Spectral analysis of large reflexive generalized inverse and Moore-Penrose inverse matrices

Taras Bodnar

Nestor Parolya

Postal address:
Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:
http://www.math.su.se

Spectral analysis of large reflexive generalized inverse and Moore-Penrose inverse matrices

Taras Bodnar[a] and Nestor Parolya[b]

[a]Department of Mathematics, Stockholm University, Stockholm, Sweden

[b]Delft Institute of Applied Mathematics, Delft University of Technology, Delft, The Netherlands

April 2020

Abstract

A reflexive generalized inverse and the Moore-Penrose inverse are often confused in statistical literature but in fact they have completely different behaviour in case the population covariance matrix is not a multiple of identity. In this paper, we study the spectral properties of a reflexive generalized inverse and of the Moore-Penrose inverse of the sample covariance matrix. The obtained results are used to assess the difference in the asymptotic behaviour of their eigenvalues.

Keywords: reflexive generalized inverse; Moore-Penrose inverse; random matrix theory; free probability.

1 Introduction

Let $\mathbf{Y}_n = (\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n)$ be the $p \times n$ data matrix which consists of n column vectors of dimension p with $E(\mathbf{y}_i) = \mathbf{0}$ and $Cov(\mathbf{y}_i) = \mathbf{\Sigma}$ for $i \in 1, ..., n$. We assume that $p/n \to c \in (1, +\infty)$ as $n \to \infty$. This type of limiting behavior is also referred to a "large dimensional asymptotics" or "the Kolmogorov asymptotics". In this case, the traditional estimators perform very poorly and tend to over/underestimate the unknown parameters of the asset returns, i.e., the mean vector and the covariance matrix.

Throughout this paper it is assumed that there exists a $p \times n$ random matrix $\mathbf{X}_n$ which consists of independent and identically distributed (i.i.d.) real random variables with zero mean and unit variance such that

$$\mathbf{Y}_n = \mathbf{\Sigma}^{\frac{1}{2}} \mathbf{X}_n \,, \tag{1}$$

where the matrix $\mathbf{\Sigma}^{\frac{1}{2}}$ denotes the symmetric square root of the population covariance matrix $\mathbf{\Sigma}$. Other square roots of $\mathbf{\Sigma}$, like the lower triangular matrix of the Cholesky decomposition, can also be used. Note that the observation matrix $\mathbf{Y}_n$ consists of dependent rows although its columns are independent. It is worth mentioning that although the assumption of time independence looks quite restrictive in real-life applications, the model can be extended to dependent variables (for instance, causal AR(1) models) [1] by imposing more complicated conditions on the elements of $\mathbf{Y}_n$ (see, [2] for details) or by assuming an m-dependent structure (e.g., [11] for a finite number of dependent entries, and [9] and [16] for a possibly increasing number). Nevertheless, this will not change the main results of our paper and would only make the proofs more technical. That is why we assume independence for the sake of brevity and transparency.

The sample covariance matrix is given by (e.g., [13])

$$\mathbf{S}_n = \frac{1}{n}\mathbf{Y}_n\mathbf{Y}_n' = \frac{1}{n}\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{X}_n\mathbf{X}_n'\mathbf{\Sigma}^{\frac{1}{2}}. \tag{2}$$

Throughout the paper it is assumed that the sample size n is smaller than the dimension p of random vectors $\mathbf{y}_i$, $i = 1, ..., n$, that is $p/n \to c > 1$ as $n \to \infty$. In this case the sample covariance matrix is singular and its inverse does not exist (cf., [6, 8, 12]). On the other side, the inverse of the population covariance matrix $\mathbf{\Sigma}$ is present in many applications from finance, signal processing, biostatistics, environmentrics, etc. (see, e.g., [5, 7, 10]). In practice, the Moore-Penrose inverse is usually employed (cf., [4, 12]), while the other types of the generalize inverse (see, [15]) can also be used.

In this paper we compare the spectral properties of two generalized inverses of the singular sample covariance matrix given by:

- Moore-Penrose inverse:

$$\mathbf{S}_n^+ = \left(\frac{1}{n}\mathbf{Y}_n\mathbf{Y}_n'\right)^+ = \frac{1}{n}\mathbf{Y}_n\left(\frac{1}{n}\mathbf{Y}_n'\mathbf{Y}_n\right)^{-2}\mathbf{Y}_n' = \frac{1}{n}\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{X}_n\left(\frac{1}{n}\mathbf{X}_n'\mathbf{\Sigma}\mathbf{X}_n\right)^{-2}\mathbf{X}_n'\mathbf{\Sigma}^{\frac{1}{2}} \tag{3}$$

- Reflexive inverse

$$\mathbf{S}_n^- = \mathbf{\Sigma}^{-\frac{1}{2}}\left(\frac{1}{n}\mathbf{X}_n\mathbf{X}_n'\right)^+\mathbf{\Sigma}^{-\frac{1}{2}} = \frac{1}{n}\mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{X}_n\left(\frac{1}{n}\mathbf{X}_n'\mathbf{X}_n\right)^{-2}\mathbf{X}_n'\mathbf{\Sigma}^{-\frac{1}{2}} \tag{4}$$

Although the Moore-Penrose inverse $\mathbf{S}^+$ can directly be computed from the observation matrix $\mathbf{Y}_n$, the derivation of its stochastic properties might be challenging in some practical problems. On the other side, the computation of the considered reflexive inverse $\mathbf{S}^-$ is not possible in practice, but the derivation of the stochastic properties is considerably simplified. The goal of this paper is to quantify the difference between the two

---

[1]Bai and Zhou [2] define a general dependence structure in the following way: for all k, $E(Y_{jk}Y_{lk}) = \sigma_{lj}$, and for any non-random matrix $\mathbf{B}$ with bounded norm, it holds that $E\left|\mathbf{y}_k^\top\mathbf{B}\mathbf{y}_k - \mathrm{tr}(\mathbf{B}\mathbf{\Sigma})\right|^2 = o\left(n^2\right)$ where $\mathbf{\Sigma} = (\sigma_{lj})$.

generalized inverse matrices with the aim to develop a statistical methodology to assess the estimation errors when the Moore-Penrose inverse is used.

The rest of the paper is structured as follows. In the next section, we study the asymptotic properties of $\mathbf{S}^+$ and $\mathbf{S}^-$ by deriving two integral equations whose solutions are the Stieltjes transforms of the limiting spectral distributions of $\mathbf{S}^+$ and $\mathbf{S}^-$. These findings are used in Section 3, where the asymptotic behaviour of the Frobenius norm of the difference between $\mathbf{S}^+$ and $\mathbf{S}^-$ is investigated. Section 4 presents the results of a numerical illustration, while concluding remarks are provided in Section 5.

2   Asymptotic properties of $\mathbf{S}^+$ and $\mathbf{S}^-$

For a symmetric matrix $\mathbf{A}$ we denote by $\lambda_1(\mathbf{A}) \geq \ldots \geq \lambda_p(\mathbf{A})$ its ordered eigenvalues and by $F^{\mathbf{A}}(t)$ the corresponding empirical distribution function (e.d.f.), that is

$$F^{\mathbf{A}}(t) = \frac{1}{p} \sum_{i=1}^{p} \mathbb{1}\{\lambda_i(\mathbf{A}) \leq t\},$$

where $\mathbb{1}\{\cdot\}$ is the indicator function. Furthermore, for a function $G : \mathbb{R} \to \mathbb{R}$ of bounded variation the Stieltjes transform is introduced by

$$m_G(z) = \int_{-\infty}^{+\infty} \frac{1}{\lambda - z} dG(\lambda); \quad z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C} : \Im z > 0\}.$$

Note that there is a direct connection between $m_G(z)$ and moment generating function of G, $\Psi_G(z)$ given by

$$\Psi_G(z) = -\frac{1}{z} m_G\left(\frac{1}{z}\right) - 1.$$

In Theorem 1 we present the expressions of the Stieltjes transform for $\mathbf{S}^+$ and $\mathbf{S}^-$.

Theorem 1. *Let the $p \times n$ noise matrix $\mathbf{X}_n$ consist of i.i.d. real random variables with zero mean and unit variance. Assume that $\mathbf{\Sigma}_n$ is nonrandom symmetric and positive definite with a bounded spectral norm and the e.d.f. $H_n = F^{\mathbf{\Sigma}_n}$ converges weakly to a nonrandom distribution function H. Then*

(i) *the e.d.f. $F^{\mathbf{S}_n^+}$ converges weakly almost surely to some deterministic c.d.f. P^+ whose Stieltjes transformation m_P satisfies the following equation:*

$$m_{P^+}(z) = -\frac{1}{z}\left(2 - c^{-1} + \int_{-\infty}^{+\infty} \frac{dH(\tau)}{z\tau c(zm_{P^+}(z) + 1) - 1}\right);$$

(ii) *the e.d.f. $F^{\mathbf{S}_n^-}$ converges weakly almost surely to some deterministic c.d.f. P^- whose Stieltjes transformation m_{P^-} satisfies the following equation:*

$$m_{P^-}(z) = -\frac{1}{z} - \frac{1}{z} \int_{-\infty}^{+\infty} \frac{dH(\tau)}{\tau c z^2 m_{P^-}(z)\left(1 - \frac{c}{1 - c - czm_{P^-}(z)}\right) - 1}.$$

3

Proof. (i) The result of part (i) of the theorem was derived in Theorem 2.1 of [4].

(ii) In order to prove the asymptotic result in the case of $\mathbf{S}_n^-$, whose eigenvalues are the same as those of the product of two matrices $\boldsymbol{\Sigma}^{-1}$ and $\tilde{\mathbf{S}}_n^+ = 1/n\mathbf{X}_n \left(\frac{1}{n}\mathbf{X}_n'\mathbf{X}_n \right)^{-2} \mathbf{X}_n'$, we use subordination results of free probability (see, e.g., [3]). Namely for two Borel probability measures μ and ν on $[0, +\infty)$ (both are limits of $F^{\tilde{S}_n^+}$ and $F^{\boldsymbol{\Sigma}^{-1}}$, respectively), there exist two unique analytic functions $F_1, F_2 : \mathbb{C} \setminus [0, +\infty) \to \mathbb{C} \setminus [0, +\infty)$ such that

$$\frac{F_1(z)F_2(z)}{z} = \eta_\mu(F_1(z)) = \eta_\nu(F_2(z)) = \eta(z) \,, \tag{5}$$

where $\eta(z) = 1 - \frac{z}{g_{P^-}(1/z)}$ is the eta transform and g_{P^-} is the Cauchy transform (the negative Stieltjes transform), i.e., $g_{P^-}(z) = -m_{P^-}(z)$, for the limit of $F^{\mathbf{S}_n^-}$ denoted by P^-. In particular, from the first and the second equalities in (5) we get that for any $z \in \mathbb{C}^+ = \{z : \Im(z) > 0\}$ such that F_2 is analytic at z, the function $F_2(z)$ satisfies the following equation:[2]

$$\eta_\nu(F_2(z)) = \eta_\mu \left(\frac{\eta_\nu(F_2(z))z}{F_2(z)} \right) \,. \tag{6}$$

On the other hand, from the last equality in (5) we have that $g_{P^-}(z)$ satisfies the equality

$$g_{P^-}(z) = \frac{1}{zF_2(1/z)} g_\nu \left(\frac{1}{F_2(1/z)} \right) \,. \tag{7}$$

Thus, we need first to find the so called subordination function $F_2(z)$ from (6) and plug it into (7). For simplicity we further suppress the subindex of g_{P^-}.

Let

$$\Theta(z) = \frac{F_2(z)}{z\eta(z)} \tag{8}$$

and rewrite (6) using $\eta(z) = \eta_\nu(F_2(z))$ in the following way

$$\eta(z) = \eta_\mu \left(\frac{\eta(z)z}{F_2(z)} \right) = \eta_\mu \left(\frac{1}{\Theta(z)} \right) \,. \tag{9}$$

Using the definition of the eta transform we get

$$1 - \frac{1}{\frac{1}{z}g(1/z)} = 1 - \frac{1}{\Theta(z)} \frac{1}{g_\mu(\Theta(z))}$$

and, hence,

$$\frac{1}{z}g(1/z) = \Theta(z)g_\mu(\Theta(z)) \,. \tag{10}$$

---

[2]In fact, a similar equation also holds for F_1 when one replaces μ and ν in (6) but for our purposes it is enough to know one of them.

From [4] we get that

$$g_\mu(z) = -m_{P^-}(z) = \frac{1}{z}\left(2 - c^{-1} + \frac{m_{MP}(1/z)}{z}\right), \tag{11}$$

where $m_{MP}(z)$ is the Stieltjes transformation of the Marchenko-Pastur law given by (see, e.g., [1])

$$m_{MP}(z) = \frac{1}{2cz}\left(1 - c - z + \sqrt{(1 + c - z)^2 - 4c}\right). \tag{12}$$

Thus, the equation (10) becomes

$$
\begin{aligned}
\frac{1}{z}g(1/z) &= \Theta(z)\frac{1}{\Theta(z)}\left(2 - c^{-1} + \frac{m_{MP}(1/\Theta(z))}{\Theta(z)}\right) \\
&= 2 - c^{-1} + \frac{1 - c - \frac{1}{\Theta(z)} + \sqrt{(1 + c - \frac{1}{\Theta(z)})^2 - 4c}}{\Theta(z)2c\frac{1}{\Theta(z)}} \\
&= \frac{1}{2c}\left(2c - 2 + (1 + c - \frac{1}{\Theta(z)}) + \sqrt{(1 + c - \frac{1}{\Theta(z)})^2 - 4c}\right),
\end{aligned}
$$

or, equivalently, by rearranging terms we obtain

$$2\left(c(\frac{1}{z}g(1/z) - 1) + 1\right) - (1 + c - \frac{1}{\Theta(z)}) = \sqrt{(1 + c - \frac{1}{\Theta(z)})^2 - 4c}, \tag{13}$$

where by squaring of both sides, we get

$$\left(c(\frac{1}{z}g(1/z) - 1) + 1\right)^2 - \left(c(\frac{1}{z}g(1/z) - 1) + 1\right)(1 + c - \frac{1}{\Theta(z)}) + c = 0$$

or,

$$
\begin{aligned}
&1 + c^2(\frac{1}{z}g(1/z) - 1)^2 + 2c(\frac{1}{z}g(1/z) - 1) - 1 + \frac{1}{\Theta(z)} \\
&- c(\frac{1}{z}g(1/z) - 1)(1 + c - \frac{1}{\Theta(z)}) = 0,
\end{aligned}
$$

which yields

$$c(\frac{1}{z}g(1/z) - 1) + (1 - c) + \frac{1}{\Theta(z)c(\frac{1}{z}g(1/z) - 1)} + \frac{1}{\Theta(z)} = 0. \tag{14}$$

From (14) we find $\Theta(z)$ as a function of $g(1/z)$ expressed as

$$\Theta(z) = \frac{1 + c(\frac{1}{z}g(1/z) - 1)}{c(\frac{1}{z}g(1/z) - 1)(c - 1 - c(\frac{1}{z}g(1/z) - 1))} \tag{15}$$

or, in terms of $F_2(z)$ given by

$$
\begin{aligned}
F_2(z) &= \frac{z - cz + cg(1/z)}{c(\frac{1}{z}g(1/z) - 1)(c - 1 - c(\frac{1}{z}g(1/z) - 1))}\frac{\frac{1}{z}g(1/z) - 1}{\frac{1}{z}g(1/z)} \\
&= \frac{z - cz + cg(1/z)}{c(-z + 2cz - cg(1/z))}\frac{z^2}{g(1/z)}. \tag{16}
\end{aligned}
$$

5

At last, we use the property $g_\nu(z) = 1/z - 1/z^2 g_H(1/z)$ and plug $F_2(z)$ in (7). This leads to

$$
\begin{aligned}
g(z) &= \frac{1}{zF_2(1/z)}\left(F_2(1/z) - F_2^2(1/z)g_H(F_2(1/z))\right) \\
&= \frac{1}{z} - \frac{F_2(1/z)}{z}g_H(F_2(1/z)) = \frac{1}{z} - \frac{F_2(1/z)}{z}\int_{-\infty}^{+\infty}\frac{dH(\tau)}{F_2(1/z) - \tau} \\
&= \frac{1}{z} + \int_{-\infty}^{+\infty}\frac{dH(\tau)}{\tau z/F_2(1/z) - z} = \frac{1}{z} + \int_{-\infty}^{+\infty}\frac{dH(\tau)}{-\tau cz^3 g(z)\left(1 - \frac{c}{1-c+czg(z)}\right) - z}.
\end{aligned}
$$

The latter can be rewritten in terms of Stieltjes transform $m_{P^-}(z)$ by

$$
m_{P^-}(z) = -\frac{1}{z} - \frac{1}{z}\int_{-\infty}^{+\infty}\frac{dH(\tau)}{\tau cz^2 m_{P^-}(z)\left(1 - \frac{c}{1-c-czm_{P^-}(z)}\right) - 1} \tag{17}
$$

and the second part of theorem is proved.

$\square$

Although the Stieltjes transforms of two generalize inverse matrices are quite different, they have one in common: besides the fact that they are equal for $\mathbf{\Sigma} = \mathbf{I}$, we also observe that they become close to each other in case c tends to one from the right. Thus, if p/n is close to one from the right, then there should be no large difference in using the Moore-Penrose inverse $\mathbf{S}^+$ or the generalized reflexive inverse $\mathbf{S}^-$ asymptotically.

3 Quantification the difference between the asymptotic behaviour of $\mathbf{S}^+$ and $\mathbf{S}^-$

Using the Stieltjes transforms computed in the case of $\mathbf{S}^+$ and $\mathbf{S}^-$, we are able to compare their moments. In order to quantify this difference more carefully we consider the quadratic or the so-called Frobenius norm of the difference between $\mathbf{S}^+$ and $\mathbf{S}^-$ given by

$$
\begin{aligned}
\|\mathbf{S}^- - \mathbf{S}^+\|_F^2 &= \mathrm{tr}\left((\mathbf{S}^- - \mathbf{S}^+)(\mathbf{S}^- - \mathbf{S}^+)^\top\right) \\
&= \mathrm{tr}\left(\mathbf{S}^-\mathbf{S}^-\right) - 2\mathrm{tr}\left(\mathbf{S}^+\mathbf{S}^-\right) + \mathrm{tr}\left(\mathbf{S}^+\mathbf{S}^+\right) \\
&= \mathrm{tr}\left(\mathbf{S}^-\mathbf{S}^-\right) - \mathrm{tr}\left(\mathbf{S}^+\mathbf{S}^+\right) \\
&= \|\mathbf{S}^-\|_F^2 - \|\mathbf{S}^+\|_F^2.
\end{aligned}
$$

In Theorem 2 the asymptotic equivalents for both Frobenius norms are provided.

6

Theorem 2. *Under assumptions of Theorem 1 it holds almost surely as $p/n \to c > 1$*

$$1/p\|\mathbf{S}^+\|_F^2 \longrightarrow c^{-1} \left(\frac{1}{m_{\underline{F}}^2(0)} - c \int\limits_{-\infty}^{+\infty} \frac{\tau^2 dH(\tau)}{(1 + \tau m_{\underline{F}}(0))^2} \right)^{-1}, \tag{18}$$

$$1/p\|\mathbf{S}^-\|_F^2 \longrightarrow \frac{(1 + c(c-1))}{c^2(c-1)^3} \left(\int\limits_{-\infty}^{+\infty} \frac{dH(\tau)}{\tau} \right)^2 + \frac{1}{(c(c-1))^2} \int\limits_{-\infty}^{+\infty} \frac{dH(\tau)}{\tau^2}, \tag{19}$$

where $m_{\underline{F}}(0)$ satisfies the following equation

$$\frac{1}{m_{\underline{F}}(0)} = c \int\limits_{-\infty}^{+\infty} \frac{\tau dH(\tau)}{1 + \tau m_{\underline{F}}(0)}.$$

Proof. In order to prove the statement of the theorem, we rewrite the Frobenius norm of random matrix $\mathbf{A}$ for all $z \in \mathbb{C}^+$ in the following way

$$1/p\|\mathbf{A}\|_F^2 = 1/p\mathrm{tr}(\mathbf{A}^2) = -\frac{1}{2}\frac{\partial^2}{\partial z^2}\frac{1}{p}\frac{\mathrm{tr}(\mathbf{A} - 1/z\mathbf{I})^{-1}}{z}\bigg|_{z=0} = -\frac{1}{2}\frac{\partial^2}{\partial z^2}\frac{m_{F^{\mathbf{A}}}(1/z)}{z}\bigg|_{z=0}. \tag{20}$$

We start with $\mathbf{A} = \mathbf{S}^+$. Let

$$\Gamma_n(z) = \frac{m_{F^{\mathbf{S}^+}}(1/z)}{z}$$

Due to Theorem 1.i we get that $\Gamma_n^+(z)$ converges almost surely to deterministic $\Gamma^+(z)$, which satisfies the following asymptotic equation

$$\Gamma^+(z) = -\left(2 - c^{-1} + \int_{-\infty}^{+\infty} \frac{z dH(\tau)}{\tau c(\Gamma^+(z) + 1) - z} \right). \tag{21}$$

For the computation of $1/p\|\mathbf{S}^+\|_F^2$, the quantities $\Gamma^+(z)$, $\frac{\partial}{\partial z}\Gamma^+(z)$ and $\frac{\partial^2}{\partial z^2}\Gamma^+(z)$ should be evaluated at zero. We rewrite (21) in an equivalent way

$$\Gamma^+(z) = -1 - c^{-1}z m_{\underline{F}}(z), \tag{22}$$

where $m_{\underline{F}}(z)$ is the limiting Stieltjes transform of $m_{F^{1/n}\mathbf{Y}^\top\mathbf{Y}}$, which satisfies the following equation (see, e.g., [14]

$$m_{\underline{F}}(z) = \left(c \int\limits_{-\infty}^{+\infty} \frac{\tau dH(\tau)}{1 + \tau m_{\underline{F}}(z)} - z \right)^{-1}. \tag{23}$$

The advantage of (22) over (21) lies in the calculation of the derivatives, which can be done easier by the former. Since the quantity $m_{\underline{F}}(z)$ is bounded at zero for $c > 1$, we immediately obtain

$$\Gamma^+(0) \equiv \lim_{z \to 0^+} \Gamma^+(z) = -1. \tag{24}$$

7

The first derivative of $\Gamma^+(z)$ is given by

$$\frac{\partial}{\partial z}\Gamma^+(z) = -c^{-1}m_{\underline{F}}(z) - c^{-1}z\frac{\partial}{\partial z}m_{\underline{F}}(z) \tag{25}$$

and, thus, taking the limit $z \to 0^+$ and using that $\frac{\partial}{\partial z}m_{\underline{F}}(z) = O(1)$ as $z \to 0^+$, we get

$$\Gamma'^{+}(0) \equiv \lim_{z\to 0^+} \frac{\partial}{\partial z}\Gamma^+(z) = -c^{-1}m_{\underline{F}}(0), \tag{26}$$

where $m_{\underline{F}}(0)$ satisfies the equation

$$\frac{1}{m_{\underline{F}}(0)} = c\int_{-\infty}^{+\infty}\frac{\tau dH(\tau)}{1 + \tau m_{\underline{F}}(0)}. \tag{27}$$

The second derivative of $\Gamma^+(z)$ is equal to

$$\frac{\partial^2}{\partial z^2}\Gamma^+(z) = -2c^{-1}\frac{\partial}{\partial z}m_{\underline{F}}(z) - c^{-1}z\frac{\partial^2}{\partial z^2}m_{\underline{F}}(z). \tag{28}$$

Denoting $m'_{\underline{F}}(z) = \frac{\partial}{\partial z}m_{\underline{F}}(z)$ and using (23) together with (27), we obtain

$$m'_{\underline{F}}(0) \equiv \lim_{z\to 0^+}m'_{\underline{F}}(z) = \lim_{z\to 0^+}\frac{c\int_{-\infty}^{+\infty}\frac{\tau^2 m'_{\underline{F}}(z)dH(\tau)}{(1+\tau m_{\underline{F}}(z))^2} + 1}{\left(c\int_{-\infty}^{+\infty}\frac{\tau dH(\tau)}{1+\tau m_{\underline{F}}(z)}\right)^2} = \frac{m'_{\underline{F}}(0)c\int_{-\infty}^{+\infty}\frac{\tau^2 dH(\tau)}{(1+\tau m_{\underline{F}}(0))^2} + 1}{\frac{1}{m_{\underline{F}}^2(0)}}$$

or, equivalently,

$$m'_{\underline{F}}(0) = \left(\frac{1}{m_{\underline{F}}^2(0)} - c\int_{-\infty}^{+\infty}\frac{\tau^2 dH(\tau)}{(1 + \tau m_{\underline{F}}(0))^2}\right)^{-1}. \tag{29}$$

Finally, the application of (28) and (29) together with $\frac{\partial^2}{\partial z^2}m_{\underline{F}}(z) = O(1)$ as $z \to 0^+$ leads to

$$\Gamma''^{+}(0) \equiv \lim_{z\to 0^+}\frac{\partial^2}{\partial z^2}\Gamma^+(z) = -2\frac{c^{-1}}{\frac{1}{m_{\underline{F}}^2(0)} - c\int_{-\infty}^{+\infty}\frac{\tau^2 dH(\tau)}{(1+\tau m_{\underline{F}}(0))^2}}.$$

Now, the first result of the theorem follows from (20).

Similarly, for the second identity of Theorem 2 we denote $\Gamma^-(z)$ as a limit of

$$\Gamma_n^-(z) = \frac{m_{F\mathbf{s}^-}(1/z)}{z}. \tag{30}$$

Then using Theorem 1, the limiting function $\Gamma^-(z)$ satisfies the following asymptotic equation

$$\Gamma^-(z) = -1 - \int_{-\infty}^{+\infty}\frac{zdH(\tau)}{\tau c\Gamma^-(z)\left(1 - \frac{c}{1-c-c\Gamma^-(z)}\right) - z}. \tag{31}$$

8

Here we immediately get that $\Gamma^-(z)$ is bounded at zero and it holds that

$$\lim_{z\to 0^+}\Gamma^-(z)=-1,\tag{32}$$

$$\lim_{z\to 0^+}\frac{\partial}{\partial z}\Gamma^-(z)=-\frac{1}{c(c-1)}\int_{-\infty}^{+\infty}\frac{dH(\tau)}{\tau},\tag{33}$$

where, the result for the first derivative of $\Gamma^-(z)$ follows from the identity

$$\left(1-z\int_{-\infty}^{+\infty}\frac{\tau c\left(1-\frac{c(1-c)}{(1-c-c\Gamma^-(z))^2}\right)dH(\tau)}{\left(\tau c\Gamma^-(z)\left(1-\frac{c}{1-c-c\Gamma^-(z)}\right)-z\right)^2}\right)\frac{\partial}{\partial z}\Gamma^-(z)$$

$$=-\int_{-\infty}^{+\infty}\frac{\tau c\Gamma^-(z)\left(1-\frac{c}{1-c-c\Gamma^-(z)}\right)dH(\tau)}{\left(\tau c\Gamma^-(z)\left(1-\frac{c}{1-c-c\Gamma^-(z)}\right)-z\right)^2}$$

$$=-\int_{-\infty}^{+\infty}\frac{dH(\tau)}{\left(\tau c\Gamma^-(z)\left(1-\frac{c}{1-c-c\Gamma^-(z)}\right)-z\right)}-\int_{-\infty}^{+\infty}\frac{zdH(\tau)}{\left(\tau c\Gamma^-(z)\left(1-\frac{c}{1-c-c\Gamma^-(z)}\right)-z\right)^2}.$$

For the second derivative we get the following equality

$$(1+O(z))\frac{\partial^2}{\partial z^2}\Gamma^-(z)=2\frac{\partial}{\partial z}\Gamma^-(z)\int_{-\infty}^{+\infty}\frac{\tau c\left(1-\frac{c(1-c)}{(1-c-c\Gamma^-(z))^2}\right)dH(\tau)}{\left(\tau c\Gamma^-(z)\left(1-\frac{c}{1-c-c\Gamma^-(z)}\right)-z\right)^2}$$

$$-2\int_{-\infty}^{+\infty}\frac{dH(\tau)}{\left(\tau c\Gamma^-(z)\left(1-\frac{c}{1-c-c\Gamma^-(z)}\right)-z\right)^2}+O(z)$$

and taking the limit $z\to 0^+$ from both sides leads to

$$\lim_{z\to 0^+}\frac{\partial^2}{\partial z^2}\Gamma^-(z)=2\frac{\partial}{\partial z}\Gamma^-(z)\frac{c(1-c(1-c))}{c^2(c-1)^2}\int_{-\infty}^{+\infty}\frac{dH(\tau)}{\tau}-2\frac{1}{(c(c-1))^2}\int_{-\infty}^{+\infty}\frac{dH(\tau)}{\tau^2}$$

$$=-2\frac{(1+c(c-1))}{c^2(c-1)^3}\left(\int_{-\infty}^{+\infty}\frac{dH(\tau)}{\tau}\right)^2-2\frac{1}{(c(c-1))^2}\int_{-\infty}^{+\infty}\frac{dH(\tau)}{\tau^2}.\tag{34}$$

The second statement of the theorem follows now from (34) and (20). $\qquad\square$

For a better visualization of the results of Theorem 2 from the statistical point of view, we present its empirical counterpart. Here, the almost sure convergence of the asymptotic equivalents are summarize in Corollary 1.

Corollary 1. *Under assumptions of Theorem 1 for $p/n\to c>1$ holds almost surely*

$$1/p\left|\|\mathbf{S}^+\|_F^2-c^{-1}\left(\frac{p}{m_{\underline{F}}^2(0)}-c\|(\boldsymbol{\Sigma}+m_{\underline{F}}(0)\mathbf{I})^{-1}\|_F^2\right)^{-1}\right|\to 0\tag{35}$$

$$1/p\left|\|\mathbf{S}^-\|_F^2-\frac{(1+c(c-1))}{c^2(c-1)^3}\frac{1}{p}\left(tr(\boldsymbol{\Sigma}^{-1})\right)^2-\frac{1}{(c(c-1))^2}\|\boldsymbol{\Sigma}^{-1}\|_F^2\right|\to 0\tag{36}$$

where $m_{\underline{F}}(0)$ satisfies asymptotically (approximately) the following equation

$$\frac{p}{m_{\underline{F}}(0)}=c\cdot tr(\boldsymbol{\Sigma}+m_{\underline{F}}(0)\mathbf{I})^{-1}.$$

9

The results of Corollary 1 show how the Frobenius norms of both inverses are connected with their population counterpart $||\mathbf{\Sigma}^{-1}||_F^2$. Namely, looking on the asymptotic behavior of $||\mathbf{S}^+||_F^2$ one can easily deduce that it is not possible to estimate $||\mathbf{\Sigma}^{-1}||_F^2$ consistently using the Moore-Penrose inverse because of the nonlinearity which is present in $(\mathbf{\Sigma} + m_{\underline{F}}(0)\mathbf{I})^{-1}$. On the other side, it is doable by reflexive generalized inverse $\mathbf{S}_n^-$. Indeed, using the proof of Theorem 2 one can find that

$$1/p \left| \text{tr}(\mathbf{S}_n^-) - \frac{1}{c(c-1)} \text{tr}(\mathbf{\Sigma}^{-1}) \right| \to 0\,,$$

which together with (36) implies that

$$1/p \left| (c(c-1))^2 \left[||\mathbf{S}^-||_F^2 - \left(\frac{1}{(c-1)} + c \right) \frac{1}{p} \left(\text{tr}(\mathbf{S}^-) \right)^2 \right] - ||\mathbf{\Sigma}^{-1}||_F^2 \right| \to 0 \qquad (37)$$

almost surely.

4   Numerical illustration

Using the results of the previous section, we present several numerical results to quantify the difference $||\mathbf{S}^- - \mathbf{S}^+||_F^2 = ||\mathbf{S}^-||_F^2 - ||\mathbf{S}^+||_F^2$ for some $\mathbf{\Sigma}$. In order to avoid the normalization $1/p$ we will use the normalized Frobenius loss (NFL) expressed as

$$NFL = \frac{||\mathbf{S}^- - \mathbf{S}^+||_F^2}{||\mathbf{S}^+||_F^2} = \frac{||\mathbf{S}^-||_F^2}{||\mathbf{S}^+||_F^2} - 1, \qquad (38)$$

which measures the difference between $\mathbf{S}^-$ and $\mathbf{S}^+$ normalized by the Frobenius norm of the Moore-Penrose inverse. The application of (38) with $\mathbf{S}^+$ and $\mathbf{S}^-$ as in (3) and (4) leads to the so called empirical NFL, while the usage of the results of Theorem 2 corresponds to the asymptotic NFL. The latter can be interpreted how much both Frobenius norms differ asymptotically.

In Figure 1 we present the results of a simulation study where the normalized Frobenius losses are computed for several values of the concentration ratio $c > 1$ as a function of dimension p. For the sake of illustration, we provide the results obtained for the samples generated from the multivariate normal distribution, while similar values were also obtained for other multivariate distributions. We set the mean vector to zero and, without loss of generality, use the diagonal covariance matrix $\mathbf{\Sigma}$ with 20% of eigenvalues equal to one, 40% equal to three and the rest equal to ten. The results for $c = 1.07$ confirm our expectations discussed after the proof of Theorem 1, namely there is no large difference between both norms: NFL is small and the empirical NFL converges quite fast to its asymptotic counterpart. By increasing c we observe that the NFL increases indicating that both norms deviate from each other. For example, in case $c = 2$ the asymptotic and empirical NFLs are close to 1.2 what means that the Frobenius norm

for reflexive inverse is more than a double of the Frobenius norm of the Moore-Penrose inverse. This observation becomes more severe for larger values of c. Moreover, for $c = 10$ we observe a bit slower convergence of the sample NFL to its asymptotic value.
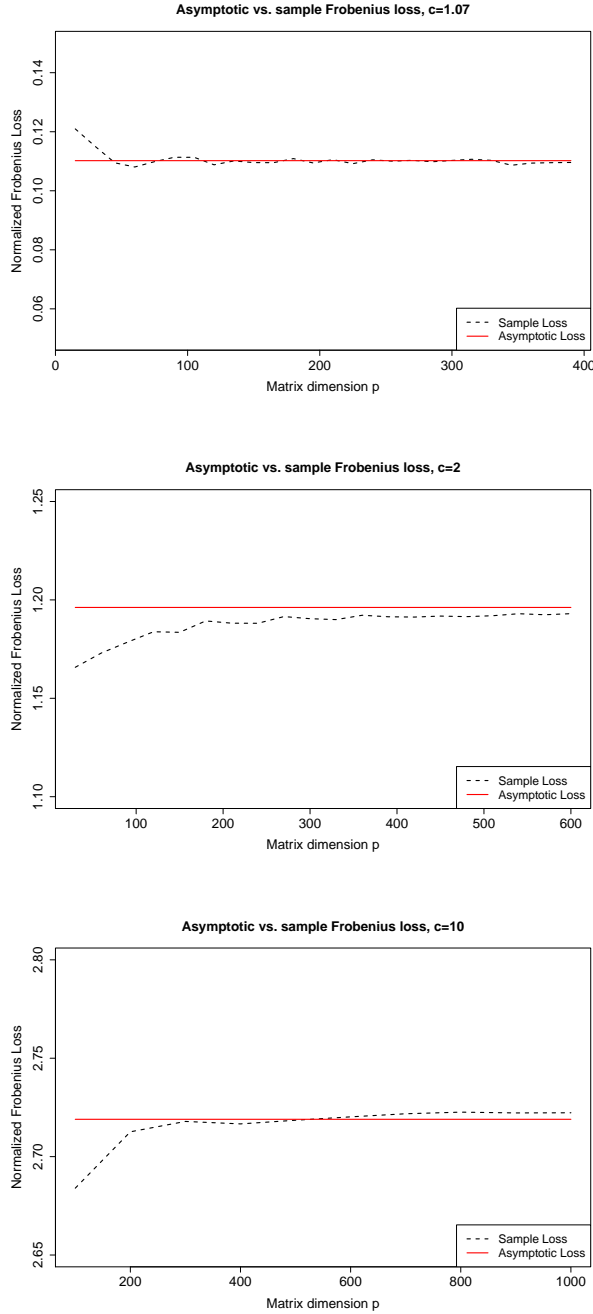

Figure 1: Normalized Frobenius losses $\frac{\|\mathbf{S}^- - \mathbf{S}^+\|_F^2}{\|\mathbf{S}^+\|_F^2}$ for several values of $c > 1$ as function of p.

The obtained findings indicate that the usage of $\mathbf{S}_n^+$ instead of $\mathbf{S}_n^-$ must be done with much care and are only reliable if the concentration ratio p/n is close to one. Otherwise, one can not expect neither a good approximation of functionals depending on the inverse

covariance matrix nor consistent estimators for them.

5 Summary

In many statistical applications the dimension of the data-generating process is larger than the sample size. However, one still needs to compute the inverse of the sample covariance matrix, which is present in many expressions. There are a plenty of ways how to define a generalized inverse with the Moore-Penrose inverse and reflexive generalized inverse matrices be the mostly used ones. While the former is easily computable and is unique, the latter can not be calculated from the data. On the other side, the behavior of the Moore-Penrose inverse in many high-dimensional applications is far away from a satisfactory one, whereas the reflexive inverse obeys very convenient statistical and asymptotic properties. Namely, the application of the reflexive inverse allows to estimate functionals of the precision matrix (the inverse of population covariance matrix) consistently in high-dimension.

In this work we study spectral properties of both inverses in detail. The almost sure limits of the Stieltjes transforms of their empirical distribution functions of eigenvalues (so-called limiting spectral distributions) is provided in the case the dimension p and the sample size n increase to infinity simultaneously such that their ratio p/n tends to a positive constant c greater than one. We discover that both limiting spectral distributions differ considerably and only coincide when c tends to one. The results are illustrated via the calculation of asymptotic normalized Frobenius loss of both inverse matrices. Finally, we justify the obtained theoretical findings via a simulation study.

References

[1] Bai, Z.D., Silverstein, J.W.: Spectral Analysis of Large Dimensional Random Matrices. Springer, New York (2010).

[2] Bai, Z.D., Zhou, W.: Large sample covariance matrices without independence structures in columns. Statistica Sinica **18**, 425–442 (2008)

[3] Belinschi, S.T., Bercovici, H.: A new approach to subordination results in free probability. Journal d'Analyse Mathématique **101**, 357–365 (2007)

[4] Bodnar, T., Dette, H., Parolya, N.: Spectral analysis of the Moore-Penrose inverse of a large dimensional sample covariance matrix. Journal of Multivariate Analysis **148**, 160–172 (2016)

[5] Bodnar, T., Dmytriv, S., Parolya, N., Schmid, W.: Tests for the weights of the

global minimum variance portfolio in a high-dimensional setting. IEEE Transactions on Signal Processing **67**, 4479–4493 (2019)

[6] Bodnar, T., Gupta, A.K., Parolya, N.: Direct shrinkage estimation of large dimensional precision matrix. Journal of Multivariate Analysis **146**, 223–236 (2016)

[7] Bodnar, T., Parolya, N. Schmid, W.: Estimation of the global minimum variance portfolio in high dimensions. European Journal of Operational Research **266**, 371–390 (2018)

[8] Bodnar, T., Okhrin, Y.: Properties of the singular, inverse and generalized inverse partitioned Wishart distributions. Journal of Multivariate Analysis **99**, 2389–2405 (2008)

[9] Friesen, O., Löwe, M., Stolz, M.: Gaussian uctuations for sample covariance matrices with dependent data. Journal of Multivariate Analysis **114**, 270–287 (2013)

[10] Holgersson, T., Karlsson, P., Stephan, A: A risk perspective of estimating portfolio weights of the global minimum-variance portfolio. AStA Advances in Statistical Analysis **104**, 59–80 (2020)

[11] Hui, J., Pan, G.: Limiting spectral distribution for large sample covariance matrices with m-dependent elements. Communications in Statistics – Theory and Methods **39**, 935–941 (2010)

[12] Imori, S., von Rosen, D.: On the mean and dispersion of the Moore-Penrose generalized inverse of a Wishart matrix. The Electronic Journal of Linear Algebra **36**, 124-133 (2020)

[13] Kollo, T., von Rosen, D.: Advanced Multivariate Statistics with Matrices. Springer, Dordrecht (2005).

[14] Silverstein, J.W., Bai, Z.D.: On the empirical distribution of eigenvalues of a class of large dimensional random matrices. Journal of Multivariate Analysis **54**, 175–192 (1995)

[15] Wang, G., Wei, Y., Qiao, S., Lin, P., Chen, Y.: Generalized Inverses: Theory and Computations. Springer, Singapore (2018).

[16] Wei, M., Yang, G., Yang, L.: The limiting spectral distribution for large sample covariance matrices with unbounded m-dependent entries. Communications in Statistics – Theory and Methods **45**, 6651–6662 (2016)