Appendix to "JADE for Tensor-Valued Observations"

Joni Virta, Bing Li, Klaus Nordhausen and Hannu Oja

November 16, 2017

The proofs of theorems

The proof of Theorem 1. Consider first the case $c = 1$ and the four terms in (11) separately fixing the choice of m. Denoting the first term of (11) by $\mathbf{B}^{ij}_{1,m}(\mathcal{X})$, then according to Lemma 5.4.1 in Virta et al. (2017) we have

$$\mathbf{B}^{ij}_{1,m}(\mathcal{X}_{st}) = \frac{\tau^4}{\rho_m}\mathbf{U}_m E\left[(\mathbf{u}^{(m)}_i)^T\mathbf{Z}_{(m)}\mathbf{Z}^T_{(m)}\mathbf{u}^{(m)}_j \cdot \mathbf{Z}_{(m)}\mathbf{Z}^T_{(m)}\right]\mathbf{U}^T_m,$$

where $\mathbf{Z}_{(m)}$ is the flattened matrix defined in Section 4 and $(\mathbf{u}^{(m)}_i)^T$ is the ith row of $\mathbf{U}_m$. Using the standard properties of expected value and independent random variables the (k, k') element of the inner expectation can be shown to be for $k \neq k'$ equal to $u^{(m)}_{ik}u^{(m)}_{jk'} + u^{(m)}_{jk}u^{(m)}_{ik'}$ and for $k = k'$ equal to $\delta_{ij}\rho_m + u^{(m)}_{ik}u^{(m)}_{jk}(\bar{\kappa}^{(m)}_k + 2)$. Using these to construct a matrix form for the expectation we have

$$\mathbf{B}^{ij}_{1,m}(\mathcal{X}_{st}) = \tau^4\mathbf{U}_m\left(\sum_{k=1}^{p} u^{(m)}_{ik}u^{(m)}_{jk}\bar{\kappa}^{(m)}_k\mathbf{E}^{kk}\right)\mathbf{U}^T_m + \tau^4\delta_{ij}\rho_m\mathbf{I} + \tau^4\mathbf{E}^{ij} + \tau^4\mathbf{E}^{ji}.$$

The second, third and fourth terms in (11) then serve to remove the extra constant terms above. That they indeed cancel one-by-one the final terms can easily be shown by examining

J. Virta and H. Oja are with the Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland (e-mail: joni.virta@utu.fi).

K. Nordhausen is with the Institute of Statistics & Mathematical Methods in Economics, Vienna University of Technology, Wiedner Hauptstr. 1040 Vienna, Austria.

B. Li is with the Department of Statistics, Pennsylvania State University, 326 Thomas Building, University Park, Pennsylvania 16802, USA.

them in the above manner using the independence of $\mathcal{X}$ and $\mathcal{X}^*$. This concludes the proof for $c = 1$ and the corresponding result for $c = 2$ can be proven in precisely the same manner. $\square$

The proof of Theorem 2. The consistency of the TJADE estimator is proven similarly as the consistency of the TFOBI estimator in the proof of Theorem 5.2.1 in Virta et al. (2017).

In the following we assume that $r = 2$ and we are interested in the asymptotical behavior of the 1-mode unmixing matrix. As discussed in Section 4, for the general case of arbitrary r and m-mode unmixing matrix, it suffices to m-flatten the tensor and replace in the following $\hat{\boldsymbol{\Sigma}}_1^{-1/2}$ with $\hat{\boldsymbol{\Sigma}}_m^{-1/2}$, $\hat{\boldsymbol{\Sigma}}_2^{-1/2}$ with $\hat{\boldsymbol{\Sigma}}_{m+1}^{-1/2} \otimes \cdots \otimes \hat{\boldsymbol{\Sigma}}_r^{-1/2} \otimes \hat{\boldsymbol{\Sigma}}_1^{-1/2} \otimes \cdots \otimes \hat{\boldsymbol{\Sigma}}_{m-1}^{-1/2}$, p_2 with ρ_m and use the corresponding row mean quantities.

For the asymptotic expressions of the diagonal elements of $\sqrt{n}(\hat{\boldsymbol{\Phi}} - \mathbf{I})$ it suffices to use the same arguments as in the proof of Theorem 5.2.1 in Virta et al. (2017) and for the off-diagonal elements we aim to use Lemma 2 from Miettinen et al. (2015).

But first, define the *symmetric* standardization functionals $\hat{\mathbf{L}} = (\hat{l}_{kk'}) := \hat{\boldsymbol{\Sigma}}_1^{-1/2}$ and $\hat{\mathbf{R}} = (\hat{r}_{ll'}) := \hat{\boldsymbol{\Sigma}}_2^{-1/2}$ giving the standardized identity-mixed observations as $\mathbf{X}_{st,i} = \hat{\mathbf{L}}\tilde{\mathbf{Z}}_i\hat{\mathbf{R}}^T$, where $\tilde{\mathbf{Z}}_i = \mathbf{Z}_i - \bar{\mathbf{Z}}$. We then have

$$\sqrt{n}(\hat{l}_{kk'} - \delta_{kk'}) = -(1/2)\sqrt{n}(\hat{s}_{kk'} - \delta_{kk'}) + o_P(1),$$

see Virta et al. (2017), and as simple moment-based estimators we have both $\sqrt{n}(\hat{\mathbf{L}} - \mathbf{I}) = O_P(1)$ and $\sqrt{n}(\hat{\mathbf{R}} - \mathbf{I}) = O_P(1)$, regardless of whether we really have $r = 2$ or use flattened tensors of higher order.

Assume then first that $c = 1$. The matrices $\hat{\mathbf{C}}_{1,1}^{kk'}$, $k, k' \in \{1, \ldots, p\}$, in (14) to be simultaneously diagonalized satisfy $\hat{\mathbf{C}}^{kk'} := \hat{\mathbf{C}}_{1,1}^{kk'} \rightarrow_P \mathbf{C}_{1,1}^{kk'}(\mathbf{Z}_i) = \delta_{kk'}\bar{\kappa}_k^{(1)}\mathbf{E}^{kk}$. In the view of Lemma 2 in Miettinen et al. (2015) this means that the only matrices $\mathbf{C}_{1,1}^{rs}(\mathbf{Z}_i)$, $r, s \in \{1, \ldots, p\}$, having non-zero kth or k'th diagonal elements are $\mathbf{C}_{1,1}^{kk}(\mathbf{Z}_i)$ and $\mathbf{C}_{1,1}^{k'k'}(\mathbf{Z}_i)$, respectively, yielding the following form for the (k, k'), $k \neq k'$, element of $\hat{\mathbf{U}} := \hat{\mathbf{U}}_1^T$ estimated by (16).

$$\sqrt{n}\hat{u}_{kk'} = \frac{\bar{\kappa}_k^{(1)}\sqrt{n}\hat{\mathbf{C}}_{kk'}^{kk} - \bar{\kappa}_{k'}^{(1)}\sqrt{n}\hat{\mathbf{C}}_{kk'}^{k'k'}}{(\bar{\kappa}_k^{(1)})^2 + (\bar{\kappa}_{k'}^{(1)})^2} + o_P(1),$$

where $\hat{\mathbf{C}}_{rs}^{kk}$ is the (r, s) element of $\hat{\mathbf{C}}^{kk}$. The above expression then together with the (k, k'), $k \neq k'$, element of the left standardization matrix $\hat{\mathbf{L}}$ gives an asymptotic expression for the off-diagonal elements of the estimated left TJADE matrix, see Virta et al. (2017):

$$\sqrt{n}\hat{\phi}_{kk'} = \sqrt{n}\hat{u}_{kk'} + \sqrt{n}\hat{l}_{kk'} + o_P(1), \tag{19}$$

reducing the problem of finding the asymptotics of TJADE into the task of finding the asymptotic behaviors of $\sqrt{n}\hat{\mathbf{C}}^{kk}_{kk'}$ and $\sqrt{n}\hat{\mathbf{C}}^{k'k'}_{kk'}$. Dropping the subscripts for clarity, note that $\hat{\mathbf{C}}^{aa} = \hat{\mathbf{B}}^{aa} - \hat{\mathbf{\Xi}}(p_2\mathbf{I} + 2\mathbf{E}^{aa})\hat{\mathbf{\Xi}}^T$ and starting from $\hat{\mathbf{B}}^{aa}$ write it out as

$$\hat{\mathbf{B}}^{aa} = \frac{1}{p_2 n}\sum_{i=1}^{n}(\hat{\mathbf{L}}_a^T\tilde{\mathbf{Z}}_i\hat{\mathbf{R}}^*\tilde{\mathbf{Z}}_i^T\hat{\mathbf{L}}_a)\cdot\hat{\mathbf{L}}\tilde{\mathbf{Z}}_i\hat{\mathbf{R}}^*\tilde{\mathbf{Z}}_i^T\hat{\mathbf{L}}^T,$$

where $\hat{\mathbf{L}}_a^T$ is the ath row of $\hat{\mathbf{L}}$ and $\hat{\mathbf{R}}^* := \hat{\mathbf{R}}^T\hat{\mathbf{R}}$. An arbitrary off-diagonal element of $\sqrt{n}(\hat{\mathbf{B}}^{aa} - \mathbf{B}^{aa}(\mathbf{Z}_i))$ then has after the matrix multiplication the form

$$\sqrt{n}\hat{\mathbf{B}}^{aa}_{kk'} = \frac{1}{p_2 n}\sum_{defgstuv}\sqrt{n}\hat{r}^*_{ef}\hat{r}^*_{tu}\hat{l}_{ad}\hat{l}_{ag}\hat{l}_{ks}\hat{l}_{k'v}\hat{H}_{de,gf,st,vu}, \tag{20}$$

where $\hat{H}_{de,gf,st,vu} = (1/n)\sum_{i=1}^n \tilde{z}_{i,de}\tilde{z}_{i,gf}\tilde{z}_{i,st}\tilde{z}_{i,vu} \to_P E(z_{i,de}z_{i,gf}z_{i,st}z_{i,vu})$. Next we expand the multiplicands $\hat{r}^*_{..}$ and $\hat{l}_{..}$ in (20) one-by-one such as $\hat{l}_{ab} = (\hat{l}_{ab} - \delta_{ab}) + \delta_{ab}$, the first term of which is $O_P(1)$ when combined with $\sqrt{n}$ allowing the use of Slutsky's theorem to the whole multiple sum and the second term of which produces an expression like (20) only with one summation index less.

Starting from left this process then produces the terms $o_P(1)$; $o_P(1)$; $\delta_{ak}\sqrt{n}\hat{l}_{kk'} + \delta_{ak'}\sqrt{n}\hat{l}_{k'k} + o_P(1)$; $\delta_{ak}\sqrt{n}\hat{l}_{kk'} + \delta_{ak'}\sqrt{n}\hat{l}_{k'k} + o_P(1)$; $\delta_{ak'}(\bar{\kappa}^{(1)}_{k'} + p_2 + 2)\sqrt{n}\hat{l}_{kk'} + (1-\delta_{ak'})p_2\sqrt{n}\hat{l}_{kk'} + o_P(1)$ and $\delta_{ak}(\bar{\kappa}^{(1)}_k + p_2 + 2)\sqrt{n}\hat{l}_{k'k} + (1-\delta_{ak})p_2\sqrt{n}\hat{l}_{k'k} + o_P(1)$ finally leaving us with the expression

$$\frac{1}{p_2}\sum_{et}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{z}^2_{i,ae}\tilde{z}_{i,kt}\tilde{z}_{i,k't} + o_P(1). \tag{21}$$

Substituting now either $a = k$ or $a = k'$, expanding $\tilde{z}_{i,ab} = z_{i,ab} - \bar{z}_{ab}$ and using the quantities defined in Section 4 the expression in (21) gets the forms $\sqrt{n}\hat{r}_{kk'} + \sqrt{n}\hat{q}_{kk'} + o_P(1)$ and $\sqrt{n}\hat{r}_{k'k} + \sqrt{n}\hat{q}_{k'k} + o_P(1)$, respectively.

Using the above, e.g. $\sqrt{n}\hat{\mathbf{B}}^{kk}_{kk'}$ gets the form

$$(p_2+2)\sqrt{n}\hat{l}_{kk'} + (\bar{\kappa}^{(1)}_k + p_2 + 2)\sqrt{n}\hat{l}_{k'k} + \sqrt{n}\hat{r}_{kk'} + \sqrt{n}\hat{q}_{kk'} + o_P(1).$$

For the asymptotic behavior of the remaining term $\hat{\mathbf{\Xi}}(p_2\mathbf{I} + 2\mathbf{E}^{aa})\hat{\mathbf{\Xi}}^T$ one can first use techniques similar to the above to show for $\hat{\mathbf{\Xi}} = (\hat{\xi}_{kk'})$ that $\sqrt{n}(\hat{\xi}_{kk'} - \delta_{kk'}) = o_P(1)$ for $k \neq k'$.

Consequently an arbitrary off-diagonal element of $\sqrt{n}(\hat{\mathbf{\Xi}}(p_2\mathbf{I} + 2\mathbf{E}^{aa})\hat{\mathbf{\Xi}}^T - p_2\mathbf{I} - 2\mathbf{E}^{aa})$ is also $o_P(1)$ implying that the term actually contributes nothing to the asymptotic variances of the estimator. Thus $\sqrt{n}\hat{\mathbf{C}}^{aa}_{kk'} = \sqrt{n}\hat{\mathbf{B}}^{aa}_{kk'} + o_P(1)$ and the result of Theorem 2 is obtained by plugging everything in into (19) and using the fact that the standardization functionals are symmetric. The asymptotic variances of Corollary 1 are then straightforward to obtain, e.g. using the table of covariances in the proof of Theorem 5.2.1 in Virta et al. (2017).

Although the starting expressions for $c = 1$ and $c = 2$ are different the final expressions for both $\sqrt{n}\hat{\mathbf{C}}^{kk}_{kk'}$ and $\sqrt{n}\hat{\mathbf{C}}^{k'k'}_{kk'}$ actually match exactly. The corresponding proof for $c = 2$ is obtained in exactly likewise manner, expanding the terms suitably and using Slustky's theorem and is thus omitted here. $\square$

Acknowledgments

The work of Virta and Oja was supported by the Academy of Finland Grant 268703. The work of Li was supported in part by the National Science Foundation Grants DMS-1407537 and DMS-1713078. The work of Nordhausen was partly supported by the Academy of Finland Grant 268703 and CRoNoS COST Action IC1408.

References

Miettinen, J., S. Taskinen, K. Nordhausen, and H. Oja (2015). Fourth moments and independent component analysis. *Statistical Science 30*(3), 372–390.

Virta, J., B. Li, K. Nordhausen, and H. Oja (2017). Independent component analysis for tensor-valued data. *Journal of Multivariate Analysis 162*, 172–192.