James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Shrinkage estimation of the mean matrix of multivariate complex normal distributions

Yoshihiko KONNO (   )

Japan Women's University

National Chaio Tung University

5 Dec. 2012

James-Stein estimator
Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

James-Stein estimator

Preliminaries **Complex normal and Wishart distributions**
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

★ Assume that $X : p \times 1 \sim N_p(\xi, I_p)$.

★ Consider the problem of estinmating ξ under a risk function

$$R(\xi, \hat{\xi}) = \mathbb{E}[\sum_{i=1}^{p}(\hat{\xi}_i - \xi_i)^2] = \mathbb{E}[(\hat{\xi} - \xi)'(\hat{\xi} - \xi)],$$

where $\hat{\xi}$ is an estimator of $\xi' = (\xi_1, \ldots, \xi_p)$ and ξ' is the transpose of the column vector ξ,

★ When $p \leq 2$, there is no estimators $\hat{\xi}$ such that $R(\xi, \hat{\xi}) \leq R(\xi, X)$ for $\forall \xi$ and $R(\xi, \hat{\xi}) < R(\xi, X)$ for some ξ.

★ However, when $p \geq 3$, the James-Stein estimator $\hat{\xi}_{JS} = \{1 - (p-2)/\sum_{1=1}^{p} X_i^2\}X$ improves X, i.e., $R(\xi, \hat{\xi}_{JS}) < R(\xi, X)$.

James-Stein estimator

Preliminaries **Complex normal and Wishart distributions**
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

★ In this talk, we construct estimators which improve the usual estimator(the maximum likelihood estimator) for the mean matrix of the matrix-variate complex normal distributions.

James-Stein estimator

Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Complex normal distribution
Complex Wishart distribution

★ Good reference is Andersen et al. (1995, Lecture note in Statistics vol. 101).

★ We say that a random variable $Z = \mathrm{Re}Z + \sqrt{-1}\mathrm{Im}Z$ has a standard complex normal distributions $\mathbb{C}N(0, 1)$ if

$$[Z] = \begin{pmatrix} \mathrm{Re}Z \\ \mathrm{Im}Z \end{pmatrix} \sim N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right].$$

★ Then the density (w.r.t. Lebesgue measure on $\mathbb{C}$) of Z is given as

$$f_Z(z) = \frac{1}{\pi} \exp(-\bar{z}z)$$

where $\bar{z}$ is the conjugate of a complex z ($z \in \mathbb{C}$).

James-Stein estimator

Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Complex normal distribution
Complex Wishart distribution

★ Assume that $Z \sim \mathbb{C}N(0, 1)$. For $\xi \in \mathbb{C}$, $\sigma \in \mathbb{R}_+$,

$$X := \xi + \sigma Z \sim \mathbb{C}N(\xi, \sigma^2).$$

★ Let X be $\mathbb{C}^p$-valued random vector. For $\forall c \in \mathbb{C}^p$ $\xi \in \mathbb{C}^p$ and $\Sigma \in \mathbb{C}_+^{p \times p}$

$$c^* X \sim \mathbb{C}N(c^* \xi, c^* \Sigma c) \iff X \sim \mathbb{C}N_p(\xi, \Sigma),$$

where c^* is the transpose complex conjugate of a column vector c and $\mathbb{C}_+^{p \times p}$ is the set of all p.d. complex matrices.

★ Let $X \sim \mathbb{C}N_p(\xi, \Sigma)$. Then the density of X(w.r.t. Lebesgue measure on $\mathbb{C}^p$) is given as

$$f_X(x) = \frac{1}{\pi^p}(\text{Det }\Sigma)^{-1} \exp\{-(x - \xi)^* \Sigma^{-1}(x - \xi)\}.$$

James-Stein estimator
Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Complex normal distribution
Complex Wishart distribution

★ Let $X = \mathrm{Re}\, X + \sqrt{-1}\mathrm{Im}\, X \sim \mathbb{C}N_p(\xi, \Sigma)$. Then the distribution of $[X]$(w.r.t Lebesgue measure on $\mathbb{R}^{2p}$) is given as

$$[X] := \begin{pmatrix} \mathrm{Re}\, X \\ \mathrm{Im}\, X \end{pmatrix} \sim N_{2p}\left[\begin{pmatrix} \mathrm{Re}\, \xi \\ \mathrm{Im}\, \xi \end{pmatrix}, \frac{1}{2}\begin{pmatrix} \mathrm{Re}\, \Sigma & -\mathrm{Im}\, \Sigma \\ \mathrm{Im}\, \Sigma & \mathrm{Re}\, \Sigma \end{pmatrix}\right],$$

where $\mathrm{Re}\, \Sigma$ and $\mathrm{Im}\, \Sigma$ are symmetric and skew-symmetric.

James-Stein estimator
Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Complex normal distribution
Complex Wishart distribution

★ Let $X_1, X_2, \ldots, X_n$ be mutually independent $\mathbb{C}N_p(0, \Sigma)$ random column vectors. Then the distribution of

$$S := \sum_{i=1}^{n} X_i X_i^*$$

is called a complex Wishart distribution with parameters n, Σ (this is denoted by $\mathbb{C}W_p(n, \Sigma)$)

★ If $n \geq p$, then $\mathbb{P}(S$ is p.d.$) = 1$ and the density of S(w.r.t. Lebesgue measure on $\mathbb{C}_+^{p \times p}$) is given as

$$f_S(w) = \frac{\mathrm{Det}\,(w)^{n-p}\exp(-\mathrm{Tr}\,(w\Sigma^{-1}))}{\mathrm{Det}\,(\Sigma)^n \pi^{p(p-1)/2} \Pi_{j=1}^{p} \Gamma(n+1-j)},$$

where $w \in \mathbb{C}_+^{p \times p}$, the set of all p.d. complex matrices, and $\Gamma(\cdot)$ is Euler's gamma function.

James-Stein estimator
Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Observation for the problem of estimating mean matrix Ξ

★ Two random matrices $X : m \times p$ and $S : p \times p$ are independent and distributed as follows.

★ $X : m \times p \sim \mathbb{C}N_{m \times p}(\Xi, I_m \otimes \Sigma)$ and $S : p \times p \sim \mathbb{C}W_p(n, \Sigma)$, where Ξ is unknown $m \times p$ matrix with complex entries and Σ is unknown Hermitian p.d. $p \times p$ matrix.
Write $X' := (X_1, X_2, \ldots, X_m)$ and $\Xi' = (\xi_1, \xi_2, \ldots, \xi_m)$, the transpose of the matrix Ξ. Then X_i ($i = 1, 2, \ldots, m$) are independently distributed as $\mathbb{C}N_p(\xi_i', \Sigma)$.

★ Assume that $n > p$.

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Loss function

★ We consider the problem of estimating unknown mean matrix Ξ based on observation **X** and **S** under a loss function:

$$L((\Xi, \Sigma), \widehat{\Xi}) = \mathrm{Tr}\,\{\Sigma^{-1}(\widehat{\Xi} - \Xi)^*(\widehat{\Xi} - \Xi)\},$$

where $\widehat{\Xi}$ is an estimator for Ξ.

★ Risk function

$$R((\Xi, \Sigma), \widehat{\Xi}) = \mathbb{E}[L((\Xi, \Sigma), \widehat{\Xi})],$$

Goal
Find estimators $\widehat{\Xi}$ s.t.

$$R((\Xi, \Sigma), \widehat{\Xi}) \leq R((\Xi, \Sigma), X) = mp \qquad \text{for } \forall(\Xi, \Sigma).$$

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Key ingredients

1. We consider a class of estimators $\widehat{\Xi}$ whose modification depends on the eigenvalues of $\mathbf{X}^*\mathbf{X}\mathbf{S}^{-1}$.
   Then $R((\Xi, \Sigma), \widehat{\Xi})$, the risk of estimators $\widehat{\Xi}$, depends on unknown parameters (Ξ, Σ).

2. To compare $R((\Xi, \Sigma), \widehat{\Xi})$ with $R((\Xi, \Sigma), \mathbf{X}) = mp$, we derive the unbiased estimate of the risk $R((\Xi, \Sigma), \widehat{\Xi})$.

$$R((\Xi, \Sigma), \widehat{\Xi}) = mp + \mathbb{E}\left[\widehat{\Delta}\right]$$

   Here $\widehat{\Delta}$ is a function of the eigenvalues of $\mathbf{X}^*\mathbf{X}\mathbf{S}^{-1}$.

3. If $\widehat{\Delta} \leq 0$, then $\widehat{\Xi}$ improves upon $\mathbf{X}$, i.e., $R((\Xi, \Sigma), \widehat{\Xi}) \leq mp$ for $\forall (\Xi, \Sigma)$.

4. To derive $\widehat{\Delta}$, we use integration-by-parts formulae (Stein identity for complex normal distribution and Stein-Haff identity for complex Wishart distribution).

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Stein identity for complex normal distribution
Stein-Haff identity for the complex Wishart distribution

Notation 1

★ $\mathbf{Rm}\ \boldsymbol{C}$ and $\mathbf{Im}\ \boldsymbol{c}$: real and imaginary parts of a complex number $\boldsymbol{c}$.

★ $\mathbb{R}_{>}^{\boldsymbol{p}} = \{(f_1,\ f_2,\ \ldots,\ f_p) \in \mathbb{R}^{\boldsymbol{p}} :\ f_1 > f_2 > \cdots > f_p > 0\}$

★ $\mathbb{C}^{\boldsymbol{m \times p}}$: the sets of all $\boldsymbol{m \times p}$ matrices of complex entries.

★ $\mathbf{C}'$ and $\overline{\mathbf{C}}$: the transpose and the conjugate of $\mathbf{C} \in \mathbb{C}^{\boldsymbol{m \times p}}$. $\mathbf{C}^* = \overline{\mathbf{C}'}$.

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Stein identity for complex normal distribution
Stein-Haff identity for the complex Wishart distribution

Notation 2 for Stein identity for complex normal distribution

★ For $z = x + \sqrt{-1}\,y$ $(x,\,y \in \mathbb{R})$ and differentiable function
$g(z) = u(z) + \sqrt{-1}v(z)$ ($u(z)$ and $v(z)$ are real-valued
differential functions), define

$$
\begin{aligned}
\frac{\partial}{\partial z} g &= \frac{1}{2}\left(\frac{\partial}{\partial x} - \sqrt{-1}\,\frac{\partial}{\partial y}\right) g \\
&= \frac{1}{2}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + \frac{\sqrt{-1}}{2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right).
\end{aligned}
$$

Therefore

$$
\mathrm{Re}\left(\frac{\partial g}{\partial z}\right) = \frac{1}{2}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = \frac{1}{2}\left(\frac{\partial (\mathrm{Re}\,g)}{\partial (\mathrm{Re}\,z)} + \frac{\partial (\mathrm{Im}\,g)}{\partial (\mathrm{Im}\,z)}\right)
$$

James-Stein estimator
Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Stein identity for complex normal distribution
Stein-Haff identity for the complex Wishart distribution

Notation 3 for Stein identity for complex normal distribution

★ $\mathbf{G} = (\mathbf{g}_{ij})_{i=1, 2, ..., m, j=1, 2, ..., p}$ be an $\mathbf{m} \times \mathbf{p}$ matrix
($\mathbf{g}_{ij}$'s are complex-valued differentiable functions on $\mathbb{C}^{m \times p}$).

★ For $\mathbf{x} = (\mathbf{x}_{ij})_{i=1, 2, ..., m, j=1, 2, ..., p} \in \mathbb{C}^{m \times p}$, set

$$\nabla_{\mathbf{x}} = \left(\frac{\partial}{\partial \mathbf{x}_{ij}}\right)_{i=1, 2, ..., m, j=1, 2, ..., p},$$

and define

$$\begin{aligned}
\mathbf{Re(Tr} \left(\nabla'_{\mathbf{x}}\mathbf{G}\right)) &= \mathbf{Tr} \left(\mathbf{Re}(\nabla'_{\mathbf{x}}\mathbf{G})\right) \\
&= \frac{1}{2} \sum_{j=1}^{p} \sum_{i=1}^{m} \left\{\frac{\partial(\mathbf{Re}\, \mathbf{g}_{ij})}{\partial(\mathbf{Re}\, \mathbf{x}_{ij})} + \frac{\partial(\mathbf{Im}\, \mathbf{g}_{ij})}{\partial(\mathbf{Im}\, \mathbf{x}_{ij})}\right\}.
\end{aligned}$$

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Stein identity for complex normal distribution
Stein-Haff identity for the complex Wishart distribution

Lemma (Stein's identity for complex normal distribution)

Let X be a $m \times p$ matrix having $\mathbb{C}N_{m \times p}(\Xi, I_m \otimes \Sigma)$ and let $G = (g_{ij})_{i=1, 2, \dots, m, j=1, 2, \dots, p}$ with

$$\mathbb{E}\left|\frac{\partial (\mathrm{Re}\, g_{ij})}{\partial (\mathrm{Re}\, x_{ij})}\right|_{x=X} < \infty, \quad \mathbb{E}\left|\frac{\partial (\mathrm{Im}\, g_{ij})}{\partial (\mathrm{Im}\, x_{ij})}\right|_{x=X} < \infty,$$

and $x = (x_{ij})_{i=1, 2, \dots, m, j=1, 2, \dots, p}$ in $\mathbb{C}^{m \times p}$. Then we have

$$\mathbb{E}[\mathrm{Tr}\,\{\Sigma^{-1}(X - \Xi)^* G(X) + \mathrm{Tr}\,\{\Sigma^{-1}G^*(X)(X - \Xi)\}]$$
$$= 2\mathbb{E}[\mathrm{Tr}\,(\mathrm{Re}(\nabla'_x G))|_{x=X}]$$

James-Stein estimator

Preliminaries   Complex normal and Wishart distributions

Set-up of a problem

Key ingredients

Integration-by-parts formulae

Class of estimators and unbiased estimate of the risk

Concluding remarks

Stein identity for complex normal distribution

Stein-Haff identity for the complex Wishart distribution

Sketch of proof $(m = 1)$

Recall that, for $X \sim \mathbb{C}N_p(\xi, \Sigma)$, we have

$$[X] := \begin{pmatrix} \mathrm{Re}\, X \\ \mathrm{Im}\, X \end{pmatrix} \sim N_{2p}\Big[\underbrace{\begin{pmatrix} \mathrm{Re}\, \xi \\ \mathrm{Im}\, \xi \end{pmatrix}}_{[\xi]:\, 2p \times 1}, \frac{1}{2}\underbrace{\begin{pmatrix} \mathrm{Re}\, \Sigma & -\mathrm{Im}\, \Sigma \\ \mathrm{Im}\, \Sigma & \mathrm{Re}\, \Sigma \end{pmatrix}}_{\{\Sigma\}:\, 2p \times 2p}\Big],$$

and note that

$$(X - \xi)^* \Sigma^{-1} g(X) + g^*(X) \Sigma^{-1}(X - \theta) = 2[X - \xi]'\{\Sigma\}^{-1}[g].$$

From Stein's identity for the multivariate real normal distribution, we have

$$\mathbb{E}[[X - \xi]'(\{\Sigma\}/2)^{-1}[g]] = \mathbb{E}\left[\sum_{i=1}^{p}\left(\frac{\partial(\mathrm{Re}\, g_i)}{\partial(\mathrm{Re}\, x_i)} + \frac{\partial(\mathrm{Im}\, g_i)}{\partial(\mathrm{Im}\, x_i)}\right)\Big|_{x = X}\right],$$

where $x' = (x_1, x_2, \ldots, x_p)$ and $g' = (g_1, g_2, \ldots, g_p)$.

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Stein identity for complex normal distribution
Stein-Haff identity for the complex Wishart distribution

We consider a real normal case for simplicity.

For $Z \sim N(\xi, 1)$ and a function $g : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}[\dot{g}(Z)] < \infty$ ($\dot{g}(z) = dg/dz$), we have

$$\int_{-\infty}^{\infty} (z - \xi) g(z) \exp(-(z - \xi)^2/2) \, dz$$
$$= \underbrace{\left[-g(z) \exp(-(z - \xi)^2/2)\right]_{-\infty}^{\infty}}_{0} + \int_{-\infty}^{\infty} \dot{g}(z) \exp(-(z - \xi)^2/2) \, dz$$

because $\int_{-\infty}^{\infty} g(z) \exp(-(z - \xi)^2/2) \, dz < \infty$. This shows that

$$\mathbb{E}[(Z - \xi)g(Z)] = \mathbb{E}[\dot{g}(Z)].$$

Note that $\mathbb{E}[\dot{g}(Z)] < \infty$ implies $\mathbb{E}[Zg(Z)] < \infty$ and $\mathbb{E}[g(Z)] < \infty$. See Sourav Chaterjee's lecture note(Stein's method and applications).

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Stein identity for complex normal distribution
Stein-Haff identity for the complex Wishart distribution

Notation Stein-Haff identity for complex Wishart distribution

★ Let $\mathbf{s} : p \times p = (s_{jk})$ and $\mathbf{G(s)} = (g_{ij}(\mathbf{s}))$.

★
$$\frac{\partial}{\partial s_{jk}} = \frac{1}{2}(1 + \delta_{jk}) \left\{ \frac{\partial}{\partial(\mathrm{Re}\ s_{jk})} + (1 - \delta_{jk}) \sqrt{-1} \frac{\partial}{\partial(\mathrm{Im}\ s_{jk})} \right\},$$

for $j,\ k = 1, 2, \ldots, p$ and δ_{jk} is the Kronecker delta.

★ Let $\mathbf{D_S} = (\partial/\partial s_{jk})$ be a $p \times p$ operator matrix given by

$$\{\mathbf{D_S G(S)}\}_{jk} = \sum_{l=1}^{p} \frac{\partial g_{lk}}{\partial s_{jl}}(\mathbf{S})$$
$$= \frac{1 + \delta_{jl}}{2} \sum_{l=1}^{p} \left\{ \frac{\partial g_{lk}}{\partial(\mathrm{Re}\ s_{jl})} + (1 - \delta_{j\ell}) \sqrt{-1} \frac{\partial g_{lk}}{\partial(\mathrm{Im}\ s_{jl})} \right\} \Bigg|_{\mathbf{s=S}}.$$

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Stein identity for complex normal distribution
Stein-Haff identity for the complex Wishart distribution

Lemma (Stein-Haff identity for complex Wishart distribution)

Assume that each entry of $\mathbf{G(S)}$ is a partially differentiable function with respect to $\mathrm{Re}\ s_{jk}$ and $\mathrm{Im}\ s_{jk}$, j, $k = 1, 2, \ldots, p$. Under some regularity conditions on $\mathbf{G(S)}$, the following identity holds:

$$\mathbb{E}[\mathrm{Tr}(\mathbf{G(S)\Sigma^{-1}})] = \mathbb{E}[(n-p)\mathrm{Tr}(\mathbf{G(S)S^{-1}}) + \mathrm{Tr}(\mathbf{D_S G(S)})]. \quad (1)$$

James-Stein estimator
Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Stein identity for complex normal distribution
Stein-Haff identity for the complex Wishart distribution

Sketch of proof

For $\boldsymbol{x} = (x_{ij})_{i=1,\ldots,n;\, j=1,\ldots,p}$, note that

$$\frac{\partial}{\partial x_{ij}} \exp(-\operatorname{Tr} x^* x) = -\overline{x}_{ij} \exp(-\operatorname{Tr} x^* x),$$

Assume that $\boldsymbol{X} \sim \mathbb{C}N_{n\times p}(\boldsymbol{0}, \boldsymbol{I}_n \otimes \boldsymbol{I}_p)$ and put
$\boldsymbol{S} = (s_{ij}) = \boldsymbol{X}^*\boldsymbol{X} \sim \mathbb{C}W(n, \boldsymbol{I}_p)$. Then we have

$$
\begin{aligned}
\mathbb{E}\left[\sum_{j=1}^{p} s_{ij} g_{jl}\right] &= \mathbb{E}\left[\sum_{j=1}^{p}\sum_{k=1}^{n} \overline{x}_{ki} x_{kj} g_{jl}\right] = \mathbb{E}\left[\sum_{j=1}^{p}\sum_{k=1}^{n} \frac{\partial}{\partial x_{ki}}(x_{kj} g_{jl})\right] \\
&= \mathbb{E}\left[n g_{il} + \sum_{j=1}^{p}\sum_{k=1}^{n} x_{kj} \frac{\partial g_{jl}}{\partial x_{ki}}\right].
\end{aligned}
$$

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Stein identity for complex normal distribution
Stein-Haff identity for the complex Wishart distribution

In matrix form

$$\mathbb{E}[\mathbf{SG}] = \mathbb{E}[n\mathbf{G} + (\mathbf{X}'\nabla_X)'\mathbf{G}].$$

Transform $\mathbf{X} \to \mathbf{X}(\mathbf{P}^*)^{-1}$ ($\mathbf{S} \to \mathbf{P}^{-1}\mathbf{S}(\mathbf{P}^*)^{-1}$) and $\mathbf{G} \to \mathbf{P}^*\mathbf{G}(\mathbf{P}^*)^{-1}$,
where $\mathbf{\Sigma} = \mathbf{PP}^*$ and $\mathbf{P}$ is a $p \times p$ nonsingular matrix. Then

$$\mathbb{E}[\mathbf{\Sigma}^{-1}\mathbf{SG}] = \mathbb{E}[n\mathbf{G} + (\mathbf{X}'\nabla_X)'\mathbf{G}].$$

Replacing $\mathbf{S}$ with $\mathbf{S}^{-1}\mathbf{G}$, taking the trace, and using the chain-rule
of differentiation, we have

$$\mathbb{E}[\mathrm{Tr}\,\mathbf{\Sigma}^{-1}\mathbf{G}] = \mathbb{E}[(n - p)\mathrm{Tr}\,(\mathbf{GS}^{-1}) + \mathrm{Tr}\,(D_S\mathbf{G})].$$

□

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Class of estimators
Unbiased risk estimate
Improved estimators

Class of estimators

★ Let $\mathbf{F} = \mathrm{Diag}\,(f_1,\, f_2,\, \ldots,\, f_{\min(m,\, p)})$ be the ordered eigenvalues of $\mathbf{X}^*\mathbf{X}\mathbf{S}^{-1}$, $f_1 \geq \cdots \geq f_{\min(m,\, p)}$.

★ For $p > m$ decompose $\mathbf{X}\mathbf{S}^{-1}\mathbf{X}^* = \mathbf{U}\mathbf{F}\mathbf{U}^*$, where $\mathbf{U}$ is an $m \times m$ unitary matrix.

★ For $m > p$ we decompose $\mathbf{S} = (\mathbf{A}^*)^{-1}\mathbf{A}^{-1}$ and $\mathbf{X}^*\mathbf{X} = (\mathbf{A}^*)^{-1}\mathbf{F}\mathbf{A}^{-1}$, where $\mathbf{A}$ is a $p \times p$ non-singular matrix.

★ Consider a class of estimators of the form

$$\widehat{\Xi}_H := \widehat{\Xi}_H(\mathbf{X},\, \mathbf{S}) = \begin{cases} \mathbf{X}\{\mathbf{I}_p + \mathbf{A}\mathbf{H}(\mathbf{F})\mathbf{A}^{-1}\} & \text{if } m > p \\ \{\mathbf{I}_m + \mathbf{U}\mathbf{H}(\mathbf{F})\mathbf{U}^*\}\mathbf{X} & \text{if } m < p \end{cases},$$

(2)

where $\mathbf{H} := \mathbf{H}(\mathbf{F}) = \mathrm{Diag}\,(h_1(\mathbf{F}),\, h_2(\mathbf{F}),\, \ldots,\, h_{\min(m,\, p)}(\mathbf{F}))$ whose i-th element $h_i := h_i(\mathbf{F})$, $i = 1,\, 2,\, \ldots,\, \min(m,\, p)$, is a real-valued function on $\mathbb{R}_{>}^{\min(m,\, p)}$.

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Class of estimators
Unbiased risk estimate
Improved estimators

Theorem

Under the suitable conditions, we have

$$R((\Xi, \Sigma), \widehat{\Xi}_H) = \begin{cases} mp + \mathbb{E}\Big[\widehat{\Delta}(n, m, p; H)\Big] & \text{if } m > p \\ mp + \mathbb{E}\Big[\widehat{\Delta}(n+m-p, p, m; H)\Big] & \text{if } m < p \end{cases}$$

where $h_{kk}(F) = (\partial h_k / \partial f_k)(F)$ and

$$\widehat{\Delta}(n, m, p; H) = \sum_{k=1}^{p} \Bigg\{ 2(m-p+1)h_k(F) + 2f_k h_{kk}(F)$$

$$+ 4 \sum_{b>k} \frac{f_k h_k(F) - f_b h_b(F)}{f_k - f_b} + (n+p-2)f_k h_k^2(F)$$

$$- 2f_k^2 h_{kk}(F)h_k(F) - 2 \sum_{b>k} \frac{f_k^2 h_k^2(F) - f_b^2 h_b^2(F)}{f_k - f_b} \Bigg\} \mathbb{I}\{f_1 > \cdots > f_p > 0\}$$

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Class of estimators
Unbiased risk estimate
Improved estimators

Evaluation of risk

★ We have

$$
R((\Xi, \Sigma), \widehat{\Xi}_H) = \mathbb{E}\Big[\mathrm{Tr}\,\{\Sigma^{-1}(X + G - \Xi)^*(X + G - \Xi)\}\Big]
$$

$$
= \underbrace{\mathbb{E}\Big[\mathrm{Tr}\,\{\Sigma^{-1}(X - \Xi)^*(X - \Xi)\}\Big]}_{mp}
$$

$$
+ \underbrace{\mathbb{E}\Big[\mathrm{Tr}\,\{\Sigma^{-1}G^*(X - \Xi)\} + \mathrm{Tr}\,\{\Sigma^{-1}(X - \Xi)^*G\}\Big]}_{\text{Stein's identity}}
$$

$$
+ \underbrace{\mathbb{E}\Big[\mathrm{Tr}\,\{\Sigma^{-1}G^*G\}\Big]}_{\text{Stein-Haff identity}}
$$

$$
= mp + \mathbb{E}\Big[2\,\mathrm{Tr}\,\{\mathrm{Re}\,(\nabla'_X G)\}\Big]
$$

$$
+ \mathbb{E}\Big[\mathrm{Tr}\,\{D_S G^* G\} + (n - p)\mathrm{Tr}\,\{S^{-1}G^*G\}\Big]
$$

James-Stein estimator
Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Class of estimators
Unbiased risk estimate
Improved estimators

★ When $m > p$, recall that $\widehat{\Xi}_H = X + G = X + XAH(F)A^{-1}$, where A is a $p \times p$ nonsingular matrix s.t. $S = (A^*)^{-1}A^{-1}$ and $A^*X^*XA = F = \mathrm{Diag}(f_1, \ldots, f_p)$, and that the modification H depends only on the eigenvalues of X^*XS^{-1}, .

★

$$
\begin{aligned}
R((\Xi, \Sigma), \widehat{\Xi}_H) &= mp + \mathbb{E}\Big[2\,\mathrm{Tr}\,\{\mathrm{Re}\,(\nabla'_X G)\}\Big] \\
&\quad + \mathbb{E}\Big[\mathrm{Tr}\,\{D_S G^* G\} + (n-p)\mathrm{Tr}\,\{S^{-1}G^* G\}\Big]. \\
&= mp + \mathbb{E}\Big[2\,\mathrm{Tr}\,\{\mathrm{Re}\,(\nabla'_X XAHA^{-1})\}\Big] \\
&\quad + \mathbb{E}\Big[\mathrm{Tr}\,\{D_S \underbrace{(XAHA^{-1})^*(XAHA^{-1})}_{(A^*)^{-1}HFH(F)A^{-1}}\} \\
&\quad + (n-p)\underbrace{\mathrm{Tr}\,\{S^{-1}(XAHA^{-1})^*(XAHA^{-1})\}}_{\mathrm{Tr}\,\{FH^2\}}\Big].
\end{aligned}
$$

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Class of estimators
Unbiased risk estimate
Improved estimators

★ The terms inside the first and second expectations are written as

$$\mathrm{Tr}\,\{\mathrm{Re}\,(\nabla'_X \mathbf{X}\mathbf{A}\mathbf{H}\mathbf{A}^{-1})\}$$
$$= \sum_{k=1}^{p} \left\{ f_k h_{kk} + (m - p + 1)h_{kk} + 2\sum_{b>k} \frac{f_k h_k - f_b h_b}{f_k - f_b} \right\}$$

$$\mathrm{Tr}\,\{\mathbf{D}_S((\mathbf{A}^*)^{-1}\mathbf{H}\mathbf{F}\mathbf{H}\mathbf{A}^{-1})\}$$
$$= \sum_{k=1}^{p} \left\{ 2(p - 1)f_k h_k^2 - 2f_k h_{kk} h_k - 2\sum_{b>k} \frac{f_k^2 h_k^2 - f_b^2 h_b^2}{f_k - f_b} \right\},$$

where $\mathbf{H} = \mathrm{Diag}(h_1, \ldots, h_p)$ and $h_{kk} = \partial h_k/\partial f_k$, and $\mathbf{A}$ is a $p \times p$ nonsingular matrix such that $\mathbf{S} = (\mathbf{A}^*)^{-1}\mathbf{A}^{-1}$, and that the modification $\mathbf{H}$ depends only on the eigenvalues of $\mathbf{X}^*\mathbf{X}\mathbf{S}^{-1}$, $\mathbf{F} = \mathrm{Diag}(f_1, f_2, \ldots, f_p)$.

James-Stein estimator
Preliminaries   Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Class of estimators
Unbiased risk estimate
Improved estimators

★ When $m < p$, recall that $\widehat{\Xi}_H := \mathbf{X} + \mathbf{G}$ where $\mathbf{G} = \mathbf{UHU}^*\mathbf{X}$, where $\mathbf{U}$ is an $m \times m$ unitary matrix such that $\mathbf{XS}^{-1}\mathbf{X}^* = \mathbf{UFU}^*$, and recall that an $m \times m$ diagonal matrix $\mathbf{H}$ depends only on $\mathbf{F} = \mathrm{Diag}(f_1, f_2, \ldots, f_m)$.

★ For $m < p$, we have

$$
\begin{aligned}
R((\Xi, \Sigma), \widehat{\Xi}_H) &= mp + \mathbb{E}\Big[2\,\mathrm{Tr}\,\{\mathrm{Re}\,(\nabla'_X \mathbf{G})\}\Big] \\
&\quad + \mathbb{E}\Big[\mathrm{Tr}\,\{\mathbf{D}_S \mathbf{G}^*\mathbf{G} + (n-p)\,\mathrm{Tr}\,\{\mathbf{S}^{-1}\mathbf{G}^*\mathbf{G}\}\Big]. \\
&= mp + \mathbb{E}\Big[2\,\mathrm{Tr}\,\{\mathrm{Re}\,(\mathbf{UH}(\mathbf{F})\mathbf{U}^*\mathbf{X})\}\Big] \\
&\quad + \mathbb{E}\Big[\mathrm{Tr}\,\{\mathbf{D}_S \underbrace{(\mathbf{UH}(\mathbf{F})\mathbf{U}^*\mathbf{X})^*(\mathbf{UH}(\mathbf{F})\mathbf{U}^*\mathbf{X})\}}_{\mathbf{X}^*\mathbf{UH}^2\mathbf{U}^*\mathbf{X}} \\
&\quad + (n-p)\,\underbrace{\mathrm{Tr}\,\{\mathbf{S}^{-1}(\mathbf{UH}(\mathbf{F})\mathbf{U}^*\mathbf{X})^*(\mathbf{UH}(\mathbf{F})\mathbf{U}^*\mathbf{X})\}}_{\mathrm{Tr}\,\{\mathbf{FH}^2\}}\Big].
\end{aligned}
$$

James-Stein estimator

Preliminaries Complex normal and Wishart distributions

Set-up of a problem

Key ingredients

Integration-by-parts formulae

Class of estimators and unbiased estimate of the risk

Concluding remarks

Class of estimators

Unbiased risk estimate

Improved estimators

★ The terms inside the first and second expectations are written as

$$\mathrm{Tr}\left\{\mathrm{Re}\left(\nabla_X' \mathbf{U} \mathbf{H} \mathbf{U}^* \mathbf{X}\right)\right\}$$
$$= \sum_{k=1}^{m} \left\{ f_k h_{kk} + (p - m + 1) h_{kk} + 2 \sum_{b>k} \frac{f_k h_k - f_b h_b}{f_k - f_b} \right\}$$

$$\mathrm{Tr}\left\{\mathbf{D}_S(\mathbf{X}^* \mathbf{U} \mathbf{H}^2 \mathbf{U} \mathbf{X})\right\}$$
$$= \sum_{k=1}^{m} \left\{ 2(m - 1) f_k h_k^2 - 2 f_k^2 h_{kk} h_k - 2 \sum_{b>k} \frac{f_k^2 h_k - f_b^2 h_b}{f_k - f_b} \right\},$$

where $\mathbf{H} = \mathrm{Diag}(h_1, \ldots, h_m)$ and $h_{kk} = \partial h_k / \partial f_k$, and $\mathbf{U}$ is an $m \times m$ unitary matrix s.t.
$\mathbf{X} \mathbf{S}^{-1} \mathbf{X}^* = \mathbf{U} \mathrm{Diag}(f_1, f_2, \ldots, f_m) \mathbf{U}^*$.

□

James-Stein estimator
Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Class of estimators
Unbiased risk estimate
Improved estimators

Theorem (Efron-Morris type estimator)

Let

$$c = \begin{cases} \frac{m-p}{n+p} & \text{if } m > p \\ \frac{p-m}{n+2m-p} & \text{if } m < p \end{cases}$$

and put

$$H(F) = -\text{Diag}\left(\frac{c}{f_1}, \frac{c}{f_2}, \cdots, \frac{c}{f_{\min(m,p)}}\right)$$

Then the estimator

$$\widehat{\Xi}_H = \begin{cases} X\{I_p - c(X^*X)^{-1}S\} & \text{if } m > p \\ \{I_m - c(XS^{-1}X^*)^{-1}\}X & \text{if } m < p \end{cases}$$

improves upon X*, i.e.,* $R((\Xi, \Sigma), \widehat{\Xi}_H) \leq R((\Xi, \Sigma), X)$ *for* $\forall(\Xi, \Sigma)$*.*

James-Stein estimator
Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

Class of estimators
Unbiased risk estimate
Improved estimators

Theorem

Let $c_k = \frac{m+p-2k}{n-p+2k}$ $(k = 1, 2, \ldots, \min(m, p))$ and put

$$\mathbf{H}(\mathbf{F}) = -\mathrm{Diag}\left(\frac{c_1}{f_1}, \frac{c_2}{f_2}, \ldots, \frac{c_{\min(m,p)}}{f_{\min(m,p)}}\right).$$

Then the estimator

$$\widehat{\Xi} = \begin{cases} \mathbf{X}\{\mathbf{I}_p + \mathbf{A}\mathbf{H}(\mathbf{F})\mathbf{A}^{-1}\} & \text{if } m > p \\ \{\mathbf{I}_m + \mathbf{U}\mathbf{H}(\mathbf{F})\mathbf{U}^*\}\mathbf{X} & \text{if } m < p \end{cases}$$

improves upon $\mathbf{X}$, i.e., $R((\Xi, \Sigma), \widehat{\Xi}_H) \leq R((\Xi, \Sigma), \mathbf{X})$ for $\forall(\Xi, \Sigma)$.

Recall that $\mathbf{U}$ is an $m \times m$ unitary matrix s.t. $\mathbf{X}\mathbf{S}^{-1}\mathbf{X}^* = \mathbf{U}\mathbf{F}\mathbf{U}^*$ and that $\mathbf{A}$ is a $p \times p$ nonsingular matrix s.t. $\mathbf{S} = (\mathbf{A}^*)^{-1}\mathbf{A}^{-1}$.

James-Stein estimator
Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

1 The results are complex version of those for estimation problem for the mean matrix of the real multivariate normal distribution (see Konno (JMVA, 1991) and Kariya et al. (JMVA, 1999)).

2 However the results are parallel to real case.

3 Note that the substitution rule to get the second part from the first part is $(p, m, n) \rightarrow (m, p, n + m - p)$, i.e.,

$$R((\Xi, \Sigma), \widehat{\Xi}_H) = \begin{cases} mp + \mathbb{E}\left[\widehat{\Delta}(n, m, p; H)\right] & \text{if } m > p \\ mp + \mathbb{E}\left[\widehat{\Delta}(n + m - p, p, m; H)\right] & \text{if } m < p \end{cases}$$

This is the same as the rule substitution for the joint distribution of $\mathbf{X}^*\mathbf{X}\mathbf{S}^{-1}$.

James-Stein estimator
Preliminaries Complex normal and Wishart distributions
Set-up of a problem
Key ingredients
Integration-by-parts formulae
Class of estimators and unbiased estimate of the risk
Concluding remarks

When $\Xi = 0$, the joint distribution of X^*XS^{-1} is, apparting from normalizing constants,

$$\Pi_{k=1}^{p} \frac{f_k^{m-p}}{(1+f_k)^{n+m}} \Pi_{k=1}^{p-1} \Pi_{j=k+1}^{p} (f_k - f_j)^2 \Pi_{k=1}^{p} df_k$$

if $m > p$ while it is

$$\Pi_{k=1}^{m} \frac{f_k^{p-m}}{(1+f_k)^{n+m}} \Pi_{k=1}^{m-1} \Pi_{j=k+1}^{m} (f_k - f_j)^2 \Pi_{k=1}^{m} df_k,$$

if $p > m$. The second one is obtained by the substitution rule $(p, m, n) \to (m, p, n + m - p)$. Good reference is Khatri(1965, Ann. ,Math. Statist).