1 Appendix: Statistical Model and Methods

1.1 Multivariate Repeated Measures Design

The analyses have been performed using a multivariate Repeated Measures design as implemented in the R-package **MANOVA.RM** (Friedrich et al., 2019a,b). The model is applicable for non-normal data with covariance heterogeneity and the implemented bootstrap approaches provide an accurate type-I error rate control even for small samples. The methods are briefly described in the following:

In order to establish the general model, let

$$\boldsymbol{X}_{ik} \;= (\boldsymbol{X}'_{ik1}, \ldots, \boldsymbol{X}'_{ikt})' = (X_{ik11}, \ldots, X_{iktd})'$$

denote independent random vectors for group $i = 1, \ldots, a$ (between-subjects factor) and individual $k = 1, \ldots, n_i$. For each individual, we measure d-variate observations $X_{iks1}, \ldots, X_{iksd}$ for $s = 1, \ldots, t$ repeated measurements (levels of the within-subjects factor). Moreover, we assume existence of the expectation vectors $\boldsymbol{\mu}_{is} = (\mu_{is1}, \ldots, \mu_{isd})' = \mathbb{E}(\boldsymbol{X}_{iks})$ and the group-specific covariance matrices $\boldsymbol{V}_i = \mathrm{Cov}(\boldsymbol{X}_{i1}) \in \mathbb{R}^{td \times td}$, which may differ between groups. Note that we do not assume any special structure for $\boldsymbol{V}_i$. A more complex factorial structure can be incorporated in this model by splitting up indices; see, e.g., Konietschke et al. (2015). For convenience, we aggregate the individual vectors into $\boldsymbol{X} = (\boldsymbol{X}'_{111}, \ldots, \boldsymbol{X}'_{an_at})'$ as well as $\boldsymbol{\mu} = (\boldsymbol{\mu}'_{11}, \ldots, \boldsymbol{\mu}'_{at})' \in \mathbb{R}^{adt}$.

An estimator for $\boldsymbol{\mu}$ is given by the vector of pooled group means $\overline{\boldsymbol{X}}_{i\cdot s} = (\boldsymbol{X}_{i1s} + \cdots + \boldsymbol{X}_{in_is})/n_i$ for all $i = 1, \ldots, a,\ s = 1, \ldots, t$, which we denote by $\overline{\boldsymbol{X}}_{\cdot} = (\overline{\boldsymbol{X}}'_{1\cdot 1}, \ldots, \overline{\boldsymbol{X}}'_{a\cdot t})'$. Moreover, the covariance matrix of $\sqrt{N}\,\overline{\boldsymbol{X}}_{\cdot}$ is given by

$$\boldsymbol{\Sigma}_N = \mathrm{Cov}(\sqrt{N}\,\overline{\boldsymbol{X}}_{\cdot}) = \mathrm{diag}\left(N\boldsymbol{V}_i/n_i : 1 \leq i \leq a\right),$$

where $N = n_1 + \cdots + n_a$ is the total number of subjects and the group-specific covariance matrices $\boldsymbol{V}_i$ are estimated by the empirical covariance matrices, resulting in $\widehat{\boldsymbol{\Sigma}}_N = \mathrm{diag}(N\widehat{\boldsymbol{V}}_i/n_i : 1 \leq i \leq a)$.

1.2 Hypotheses

Now, null hypotheses can be formulated in terms of the mean vector as $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$, where $\boldsymbol{T}$ is a suitable contrast matrix. More precisely, the null hypotheses of interest about main and interaction effects in the above model can be tested by the following hypotheses matrices:

- No effect of the between-subjects factor A: $H_0(A) : \boldsymbol{T}_A = \boldsymbol{P}_a \otimes \frac{1}{t}\boldsymbol{J}_t \otimes \boldsymbol{I}_d$
- No effect of the within-subjects factor T: $H_0(T) : \boldsymbol{T}_T = \frac{1}{a}\boldsymbol{J}_a \otimes \boldsymbol{P}_t \otimes \boldsymbol{I}_d$

- No interaction effect between A and T: $H_0(AT) : \boldsymbol{T}_{AT} = \boldsymbol{P}_a \otimes \boldsymbol{P}_t \otimes \boldsymbol{I}_d$

Here, we denote by $\boldsymbol{I}_d$ the d-dimensional unit matrix and by $\boldsymbol{J}_d$ the $d \times d$ matrix of 1s, i.e., $\boldsymbol{J}_d = \mathbf{1}_d \mathbf{1}_d'$, where $\mathbf{1}_d = (1, \ldots, 1)'$ is the d-dimensional column vector of 1s. Furthermore, $\boldsymbol{P}_d = \boldsymbol{I}_d - d^{-1}\boldsymbol{J}_d$ denotes the d-dimensional centering matrix and $\otimes$ denotes the Kronecker product.

1.3 Test Statistics

A commonly used test statistic for multivariate data is the Wald-type statistic (WTS)

$$T_N = N\overline{\boldsymbol{X}}_{\cdot}^{\top}\boldsymbol{T}(\boldsymbol{T}\widehat{\boldsymbol{\Sigma}}_N\boldsymbol{T})^{+}\boldsymbol{T}\overline{\boldsymbol{X}}_{\cdot}, \tag{1}$$

which requires the additional assumption $\boldsymbol{V}_i > 0$ for all $i \in \{1, \ldots, a\}$.

Under $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \mathbf{0}$, the WTS has asymptotically a χ^2_f distribution with $f = \operatorname{rank}(\boldsymbol{T})$ degrees of freedom. However, large sample sizes are necessary to maintain a pre-assigned level α using quantiles of the limiting χ^2 distribution. Thus, bootstrap approaches are used to improve its small sample behavior.

In situations with singular covariance matrices, the WTS cannot be applied. Thus, we additionally consider the modified ANOVA-type statistic (MATS) proposed by Friedrich and Pauly (2018), which is given as:

$$Q_N = N\overline{\boldsymbol{X}}_{\cdot}^{\top}\boldsymbol{T}(\boldsymbol{T}\widehat{\boldsymbol{D}}_N\boldsymbol{T})^{+}\boldsymbol{T}\overline{\boldsymbol{X}}_{\cdot}. \tag{2}$$

Here, $\widehat{\boldsymbol{D}}_N = \operatorname{diag}(N/n_i \cdot \widehat{\sigma}^2_{isj})$ for all $i = 1, \ldots, a$, $s = 1, \ldots, t$ and $j = 1, \ldots, d$, where $\widehat{\sigma}^2_{isj}$ is the empirical variance of component j in group i at time point s. The MATS is invariant under scale transformations of the data for null hypotheses of the form $\boldsymbol{T} = \boldsymbol{M} \otimes \boldsymbol{I}_d$, i.e., hypotheses as formulated above.

Since the asymptotic distribution of the MATS depends on the unknown variances σ^2_{isj}, inference is based on a bootstrap approach, see Friedrich and Pauly (2018) for more details and formal proofs of the validity of the proposed approaches.

Contrary to classical multivariate repeated measures analyses, our approach incorporates general heteroscedastic designs and allows for singular covariance matrices and non-normal data. Moreover, the proposed MATS statistic is invariant under linear transformations of the response variables for the null hypotheses of interest. Inference is based on p-values obtained by a parametric bootstrap approach, which has shown accurate type-I error rate control even for small sample sizes. The proposed tests along with the calculation of confidence regions and post-hoc comparisons are implemented in the R package `MANOVA.RM`.

References

Friedrich, S., Konietschke, F., and Pauly, M. (2019a). *MANOVA.RM: Resampling-Based Analysis of Multivariate Data and Repeated Measures Designs*. R package version 0.4.1.

Friedrich, S., Konietschke, F., and Pauly, M. (2019b). Resampling-Based Analysis of Multivariate Data and Repeated Measures Designs with the R Package MANOVA.RM. *The RJournal*.

Friedrich, S. and Pauly, M. (2018). MATS: Inference for potentially singular and heteroscedastic MANOVA. *Journal of Multivariate Analysis*, 165:166–179.

Konietschke, F., Bathke, A., Harrar, S., and Pauly, M. (2015). Parametric and nonparametric bootstrap methods for general MANOVA. *Journal of Multivariate Analysis*, 140:291–301.