

Approximate Distributions of Quadratic Forms in High-Dimensional Repeated-Measures Designs

Edgar Brunner, Benjamin Becker, and Carola Werner

Department Medizinische Statistik
Georg-August-Universität Göttingen

1 Introduction

Repeated measures are data which are observed repeatedly on the same subjects. Such data appear in many medical or biological trials and are called *high-dimensional* if the dimension d (i.e. the number of repeated measures per subject) is larger (or even much larger) than the number n of the independent subjects on which they are observed. Most of the classical procedures require that $n > d$. If this is not the case then they break down or cannot even be computed.

For example, Wald-type statistics can no longer be computed since they require an inverse of the sample covariance matrix which becomes singular if $n < d$. The ANOVA-type statistic (ATS), first mentioned by Box (1954) and further developed by Geisser and Greenhouse (1958) and Greenhouse and Geisser (1959), however, can be computed. But it may be noted that a test based on this statistic using a plug-in estimator of the sample covariance matrix becomes conservative if $n < d$.

Meanwhile a plethora of papers considers procedures for high-dimensional data. The ideas underlying the procedures developed may be classified as

- (1) approximations for fixed n and fixed (but arbitrary large) d assuming a multivariate normal distribution but admitting any type of covariance matrices, so-called unstructured covariance matrices,
- (2) procedures assuming $n, d \rightarrow \infty$ while $d/n \rightarrow \kappa \in (0, 1)$ - assuming or not a multivariate normal distribution,
- (3) procedures assuming n fixed while $d \rightarrow \infty$ with different assumptions on the structure of the covariance matrix and the underlying distribution of the repeated measures.

In this technical report we will focus on the first idea, i.e. we assume a multivariate normal distribution and want to provide approximations based on the ANOVA-type statistic considered by Box (1954). We note that it is our intention to provide approximations to the distribution of the ATS while the quality of the approximation is uniform with respect to the dimension d .

¹First draft (German): October 2004, update (English): November 2009

2 The ANOVA-Type Statistic

2.1 Box-Approximation

We consider the simple repeated measures model and assume that

$$\mathbf{X}_k = (X_{k1}, \dots, X_{kd})' \sim N(\boldsymbol{\mu}, \mathbf{V}), \quad k = 1, \dots, n, \quad (2.1)$$

are independent vectors representing the subjects. Let \mathbf{H} denote an appropriate contrast matrix for testing the hypothesis $H_0 : \mathbf{H}\boldsymbol{\mu} = \mathbf{0}$ and let $\mathbf{T} = \mathbf{H}'(\mathbf{H}\mathbf{H}')^{-1}\mathbf{H}$ denote the unique projection matrix derived from \mathbf{H} . We allow for any suitable factorial structure of the repeated measures by partitioning the index $s = 1, \dots, d$ in sub-indices, for example, $s_1 = 1, \dots, b$ and $s_2 = 1, \dots, t$ such that $d = bt$. One special example may be that the d repeated measures are obtained at t time points under b consecutive treatments or conditions.

We consider the projection $\mathbf{Y}_k = \mathbf{T}\mathbf{X}_k \sim N(\mathbf{T}\boldsymbol{\mu}, \mathbf{S})$, where $\mathbf{S} = \mathbf{T}\mathbf{V}\mathbf{T}$ and $\mathbf{T}\boldsymbol{\mu} = \mathbf{0}$ under the hypothesis. To test the hypothesis $H_0 : \mathbf{H}\boldsymbol{\mu} = \mathbf{0} \iff \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$, we consider the statistic $Q_n = n \bar{\mathbf{Y}}' \bar{\mathbf{Y}} = n \bar{\mathbf{X}}' \mathbf{T} \bar{\mathbf{X}}$, where $\bar{\mathbf{Y}} = \frac{1}{n} \sum_{k=1}^n \mathbf{Y}_k$ and $\bar{\mathbf{X}} = \frac{1}{n} \sum_{k=1}^n \mathbf{X}_k$ denote the means. Note that $\sqrt{n} \bar{\mathbf{Y}} \sim \mathcal{N}(\mathbf{0}, \mathbf{S})$ under H_0 . Then, using the representation theorem of quadratic forms it follows under $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$ that

$$Q_n = n \bar{\mathbf{Y}}' \bar{\mathbf{Y}} = n \bar{\mathbf{X}}' \mathbf{T} \bar{\mathbf{X}} = \sum_{i=1}^d \lambda_i C_i, \quad (2.2)$$

where $C_i \sim \chi_1^2$ are independent and the λ_i are the eigenvalues of \mathbf{S} . Based on an idea of Patniak (1949), Box (1954) suggested to approximate the distribution of $\sum_{i=1}^d \lambda_i C_i$ by a scaled χ^2 -distribution $g \cdot \chi_f^2$ such that the first two moments of $\sum_{i=1}^d \lambda_i C_i$ and $g \cdot \chi_f^2$ coincide.

$$\begin{aligned} g f = E(g \chi_f^2) &= E\left(\sum_{i=1}^d \lambda_i C_i\right) = \sum_{i=1}^d \lambda_i = \text{tr}(\mathbf{S}) \\ 2 g^2 f = \text{Var}(g \chi_f^2) &= \text{Var}\left(\sum_{i=1}^d \lambda_i C_i\right) = 2 \sum_{i=1}^d \lambda_i^2 = 2 \text{tr}(\mathbf{S}^2). \end{aligned}$$

Straightforward computation leads to

$$\frac{Q_n}{\text{tr}(\mathbf{S})} = \frac{n \bar{\mathbf{X}}' \mathbf{T} \bar{\mathbf{X}}}{\text{tr}(\mathbf{T}\mathbf{V})} \dot{\sim} \chi_f^2 / f = F(f, \infty), \quad (2.3)$$

where $f = [\text{tr}(\mathbf{S})]^2 / \text{tr}(\mathbf{S}^2)$. In practice, however, \mathbf{S} is unknown and must be estimated from the data. It is important to note that two quantities are involved in the estimation of \mathbf{S} , namely the sample size n and the number d of the repeated measures. This shall be considered in detail in the next subsection.

2.2 Consistency of Estimators

We consider an array of estimators $\widehat{\theta}_{n,d}$ of a quantity $\theta_{n,d}$ which becomes larger with increasing d . Thus we consider the ratio $\widehat{\theta}_{n,d}/\theta_{n,d}$. Then it follows from Tschebychev's inequality

$$\begin{aligned} P\left(\left|\widehat{\theta}_{n,d}/\theta_{n,d} - 1\right| > \varepsilon\right) &\leq \frac{1}{\varepsilon^2} E\left(\left|\widehat{\theta}_{n,d}/\theta_{n,d} - 1\right|^2\right) \\ &\leq \frac{1}{\varepsilon^2} E\left(\left|\widehat{\theta}_{n,d}/\theta_{n,d} - E\left(\widehat{\theta}_{n,d}/\theta_{n,d}\right) + E\left(\widehat{\theta}_{n,d}/\theta_{n,d}\right) - 1\right|^2\right) \\ &\leq \frac{2}{\varepsilon^2} \left\{ \text{Var}\left(\widehat{\theta}_{n,d}/\theta_{n,d}\right) + \left[E\left(\widehat{\theta}_{n,d}/\theta_{n,d}\right) - 1\right]^2 \right\}. \end{aligned}$$

We note that the right-hand side is uniformly bounded in d , if $E(\widehat{\theta}_{n,d}) = \theta_{n,d}$ and $\text{Var}(\widehat{\theta}_{n,d}/\theta_{n,d}) \leq K_n$, where K_n denotes a sequence converging to 0 which is uniformly bounded with respect to d . This leads to the following definition.

DEFINITION 2.1 *An array of estimators $\widehat{\theta}_{n,d}$ of a quantity $\theta_{n,d}$ is called dimensionally stable if*

1. $\left|E(\widehat{\theta}_{n,d}/\theta_{n,d}) - 1\right| \leq D_n$,
2. $\text{Var}(\widehat{\theta}_{n,d}/\theta_{n,d}) \leq K_n$,

where $D_n \rightarrow 0$ and $K_n \rightarrow 0$ are uniformly bounded with respect to d .

As an example, consider the sample covariance matrix

$$\widehat{\mathbf{S}}_n = \frac{1}{n-1} \sum_{k=1}^n (\mathbf{Y}_k - \bar{\mathbf{Y}})(\mathbf{Y}_k - \bar{\mathbf{Y}})',$$

where $\bar{\mathbf{Y}} = \frac{1}{n} \sum_{k=1}^n \mathbf{Y}_k$ denotes the mean vector of the observations \mathbf{Y}_k , $k = 1, \dots, n$. Then it is easily seen that the plug-in estimator $\text{tr}(\widehat{\mathbf{S}}_n^2)$ of $\text{tr}(\mathbf{S}^2)$ is not dimensionally stable since it is biased and the bias increases with d . Therefore it is our aim to derive unbiased estimators of the traces involved in (2.3) and then show that their variances fulfill the requirement (2) of Definition 2.1.

We note that Bai and Saranadasa (1996) have derived an unbiased estimator of $\text{tr}(\mathbf{S}^2)$ using the sample covariance matrix $\widehat{\mathbf{S}}_n$ and showed its ratio consistency, i.e. if $n, d \rightarrow \infty$ while $d/n \rightarrow \kappa \in (0, 1)$. As we intend, however, to derive an approximation for fixed n and fixed d , we will derive other unbiased estimators of the quantities $\text{tr}(\mathbf{S})$, $[\text{tr}(\mathbf{S})]^2$, and $\text{tr}(\mathbf{S}^2)$ which are dimensionally stable. This will be worked out in the next section.

3 Unbiased Estimators

3.1 Derivation of the estimators

In the sequel we will derive unbiased estimators of $[tr(\mathbf{TV})]^2$ and $tr[(\mathbf{TV})^2]$. Numerator and denominator of $f = [tr(\mathbf{S})]^2 / tr(\mathbf{S}^2)$ will be separately estimated consistently and then the ratio is taken as an estimator of f . The bias of this estimator is obtained from a Taylor expansion.

For Model (2.1) we will provide procedures for testing the linear hypothesis $H_0 : \mathbf{H}\boldsymbol{\mu} = \mathbf{0}$, where \mathbf{H} denotes a suitable hypothesis matrix. As we will need the properties that \mathbf{H} is symmetric and idempotent, we will equivalently work with the unique projection matrix $\mathbf{T} = \mathbf{H}'(\mathbf{H}\mathbf{H}')^{-1}\mathbf{H}$ and note that $\mathbf{H}\boldsymbol{\mu} = \mathbf{0} \iff \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$. For convenience, let $\mathbf{Y}_k = \mathbf{T}\mathbf{X}_k$ denote the projection of the observation vectors and let $\mathbf{S}_k = \mathbf{Y}_k\mathbf{Y}_k'$. Then, under $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$, it follows that $E(\mathbf{Y}_k) = \mathbf{0}$ and it is easily seen that

$$\tilde{\mathbf{S}}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{S}_k \quad (3.4)$$

is an unbiased estimator of $\mathbf{S} = \mathbf{T}\mathbf{V}\mathbf{T}$ under H_0 . Using the invariance of the trace under cyclic permutations, a natural estimator of $tr(\mathbf{TV}) = tr(\mathbf{S})$ is given by

$$B_0 = \frac{1}{n} \sum_{k=1}^n A_k = \bar{A}, \quad (3.5)$$

where $A_k = \mathbf{Y}_k'\mathbf{Y}_k$. The estimator B_0^2 , however, is a biased estimator of $[tr(\mathbf{TV})]^2$ and the bias τ^2 follows from

$$\begin{aligned} E\left([tr(\tilde{\mathbf{S}}_n)]^2\right) &= Var(tr(\tilde{\mathbf{S}}_n)) + \left(E[tr(\tilde{\mathbf{S}}_n)]\right)^2 \\ &= \underbrace{Var(tr(\tilde{\mathbf{S}}_n))}_{:= \tau^2} + [tr(\mathbf{S})]^2. \end{aligned} \quad (3.6)$$

Since by independence, $\tau^2 = Var\left(tr(\tilde{\mathbf{S}}_n)\right) = Var\left(\frac{1}{n} \sum_{k=1}^n A_k\right) = \frac{1}{n} Var(A_1)$, it follows that

$$\hat{\tau}^2 = \frac{1}{n(n-1)} \sum_{k=1}^n (A_k - \bar{A})^2$$

is an unbiased estimator of τ^2 . Combining this result with (3.6), one obtains an unbiased estimator B_1 of $[tr(\mathbf{TV})]^2$, namely

$$B_1 = [tr(\tilde{\mathbf{S}}_n)]^2 - \frac{1}{n(n-1)} \sum_{k=1}^n (A_k - \bar{A})^2. \quad (3.7)$$

To derive an unbiased estimator B_2 of $tr[(\mathbf{TV})^2]$, a simple result from matrix algebra is used. Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$, then $tr(\mathbf{AB}') = \mathbf{1}'_n(\mathbf{A} \# \mathbf{B})\mathbf{1}_n$, where $\mathbf{A} \# \mathbf{B}$ denotes the Hadamard product of \mathbf{A} and \mathbf{B} . Using this result and (3.4), the bias of $tr(\tilde{\mathbf{S}}_n^2)$ can be obtained from

$$\begin{aligned} E[tr(\tilde{\mathbf{S}}_n^2)] &= \mathbf{1}'_d E[\tilde{\mathbf{S}}_n \# \tilde{\mathbf{S}}_n] \mathbf{1}_d = \mathbf{1}'_d E \left[\left(\frac{1}{n} \sum_{k=1}^n \mathbf{S}_k \right) \# \left(\frac{1}{n} \sum_{k=1}^n \mathbf{S}_k \right) \right] \mathbf{1}_d \\ &= \mathbf{1}'_d \left[\frac{1}{n^2} \sum_{k=1}^n \sum_{k'=1}^n E(\mathbf{S}_k \# \mathbf{S}_{k'}) \right] \mathbf{1}_d \\ &= \mathbf{1}'_d \left[\frac{n-1}{n} \mathbf{S} \# \mathbf{S} + \underbrace{\frac{1}{n^2} \sum_{k=1}^n E(\mathbf{S}_k \# \mathbf{S}_k)}_{:= \pi^2} \right] \mathbf{1}_d. \end{aligned}$$

To estimate the unknown quantity π^2 , again the above quoted result from matrix algebra is used and one obtains $\mathbf{1}'_d(\mathbf{S}_k \# \mathbf{S}_k)\mathbf{1}_d = tr(\mathbf{S}_k^2)$. Finally, this leads to the desired unbiased estimator B_2 of $tr[(\mathbf{TV})^2]$, namely

$$B_2 = tr \left(\frac{n}{n-1} \tilde{\mathbf{S}}_n^2 - \frac{1}{n(n-1)} \sum_{k=1}^n \mathbf{S}_k^2 \right). \quad (3.8)$$

To reduce memory space and computation time (which is important for simulations), the traces of the $d \times d$ matrices \mathbf{S}_k , \mathbf{S}_k^2 , and $\tilde{\mathbf{S}}_n^2$ are computed from simple bilinear and quadratic forms. This is summarized in the following lemma.

LEMMA 3.1 *Let $\mathbf{X}_k = (X_{k1}, \dots, X_{kd})'$, $k, l = 1, \dots, n$ be i.i.d. random vectors, let \mathbf{T} denote a projection matrix and denote $\mathbf{Y}_k = \mathbf{TX}_k$. Further let $A_{kl} = \mathbf{X}'_k \mathbf{T} \mathbf{X}_l$ denote a symmetric bilinear form in \mathbf{X}_k and \mathbf{X}_l and for $k = l$ let $A_k := A_{kk}$ denote a quadratic form. Then*

$$\begin{aligned} 1. B_1 &= \left[tr(\tilde{\mathbf{S}}_n) \right]^2 - \frac{1}{n(n-1)} \sum_{k=1}^n (A_k - \bar{A})^2 = \frac{1}{n(n-1)} \underbrace{\sum_{k=1}^n \sum_{l=1}^n}_{k \neq l} A_k A_l \\ 2. B_2 &= tr \left(\frac{n}{n-1} \tilde{\mathbf{S}}_n^2 - \frac{1}{n(n-1)} \sum_{k=1}^n \mathbf{S}_k^2 \right) = \frac{1}{n(n-1)} \underbrace{\sum_{k=1}^n \sum_{l=1}^n}_{k \neq l} A_{kl}^2. \end{aligned}$$

PROOF:

1. First note that $\mathbf{Y}_k = \mathbf{TX}_k$ and $\mathbf{S}_k = \mathbf{Y}_k \mathbf{Y}'_k$. Thus,

$$\begin{aligned} tr(\tilde{\mathbf{S}}_n) &= tr \left(\frac{1}{n} \sum_{k=1}^n \mathbf{S}_k \right) = \frac{1}{n} \sum_{k=1}^n tr(\mathbf{S}_k) = \frac{1}{n} \sum_{k=1}^n tr(\mathbf{Y}_k \mathbf{Y}'_k) \\ &= \frac{1}{n} \sum_{k=1}^n (\mathbf{Y}'_k \mathbf{Y}_k) = \frac{1}{n} \sum_{k=1}^n \underbrace{\mathbf{X}'_k \mathbf{T} \mathbf{X}_k}_{= A_k} = \frac{1}{n} \sum_{k=1}^n A_k = \bar{A}. \end{aligned}$$

using invariance under cyclic permutations. Replacing $tr(\tilde{\mathbf{S}}_n)$ with \bar{A} leads to

$$\begin{aligned}
B_1 &= \left[tr(\tilde{\mathbf{S}}_n^2) \right]^2 - \frac{1}{n(n-1)} \sum_{k=1}^n (A_k - \bar{A})^2 \\
&= \bar{A}^2 - \frac{1}{n(n-1)} \left[\sum_{k=1}^n A_k^2 - n\bar{A}^2 \right] = \frac{n}{n-1} \bar{A}^2 - \frac{1}{n(n-1)} \sum_{k=1}^n A_k^2 \\
&= \frac{1}{n(n-1)} \left[\left(\sum_{k=1}^n A_k \right)^2 - \sum_{k=1}^n A_k^2 \right] = \frac{1}{n(n-1)} \underbrace{\sum_{k=1}^n \sum_{l=1}^n}_{k \neq l} A_k A_l.
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
\frac{n}{n-1} \left(\frac{1}{n} \sum_{k=1}^n \mathbf{S}_k \right)^2 - \frac{1}{n(n-1)} \sum_{k=1}^n \mathbf{S}_k^2 &= \\
\frac{1}{n(n-1)} \left[\sum_{k=1}^n \mathbf{S}_k \mathbf{S}_l - \sum_{k=1}^n \mathbf{S}_k^2 \right] &= \frac{1}{n(n-1)} \sum_{k \neq l} \mathbf{S}_k \mathbf{S}_l.
\end{aligned}$$

Then, taking traces and observing the invariance under cyclic permutations one obtains for $k \neq l$

$$tr(\mathbf{S}_k \mathbf{S}_l) = tr(\mathbf{Y}_k \mathbf{Y}_k' \mathbf{Y}_l \mathbf{Y}_l') = \mathbf{Y}_l' \mathbf{Y}_k \mathbf{Y}_k' \mathbf{Y}_l = A_{lk} A_{kl} = A_{kl}^2$$

since $A_{kl} = A_{lk}$.

This leads to an unbiased estimator B_2 of $tr(\mathbf{S}^2)$

$$B_2 = \frac{1}{n(n-1)} \underbrace{\sum_{k=1}^n \sum_{l=1}^n}_{k \neq l} A_{kl}^2.$$

Combining all the results above, one obtains the new estimator \tilde{f} of $f = [tr(\mathbf{S})]^2 / tr(\mathbf{S}^2)$

$$\tilde{f} = \frac{B_1}{B_2} = \frac{\sum_{k \neq l} A_k A_l}{\sum_{k \neq l} A_{kl}^2}. \quad (3.9)$$

3.2 Quadratic and Bilinear Forms

In order to derive the properties of the estimators B_0 in (3.5), B_1 in (3.7), and B_2 in (3.8), some results on bilinear and quadratic forms are needed.

Let \mathbf{T} be a symmetric matrix. Then the quantity $A_{kl} = \mathbf{X}_k' \mathbf{T} \mathbf{X}_l$ is called a *bilinear form* if $k \neq l$ and if $k = l$ the quantity $A_k = \mathbf{X}_k' \mathbf{T} \mathbf{X}_k$ is called a *quadratic form*. First we consider the representation of a quadratic form in a random vector \mathbf{X} .

LEMMA 3.2 (REPRESENTATION OF A QUADRATIC FORM) Let $\mathbf{X} = (X_1, \dots, X_d)'$ denote a random vector with $E(\mathbf{X}) = \boldsymbol{\mu}$ and $\text{Cov}(\mathbf{X}) = \mathbf{V} > 0$ and let \mathbf{T} be a symmetric matrix. Then,

$$A = \mathbf{X}'\mathbf{T}\mathbf{X} = \begin{cases} \sum_{i=1}^d \lambda_i (U_i + b_i)^2 & \text{if } E(\mathbf{X}) = \boldsymbol{\mu} \neq \mathbf{0} \\ \sum_{i=1}^d \lambda_i U_i^2 & \text{if } E(\mathbf{X}) = \mathbf{0} \end{cases}$$

where the λ_i are the eigenvalues of $\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}$. Moreover, $\mathbf{U} = (U_1, \dots, U_d)$ is a random vector with $E(\mathbf{U}) = \mathbf{0}$, $\text{Cov}(\mathbf{U}) = \mathbf{I}$, and $(b_1, \dots, b_d) = (\mathbf{P}'\mathbf{V}^{1/2}\boldsymbol{\mu})'$ where $\mathbf{P}\mathbf{P}' = \mathbf{I}$ denotes an orthogonal matrix.

PROOF: see MATHAI and PROVOST (1992), p. 28f.

If \mathbf{X} has a multivariate normal distribution, then the distribution of a quadratic form is given in the following corollary.

COROLLARY 3.3 (DISTRIBUTION OF A QUADRATIC FORM) Let $\mathbf{X} = (X_1, \dots, X_d)' \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$, $|\mathbf{V}| \neq 0$, $\mathbf{V} = \mathbf{V}'$ and let \mathbf{T} denote a symmetric matrix. Then,

$$A = \mathbf{X}'\mathbf{T}\mathbf{X} = \sum_{i=1}^d \lambda_i C_i,$$

where $C_i \sim \chi_1^2$ are i.i.d. random variables and λ_i are the eigenvalues of $\mathbf{T}\mathbf{V}$, $i = 1, \dots, d$.

PROOF: Note that $|\mathbf{V}| \neq 0$ and $\mathbf{V} = \mathbf{V}' \Rightarrow \exists \mathbf{V}^{1/2}$ and $\mathbf{V}^{-1/2}$ is symmetric and regular $\Rightarrow \mathbf{V}^{1/2}\mathbf{V}^{-1/2} = \mathbf{I}$.

$$\begin{aligned} \mathbf{X}'\mathbf{T}\mathbf{X} &= \mathbf{X}'\mathbf{V}^{-1/2}\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}\mathbf{V}^{-1/2}\mathbf{X} \\ &= \mathbf{X}'\mathbf{V}^{-1/2}\mathbf{P}'\mathbf{P}\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}\mathbf{P}'\mathbf{P}\mathbf{V}^{-1/2}\mathbf{X} \end{aligned}$$

Since $\mathbf{V}^{1/2}$ and \mathbf{T} are symmetric it follows that $\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}$ is also symmetric. Thus, by singular value decomposition theorem it follows that there exists an orthogonal matrix \mathbf{P} such that $\mathbf{P}\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}\mathbf{P}' = \text{diag}\{\lambda'_1, \dots, \lambda'_d\} = \Delta$. The quantities λ'_i are the eigenvalues of $\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}$. Thus

$$\mathbf{X}'\mathbf{T}\mathbf{X} = (\mathbf{P}\mathbf{V}^{-1/2}\mathbf{X})' \underbrace{\mathbf{P}\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}\mathbf{P}'}_{\Delta} (\mathbf{P}\mathbf{V}^{-1/2}\mathbf{X}) = \mathbf{Z}'\Delta\mathbf{Z} = \sum_{i=1}^d \lambda'_i Z_i^2,$$

where $\mathbf{Z} = \mathbf{P}\mathbf{V}^{-1/2}\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}\mathbf{V}^{-1/2}\mathbf{V}\mathbf{V}^{-1/2}\mathbf{P}') = \mathcal{N}(\mathbf{0}, \mathbf{I}_d)$. Thus it remains to show that $\lambda_i = \lambda'_i$. To this end let λ denote an arbitrary eigenvalue of $\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}$. Then,

$$\begin{aligned} \mathbf{0} = |\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2} - \lambda\mathbf{I}| &= |\mathbf{I}||\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2} - \lambda\mathbf{I}| \\ &= |\mathbf{V}^{-1/2}||\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2} - \lambda\mathbf{I}||\mathbf{V}^{1/2}| \\ &= |\mathbf{V}^{-1/2}\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}\mathbf{V}^{1/2} - \lambda\mathbf{V}^{-1/2}\mathbf{I}\mathbf{V}^{1/2}| \\ &= |\mathbf{T}\mathbf{V} - \lambda\mathbf{I}| \end{aligned}$$

Thus, it follows that λ is also an eigenvalue of \mathbf{TV} . □

This corollary shows that a quadratic form is distributed as a sum of uncorrelated (in case of a normal distribution independent) χ_1^2 -distributed random variables. With this in mind, we will derive the central moments of quadratic forms.

LEMMA 3.4 (MOMENTS OF A QUADRATIC FORM)

Let $\mathbf{X}_k = (X_{k1}, \dots, X_{kd}) \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{V})$, $k = 1, \dots, n$, *i.i.d.* random variables and let \mathbf{T} denote a projection matrix. Then for the quadratic form $A_k = \mathbf{X}_k' \mathbf{T} \mathbf{X}_k$, it holds in general that

1. $E(A_k) = \text{tr}(\mathbf{TV}) + \boldsymbol{\mu}' \mathbf{T} \boldsymbol{\mu}$
2. $\text{Var}(A_k) = 2 \text{tr}[(\mathbf{TV})^2] + 4 \boldsymbol{\mu}' \mathbf{T} \mathbf{V} \mathbf{T} \boldsymbol{\mu}$

Moreover, if $\mathbf{T} \boldsymbol{\mu} = \mathbf{0}$ then,

3. $E(A_k) = \text{tr}(\mathbf{TV}) =: \nu$
4. $\text{Var}(A_k) = 2 \text{tr}[(\mathbf{TV})^2] =: \tau^2$
5. For the pairs $A_k A_l$ and $A_r A_s$ (where $k \neq l$ and $r \neq s$) it holds that

$$\text{Cov}(A_k A_l, A_r A_s) = \begin{cases} \tau^4 + 2 \tau^2 \nu^2 & : \text{ if } (k, l) = (r, s) \\ & \text{ or if } (k, l) = (s, r) \\ \tau^2 \nu^2 & : \text{ if } k = r \text{ and } l \neq s \\ & \text{ or if } k \neq r \text{ and } l = s \\ 0 & : \text{ otherwise} \end{cases}$$

PROOF:

1. this is the well-known Lancaster theorem.
2. see MATHAI and PROVOST (1992), p.53
3. follows immediately from (1)
4. follows immediately from (2)
5.
 - $\text{Var}(A_k A_l) = E(A_k A_l A_k A_l) - [E(A_k A_l)]^2$
 $= [E(A_k^2)]^2 - [E(A_k)]^4 = (\tau^2 + \nu^2)^2 - \nu^4$
 $= \tau^4 + 2 \tau^2 \nu^2$
 - $\text{Cov}(A_k A_l, A_k A_r) = E(A_k A_l A_k A_r) - E(A_k A_l) E(A_k A_r)$
 $= (\tau^2 + \nu^2) \nu^2 - \nu^4 = \tau^2 \nu^2$
 - $\text{Cov}(A_k A_l, A_r A_s) = \nu^4 - \nu^4 = 0$

□

Next we consider the representation, distribution and moments of bilinear forms.

LEMMA 3.5 (REPRESENTATION OF A BILINEAR FORM)

Let $\mathbf{X} = (X_1, \dots, X_d)'$ and $\mathbf{Y} = (Y_1, \dots, Y_d)'$ be independent identically distributed random vectors satisfying $E(\mathbf{X}) = E(\mathbf{Y}) = \boldsymbol{\mu}$ and $\text{Cov}(\mathbf{X}) = \text{Cov}(\mathbf{Y}) = \mathbf{V} > 0$ and let \mathbf{T} denote a symmetric matrix. Then,

$$A = \mathbf{X}'\mathbf{T}\mathbf{Y} = \begin{cases} \sum_{i=1}^d \lambda_i (U_i + b_i)(W_i + b_i) & \text{if } E(\mathbf{X}) = E(\mathbf{Y}) = \boldsymbol{\mu} \neq \mathbf{0} \\ \sum_{i=1}^d \lambda_i U_i W_i & \text{if } E(\mathbf{X}) = E(\mathbf{Y}) = \mathbf{0} \end{cases}$$

where the λ_i are the eigenvalues of $\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}$. Moreover, $\mathbf{U} = (U_1, \dots, U_d)$, $\mathbf{W} = (W_1, \dots, W_d)$ are random vectors with $E(\mathbf{U}) = E(\mathbf{W}) = \mathbf{0}$, $\text{Cov}(\mathbf{U}, \mathbf{W}) = \mathbf{0}$ and $\text{Cov}(\mathbf{U}) = \text{Cov}(\mathbf{W}) = \mathbf{I}$ and $(b_1, \dots, b_d) = (\mathbf{P}'\mathbf{V}^{\frac{1}{2}}\boldsymbol{\mu})'$, where $\mathbf{P}\mathbf{P}' = \mathbf{I}$ denotes an orthogonal matrix.

PROOF: Let $\mathbf{Z}_X = (\mathbf{V}^{-1/2}\mathbf{X} - \mathbf{V}^{-1/2}\boldsymbol{\mu})$ and $\mathbf{Z}_Y = (\mathbf{V}^{-1/2}\mathbf{Y} - \mathbf{V}^{-1/2}\boldsymbol{\mu})$. Then, by the singular value decomposition theorem, there exists an orthogonal matrix \mathbf{P} such that

$$\mathbf{P}\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}\mathbf{P}' = \text{diag}\{\lambda'_1, \dots, \lambda'_d\} = \Delta,$$

where $\mathbf{P}\mathbf{P}' = \mathbf{I}$. Then set $\mathbf{U} = \mathbf{P}'\mathbf{Z}_X$ and $\mathbf{W} = \mathbf{P}'\mathbf{Z}_Y$. Thus,

$$E(\mathbf{U}) = E(\mathbf{W}) = \mathbf{0} \quad \text{Cov}(\mathbf{U}) = \text{Cov}(\mathbf{W}) = \mathbf{I}.$$

Then it follows for the bilinear form $A = \mathbf{X}'\mathbf{T}\mathbf{Y}$ that

$$\begin{aligned} \mathbf{X}'\mathbf{T}\mathbf{Y} &= \mathbf{X}'\mathbf{V}^{-1/2}\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}\mathbf{V}^{-1/2}\mathbf{Y} \\ &= (\mathbf{Z}_X + \mathbf{V}^{-1/2}\boldsymbol{\mu})'\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}(\mathbf{Z}_Y + \mathbf{V}^{-1/2}\boldsymbol{\mu}) \\ &= (\mathbf{Z}_X + \mathbf{V}^{-1/2}\boldsymbol{\mu})'\mathbf{P}'\underbrace{\mathbf{P}\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}\mathbf{P}'}_{\Delta}\mathbf{P}(\mathbf{Z}_Y + \mathbf{V}^{-1/2}\boldsymbol{\mu}) \\ &= (\mathbf{U} + \mathbf{b})'\Delta(\mathbf{W} + \mathbf{b}). \end{aligned}$$

□

If \mathbf{X} and \mathbf{Y} have multivariate normal distributions, then the distribution of a bilinear form is given in the following corollary.

COROLLARY 3.6 (DISTRIBUTION OF A BILINEAR FORM)

Let $\mathbf{X} = (X_1, \dots, X_d)'$ and $\mathbf{Y} = (Y_1, \dots, Y_d)' \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$ be independent random vectors with $|\mathbf{V}| \neq 0$, $\mathbf{V} = \mathbf{V}'$ and let \mathbf{T} denote a symmetric matrix. Then,

$$A = \mathbf{X}'\mathbf{T}\mathbf{Y} \sim \sum_{i=1}^d \lambda_i C_i D_i,$$

where $C_i, D_i \sim \mathcal{N}(0, 1)$ are independent and the λ_i are the eigenvalues of $\mathbf{T}\mathbf{V}$.

PROOF: From $|\mathbf{V}| \neq 0$ and $\mathbf{V} = \mathbf{V}' \Rightarrow \exists \mathbf{V}^{1/2}$ and $\mathbf{V}^{-1/2}$, symmetric and regular
 $\Rightarrow \mathbf{V}^{1/2}\mathbf{V}^{-1/2} = \mathbf{I}$.

$$\begin{aligned}\mathbf{X}'\mathbf{T}\mathbf{Y} &= \mathbf{X}'\mathbf{V}^{-1/2}\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}\mathbf{V}^{-1/2}\mathbf{Y} \\ &= \mathbf{X}'\mathbf{V}^{-1/2}\mathbf{P}'\mathbf{P}\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}\mathbf{P}'\mathbf{P}\mathbf{V}^{-1/2}\mathbf{Y}.\end{aligned}$$

Since $\mathbf{V}^{1/2}$ and \mathbf{T} are symmetric it follows that $\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}$ is also symmetric. Thus, by singular value decomposition theorem it follows that there exists an orthogonal matrix \mathbf{P} such that $\mathbf{P}\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}\mathbf{P}' = \text{diag}\{\lambda'_1, \dots, \lambda'_d\} = \Delta$. The quantities λ'_i are the eigenvalues of $\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}$. Thus

$$\mathbf{X}'\mathbf{T}\mathbf{Y} = (\mathbf{P}\mathbf{V}^{-1/2}\mathbf{X})' \underbrace{\mathbf{P}\mathbf{V}^{1/2}\mathbf{T}\mathbf{V}^{1/2}\mathbf{P}'}_{\Delta} (\mathbf{P}\mathbf{V}^{-1/2}\mathbf{Y}) = \mathbf{C}'\Delta\mathbf{D} = \sum_{i=1}^d \lambda'_i C_i D_i,$$

where

$$\begin{aligned}\mathbf{C} &= \mathbf{P}\mathbf{V}^{-1/2}\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}\mathbf{V}^{-1/2}\mathbf{V}\mathbf{V}^{-1/2}\mathbf{P}') = \mathcal{N}(\mathbf{0}, \mathbf{I}_n), \\ \mathbf{D} &= \mathbf{P}\mathbf{V}^{-1/2}\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}\mathbf{V}^{-1/2}\mathbf{V}\mathbf{V}^{-1/2}\mathbf{P}') = \mathcal{N}(\mathbf{0}, \mathbf{I}_n).\end{aligned}$$

Thus it remains to show that $\lambda_i = \lambda'_i$. This has already been shown in the proof of Corollary 3.3. \square

By this corollary it is possible to derive the central moments of bilinear forms.

LEMMA 3.7 (MOMENTS OF A BILINEAR FORM)

Let $\mathbf{X}_k = (X_{k1}, \dots, X_{kd})'$, $\mathbf{X}_l = (X_{l1}, \dots, X_{ld})' \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$, be i.i.d. for $(k \neq l \in \{1, \dots, n\})$ and let \mathbf{T} denote a projection matrix. Then,

1. $E(A_{kl}) = 0$
2. $\text{Var}(A_{kl}) = E(A_{kl}^2) = \text{tr}[(\mathbf{T}\mathbf{V})^2]$
3. $E(A_{kl}^4) = 6 \text{tr}[(\mathbf{T}\mathbf{V})^4] + 3 (\text{tr}[(\mathbf{T}\mathbf{V})^2])^2$
4. $E(A_{kl}^2 A_{kn}^2) = 2 \text{tr}[(\mathbf{T}\mathbf{V})^4] + (\text{tr}[(\mathbf{T}\mathbf{V})^2])^2$

PROOF:

$$\begin{aligned}1. \quad E(A_{kl}) &= E\left(\sum_{i=1}^d \lambda_i C_i D_i\right) = \sum_{i=1}^d \lambda_i \underbrace{E(C_i)E(D_i)}_{=0} = 0 \\ 2. \quad \text{Var}(A_{kl}) &= E(A_{kl}^2) = E\left(\sum_{i=1}^d \lambda_i C_i D_i \sum_{j=1}^d \lambda_j C_j D_j\right) \\ &= E\left(\sum_{i=1}^d \lambda_i^2 C_i^2 D_i^2\right) = \sum_{i=1}^d \lambda_i^2 \underbrace{E(C_i^2)}_{=1} \underbrace{E(D_i^2)}_{=1} \\ &= \sum_{i=1}^d \lambda_i^2 = \text{tr}[(\mathbf{T}\mathbf{V})^2]\end{aligned}$$

$$\begin{aligned}
3. E(A_{kl}^4) &= E \left(\left[\sum_{i=1}^d \lambda_i C_i D_i \right]^4 \right) \\
&= \sum_{i=1}^d \sum_{j=1}^d \sum_{r=1}^d \sum_{s=1}^d \lambda_i \lambda_j \lambda_r \lambda_s E(C_i D_i C_j D_j C_r D_r C_s D_s) \\
&= \sum_{i=1}^d \lambda_i^4 \underbrace{E(C_i^4 D_i^4)}_9 + 3 \sum_{i \neq j}^d \lambda_i^2 \lambda_j^2 \underbrace{E(C_i^2 D_i^2 C_j^2 D_j^2)}_1 \\
&= 6 \sum_{i=1}^d \lambda_i^4 + 3 \sum_{i,j}^d \lambda_i^2 \lambda_j^2 = 6 \operatorname{tr}[(\mathbf{TV})^4] + 3 (\operatorname{tr}[(\mathbf{TV})^2])^2
\end{aligned}$$

$$\begin{aligned}
4. E(A_{kl}^2 A_{kr}^2) &= E \left[\left(\sum_{i=1}^d \lambda_i C_i D_i \right)^2 \left(\sum_{i=1}^d \lambda_i C_i E_i \right)^2 \right] \\
&= \sum_{i=1}^d \sum_{j=1}^d \sum_{r=1}^d \sum_{s=1}^d \lambda_i \lambda_j \lambda_r \lambda_s E(C_i D_i C_j D_j C_r E_r C_s E_s) \\
&= \sum_{i=1}^d \lambda_i^4 \underbrace{E(C_i^4 D_i^2 E_i^2)}_3 + \sum_{i \neq j}^d \lambda_i^2 \lambda_j^2 \underbrace{E(C_i^2 D_i^2 C_j^2 E_j^2)}_1 \\
&= 2 \sum_{i=1}^d \lambda_i^4 + \sum_{i=1}^d \sum_{j=1}^d \lambda_i^2 \lambda_j^2 = 2 \operatorname{tr}[(\mathbf{TV})^4] + (\operatorname{tr}[(\mathbf{TV})^2])^2 \quad \square
\end{aligned}$$

3.3 Consistency and dimensional stability of the estimators

In the following lemma, a regularity condition is given from which it follows that the new estimators are consistent and dimensionally stable.

LEMMA 3.8 *Let λ_i , ($\lambda_i \geq \lambda_0 > 0$) denote the eigenvalues of \mathbf{TVT} and let $m := \max_i \lambda_i < \infty$. Then,*

$$\begin{aligned}
1. \lim_{d \rightarrow \infty} \frac{m}{\sum_{i=1}^d \lambda_i} = 0 &\Rightarrow \lim_{d \rightarrow \infty} \frac{\operatorname{tr}[(\mathbf{TV})^2]}{[\operatorname{tr}(\mathbf{TV})]^2} = 0, \\
2. \lim_{d \rightarrow \infty} \frac{m^2}{\sum_{i=1}^d \lambda_i^2} = 0 &\Rightarrow \lim_{d \rightarrow \infty} \frac{\operatorname{tr}[(\mathbf{TV})^4]}{(\operatorname{tr}[(\mathbf{TV})^2])^2} = 0.
\end{aligned}$$

PROOF:

$$1. \frac{tr[(\mathbf{TV})^2]}{[tr(\mathbf{TV})]^2} = \frac{\sum_{i=1}^d \lambda_i^2}{(\sum_{i=1}^d \lambda_i)^2} \leq \frac{m \sum_{i=1}^d \lambda_i}{(\sum_{i=1}^d \lambda_i)^2} = \frac{m}{\sum_{i=1}^d \lambda_i} \xrightarrow{d \rightarrow \infty} 0 .$$

$$2. \frac{tr[(\mathbf{TV})^4]}{(tr[(\mathbf{TV})^2])^2} = \frac{\sum_{i=1}^d \lambda_i^4}{(\sum_{i=1}^d \lambda_i^2)^2} \leq \frac{m^2 \sum_{i=1}^d \lambda_i^2}{(\sum_{i=1}^d \lambda_i^2)^2} = \frac{m^2}{\sum_{i=1}^d \lambda_i^2} \xrightarrow{d \rightarrow \infty} 0 .$$

REMARK 3.1 *The first condition follows from the second condition since $\lambda_i \geq 0$ and*

$$\frac{m^2}{(\sum_{i=1}^d \lambda_i)^2} \leq \frac{m^2}{\sum_{i=1}^d \lambda_i^2}$$

which follows from $(\sum_{i=1}^d \lambda_i)^2 \geq \sum_{i=1}^d \lambda_i^2$.

In the sequel it will be shown by Lemma 3.7 that the unbiased estimators B_1 and B_2 are consistent and dimensionally stable - provided that the conditions of Lemma 3.8 are fulfilled. To this end it has to be shown that the variance of the ratio of the estimator and the quantity to be estimated is uniformly bounded in d and that this bound does not depend on d .

The following trivial properties of the variance of a sum of random variables will be used in the sequel:

$$Var \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n Var(X_i) + \sum_{i \neq j} Cov(X_i, X_j)$$

$$Var \left(\sum_{i=1}^n X_i \right) = E \left(\sum_{i,j} X_i X_j \right) - \left[E \left(\sum_{i=1}^n X_i \right) \right]^2 .$$

Thus, one obtains for $V_1 = Var(B_1)$

$$n^2(n-1)^2 V_1 = Var \left(\sum_{k \neq l} A_k A_l \right) = \sum_{k \neq l} Var(A_k A_l) + \underbrace{\sum_{k \neq l} \sum_{r \neq s} Cov(A_k A_l, A_r A_s)}_{(k,l) \neq (r,s)} .$$

Moreover, by Lemma 3.4,

$$\begin{aligned} n^2(n-1)^2 V_1 &= n(n-1) Var(A_k A_l) \\ &\quad + n(n-1)(n-2)(n-3) Cov(A_k A_l, A_r A_s) \\ &\quad + 4n(n-1)(n-2) Cov(A_k A_l, A_k A_r) \\ &\quad + n(n-1) \underbrace{Cov(A_k A_l, A_l A_k)}_{= Var(A_k A_l)} \\ &= 2n(n-1)(\sigma^4 + 2\sigma^2\mu^2) + 4n(n-1)(n-2)\sigma^2\mu^2 \\ &= 2n(n-1)\sigma^4 + (4n^3 - 8n^2 + 4n)\sigma^2\mu^2 . \end{aligned}$$

Note that $\mu^4 = [\text{tr}(\mathbf{TV})]^4$ does no longer appear in the variance. Thus, it can be shown by Lemma 3.8 that $V_1/[\text{tr}(\mathbf{TV})]^4 \rightarrow 0$ if $d \rightarrow \infty$.

$$\frac{V_1}{[\text{tr}(\mathbf{TV})]^4} = \frac{2}{n(n-1)} \underbrace{\frac{(\text{tr}[(\mathbf{TV})^2])^2}{[\text{tr}(\mathbf{TV})]^4}}_{\rightarrow 0} + \frac{(4n^3 - 8n^2 + 4n)}{n^2(n-1)^2} \underbrace{\frac{\text{tr}[(\mathbf{TV})^2]}{[\text{tr}(\mathbf{TV})]^2}}_{\rightarrow 0}. \quad (3.10)$$

Further note that this ratio is bounded for all $d, n > 1$ by $\frac{4n}{(n-1)^2}$.

$$\begin{aligned} \frac{V_1}{[\text{tr}(\mathbf{TV})]^4} &= \frac{2}{n(n-1)} \underbrace{\frac{(\text{tr}[(\mathbf{TV})^2])^2}{[\text{tr}(\mathbf{TV})]^4}}_{\leq 1} + \frac{(4n^3 - 8n^2 + 4n)}{n^2(n-1)^2} \underbrace{\frac{\text{tr}[(\mathbf{TV})^2]}{[\text{tr}(\mathbf{TV})]^2}}_{\leq 1} \\ &\leq \frac{2}{n(n-1)} + \frac{(4n^2 - 8n + 4)}{n(n-1)^2} \leq \frac{4n^2 - 6n + 2}{n(n-1)^2} \leq \frac{4n}{(n-1)^2}. \end{aligned}$$

For the variance $V_2 = \text{Var}(B_2)$ of the second estimator it follows that

$$\begin{aligned} n^2(n-1)^2 V_2 &= \text{Var} \left(\sum_{k \neq l}^n A_{kl}^2 \right) \\ &= E \left(\sum_{k \neq l}^n \sum_{r \neq s}^n A_{kl}^2 A_{rs}^2 \right) - \left[E \left(\sum_{k \neq l}^n A_{kl}^2 \right) \right]^2. \end{aligned} \quad (3.11)$$

Moreover, by Lemma 3.7,

$$\begin{aligned} n^2(n-1)^2 V_2 &= 2n(n-1)E(A_{kl}^4) \\ &\quad + 4n(n-1)(n-2)E(A_{kl}^2 A_{kn}^2) \\ &\quad + n(n-1)(n-2)(n-3)E(A_{kl}^2)^2 \\ &\quad - n^2(n-1)^2 \text{Var}(A_{kl})^2 \\ &= 2n(n-1)(6 \text{tr}[(\mathbf{TV})^4] + 3(\text{tr}[(\mathbf{TV})^2])^2) \\ &\quad + 4n(n-1)(n-2)(2 \text{tr}[(\mathbf{TV})^4] + (\text{tr}[(\mathbf{TV})^2])^2) \\ &\quad + n(n-1)(n-2)(n-3)(\text{tr}[(\mathbf{TV})^2])^2 \\ &\quad - n^2(n-1)^2 (\text{tr}[(\mathbf{TV})^2])^2 \\ &= n(n-1)[(8n-4) \text{tr}[(\mathbf{TV})^4] + 4(\text{tr}[(\mathbf{TV})^2])^2]. \end{aligned}$$

Unfortunately, the term $V_2/(\text{tr}[(\mathbf{TV})^2])^2$ does not vanish for $d \rightarrow \infty$ and for fixed n . One obtains

$$\frac{V_2}{(\text{tr}[(\mathbf{TV})^2])^2} = \frac{8n-4}{n(n-1)} \underbrace{\frac{\text{tr}[(\mathbf{TV})^4]}{(\text{tr}[(\mathbf{TV})^2])^2}}_{\rightarrow 0} + \frac{4}{n(n-1)}. \quad (3.12)$$

This ratio is bounded, however, for all $d, n > 1$ by $\frac{8}{n-1}$

$$\frac{V_2}{(tr[(\mathbf{TV})^2])^2} = \frac{8n-4}{n(n-1)} \underbrace{\frac{tr[(\mathbf{TV})^4]}{(tr[(\mathbf{TV})^2])^2}}_{\leq 1} + \frac{4}{n(n-1)} \leq \frac{8n-4+4}{n(n-1)} \leq \frac{8}{(n-1)}.$$

Here it has been shown that for fixed n and $d \rightarrow \infty$ the quantity $B_1/[tr(\mathbf{TV})]^2$ is \mathcal{L}_2 -consistent and thus it is dimensionally stable. Moreover, $B_1/[tr(\mathbf{TV})]^2$ is also \mathcal{L}_2 -consistent for $n \rightarrow \infty$. The quantity $B_2/tr[(\mathbf{TV})^2]$, however, is only \mathcal{L}_2 -consistent for $n \rightarrow \infty$ and arbitrary fixed d . For fixed n and $d \rightarrow \infty$, it is at least uniformly bounded by $8/(n-1)$ and thus, it is also dimensionally stable.

3.4 Bias of the Ratio $[tr(\mathbf{TV})]^2 / tr[(\mathbf{TV})^2]$

It has been shown that B_1 and B_2 are unbiased estimators of $[tr(\mathbf{TV})]^2$ and $tr[(\mathbf{TV})^2]$, respectively. The ratio B_1/B_2 , however may be biased. This shall be investigated in the sequel using a Taylor expansion for the ratio of two random variables.

$$E\left(\frac{B_1}{B_2}\right) \approx \frac{E(B_1)}{E(B_2)} \left(1 + \frac{Var(B_2)}{[E(B_2)]^2} - \frac{Cov(B_1, B_2)}{E(B_1)E(B_2)}\right), \quad (3.13)$$

where the sign \approx means ‘‘approximately equal’’. Using the derivation in the previous section, one obtains

$$E\left(\frac{B_1}{B_2}\right) \approx \frac{[tr(\mathbf{TV})]^2}{tr[(\mathbf{TV})^2]} \left(1 + \frac{4}{n(n-1)} - \frac{Cov(B_1, B_2)}{[tr(\mathbf{TV})]^2 tr[(\mathbf{TV})^2]}\right). \quad (3.14)$$

It is easily seen that the last term in (3.14) can be neglected for large d or for large n . From Cauchy-Schwarz-inequality and by (3.10) and (3.12), it follows under the assumptions of Lemma 3.8 that

$$\frac{Cov(B_1, B_2)}{[tr(\mathbf{TV})]^2 tr[(\mathbf{TV})^2]} = O\left(\frac{1}{n\sqrt{d}}\right).$$

Thus, the ratio $Cov(B_1, B_2)/([tr(\mathbf{TV})]^2 tr[(\mathbf{TV})^2])$ vanishes for n fixed and $d \rightarrow \infty$ as well as for d fixed and $n \rightarrow \infty$. \square

The bias of $\tilde{f} = B_1/B_2$ can be approximately determined by plugging in this result in (3.13).

$$E(\tilde{f}) \approx \frac{[tr(\mathbf{TV})]^2}{tr[(\mathbf{TV})^2]} \left(1 + \frac{4}{n(n-1)}\right). \quad (3.15)$$

Obviously, the new estimator is slightly biased, but the bias disappears rapidly with increasing n . Furthermore, numerator and denominator of f are dimensionally stable and f is asymptotically unbiased for $n \rightarrow \infty$ and arbitrary d .

Simulations show that the bias of \tilde{f} is much smaller than the bias of the simple plug-in estimator $\hat{f} = [tr(\hat{\mathbf{S}}_n)]^2 / tr(\hat{\mathbf{S}}_n^2)$ in the ANOVA-type statistic.

3.5 F -Approximation (continued)

Now all relevant estimators have been derived and their properties have been investigated. Thus, the F -approximation started in Section 2.1 can be resumed and the distribution of the statistic Q_n/B_0 shall be approximated by a F -distribution such that the first two moments coincide.

$$F_n = \frac{Q_n}{B_0} \dot{\sim} F(f_1, f_2).$$

We need the following moments:

LEMMA 3.9 *Let $B_0 = \frac{1}{n} \sum_{k=1}^n A_k$ and $Q_n = \frac{1}{n} \sum_{k=1}^n \sum_{l=1}^n A_{kl}$, where $A_{kl} = \mathbf{X}'_k \mathbf{T} \mathbf{X}_l$ are symmetric bilinear forms. Then,*

1. $E(B_0) = \text{tr}(\mathbf{T}\mathbf{V})$
2. $\text{Var}(B_0) = \frac{2}{n} \text{tr}[(\mathbf{T}\mathbf{V})^2]$
3. $\text{Cov}(Q_n, B_0) = \text{Var}(B_0)$.

Proof: The statements follow easily using the results from Lemma 3.4 on the moments of quadratic forms and from Lemma 3.7 on the moments of bilinear forms.

1.
$$E\left(\frac{1}{n} \sum_{k=1}^n A_k\right) = \frac{1}{n} \sum_{k=1}^n E(A_k) = \text{tr}(\mathbf{T}\mathbf{V})$$
2.
$$\text{Var}\left(\frac{1}{n} \sum_{k=1}^n A_k\right) = \frac{1}{n^2} \sum_{k=1}^n \text{Var}(A_k) = \frac{2}{n} \text{tr}[(\mathbf{T}\mathbf{V})^2]$$
3.
$$\begin{aligned} \text{Cov}(Q_n, B_0) &= E(Q_n B_0) - E(Q_n)E(B_0) = \\ &= E\left(\frac{1}{n^2} \sum_{k=1}^n \sum_{l=1}^n A_{kl} \sum_{m=1}^n A_m\right) - E\left(\frac{1}{n} \sum_{k=1}^n \sum_{l=1}^n A_{kl}\right) \text{tr}(\mathbf{T}\mathbf{V}) \\ &= \frac{1}{n^2} \sum_{k=1}^n \sum_{l=1}^n E(A_k A_l) - [\text{tr}(\mathbf{T}\mathbf{V})]^2 \\ &= \frac{1}{n} E(A_k^2) + \frac{n-1}{n} [E(A_k)]^2 - [\text{tr}(\mathbf{T}\mathbf{V})]^2 = \frac{2}{n} \text{tr}[(\mathbf{T}\mathbf{V})^2]. \quad \square \end{aligned}$$

Thus, $B_0 = \frac{1}{n} \sum_{k=1}^n A_k$ is an unbiased, consistent and dimensionally stable estimator of $\text{tr}(\mathbf{T}\mathbf{V})$.

The first two central moments of F_n are approximately determined by a Taylor expansion for the ratio of two random variables X and Y with $\text{Var}(X) < \infty$ and $\text{Var}(Y) < \infty$.

The approximation for the expectation $E(X/Y)$ is given in (3.13) replacing B_1 with X and B_2 with Y . For the variance $Var(X/Y)$ we have

$$Var\left(\frac{X}{Y}\right) \doteq \frac{E(X)^2}{E(Y)^2} \left(\frac{Var(X)}{[E(X)]^2} + \frac{Var(Y)}{[E(Y)]^2} - 2 \frac{Cov(X, Y)}{E(X)E(Y)} \right)$$

By noting that $E(Q_n) = tr(\mathbf{TV})$, $Var(Q_n) = 2 tr[(\mathbf{TV})^2]$, and $Cov(Q_n, B_0) = Var(B_0)$, one obtains

$$\begin{aligned} E(F_n) &\doteq \frac{E(Q_n)}{E(B_0)} \left[1 + \frac{Var(B_0)}{[E(B_0)]^2} - \frac{Cov(Q_n, B_0)}{E(Q_n)E(B_0)} \right] \\ &\doteq \frac{tr(\mathbf{TV})}{tr(\mathbf{TV})} \left[1 + \frac{Var(B_0)}{[tr(\mathbf{TV})]^2} - \frac{Var(B_0)}{[tr(\mathbf{TV})]^2} \right] \doteq 1 \\ Var(F_n) &\doteq \frac{(E(Q_n))^2}{[E(B_0)]^2} \left[\frac{Var(Q_n)}{[E(Q_n)]^2} + \frac{Var(B_0)}{[E(B_0)]^2} - \frac{2 Cov(Q_n, B_0)}{E(Q_n)E(B_0)} \right] \\ &\doteq \frac{[tr(\mathbf{TV})]^2}{[tr(\mathbf{TV})]^2} \left[\frac{Var(Q_n) - Var(B_0)}{[tr(\mathbf{TV})]^2} \right] \\ &\doteq \frac{(2 - \frac{2}{n}) tr[(\mathbf{TV})^2]}{[tr(\mathbf{TV})]^2}. \end{aligned}$$

Now, $E(F_n)$ and $Var(F_n)$ are equated with the first two moments of the $F(f_1, f_2)$ -distribution

	E	Var
$F(f_1, f_2)$	$\frac{f_2 - 2}{f_2}$	$\frac{f_2^3 + 2f_2^2 f_1 - 4f_2^2}{(f_1 f_2^2 - 4f_2 f_1 + 4f_1)(f_2 - 4)}$
F_n	1	$\frac{(2 - \frac{2}{n}) tr[(\mathbf{TV})^2]}{[tr(\mathbf{TV})]^2}$

and one obtains

$$\tilde{f} = f_1 = \frac{[tr(\mathbf{TV})]^2}{(1 - \frac{1}{n}) tr[(\mathbf{TV})^2]}, \quad f_2 = \infty.$$

The distribution of F_n can be approximated by a $\chi_{\tilde{f}}^2/\tilde{f}$ -distribution, where

$$\tilde{f} = \frac{B_1}{(1 - \frac{1}{n})B_2} = \frac{n}{n-1} \frac{B_1}{B_2}.$$

Obviously, $\tilde{f}/f \rightarrow 1$ for $n \rightarrow \infty$, where $f = B_1/B_2$. The bias in (3.15) is reduced by the factor $n/(n-1)$.

Literatur

- [1] BAI, Z. and SARANADASA, H. (1996), Effect of High Dimension: by an Example of a Two Sample Problem. *Statistica Sinica*, **6**, 311-329.
- [2] BOX, G.E.P. (1954). Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, II. Effects of Inequality of Variance and of Correlation Between Errors in the Two-Way Classification. *The Annals of Mathematical Statistics* **25**, 484-498.
- [3] GEISSER, S. and GREENHOUSE, S. W. (1958). An Extension of Box's Result on the Use of the F Distribution in Multivariate Analysis. *Annals of Mathematical Statistics*, **29**, 885-891.
- [4] GREENHOUSE, S. W. and GEISSER, S. (1959). On Methods in the Analysis of Profile Data. *Psychometrika*, **24**, 95-112.
- [5] MATHAI, A.M. and PROVOST, S.B. (1992). *Quadratic Forms in Random Variables, Theory and Applications*. Marcel Dekker, Inc.
- [6] PATNAIK, P.B. (1949). The non-central χ^2 -and F-distributions and their Applications. *Biometrika*, 202-232.