

On Weighted Multivariate Sign Functions: Supplementary Material

Subhabrata Majumdar* and Snigdhasu Chatterjee†

*School of Statistics,
University of Minnesota,
313 Ford Hall,
224 Church Street S.E.,
Minneapolis 55455, USA*
e-mail: subho@research.att.com; chatt019@umn.edu

Appendix A: Form of V_W

First observe that for F having covariance matrix $\Sigma = \Gamma\Lambda\Gamma^T$,

$$V_W = (\Gamma \otimes \Gamma)V_{W,\Lambda}(\Gamma \otimes \Gamma)^T \quad (1)$$

where $V_{W,\Lambda}$ is the covariance matrix of \mathbb{F}_Λ , the elliptic distribution with mean μ and covariance matrix Λ . Now,

$$\begin{aligned} V_{W,\Lambda} &= \mathbb{E} \left[\text{vec} \left\{ \frac{W^2(Z, \mathbb{F}_Z) \Lambda^{1/2} Z Z^T \Lambda^{1/2}}{Z^T \Lambda Z} - \tilde{\Lambda} \right\} \text{vec}^T \left\{ \frac{W^2(Z, \mathbb{F}_Z) \Lambda^{1/2} Z Z^T \Lambda^{1/2}}{Z^T \Lambda Z} - \tilde{\Lambda} \right\} \right] \\ &= \mathbb{E} \left[\text{vec} \left\{ W^2(Z, \mathbb{F}_Z) \mathbb{S}(\Lambda^{1/2} Z; \mathbf{0}) \right\} \text{vec}^T \left\{ W^2(Z, \mathbb{F}_Z) \mathbb{S}(\Lambda^{1/2} Z; \mathbf{0}) \right\} \right] \\ &\quad - \text{vec}(\tilde{\Lambda}) \text{vec}^T(\tilde{\Lambda}) \end{aligned}$$

The matrix $\text{vec}(\tilde{\Lambda}) \text{vec}^T(\tilde{\Lambda})$ consists of elements $\lambda_i \lambda_j$ at $(i, j)^{\text{th}}$ position of the $(i, j)^{\text{th}}$ block, and 0 otherwise. These positions correspond to variance and covariance components of on-diagonal elements. For the expectation matrix, all its elements are of the form $\mathbb{E}[\sqrt{\lambda_a \lambda_b \lambda_c \lambda_d} Z_a Z_b Z_c Z_d W^4(Z, \mathbb{F}_Z) / (Z^T \Lambda Z)^2]$, with $1 \leq a, b, c, d \leq p$. Since $W^4(Z, \mathbb{F}_Z) / (Z^T \Lambda Z)^2$ is even in Z , which has a spherically symmetric distribution, all such expectations will be 0 unless a, b, c, d are all equal or pairwise equal. Following a similar derivation for spatial sign covariance matrices in [2], we collect the non-zero elements and write the matrix of expectations:

$$(\mathbb{I}_{p^2} + \mathbb{K}_{p,p}) \left\{ \sum_{a=1}^p \sum_{b=1}^p \tilde{\gamma}_{ab} (e_a e_a^T \otimes e_b e_b^T) - \sum_{a=1}^p \tilde{\gamma}_{aa} (e_a e_a^T \otimes e_a e_a^T) \right\} + \sum_{a=1}^p \sum_{b=1}^p \tilde{\gamma}_{ab} (e_a e_b^T \otimes e_a e_b^T)$$

where $\mathbb{I}_k = (e_1, \dots, e_k)$, $\mathbb{K}_{m,n} = \sum_{i=1}^m \sum_{j=1}^n \mathbb{J}_{ij} \otimes \mathbb{J}_{ij}^T$ with $\mathbb{J}_{ij} \in \mathbb{R}^{m \times n}$ having 1 as (i, j) -th element and 0 elsewhere, and $\tilde{\gamma}_{mn} = \mathbb{E}[\lambda_m \lambda_n Z_m^2 Z_n^2 W^4(Z, \mathbb{F}_Z) / (Z^T \Lambda Z)^2]; 1 \leq m, n \leq p$.

*Currently at AT&T Labs Research

†Corresponding author

Putting everything together, denote by $\hat{\Lambda}$ the sample version of $\tilde{\Lambda}$, the weighted covariance matrix obtained from \mathbb{F}_Λ , i.e. $\hat{\Lambda} = \sum_{i=1}^n W_n^2(Z_i, \mathbb{F}_{Z,n}) \mathbb{S}(\Lambda^{1/2} Z_i; \hat{\mu}_n)/n$. Then the different types of elements in the matrix $\hat{\Lambda}$ are as given below ($1 \leq a, b, c, d \leq p$):

- Variance of on-diagonal elements

$$A\mathbb{V}(\sqrt{n}\hat{\Lambda}(a, a)) = \mathbb{E} \left[\frac{W^4(Z, \mathbb{F}_Z) \lambda_a^2 Z_a^4}{(Z^T \Lambda Z)^2} \right] - \tilde{\lambda}_a^2$$

- Variance of off-diagonal elements ($a \neq b$)

$$A\mathbb{V}(\sqrt{n}\hat{\Lambda}(a, b)) = \mathbb{E} \left[\frac{W^4(Z, \mathbb{F}_Z) \lambda_a \lambda_b Z_a^2 Z_b^2}{(Z^T \Lambda Z)^2} \right]$$

- Covariance of two on-diagonal elements ($a \neq b$)

$$A\mathbb{V}(\sqrt{n}\hat{\Lambda}(a, a), \sqrt{n}\hat{\Lambda}(b, b)) = \mathbb{E} \left[\frac{W^4(Z, \mathbb{F}_Z) \lambda_a \lambda_b Z_a^2 Z_b^2}{(Z^T \Lambda Z)^2} \right] - \tilde{\lambda}_a \tilde{\lambda}_b$$

- Covariance of two off-diagonal elements ($a \neq b, c \neq d$)

$$A\mathbb{V}(\sqrt{n}\hat{\Lambda}(a, b), \sqrt{n}\hat{\Lambda}(c, d)) = 0$$

- Covariance of one off-diagonal and one on-diagonal element ($a \neq b \neq c$)

$$A\mathbb{V}(\sqrt{n}\hat{\Lambda}(a, b), \sqrt{n}\hat{\Lambda}(c, c)) = 0$$

The above give all the elements of $V_{W,\Lambda}$. We plug these in (1) to recover V_W .

Appendix B: Proofs

Proof sketch of Corollary 2.4. We start with slightly modified versions of Lemmas A.4 and A.5 in [4]:

Lemma B.1. *Given that condition 1 in Corollary 2.4 holds, we have $\lambda_{\max}(\Psi_{1W}) \leq 2W_{\max} \frac{\lambda_{\max}(\Sigma)}{\text{Tr}(\Sigma)} (1 + o(1))$.*

Lemma B.2. *Define $\Psi_{3W} = \mathbb{E} \left[\frac{W^2(\epsilon, \mathbb{F}_\epsilon)}{|\epsilon|^2} (\mathbb{I}_p - S(\epsilon)S(\epsilon)^T) \right]$. Then $\lambda_{\max}(\Psi_{2W}) \leq W_{\max} \mathbb{E}(|\epsilon|^{-1})$ and $\lambda_{\max}(\Psi_{3W}) \leq (W_{\max})^2 \mathbb{E}(|\epsilon|^{-2})$. Further, if conditions 1 and 2 of Corollary 2.4 hold then $\lambda_{\min}(\Psi_{2W}) \geq \mathbb{E}(W(\epsilon, \mathbb{F}_\epsilon)/|\epsilon|)(1 + o(1))/\sqrt{3}$.*

The lemmas can be proved using similar steps as the proofs of the original lemmas in [4] and using the upper bound on the weight function. Corollary 2.4 is now proved by applying Corollary 2.2, plugging in upper bound of $\lambda_{\max}(\Psi_{1W})$ from Lemma B.1 and lower bound of $\lambda_{\max}(\Psi_{1W})$ from Lemma B.2 into the ARE expression. \square

Proof of Theorem 3.4. We suppose $G_n = (g_1, \dots, g_p)$, $L_n = \text{diag}(l_1, \dots, l_p)$. In spirit, this corollary is similar to Theorem 13.5.1 in [1]. We start with the following result, due to [3], allows us to obtain asymptotic joint distributions of eigenvectors and eigenvalues of $\hat{\hat{\Sigma}}$, provided we know the limiting distribution of $\hat{\hat{\Sigma}}$ itself:

Theorem B.3. *Let \mathbb{F}_Λ be defined as before, and \hat{C} be any positive definite symmetric $p \times p$ matrix such that at F_Λ the limiting distribution of $\sqrt{n} \text{vec}(\hat{C} - \Lambda)$ is a p^2 -variate (singular) normal distribution with mean zero. Write the spectral decomposition of \hat{C} as $\hat{C} = \hat{P}\hat{\Lambda}\hat{P}^T$. Then the limiting distributions of $\sqrt{n} \text{vec}(\hat{P} - \mathbb{I}_p)$ and $\sqrt{n} \text{vec}(\hat{\Lambda} - \Lambda)$ are multivariate (singular) normal and*

$$\sqrt{n} \text{vec}(\hat{C} - \Lambda) = [(\Lambda \otimes \mathbb{I}_p) - (\mathbb{I}_p \otimes \Lambda)] \sqrt{n} \text{vec}(\hat{P} - \mathbb{I}_p) + \sqrt{n} \text{vec}(\hat{\Lambda} - \Lambda) + o_P(1) \quad (2)$$

The first matrix picks only off-diagonal elements of the left-hand side and the second one only diagonal elements. We shall now use this as well as the form of the asymptotic covariance matrix of the vectorized $\hat{\hat{\Sigma}}$, i.e. V_W to obtain limiting variance and covariances of eigenvalues and eigenvectors.

Due to the decomposition (2) we have, for \mathbb{F}_Λ , the following relation between any off-diagonal element of $\hat{\Lambda}$ and the corresponding element in the estimate of eigenvectors, say $\hat{\hat{\Gamma}}_\Lambda$:

$$\sqrt{n} \hat{\hat{\gamma}}_{\Lambda, ij} = \sqrt{n} \frac{\hat{\hat{\Lambda}}(i, j)}{\hat{\lambda}_i - \hat{\lambda}_j}; \quad i \neq j$$

So that for eigenvector estimates of the original \mathbb{F} we have

$$\sqrt{n}(\hat{\gamma}_i - \gamma_i) = \sqrt{n}\Gamma(\hat{\hat{\gamma}}_{\Lambda, i} - e_i) = \sqrt{n} \left[\sum_{k=1; k \neq i}^p \hat{\hat{\gamma}}_{\Lambda, i, k} \gamma_k + (\hat{\hat{\gamma}}_{\Lambda, i, i} - 1) \gamma_i \right] \quad (3)$$

Now $\sqrt{n}(\hat{\hat{\gamma}}_{\Lambda, i, i} - 1) = o_P(1)$ and $A\mathbb{W}(\sqrt{n}\hat{\hat{\Lambda}}(i, k), \sqrt{n}\hat{\hat{\Lambda}}(i, l)) = 0$ for $k \neq l$, so the above equation implies

$$A\mathbb{W}(g_i) = A\mathbb{W}(\sqrt{n}(\hat{\gamma}_i - \gamma_i)) = \sum_{k=1; k \neq i}^p \frac{A\mathbb{W}(\sqrt{n}\hat{\hat{\Lambda}}(i, k))}{(\hat{\lambda}_i - \hat{\lambda}_k)^2} \gamma_k \gamma_k^T$$

For the covariance terms, from (3) we get, for $i \neq j$,

$$\begin{aligned}
 A\mathbb{V}(g_i, g_j) &= A\mathbb{V}(\sqrt{n}(\hat{\gamma}_i - \gamma_i), \sqrt{n}(\hat{\gamma}_j - \gamma_j)) \\
 &= A\mathbb{V}\left(\sum_{k=1; k \neq i}^p \sqrt{n}\hat{\gamma}_{\Lambda, ik}\gamma_k, \sum_{k=1; k \neq j}^p \sqrt{n}\hat{\gamma}_{\Lambda, jk}\gamma_k\right) \\
 &= A\mathbb{V}\left(\sqrt{n}\hat{\gamma}_{\Lambda, ik}\gamma_j, \sqrt{n}\hat{\gamma}_{\Lambda, ik}\gamma_i\right) \\
 &= -\frac{A\mathbb{V}(\sqrt{n}\hat{\Lambda}(i, j))}{(\tilde{\lambda}_i - \tilde{\lambda}_j)^2}\gamma_j\gamma_i^T
 \end{aligned}$$

The exact forms given in the statement of the corollary now follows from the Form of V_W in Section A.

For the on-diagonal elements of $\hat{\Lambda}$, using Theorem B.3 we have for the i^{th} eigenvalue of $\hat{\Lambda}$, say $\lambda_{\Lambda, i}$,

$$\sqrt{n}\hat{\lambda}_{\Lambda, i} = \sqrt{n}\hat{\Lambda}(i, i),$$

for $i = 1, \dots, p$. Hence

$$\begin{aligned}
 A\mathbb{V}(l_i) &= A\mathbb{V}(\sqrt{n}(\hat{\lambda}_{\Lambda, i} - \tilde{\lambda}_i)) \\
 &= A\mathbb{V}(\sqrt{n}(\hat{\lambda}_{\Lambda, i} - \tilde{\lambda}_{\Lambda, i})) \\
 &= A\mathbb{V}(\sqrt{n}\hat{\Lambda}(i, i))
 \end{aligned}$$

A similar derivation gives the expression for $A\mathbb{V}(l_i, l_j); i \neq j$. Finally, since the asymptotic covariance between an on-diagonal and an off-diagonal element of $\hat{\Lambda}$, it follows that the elements of G_n and diagonal elements of L_n are independent. \square

References

- [1] Anderson, T. (2009). *An Introduction to Multivariate Statistical Analysis*. Wiley, Hoboken, NJ, 3 edition.
- [2] Magyar, A. F. and Tyler, D. E. (2014). The asymptotic inadmissibility of the spatial sign covariance matrix for elliptically symmetric distributions. *Biometrika*, 101(3):673–688.
- [3] Taskinen, S., Koch, I., and Oja, H. (2012). Robustifying principal component analysis with spatial sign vectors. *Statistics & Probability Letters*, 82(4):765 – 774.
- [4] Wang, L., Peng, B., and Li, R. (2015). A high-dimensional nonparametric multivariate test for mean vector. *Journal of the American Statistical Association*, 110(512):1658 – 1669.