles conditions sur G, les estimateurs $\hat{\Sigma}_{a_o,G}$ améliorent-ils $\hat{\Sigma}_{a_0} = a_o S$?

Une nouvelle identité de type Stein-Haff

Lemme

Soit $V(z,s)$ une fonction matricielle $p \times p$ tel que, pour tout $z \in \mathbb{R}^{q \times p}$ fixé, la fonction $V(z,s)$ soit faiblement différenciable par rapport à $s \in \mathbb{R}^{p \times p}$. Supposons que $E_{\theta,\Sigma}\left[\left|\mathrm{tr}\left(\Sigma^{-1} S\, S^{+}\, V\right)\right|\right] < \infty$.

Alors

$$E_{\theta,\Sigma}\left[\mathrm{tr}\left(\Sigma^{-1} S S^{+} V\right)\right] = K^{*} E^{*}_{\theta,\Sigma}\left[\mathrm{tr}\left((n-(p\wedge n)-1)\, S^{+} V + 2\, S S^{+}\, \mathcal{D}_S\{S S^{+} V\}^{\top}\right)\right].$$

Ici, $\mathcal{D}_S$ est l'opérateur différentiel de Haff dont le terme générique est

$$d_{ij} = \frac{1}{2}(1+\delta_{ij})\frac{\partial}{\partial S_{ij}}$$

avec $\delta_{ij} = 1$ quand $i = j$ et $\delta_{ij} = 0$ quand $i \neq j$.

Application de l'identité de type Stein-Haff

Rappelons que

$$\Delta(G) = a_o^2\, E_{\theta,\Sigma}\left[\mathrm{tr}\left(\Sigma^{-1} S S^+ (2S+G) \Sigma^{-1} S S^+ G\right)\right] \\ -2\, a_o\, E_{\theta,\Sigma}\left[\mathrm{tr}\left(\Sigma^{-1} S S^+ G\right)\right].$$

On applique l'identité de type Stein-Haff

- une fois pour $E_{\theta,\Sigma}\left[\mathrm{tr}\left(\Sigma^{-1} S S^+ G\right)\right]$
- deux fois pour $E_{\theta,\Sigma}\left[\mathrm{tr}\left(\Sigma^{-1} S S^+ (2S+G) \Sigma^{-1} S S^+ G\right)\right]$

Conditions d'amélioration

Théorème 1

- *Soit une densité de la forme (1.3) telle que* $c \leq F^*(t)/f(t) \leq b$.
- *Si*

$$\mathrm{tr}\left[\, 2\, S^+ S\, \mathcal{D}_S\{SS^+ G\} - \; S^+ G \,\right] \geq 0\,,$$

- *alors les estimateurs* $\hat{\Sigma}_{a_o,G}$ *améliorent* $\hat{\Sigma}_{a_o}$ *dès que*

$$\mathrm{tr}\Big[2S^+ S\, \mathcal{D}_S\{SS^+ T^*\}^\top - S^+ T^* \\ - 2(p+n+1)\frac{c^2}{b^2}\left(2S^+ S\, \mathcal{D}_S\{SS^+ G\} - S^+ S\right)\Big] \leq 0\,,$$

où

$$T^* = 4(S+G)\,\mathcal{D}_S\{SS^+ G\} + G\left(2\, n\, I_p - (p-n+1)S^+ G\right).$$

Limites du coût quadratique

- Difficile à manipuler car on applique l'opérateur de Haff $\mathcal{D}_S$ deux fois.
- Impose de fortes conditions sur SS^+G, ce qui rend difficile la constructions d'estimateurs améliorés.

D'ou viennent ces difficultés ?

Limites du coût quadratique et remède

- Le coût quadratique peut être réécrit sous la

$$L(\Sigma, \hat{\Sigma}) = \mathrm{tr}\left((\Sigma^{-1}\,\hat{\Sigma} - I_p)^2 \right)$$

Deux fois Σ^{-1}

$$= \boxed{\mathrm{tr}(\Sigma^{-1}\,\hat{\Sigma}\,\Sigma^{-1}\,\hat{\Sigma})} - 2\,\mathrm{tr}(\Sigma^{-1}\,\hat{\Sigma}) + p\,.$$

- Nécessite une double application de l'identité de Stein-Haff.

Comment y remédier ?

- On introduit de la donnée dans ce coût, ce qui nous mène au coût basé sur les données

$$L_S(\Sigma, \hat{\Sigma}) = \mathrm{tr}(S^+\,\Sigma\,(\Sigma^{-1}\,\hat{\Sigma} - I_p)^2)$$

Une fois Σ^{-1}

$$= \boxed{\mathrm{tr}(\Sigma^{-1}\,\hat{\Sigma}\,S^+\,\hat{\Sigma})} - 2\,\mathrm{tr}\,(S^+\,\hat{\Sigma}) + \mathrm{tr}\,(S^+\,\Sigma)\,.$$

Estimateurs alternatifs

- Considérons le risque basé sur les données

$$R_S(\hat{\Sigma}, \Sigma) = E_{\theta,\Sigma}[\mathrm{tr}(S^+ \Sigma\, (\Sigma^{-1}\, \hat{\Sigma} - I_p)^2)],$$

- Pour $\hat{\Sigma} = \hat{\Sigma}_a = a\, S$, la meilleur constante a est

$$a_o = \frac{1}{K^*\,(p \vee n)}\,.$$

- Considérons des estimateurs alternatifs de la forme

$$\hat{\Sigma}_{a_o,G} = a_o\, \left(S + SS^+ G(Z, S)\right)\,,$$

  où $SS^+ G(Z, S)$ une matrice de correction non nécessairement symétrique.

- Les estimateurs $\hat{\Sigma}_{a_o,G}$ améliorent $\hat{\Sigma}_{a_o}$ si

$$\Delta(G) = R_S(\hat{\Sigma}_{a_o,G}, \Sigma) - R_S(\hat{\Sigma}_{a_o}, \Sigma) \leq 0.$$

Différences des risques

La proposition suivante assure la finitude de la diff. des risques $\Delta(G)$.

Proposition 2

Supposons $E_{\theta,\Sigma}\left[\|S^+G\|_F^2\right]$, $E_{\theta,\Sigma}\left[\|\Sigma^{-1}SS^+G\|_F^2\right]$, $E_{\theta,\Sigma}\left[\mathrm{tr}(\Sigma\, S^+)\right]$ *et* $E_{\theta,\Sigma}\left[\mathrm{tr}(\Sigma^{-1}\, S)\right]$ *soient des espérances finies.*

Alors

$$\Delta(G) = a_o^2\, E_{\theta,\Sigma}\left[\mathrm{tr}\left(\Sigma^{-1} SS^+ G \left\{ I_p + S^+G + SS^+ \right\}\right)\right] - 2\, a_o\, E_{\theta,\Sigma}\left[\mathrm{tr}\left(S^+\, G\right)\right] < \infty\,.$$

Conditions d'amélioration

Théorème 2

- *Soit une densité de la forme (1.3) telle que* $c \leq F^*(t)/f(t) \leq b$.
- *Si*

$$\mathrm{tr}(S^+ G) \geq 0 \,,$$

- *alors les estimateurs* $\hat{\Sigma}_{a_o,G}$ *améliorent* $\hat{\Sigma}_{a_o}$ *dès que*

$$\mathrm{tr}\Big[(n-(p\wedge n)-1)(S^+GSS^+ + (S^+G)^2) + \alpha\, S^+G \\ +2\, SS^+\, \mathcal{D}_S\{SS^+G + SS^+GS^+G + SS^+\,G\,SS^+\}^T\Big] \leq 0 \,,$$

où

$$\alpha = (n-(p\wedge n)-2\,\frac{c}{b}\,(p\vee n)-1)\,.$$

Estimateurs orthogonalement invariants

Estimateurs orthogonalement invariants

- Considérons la décomposition en valeurs propres de S, soit

$$S = H_1 \, L \, H_1^\top$$

où

$$H_1 \in \mathbb{R}^{p \times (p \wedge n)} \text{ est une matrice semi-orthogonale}$$

et

$$L = \mathrm{diag}(l_1, \ldots, l_{(p \wedge n)}) \quad \text{avec} \quad l_1 > l_2 > \cdots > l_{(p \wedge n)} > 0 \, .$$

- Les estimateurs orthogonalement invariants sont de la forme

$$\hat{\Sigma} = H_1 \Phi(L) H_1^T$$

où

$$\Phi(L) = \mathrm{diag}(\phi_1(L), \ldots, \phi_{(p \wedge n)}(L)) \text{ avec } \phi_1(L) > \cdots > \phi_{(p \wedge n)}(L).$$

Matrice d'échelle empirique dans le cas gaussien

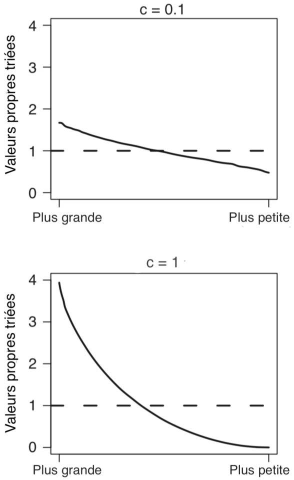


FIGURE 1 – Estim. des v.p. de $\Sigma = I_p$ par $\hat{\Sigma}_{1/n-1} = \frac{1}{n-1} H_1 L H_1^{\mathrm{T}}$ où $c = p/n$

Forme des estimateurs de type Haff

- Rappelons que

$$\hat{\Sigma}_{a_o,G} = a_o \left(S + SS^+ G(Z,S) \right).$$

- Posant

$$G(Z,S) = \nu\, t(\nu) H_1\, H_1^{\mathrm{T}}$$

  où

$$\nu = 1/\mathrm{tr}(S^+)$$

  et $t(\cdot)$ est une fonct. convexe, str. positive, décroissante et deux fois différentiable.

- Les estimateurs de type Haff sont de la forme

$$\hat{\Sigma}_{Haff} = a_o\, H_1 \operatorname{diag}\left(l_1 + \nu\, t(\nu), l_2 + \nu\, t(\nu), \ldots, l_{p\wedge n} + \nu\, t(\nu)\right) H_1^{\mathrm{T}}\ .$$

Proposition 3

Soit une densité de la forme (1.3) telle que $c \leq F^*(t)/f(t) \leq b$ *et*

$$\frac{p+n-(p\wedge n)+2}{p+n+1} \leq \frac{c^2}{b^2} \leq \frac{2\,p+2\,n-5\,(p\wedge n)-3}{p+n+1}\,.$$

Alors tout $\hat{\Sigma}_{Haff}$ *améliore* $\hat{\Sigma}_{a_o}$ *dès que*

(*i*) $(p+n-2\,(p\wedge n)+1)\,t(\nu)+2\,\nu\,t'(\nu)\geq 0\,,$

(*ii*) $0 \leq t(\nu) \leq \dfrac{2\,(p+n-2\,(p\wedge n)-1)\left((p+n+1)\frac{c^2}{b^2}-p-n+p\wedge n-2\right)}{(p+n-2\,(p\wedge n)+1)(p+n-2\,(p\wedge n)+3)}$

(*iii*) $$\{2\,(p+n-4\,(p\wedge n)+3)\,t(\nu)+2\,\nu\,t'(\nu) + \left[2\,p+2\,n-5\,(p\wedge n)+5-(p+n+1)\frac{c^2}{b^2}\right]\Big\}\,t'(\nu) + 2\,\left\{t(\nu)+(p\wedge n)^2\right\}\,\nu\,t''(\nu) \leq 0\,.$$

Haff sous coût basé sur les données

Proposition 4

Soit une densité de la forme (1.3) telle que $c \leq F^*(t)/f(t) \leq b$ *et*

$$\frac{c}{b} \geq \frac{(p \vee n) - (p \wedge n) + 1}{(p \vee n)}.$$

Alors $\hat{\Sigma}_{Haff}$ *améliore* $\hat{\Sigma}_{a_o}$ *dès que*

$$0 \leq t(\nu) \leq 2\,\frac{((p \wedge n) - 1) + (p \vee n)(c/b - 1)}{(p \vee n) - (p \wedge n) + 1}.$$

Forme de l'estimateur de Konno

- Rappelons que

$$\hat{\Sigma}_{a_o,G} = a_o \left(S + SS^{+} G(Z,S) \right) .$$

- Posant

$$G(Z,S) = \nu \, t \, H_1 \, H_1^{\mathrm{T}}$$

où

$$\nu = 1/\mathrm{tr}(S^{+})$$

et

$$t \text{ constante positive,}$$

- l'estimateur de Konno est de la forme

$$\hat{\Sigma}_{Kon.} = a_o \, H_1 \, \mathrm{diag}\big(l_1 + \nu \, t, l_2 + \nu \, t, \ldots, l_{p \wedge n} + \nu \, t \big) H_1^{\mathrm{T}} \; .$$

Konno sous coût quadratique

Proposition 5

Soit une densité de la forme (1.3) telle que $c \leq F^*(t)/f(t) \leq b$ *et*

$$\frac{p+n-(p\wedge n)+2}{p+n+1} \leq \frac{c^2}{b^2} \leq \frac{2\,p+2\,n-5\,(p\wedge n)-3}{p+n+1}\,.$$

Alors $\hat{\Sigma}_{Kon.}$ *améliore* $\hat{\Sigma}_{a_o}$ *dès que*

$$0 \leq t \leq \frac{2\,(p+n-2\,(p\wedge n)-1)\,((p+n+1)c^2/b^2-p-n+p\wedge n-2\,)}{(p+n-2\,(p\wedge n)+1)(p+n-2\,(p\wedge n)+3)}$$

Konno sous coût basé sur les données

Résultat qui n'est plus restreint à la condition $c \leq F^*(t)/f(t) \leq b$.

Proposition 6

Soit une densité de la forme (1.3).

Alors $\hat{\Sigma}_{Kon.}$ *améliore* $\hat{\Sigma}_{a_o}$ *dès que*

$$0 \leq t \leq \frac{2\,((p \wedge n) - 1)}{(p \vee n) - (p \wedge n) + 1}\,.$$

Comparaison numérique pour l'estimateur de Konno

- Pour chaque coût nous allons évaluer le pourcentage d'amélioration de $\hat{\Sigma}_{Kon.}$ par rapport à $\hat{\Sigma}_{a_o}$.
- Ce pourcentage d'amélioration [1] est définie par

$$PRIAL = \frac{\text{coût moyen de } \hat{\Sigma}_{a_o} - \text{coût moyen de } \hat{\Sigma}_{Kon.}}{\text{coût moyen de } \hat{\Sigma}_{a_o}} \times 100 \,.$$

- Notons $PRIAL_Q$ (resp. $PRIAL_S$) le pourcentage d'amélioration par rapport au coût quadratique (resp. coût basé sur les donnés).

1. Percentage Relative Improvement in Average Loss.

Comparaison numérique pour l'estimateur de Konno

Résultat de 1000 simulations de type Monte Carlo pour deux valeurs de p fixées, la valeur de n variant et

$$(\Sigma)_{ij} = 0.9^{|i-j|}.$$

p	n	$PRIAL_Q\%$	$PRIAL_S\%$
20	4	0.91	15.00
20	8	2.56	18.56
20	12	5.05	25.56
20	16	4.17	47.03
100	20	0.40	3.39
100	40	1.24	4.19
100	60	1.56	5.76
100	80	1.531	10.42

Estimateurs de type Efron et Morris

Estimateurs de type Efron et Morris

- Rappelons que

$$\hat{\Sigma}_{a_o,G} = a_o \left(S + SS^{+} G(Z,S)\right).$$

- Soit $Q \in \mathbb{R}^{(p\wedge n)\times q}$ une matrice de constantes telle que

$$Rang(Q) = q \leq p \wedge n.$$

- Soit la matrice de projection orthogonale

$$Q_o = Q\,(Q^{\mathrm{T}}\,Q)^{-1}\,Q^{\mathrm{T}}.$$

- Pour

$$G(Z,S) = \frac{t}{\mathrm{tr}(L^{-1}\,Q_o)}\,H_1\,Q_o\,H_1^{\mathrm{T}}, \qquad t > 0,$$

- les estimateurs de type Efron et Morris sont

$$\hat{\Sigma}_{EM} = a_o \left(S + \frac{t}{\mathrm{tr}(L^{-1}\,Q_o)}\,H_1\,Q_o\,H_1^{\mathrm{T}}\right)$$

-

Proposition 7

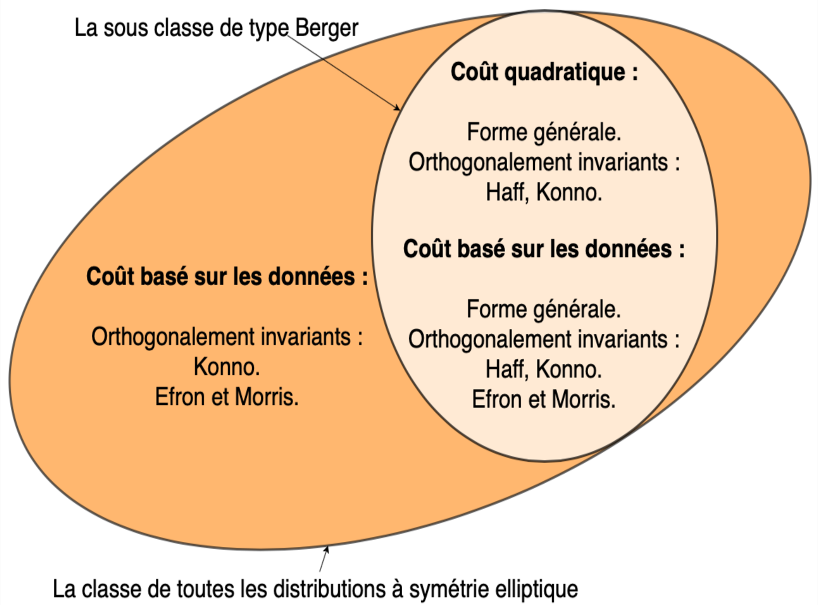
Soit une densité de la forme (1.3).

Alors tout $\hat{\Sigma}_{EM}$ *améliore* $\hat{\Sigma}_{a_o}$ *dès que*

$$0 \leq t \leq \frac{2\left((p \wedge n) - 1\right)}{(p \vee n) - (p \wedge n) + 1}\,.$$

Synthèse et perspectives

Synthèse


La sous classe de type Berger
Coût quadratique :
Forme générale.
Orthogonalement invariants :
Haff, Konno.
Coût basé sur les données :
Forme générale.
Orthogonalement invariants :
Haff, Konno.
Efron et Morris.
Coût basé sur les données :
Orthogonalement invariants :
Konno.
Efron et Morris.
La classe de toutes les distributions à symétrie elliptique

Matrice d'échelle non inversible

Considérons

$$Y = M + \mathcal{E}, \qquad \mathcal{E} \sim ES(0_{mp}, I_m \otimes \Sigma)$$

où Σ une matrice d'échelle de Rang = $r < p$.

Le bruit $\mathcal{E}$ n'admet pas de densité par rapport à la mesure de Lebesgue

[2] *J.A Díaz-García, V. Leiva-Sánchez and M. Galea. Singular Elliptical Distribution : Density and Applications.* Communications in Statistics - Theory and Methods. *5 :665–681, 2002.*

Problèmes d'estimation dans le cadre elliptique singulier

Estimation de Σ sous le coût quadratique

$$L(\hat{\Sigma}, \Sigma) = \mathrm{tr}\left(\hat{\Sigma}\Sigma^{+} - \Sigma\Sigma^{+}\right)^{2}.$$

Estimation de la matrice de précision Σ^{+} sous le coût de Frobenius

$$L(\hat{\Sigma}^{+}, \Sigma^{+}) = \| \hat{\Sigma}^{+} - \Sigma^{+} \|_F^2 .$$

Le cas gaussien

Le cas orthogonalement invariant

[1] *D. Chételat, M. T. Wells. Improved second order estimation in the singular multivariate normal model.* Journal of Multivariate Analysis , *147 :11 – 19, 2016.*

Une identité de type Stein-Haff de la forme

$$E\left[\mathrm{tr}\left(\Sigma^{+}H_1\Phi(L)H_1^{\top}\right)\right] = E\left[\sum_{i=1}^{n\wedge r}\left\{(|n-r|-1)\frac{\phi_i}{l_i} + 2\frac{\partial\phi_i}{\partial l_i} + 2\sum_{j>i}\frac{\phi_i-\phi_j}{l_i-l_j}\right\}\right]$$

où $Rang(S) = n \wedge r$.

Généralisation au cas elliptique singulier

- Comment élaborer une nouvelle identité de type Stein-Haff dans le cas des estimateurs orthogonalement invariants

- Comment élaborer une identité de la forme

$$E\left[\operatorname{tr}\left(\Sigma^{+} S S^{+} G(Z, S)\right)\right] = ?$$

Merci pour votre attention !

Pour les références et autres détails, vous pouvez m'écrire à
mohamed.haddouche@insa-rouen.fr

D. Chételat and M.T. Wells.
Improved second order estimation in the singular multivariate normal model.
Journal of Multivariate Analysis, 147(C), 2016.

J.A. Díaz-García, V. Leiva-Sánchez, and M. Galea.
Singular elliptical distribution : Density and apllications.
Communications in Statistics - Theory and Methods, 31(5) :665–681, 2002.

K.T. Fang and Y.T. Zhang.
***Generalized multivariate analysis. 1990*.**
Science Press, Springer-Verlag, Beijing, 1990.

D. Fourdrinier, F. Mezoued, and W.E. Strawderman.
Bayes minimax estimators of a location vector for densities in the Berger class.
Electronic Journal of Statistics, 6 :783–809, 2012.

D. Fourdrinier, W.E. Strawderman, and M.T. Wells.
***Shrinkage Estimation*.**

Springer, 2018.

M. Hamada and E. A. Valdez.
Capm and option pricing with elliptically contoured distributions.
Journal of Risk and Insurance, 75(2) :387–409, 2008.

W. James and C. Stein.
Estimation with quadratic loss.
In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1 : Contributions to the Theory of Statistics*, pages 361–379, Berkeley, Calif., 1961. University of California Press.