Test for the mean in a Growth Curve model for high dimensions

Muni S. Srivastava
Department of Statistics – University of Toronto, Canada

Joint work with Martin Singull, Linköping University, Sweden

LinStat2014
August 24 – 2014

- We consider the problem of estimating and testing a general linear hypothesis in a general multivariate linear model, the so called Growth Curve Model, when the $p \times N$ observation matrix is normally distributed with an unknown covariance matrix.
- The maximum likelihood estimator (MLE) for the mean is a weighted estimator with the inverse of the sample covariance matrix which is unstable for large p close to N and singular for p larger than N.
- We modify the MLE to an unweighted estimator and propose a new test which we compare with the previous likelihood ratio test (LRT) based on the weighted estimator, i.e., the MLE.

- We show that the performance of the LRT and the new test based on the unweighted estimator are similar.
- For the high-dimensional case, when p is larger than N, we construct two new tests based on the trace of the variation matrices due to the hypothesis (between sum of squares) and the error (within sum of squares).
- To compare the performance of these four tests we compute the attained significance level and the empirical power.

Let $\boldsymbol{x}_{ij}$ be independent and identically distributed vectors with p-variate normal distribution $N_p(\boldsymbol{\mu}_i, \Sigma_i)$, where $i = 1, 2$ and $j = 1, ..., N_i$.

The sample mean vectors are, respectively, given by

$$\bar{\boldsymbol{x}}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \boldsymbol{x}_{ij}, \quad i = 1, 2,$$

and the sample covariance matrices are, respectively, given by

$$\boldsymbol{S}_i = \frac{1}{n_i} \boldsymbol{X}(\boldsymbol{I} - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}')\boldsymbol{X}', \quad n_i = N_i - 1, \quad i = 1, 2.$$

When $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$, an unbiased estimator of $\boldsymbol{\Sigma}$ is given by

$$\boldsymbol{S} = \frac{n_1 \boldsymbol{S}_1 + n_2 \boldsymbol{S}_2}{n}, \quad n = n_1 + n_2 = N_1 + N_2 - 2.$$

A test if two mean vectors are equal, $H : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ has been proposed by Dempster (1958), under the assumption that the two distributions have the same covariance matrix.

Dempster's test statistic is given by

$$T_D = \left(\frac{1}{N_1} + \frac{1}{N_2}\right)^{-1} \frac{(\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)'(\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)}{\mathrm{tr}\boldsymbol{S}}$$

If $\boldsymbol{\Sigma} = \gamma^2 \boldsymbol{I}_p$ one can show that under the null hypothesis

$$T_D \sim F(p, np).$$

It may be noted that when $\boldsymbol{\Sigma} = \gamma^2 \boldsymbol{I}_p$, and under the assumption of normality Dempster's test T_D is uniformly most powerful among all tests whose power depends on $\boldsymbol{\mu}'\boldsymbol{\mu}/\gamma^2$.

For a general $\boldsymbol{\Sigma}$, under the assumption of normality and assuming

$$(\star) \qquad 0 < \lim_{p \to \infty} a_i < \infty,\ i = 1, ..., 4, \text{ where } a_i = \frac{\mathrm{tr}\boldsymbol{\Sigma}^i}{p}$$

one can show that, under the null hypothesis,

$$T_D \approx F([\hat{r}], [n\hat{r}]),$$

where $[a]$ denotes the largest integer value $\leq a$, $\hat{r} = p\hat{b}$, $\hat{b} = \frac{\hat{a}_1^2}{\hat{a}_2}$,

$$\hat{a}_1 = \frac{\mathrm{tr}\boldsymbol{S}}{p}, \quad \text{and} \quad \hat{a}_2 = \frac{1}{p}\left(\mathrm{tr}\boldsymbol{S}^2 - \frac{1}{n}(\mathrm{tr}\boldsymbol{S})^2\right).$$

Bai and Saranadasa (1996) proposed another asymptotically equivalent test which does not require the assumption of normality but have asymptotically the same power as the one proposed by Dempster (1958).

The statistic testing equal means given by Bai and Saranadasa is

$$T_{BS} = \frac{\left(\frac{1}{N_1} + \frac{1}{N_2}\right)^{-1} (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)'(\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2) - \mathrm{tr}\boldsymbol{S}}{\sqrt{2\left(\mathrm{tr}\boldsymbol{S}^2 - \frac{1}{n}(\mathrm{tr}\boldsymbol{S})^2\right)}}$$

Bai and Saranadasa also showed that under the null hypothesis T_{BS} is normally distributed with mean 0 and variance 1 for a general model that includes the normal model as a special case.

Srivastava (2007) proposed a Hotelling's T^2 type test, by using Moore-Penrose inverse of the sample covariance matrix $\boldsymbol{S}^+$ instead of the inverse when N is smaller than p.

The test statistic given by Srivastava (2007) is

$$T^{+2} = \left(\frac{1}{N_1} + \frac{1}{N_2}\right)^{-1} (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)' \boldsymbol{S}^+ (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)$$

and the asymptotic distribution, assuming $(\star)$, is proved to be

$$\frac{\hat{b}p}{n} T^{+2} \approx \chi^2(n).$$

It may be noted that all the above discussed tests are invariant under the group of orthogonal matrices.

A test that is invariant under the group of non-singular diagonal matrices has recently been proposed by Srivastava and Du (2008) under the normal distribution and Srivastava (2009) under non-normality.

It may be noted that this test is not invariant under the transformation by orthogonal matrices.

The test statistic given by Srivastava and Du is

$$T_{SD} = \frac{\left(\frac{1}{N_1} + \frac{1}{N_2}\right)^{-1} (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)' \boldsymbol{D}_{\boldsymbol{S}}^{-1} (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2) - p}{\sqrt{2\left(\mathrm{tr}\widehat{\boldsymbol{R}}^2 - \frac{p^2}{n}\right) C_{p,n}}}$$

where $\widehat{\boldsymbol{R}} = \boldsymbol{D}_{\boldsymbol{S}}^{-1/2} \boldsymbol{S} \boldsymbol{D}_{\boldsymbol{S}}^{-1/2}$, $\boldsymbol{D}_{\boldsymbol{S}} = diag(s_{11}, ..., s_{pp})$, $\boldsymbol{S} = (s_{ij})$ and

$$C_{p,n} = 1 + \frac{\mathrm{tr}\widehat{\boldsymbol{R}}^2}{p^{3/2}} \xrightarrow{p} 1 \quad \text{as } (n,p) \to \infty.$$

Assuming some conditions, similar to $(\star)$ on the correlation matrix $\boldsymbol{R}$, and under the hypothesis of equality of two mean vectors, T_{SD} has asymptotically standard normal distribution.

Definition

Let $\boldsymbol{X} = \boldsymbol{\xi}\boldsymbol{A} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\xi} : p \times m$ unknown parameter matrix, $\boldsymbol{A} : m \times N$ known design matrix such that $r = \text{rank}(\boldsymbol{A})$ and $r + p \leq N$. The Multivariate Linear Model is given by

$$\boldsymbol{X} = \boldsymbol{\xi}\boldsymbol{A} + \boldsymbol{\varepsilon},$$

where the columns of $\boldsymbol{\varepsilon}$ are assumed to be independently $p-$variate normally distributed with mean zero and an unknown positive definite covariance matrix $\boldsymbol{\Sigma}$, i.e.,

$$\boldsymbol{\varepsilon} \sim N_{p,N}\left(\mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{I}_N\right).$$

By the notation of the matrix normal distribution

$$\boldsymbol{\varepsilon} \sim N_{p,N}\left(\mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{I}_N\right)$$

we just mean that the vectorization of the matrix is multivariate normal distributes as

$$\text{vec}\ \boldsymbol{\varepsilon} \sim N_{pN}\left(\mathbf{0}, \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}\right),$$

where $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1, ..., \boldsymbol{\varepsilon}_N) : p \times N$ and vec $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1', ..., \boldsymbol{\varepsilon}_N')' : pN \times 1$.

If $\boldsymbol{A}$ has full rank, the MLEs for the multivariate linear model is given by

$$\widehat{\boldsymbol{\xi}} = \boldsymbol{X}\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1},$$
$$N\widehat{\boldsymbol{\Sigma}} = \boldsymbol{X}\left(\boldsymbol{I}_N - \boldsymbol{P}_{\boldsymbol{A}'}\right)\boldsymbol{X}' = \widehat{\boldsymbol{R}}\widehat{\boldsymbol{R}}' = \boldsymbol{V},$$

where $\boldsymbol{P}_{\boldsymbol{A}'} = \boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{A}$.

The general linear hypothesis is expressed as $H : \boldsymbol{C\xi}' = \mathbf{0}$, where $\boldsymbol{C}$ is a known $q \times m$ matrix of rank $q \leq m$.

Let the error sum of squares and products be given by the matrix

$$\boldsymbol{V} = \boldsymbol{X}\left(\boldsymbol{I}_N - \boldsymbol{P}_{\boldsymbol{A}'}\right)\boldsymbol{X}' \quad \text{and} \quad \boldsymbol{S} = \frac{1}{n}\boldsymbol{V}, \quad n = N - m,$$

and the sum of squares and products due to regression under the hypothesis H is

$$\boldsymbol{W} = \widehat{\boldsymbol{\xi}}\boldsymbol{C}'(\boldsymbol{C}(\boldsymbol{AA}')^{-1}\boldsymbol{C}')^{-1}\boldsymbol{C}\widehat{\boldsymbol{\xi}}'.$$

Fujikoshi et al. (2004) generalize the two-sample test given by Dempster (1958) to the MANOVA problem, under the assumption that $(p/n) \to c \in (0, \infty)$. The statistic given by Fujikoshi et al. is

$$\tilde{T}_D = \sqrt{p}\left(\frac{\mathrm{tr}\boldsymbol{W}}{\mathrm{tr}\boldsymbol{S}} - q\right)$$

and

$$\frac{\tilde{T}_D}{\hat{\sigma}_D} \to N(0,1),$$

where $\hat{\sigma}_D = 2q\dfrac{\hat{a}_2}{\hat{a}_1}$,

$$\hat{a}_1 = \frac{\mathrm{tr}\boldsymbol{S}}{p}, \quad \text{and} \quad \hat{a}_2 = \frac{1}{p}\left(\mathrm{tr}\boldsymbol{S}^2 - \frac{1}{n}(\mathrm{tr}\boldsymbol{S})^2\right).$$

Other tests that do not require the assumption

$$(p/n) \to c \in (0, \infty)$$

have been proposed by Srivastava and Fujikoshi (2006) with the test statistic

$$T_{SF} = \frac{\sqrt{p}(\mathrm{tr}\boldsymbol{W} - q\mathrm{tr}\boldsymbol{S})}{\sqrt{2q\hat{a}_2}}.$$

Under the general linear hypothesis the asymptotic distribution of T_{SF} is normal with mean 0 and variance 1.

Schott (2007) proposed the same test as proposed by Srivastava and Fujikoshi (2006) but required the assumption above to obtain the asymptotic distribution of the test statistic. This is a severe restriction.

The above tests are, however, not invariant under the transformation by non-singular diagonal matrices. A test that has this property for the MANOVA problem has been recently proposed by Yamada and Srivastava (2012) under normality.

The test statistic given by Yamada and Srivastava (2012) is

$$T_{YS} = \frac{\mathrm{tr}\boldsymbol{W}\boldsymbol{D}_S^{-1} - \frac{n}{n-2}pq}{\sqrt{2q\left(\mathrm{tr}\widehat{\boldsymbol{R}}^2 - \frac{p^2}{n}\right)C_{p,n}}}$$

where $\widehat{\boldsymbol{R}} = \boldsymbol{D}_{\boldsymbol{S}}^{-1/2}\boldsymbol{S}\boldsymbol{D}_{\boldsymbol{S}}^{-1/2}$ and $C_{p,n} = 1 + \frac{\mathrm{tr}\widehat{\boldsymbol{R}}^2}{p^{3/2}}$

Assuming certain conditions on the correlation matrix $\boldsymbol{R}$ the asymptotic distribution under the null hypothesis is standard normal.

The Growth Curve model was first proposed by Potthoff and Roy (1964) and is defined as follows.

Definition

Let $\boldsymbol{X} : p \times N$ and $\boldsymbol{\xi} : q \times m$ be the observation and parameter matrices, respectively, and let $\boldsymbol{B} : p \times q$ and $\boldsymbol{A} : m \times N$ be the within and between individual design matrices, respectively. Suppose that $q \leq p$ and $r + p \leq N$, where $r = \text{rank}(\boldsymbol{A})$. The Growth Curve model is given by

$$\boldsymbol{X} = \boldsymbol{B}\boldsymbol{\xi}\boldsymbol{A} + \boldsymbol{\varepsilon},$$

where

$$\boldsymbol{\varepsilon} \sim N_{p,N}\left(\mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{I}_N\right).$$

If $\boldsymbol{B}$ and $\boldsymbol{A}$ has full rank, the MLEs for the Growth Curve model is given by

$$\begin{aligned}
\widehat{\boldsymbol{\xi}}_{MLE} &= \left(\boldsymbol{B}'\boldsymbol{V}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{B}'\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{A}'\left(\boldsymbol{A}\boldsymbol{A}'\right)^{-1}, \text{ i.e.,}\\
\boldsymbol{B}\widehat{\boldsymbol{\xi}}_{MLE}\boldsymbol{A} &= \boldsymbol{P}_{\boldsymbol{B}}^{\boldsymbol{V}}\boldsymbol{X}\boldsymbol{P}_{\boldsymbol{A}'},\\
N\widehat{\boldsymbol{\Sigma}}_{MLE} &= \left(\boldsymbol{X} - \boldsymbol{B}\widehat{\boldsymbol{\xi}}_{MLE}\boldsymbol{A}\right)\left(\right)'\\
&= \left(\boldsymbol{X}(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{A}'}) + \boldsymbol{X}\boldsymbol{P}_{\boldsymbol{A}'} - \boldsymbol{B}\widehat{\boldsymbol{\xi}}_{MLE}\boldsymbol{A}\right)\left(\right)'\\
&= \boldsymbol{V} + \widehat{\boldsymbol{R}}_1\widehat{\boldsymbol{R}}_1',
\end{aligned}$$

where

$$\begin{aligned}
\widehat{\boldsymbol{R}}_1 &= \boldsymbol{X}\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{A} - \boldsymbol{B}\widehat{\boldsymbol{\xi}}_{MLE}\boldsymbol{A} = \left(\boldsymbol{I}_p - \boldsymbol{P}_{\boldsymbol{B}}^{\boldsymbol{V}}\right)\boldsymbol{X}\boldsymbol{P}_{\boldsymbol{A}'},\\
\boldsymbol{V} &= \boldsymbol{X}\left(\boldsymbol{I}_N - \boldsymbol{P}_{\boldsymbol{A}'}\right)\boldsymbol{X}',\\
\boldsymbol{V}_1 &= \boldsymbol{X}\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{A}\boldsymbol{X}' \quad \text{and}\\
\boldsymbol{P}_{\boldsymbol{B}}^{\boldsymbol{V}} &= \boldsymbol{B}\left(\boldsymbol{B}'\boldsymbol{V}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{B}'\boldsymbol{V}^{-1}.
\end{aligned}$$

The mean and covariance matrix for the estimator $\widehat{\boldsymbol{\xi}}_{MLE}$ are given in Kollo and von Rosen (2005) as

$$\mathrm{E}\left(\widehat{\boldsymbol{\xi}}_{MLE}\right) = \boldsymbol{\xi},$$

$$\mathrm{cov}\left(\widehat{\boldsymbol{\xi}}_{MLE}\right) = \frac{n-1}{n-1-(p-q)}(\boldsymbol{A}\boldsymbol{A}')^{-1} \otimes (\boldsymbol{B}'\boldsymbol{\Sigma}^{-1}\boldsymbol{B})^{-1},$$

if $n-1-(p-q) > 0$, where $n = N - m$.

Since $q \leq p$ we have $\dfrac{n-1}{n-1-(p-q)} \geq 1$.

A natural alternative to the MLE would be an unweighted estimator of $\boldsymbol{\xi}$ given by

$$\widehat{\boldsymbol{\xi}} = (\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'\boldsymbol{X}\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1}.$$

This estimator is simpler than the MLE, since we do not need to calculate the inverse of the sum of squares matrix $\boldsymbol{V}^{-1}$.

This unweighted estimator is obtained by considering

$$\boldsymbol{X}\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1} = \boldsymbol{B}\boldsymbol{\xi} + \boldsymbol{\eta},$$
$$\boldsymbol{\eta} = \boldsymbol{\varepsilon}\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1} \sim N_{p,m}\left(\mathbf{0}, \boldsymbol{\Sigma}, (\boldsymbol{A}\boldsymbol{A}')^{-1}\right).$$

The distribution of the estimator is given by

$$\widehat{\boldsymbol{\xi}} \sim N_{q,m}(\boldsymbol{\xi}, (\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'\boldsymbol{\Sigma}\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^{-1}, (\boldsymbol{A}\boldsymbol{A}')^{-1}),$$

i.e., we have

$$\begin{aligned} \mathrm{E}\left(\widehat{\boldsymbol{\xi}}\right) &= \boldsymbol{\xi}, \\ \mathrm{cov}\left(\widehat{\boldsymbol{\xi}}\right) &= (\boldsymbol{A}\boldsymbol{A}')^{-1} \otimes (\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'\boldsymbol{\Sigma}\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^{-1}. \end{aligned}$$

Note that with $\boldsymbol{P}_{\boldsymbol{B}} = \boldsymbol{B}\left(\boldsymbol{B}'\boldsymbol{B}\right)^{-1}\boldsymbol{B}'$ we have

$$\begin{aligned} &\left(\boldsymbol{X} - \boldsymbol{B}\widehat{\boldsymbol{\xi}}_{MLE}\boldsymbol{A}\right)\left(\boldsymbol{X} - \boldsymbol{B}\widehat{\boldsymbol{\xi}}_{MLE}\boldsymbol{A}\right)' \\ &\quad = \left(\boldsymbol{X} - \boldsymbol{B}\widehat{\boldsymbol{\xi}}_{MLE}\boldsymbol{A}\right)()' \\ &\quad = \left(\boldsymbol{X}(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{A}'}) + \boldsymbol{X}\boldsymbol{P}_{\boldsymbol{A}'} - \boldsymbol{B}\widehat{\boldsymbol{\xi}}_{MLE}\boldsymbol{A}\right)()' \\ &\quad = \boldsymbol{X}(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{A}'})\boldsymbol{X}' + (\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{B}})\boldsymbol{X}\boldsymbol{P}_{\boldsymbol{A}'}\boldsymbol{X}'(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{B}}) \\ &\quad = \boldsymbol{V} + (\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{B}})\boldsymbol{X}\boldsymbol{P}_{\boldsymbol{A}'}\boldsymbol{X}'(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{B}}). \end{aligned}$$

An unbiased and consistent estimator of the covariance matrix $\boldsymbol{\Sigma}$ is given by

$$n\widehat{\boldsymbol{\Sigma}} = \boldsymbol{V} \sim W_p(\boldsymbol{\Sigma}, n),$$

where $n = N - m$, irrespective of which estimator of $\boldsymbol{\xi}$ is used.

Both $\widehat{\boldsymbol{\xi}}_{MLE}$ and $\widehat{\boldsymbol{\xi}}$ are unbiased.

The covariances for $\widehat{\boldsymbol{\xi}}_{MLE}$ and $\widehat{\boldsymbol{\xi}}$ respectively are given by

$$\text{cov}\left(\widehat{\boldsymbol{\xi}}_{MLE}\right) = \frac{n-1}{n-1-(p-q)}(\boldsymbol{A}\boldsymbol{A}')^{-1} \otimes (\boldsymbol{B}'\boldsymbol{\Sigma}^{-1}\boldsymbol{B})^{-1},$$

$$\text{cov}\left(\widehat{\boldsymbol{\xi}}\right) = (\boldsymbol{A}\boldsymbol{A}')^{-1} \otimes (\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'\boldsymbol{\Sigma}\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^{-1}.$$

To compare the two estimators we want to compare their covariances, i.e., we want to compare

$$(\boldsymbol{B}'\boldsymbol{\Sigma}^{-1}\boldsymbol{B})^{-1} \quad \text{and} \quad (\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'\boldsymbol{\Sigma}\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^{-1}.$$

Following Rao (1967) (Lemma 2.c) one can show that

$$(\boldsymbol{B}'\boldsymbol{\Sigma}^{-1}\boldsymbol{B})^{-1} \leq (\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'\boldsymbol{\Sigma}\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^{-1}$$

with equality if and only if $\mathcal{C}(\boldsymbol{\Sigma}^{-1}\boldsymbol{B}) = \mathcal{C}(\boldsymbol{B})$. The inequality is with respect to the Loewner partial ordering, i.e.,

$$(\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'\boldsymbol{\Sigma}\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^{-1} - (\boldsymbol{B}'\boldsymbol{\Sigma}^{-1}\boldsymbol{B})^{-1}$$

is nonnegative definite.

Hence, for large n, the unweighted unbiased estimator of $\boldsymbol{\xi}$ has a larger covariance than the weighted one, as expected since the weighted estimator is the MLE.

But, when also p is large, but still less than n, the factor

$$(n-1)/(n-1-(p-q))$$

can be much greater than one and the covariance for the weighted estimator can actually be larger than the covariance for the unweighted estimator.

The general problem of testing the hypothesis

$$H : \boldsymbol{L\xi M} = \mathbf{0} \quad \text{vs.} \quad A : \boldsymbol{L\xi M} \neq \mathbf{0}$$

is equivalent testing the hypothesis

$$H : \boldsymbol{\xi} = \boldsymbol{L}_1\boldsymbol{\delta}\boldsymbol{M}_1 \quad \text{vs.} \quad A : \boldsymbol{\xi} \neq \boldsymbol{L}_1\boldsymbol{\delta}\boldsymbol{M}_1,$$

for some $\boldsymbol{L}_1$ and $\boldsymbol{M}_1$ depending on $\boldsymbol{L}$ and $\boldsymbol{M}$ and their dimension. Since $\boldsymbol{L}_1$ and $\boldsymbol{M}_1$ can be combined with $\boldsymbol{B}$ and $\boldsymbol{A}$, without loss of generality, we shall consider only the hypothesis

$$H : \boldsymbol{\xi} = \mathbf{0} \quad \text{vs.} \quad A : \boldsymbol{\xi} \neq \mathbf{0}.$$

Given the MLEs the LRT is given as as

$$\lambda_{MLE}^{2/N} = \frac{|\boldsymbol{V} + \widehat{\boldsymbol{R}}_1 \widehat{\boldsymbol{R}}_1'|}{|\boldsymbol{V} + \boldsymbol{V}_1|}$$

where $\widehat{\boldsymbol{R}}_1$, $\boldsymbol{V}$ and $\boldsymbol{V}_1$ are given above. Using Box's method for approximate the distribution of

$$T_1 = -r \log \lambda_{MLE}^{2/N},$$

one can show that for large N,

$$P_0 \left(T_1 > c\right) = P\left(\chi_f^2 > c\right),$$

where $r = n - p + q - (q - m + 1)/2$ and $f = qm$.

See Srivastava and Khatri (1979) for more details.

Equivalently testing the hypothesis $H : \boldsymbol{\xi} = \mathbf{0}$ one can test the hypothesis

$$H : \boldsymbol{\eta} = \mathbf{0} \quad \text{vs.} \quad A : \boldsymbol{\eta} \neq \mathbf{0},$$

where $\boldsymbol{\eta} = (\boldsymbol{B}'\boldsymbol{B})^{1/2}\boldsymbol{\xi}(\boldsymbol{A}\boldsymbol{A}')^{1/2}$.

Using the unweighted estimator $\widehat{\boldsymbol{\xi}}$ given above the distribution of $\widehat{\boldsymbol{\eta}} = (\boldsymbol{B}'\boldsymbol{B})^{1/2}\widehat{\boldsymbol{\xi}}(\boldsymbol{A}\boldsymbol{A}')^{1/2}$ is

$$\widehat{\boldsymbol{\eta}} = (\boldsymbol{B}'\boldsymbol{B})^{1/2}\widehat{\boldsymbol{\xi}}(\boldsymbol{A}\boldsymbol{A}')^{1/2} \sim N_{q,m}(\boldsymbol{\eta}, \boldsymbol{\Delta}, \boldsymbol{I}_m),$$

where

$$\boldsymbol{\Delta} = \boldsymbol{G}_1\boldsymbol{\Sigma}\boldsymbol{G}_1',$$
$$\boldsymbol{G}_1 = (\boldsymbol{B}'\boldsymbol{B})^{-1/2}\boldsymbol{B}'.$$

Furthermore, we estimate $\boldsymbol{\Delta}$ with $\widehat{\boldsymbol{\Delta}} = \frac{1}{n}\boldsymbol{V}^*$ which is distributed as

$$\boldsymbol{V}^* = \boldsymbol{G}_1 \boldsymbol{V} \boldsymbol{G}_1' \sim W_q(\boldsymbol{\Delta}, n),$$

where $n = N - m$. Also, $\widehat{\boldsymbol{\eta}}$ and $\boldsymbol{V}$ (or $\boldsymbol{V}^*$) are independently distributed.

The likelihood function of $\boldsymbol{\eta}$ and $\boldsymbol{\Delta}$ is given by

$$c|\boldsymbol{\Delta}|^{-N/2}|\boldsymbol{V}^*|^{(n-q-1)/2}\text{etr}\left\{\frac{1}{2}\boldsymbol{\Delta}^{-1}(\boldsymbol{V}^* + N(\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta})()'\right\},$$

where c is a constant. Thus, another LRT is given by

$$\lambda^{2/N} = \frac{|\boldsymbol{V}^*|}{|\boldsymbol{V}^* + \widehat{\boldsymbol{\eta}}\widehat{\boldsymbol{\eta}}'|}.$$

Again using Box's method to approximate the distribution, one can shown that for large N the distribution of

$$T_2 = -r \log \lambda^{2/N}$$

is given by

$$P_0\left(T_2 > c\right) = P\left(\chi_f^2 > c\right),$$

where $r = n - (q - m + 1)/2$ and $f = qm$.

See Srivastava and Khatri (1979) for more details.

For high dimensions, when $N - m < p$, then $\boldsymbol{V}$ is singular and none of the tests given in above are applicable.

We will propose two new tests.

First, let

$$\boldsymbol{H}_1 = \boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1/2} : N \times m,$$
$$\boldsymbol{H} = \boldsymbol{H}_1\boldsymbol{H}_1' : N \times N$$

and

$$\boldsymbol{G}_1 = (\boldsymbol{B}'\boldsymbol{B})^{-1/2}\boldsymbol{B}' : q \times p,$$
$$\boldsymbol{G} = \boldsymbol{G}_1'\boldsymbol{G}_1 : p \times p.$$

Now consider the variable

$$(\boldsymbol{B}'\boldsymbol{B})^{1/2}\widehat{\boldsymbol{\xi}}(\boldsymbol{A}\boldsymbol{A}')^{1/2} = \boldsymbol{G}_1\boldsymbol{X}\boldsymbol{H}_1 \sim N_{q,m}\left((\boldsymbol{B}'\boldsymbol{B})^{1/2}\boldsymbol{\xi}(\boldsymbol{A}\boldsymbol{A}')^{1/2}, \boldsymbol{\Delta}, \boldsymbol{I}_m\right),$$

where $\boldsymbol{\Delta} = \boldsymbol{G}_1\boldsymbol{\Sigma}\boldsymbol{G}_1'$.

Under the hypothesis $H : \boldsymbol{\xi} = \boldsymbol{0}$ we have

$$\boldsymbol{W} = \boldsymbol{G}_1\boldsymbol{X}\boldsymbol{H}\boldsymbol{X}'\boldsymbol{G}_1' \sim W_q(\boldsymbol{\Delta}, m).$$

We see that $\boldsymbol{W}$ and $\widehat{\boldsymbol{\Sigma}}$ are independently distributed.

The mean and variance of the statistic tr$\boldsymbol{W}$, under the hypothesis $H : \boldsymbol{\xi} = \mathbf{0}$, are given by

$$\mathrm{E}(\mathrm{tr}\boldsymbol{W}) = m\,\mathrm{tr}\boldsymbol{\Delta},$$
$$\mathrm{var}(\mathrm{tr}\boldsymbol{W}) = 2m\,\mathrm{tr}\boldsymbol{\Delta}^2.$$

Under the assumption of normality, unbiased and consistent estimators of tr$\boldsymbol{\Delta}$ and tr$\boldsymbol{\Delta}^2$ are given by

$$\widehat{\mathrm{tr}\boldsymbol{\Delta}} = \frac{1}{n}\mathrm{tr}\boldsymbol{V}^*,$$
$$\widehat{\mathrm{tr}\boldsymbol{\Delta}^2} = \frac{1}{(n-1)(n+2)}\left(\mathrm{tr}\boldsymbol{V}^{*2} - \frac{1}{n}(\mathrm{tr}\boldsymbol{V}^*)^2\right),$$

respectively. See Srivastava (2005) for details.

Furthermore, we know that under the hypothesis $H : \boldsymbol{\xi} = \mathbf{0}$,

$$\widetilde{T}_3 = \frac{1}{\sqrt{m}} \frac{\mathrm{tr}\boldsymbol{W} - m\,\mathrm{tr}\boldsymbol{\Delta}}{\sqrt{2\,\mathrm{tr}\boldsymbol{\Delta}^2}} \to N(0,1).$$

Substituting unbiased and consistent estimators of $\mathrm{tr}\boldsymbol{\Delta}$ and $\mathrm{tr}\boldsymbol{\Delta}^2$ we get a test statistic, proposed by Srivastava and Fujikoshi (2006) and Srivastava (2007), which is given by

$$T_3 = \frac{\mathrm{tr}\boldsymbol{W} - \dfrac{m}{n}\mathrm{tr}\boldsymbol{V}^*}{\sqrt{\dfrac{2m}{(n-1)(n+2)}\left(\mathrm{tr}\boldsymbol{V}^{*2} - \dfrac{1}{n}(\mathrm{tr}\boldsymbol{V}^*)^2\right)}} \to N(0,1).$$

The test statistic T_3 is invariant under the group of orthogonal transformations, but not invariant under the units of measurements, which is an undesirable feature.

That is, the test is not invariant under a diagonal transformation and the test statistic T_3 changes.

We will now propose a test that is invariant under diagonal transformation.

We will also show that this new test performs better than the test above.

The test statistic will be based on the quantity $\mathrm{tr}\boldsymbol{W}\boldsymbol{D}_{\widehat{\boldsymbol{\Delta}}}^{-1}$, where $\boldsymbol{D}_{\widehat{\boldsymbol{\Delta}}} = \mathrm{diag}(\widehat{\boldsymbol{\Delta}})$, the diagonal matrix with the diagonal elements of $\widehat{\boldsymbol{\Delta}}$.

More precise the test statistic is based on $\mathrm{tr}(\boldsymbol{W}\boldsymbol{D}_{\boldsymbol{V}^*}^{-1})$ and given as

$$T_4 = \frac{n\mathrm{tr}(\boldsymbol{W}\boldsymbol{D}_{\boldsymbol{V}^*}^{-1}) - nqm/(n-2)}{\sqrt{2m(\mathrm{tr}\boldsymbol{R}^2 - q^2/n)c_{q,n}}},$$

where $\boldsymbol{R} = \boldsymbol{D}_{\boldsymbol{V}^*}^{-1/2}\boldsymbol{V}^*\boldsymbol{D}_{\boldsymbol{V}^*}^{-1/2}$ and $c_{q,n} = 1 + \mathrm{tr}\boldsymbol{R}^2/q^{3/2}$ is an adjustment factor converging to 1 in probability as $(n, q) \to \infty$, $n = O(q^\delta)$, $\delta > 1/2$ proposed by Srivastava and Du (2008).

Define the population correlation matrix as

$$\mathcal{R} = \boldsymbol{D}_{\boldsymbol{\Delta}}^{-1/2} \boldsymbol{\Delta} \boldsymbol{D}_{\boldsymbol{\Delta}}^{-1/2}.$$

It has been shown by Srivastava and Du (2008) that a consistent estimator of $\mathrm{tr}\mathcal{R}^2/q$ is given by

$$\frac{1}{q}\left(\mathrm{tr}\boldsymbol{R}^2 - \frac{q^2}{n}\right).$$

Hence, for large n and q we have

$$T_4 \stackrel{d}{=} \frac{(n\mathrm{tr}(\boldsymbol{W}\boldsymbol{D}_{\boldsymbol{V}^*}^{-1}) - qm)/\sqrt{q}}{\sqrt{2m\mathrm{tr}\mathcal{R}^2/q}}.$$

We will give the asymptotic distribution of the test statistic T_4 under the following assumptions

A0: $n = O(q^\delta)$, $\delta > 1/2$,

A1: $\frac{1}{q}\mathrm{tr}\mathcal{R}^2 = O(1)$ as $q \to \infty$,

A2: $\frac{1}{q^2}\mathrm{tr}\mathcal{R}^4 = o(1)$ as $q \to \infty$.

Theorem

Suppose that the assumptions A0-A2 hold. Then under the null hypothesis $H : \boldsymbol{\xi} = \mathbf{0}$ *the following is true*

$$\lim_{(n,q)\to\infty} P(T_4 \leq x) = \Phi(x),$$

where $\Phi(\cdot)$ *denotes the cumulative standard normal distribution function.*

To compare the three tests we can compute the attained significance level (ASL) and the empirical power.

Let c be the critical value from the distribution considered for the test statistics. With 1000 simulated replications under the null hypothesis, the ASL is computed as

$$\widehat{\alpha} = \frac{(\# \text{ of } t_H \geq c)}{(\# \text{ simulated replications})},$$

where t_H is the values of the test statistics derived from the simulated data under the null hypothesis.

We set the nominal significance level to $\alpha = 0.05$.

For the simulations let

$$\boldsymbol{B} = (b_{ij}), \quad b_{ij} \sim U(0,1), \quad i = 1, ..., p, \quad j = 1, ..., q$$

$$\text{and} \quad \boldsymbol{A} = \begin{pmatrix} \mathbf{1}'_{N_1} & \mathbf{0}'_{N_2} \\ \mathbf{0}'_{N_1} & \mathbf{1}'_{N_2} \end{pmatrix},$$

i.e., with $m = 2$.

For simplicity we will put $N_1 = N_2 = N/2$.

Furthermore, N, p and q will vary depending on which asymptotic is considered.

Since the covariances for the estimators depending on $\mathbf{\Sigma}$, we will use three different covariance matrices for the simulation study.

The first one is identity, i.e., $\mathbf{\Sigma}_1 = \boldsymbol{I}_p$.

Furthermore, let $\boldsymbol{D}_j = diag(\sigma_1^{(j)}, \ldots, \sigma_p^{(j)})$, for $j = 2, 3$, be two different diagonal matrices. Define $\sigma_i^{(2)} = 2 + (p - i + 1)/p$, for $i = 1, \ldots, p$, and $\sigma_i^{(3)}$, for $i = 1, \ldots, p$, are independent observations from $\sqrt{U\,[0, 2]}$, respectively.

Also, let $\boldsymbol{R} = (\rho_{ij})$, where $\rho_{ij} = (-1)^{i+j} r^{|i-j|^f}$.

The other two covariance matrices that we will use are given as

$$\mathbf{\Sigma}_j = \boldsymbol{D}_j \boldsymbol{R} \boldsymbol{D}_j, \quad \text{with } r = 0.2, f = 0.1, \text{ for } j = 2, 3.$$

We can compute the empirical power using two different critical values.

We can either use the critical value c from the asymptotic distribution, or we can use the estimated critical value $\widehat{c}$ calculated from the simulated data under the null hypothesis, i.e., the critical value calculated from the empirical null distribution.

We will use the estimated critical value since the ASL is greatly affected for some tests.

The empirical power is calculated from 1000 new replications simulated under the alternative hypothesis when $\boldsymbol{\xi} = (\xi_{ij})$ and $\xi_{ij} \sim U(-1/5, 1/5)$ if $i + j$ is even and zero otherwise.

Let t_A be the value of the test statistic derived from the simulated data under the alternative hypothesis.

The empirical power are given as

$$\widehat{\beta} = \frac{(\#\text{ of } t_A \geq \widehat{c})}{(\#\text{ simulated replications})}.$$

	q	ASL T_1	ASL T_2	ASL T_3	ASL T_4	Power T_1	Power T_2	Power T_3	Power T_4
(I)	10	0.058	0.054	0.078	0.063	0.308	0.666	**0.746**	0.730
	14	0.063	0.042	0.063	0.048	0.852	0.997	**0.999**	**0.999**
	18	0.063	0.049	0.063	0.058	0.319	0.551	**0.709**	0.675
	22	0.094	0.073	0.068	0.058	0.284	0.413	**0.640**	0.605
	26	0.110	0.094	0.061	0.054	0.838	0.948	**0.999**	**0.999**
	30	0.132	0.132	0.064	0.061	0.852	0.852	**0.999**	0.998
(II)	10	0.054	0.057	0.079	0.061	0.236	**0.312**	0.243	0.286
	14	0.062	0.060	0.077	0.057	0.669	**0.925**	0.895	0.920
	18	0.062	0.068	0.065	0.053	0.182	0.226	**0.286**	0.250
	22	0.088	0.065	0.065	0.037	0.129	**0.201**	0.159	0.186
	26	0.108	0.097	0.084	0.074	0.574	0.638	0.643	**0.663**
	30	0.115	0.115	0.065	0.053	0.614	0.614	0.656	**0.710**
(III)	10	0.055	0.048	0.073	0.058	0.446	**0.811**	0.730	0.760
	14	0.061	0.056	0.066	0.047	0.962	0.999	**1.000**	**1.000**
	18	0.077	0.055	0.067	0.050	0.413	0.613	**0.660**	0.641
	22	0.102	0.078	0.073	0.059	0.370	**0.523**	0.423	0.441
	26	0.103	0.099	0.067	0.049	0.957	0.984	**0.998**	**0.998**
	30	0.128	0.128	0.080	0.064	0.936	0.936	0.996	**0.998**

Table : $p = 30, N = 50$

From the table above we see that for large q, the significance level of all the tests except T_4 are greatly affected.

From the simulated power we see that it seems like T_2, T_3 and T_4 perform equally well and better than T_1, which is based the MLE.

Thus, T_4 should be preferred for large q.

In the next comparisons we choose smaller q, namely

$$q = 5 \quad \text{and} \quad q = 10.$$

	p	N	ASL T_1	ASL T_2	ASL T_3	ASL T_4	Power T_1	Power T_2	Power T_3	Power T_4
(I)	10	30	0.056	0.056	0.076	0.060	0.115	0.125	0.140	**0.143**
	10	50	0.041	0.059	0.082	0.066	0.246	0.238	**0.261**	0.256
	10	100	0.043	0.056	0.080	0.050	0.505	0.499	0.519	**0.528**
	30	30	*	*	0.065	0.042	*	*	**0.424**	0.410
	30	50	0.063	0.063	0.076	0.060	0.212	0.551	**0.617**	0.594
	30	100	0.055	0.045	0.070	0.049	0.813	0.937	0.944	**0.947**
	50	30	*	*	0.089	0.071	*	*	**0.685**	0.670
	50	50	*	*	0.075	0.050	*	*	**0.944**	0.939
	50	100	0.043	0.047	0.064	0.044	0.940	0.999	**1.000**	0.999
(II)	10	30	0.045	0.048	0.086	0.064	0.095	**0.114**	0.070	0.083
	10	50	0.057	0.053	0.084	0.054	**0.125**	0.141	0.099	0.115
	10	100	0.049	0.054	0.075	0.047	**0.277**	0.272	0.169	0.202
	30	30	*	*	0.094	0.069	*	*	0.102	**0.115**
	30	50	0.048	0.055	0.076	0.054	0.175	**0.304**	0.243	0.263
	30	100	0.048	0.042	0.069	0.041	0.576	**0.695**	0.511	0.584
	50	30	*	*	0.078	0.062	*	*	0.312	**0.372**
	50	50	*	*	0.069	0.050	*	*	0.537	**0.630**
	50	100	0.044	0.053	0.050	0.035	0.807	0.943	0.932	**0.959**
(III)	10	30	0.053	0.050	0.078	0.059	0.139	**0.183**	0.108	0.111
	10	50	0.047	0.048	0.102	0.065	0.260	**0.304**	0.161	0.201
	10	100	0.049	0.046	0.070	0.041	0.612	**0.641**	0.444	0.497
	30	30	*	*	0.089	0.067	*	*	**0.342**	0.300
	30	50	0.053	0.044	0.073	0.060	0.342	**0.698**	0.633	0.593
	30	100	0.044	0.044	0.061	0.040	0.920	**0.972**	0.968	0.954
	50	30	*	*	0.091	0.063	*	*	0.642	**0.704**
	50	50	*	*	0.075	0.051	*	*	0.937	**0.952**
	50	100	0.044	0.060	0.080	0.063	0.988	**1.000**	**1.000**	**1.000**

Table : $q = 5$, * means not available

	p	N	ASL T_1	ASL T_2	ASL T_3	ASL T_4	Power T_1	Power T_2	Power T_3	Power T_4
(I)	10	30	0.057	0.057	0.065	0.056	0.187	0.187	0.233	**0.234**
	10	50	0.052	0.052	0.065	0.049	0.329	0.329	**0.421**	0.409
	10	100	0.039	0.039	0.061	0.040	0.765	0.765	0.797	**0.801**
	30	30	*	*	0.057	0.048	*	*	**0.692**	0.670
	30	50	0.049	0.058	0.070	0.053	0.480	0.822	0.922	**0.923**
	30	100	0.051	0.039	0.051	0.037	0.991	**1.000**	**1.000**	**1.000**
	50	30	*	*	0.068	0.060	*	*	**0.967**	0.960
	50	50	*	*	0.065	0.049	*	*	**1.000**	**1.000**
	50	100	0.046	0.037	0.058	0.044	**1.000**	**1.000**	**1.000**	**1.000**
(II)	10	30	0.060	0.060	0.076	0.066	**0.108**	**0.108**	0.072	0.096
	10	50	0.057	0.057	0.072	0.057	**0.195**	**0.195**	0.114	0.140
	10	100	0.062	0.062	0.065	0.038	**0.478**	**0.478**	0.248	0.341
	30	30	*	*	0.086	0.077	*	*	**0.228**	0.209
	30	50	0.052	0.058	0.086	0.065	0.292	0.441	0.405	**0.442**
	30	100	0.053	0.053	0.070	0.042	0.835	**0.889**	0.858	0.869
	50	30	*	*	0.076	0.056	*	*	0.442	**0.497**
	50	50	*	*	0.081	0.062	*	*	0.714	**0.771**
	50	100	0.065	0.042	0.072	0.049	0.947	**1.000**	0.995	0.999
(III)	10	30	0.061	0.061	0.064	0.057	0.214	0.214	0.236	**0.260**
	10	50	0.055	0.055	0.081	0.064	0.448	0.448	0.325	**0.450**
	10	100	0.050	0.050	0.071	0.052	**0.881**	**0.881**	0.761	0.848
	30	30	*	*	0.079	0.068	*	*	**0.661**	0.625
	30	50	0.050	0.045	0.073	0.053	0.640	0.915	**0.938**	0.926
	30	100	0.052	0.060	0.071	0.046	1.000	0.999	**1.000**	0.999
	50	30	*	*	0.074	0.065	*	*	0.964	**0.970**
	50	50	*	*	0.081	0.067	*	*	**0.999**	**0.999**
	50	100	0.051	0.055	0.064	0.051	**1.000**	**1.000**	**1.000**	**1.000**

Table : $q = 10$, * means not available

For $q = 5$ one can see in the table above that the significance level of test T_3 is affected, while the other tests seems to attain the significance level.

For $q = 10$ all test are good in terms of ASL.

Test T_4 seems to have similar or greater power than then the other test, even for small p and large N, for $q = 5$ and $q = 10$.

Note, that for small p and large N test T_1, based on the MLE, should perform good.

We will also see the test statistics T_3 and T_4 are robust under non-normality. Assume data is generated from the model

$$\boldsymbol{X} = \boldsymbol{B}\boldsymbol{\xi}\boldsymbol{A} + \boldsymbol{\Sigma}^{-1}\boldsymbol{Z},$$

where the elements $\boldsymbol{Z} = (z_{ij})$ are independently distributed as either of the following three distributions

$$\begin{aligned} \text{(i)}\ z_{ij} &\sim N(0,1)\ \text{(as above)}, \\ \text{(ii)}\ z_{ij} &\sim \frac{\chi^2(2)-2}{2}, \\ \text{(iii)}\ z_{ij} &\sim \frac{\chi^2(8)-8}{4}. \end{aligned}$$

Observe that the skewness and kurtosis of $\chi^2(m)$ is, respectively, $\sqrt{8/m}$ and $3 + 12/m$. Hence, $\chi^2(2)$ has higher skewness and kurtosis, 2 and 9 respectively, compare to $\chi^2(8)$ with 1 and 4.5, respectively.

			Normal		$\chi^2(2)$		$\chi^2(8)$	
	p	N	T_3	T_4	T_3	T_4	T_3	T_4
(I)	10	30	0.094	0.072	0.076	0.065	0.070	0.057
	10	50	0.073	0.058	0.055	0.045	0.071	0.058
	10	100	0.061	0.042	0.056	0.045	0.056	0.037
	30	30	0.086	0.073	0.070	0.059	0.078	0.066
	30	50	0.076	0.054	0.071	0.053	0.066	0.047
	30	100	0.068	0.053	0.082	0.058	0.055	0.032
	50	30	0.073	0.062	0.061	0.054	0.086	0.073
	50	50	0.074	0.064	0.072	0.059	0.069	0.056
	50	100	0.073	0.050	0.047	0.034	0.067	0.054
(II)	10	30	0.072	0.058	0.079	0.060	0.078	0.056
	10	50	0.064	0.046	0.073	0.050	0.057	0.046
	10	100	0.082	0.060	0.060	0.043	0.071	0.053
	30	30	0.074	0.065	0.068	0.067	0.078	0.062
	30	50	0.055	0.043	0.060	0.040	0.069	0.047
	30	100	0.060	0.041	0.053	0.050	0.074	0.048
	50	30	0.082	0.061	0.071	0.056	0.087	0.069
	50	50	0.067	0.054	0.068	0.049	0.061	0.036
	50	100	0.079	0.058	0.063	0.050	0.068	0.051
(III)	10	30	0.085	0.074	0.068	0.056	0.068	0.052
	10	50	0.077	0.052	0.062	0.051	0.073	0.052
	10	100	0.058	0.045	0.071	0.045	0.062	0.041
	30	30	0.077	0.059	0.091	0.078	0.078	0.062
	30	50	0.067	0.055	0.070	0.052	0.062	0.052
	30	100	0.055	0.044	0.062	0.049	0.061	0.044
	50	30	0.068	0.048	0.084	0.073	0.079	0.068
	50	50	0.065	0.053	0.083	0.065	0.067	0.048
	50	100	0.066	0.047	0.089	0.067	0.062	0.047

Table : Attained significance level, $q = 10$

	p	N	Normal T_3	Normal T_4	$\chi^2(2)$ T_3	$\chi^2(2)$ T_4	$\chi^2(8)$ T_3	$\chi^2(8)$ T_4
(I)	10	30	0.383	0.336	0.386	0.373	0.400	0.389
	10	50	0.696	0.661	0.692	0.670	0.641	0.633
	10	100	0.979	0.976	0.969	0.969	0.973	0.973
	30	30	0.946	0.931	0.954	0.949	0.945	0.946
	30	50	1.000	1.000	0.998	0.997	0.999	0.999
	30	100	1.000	1.000	1.000	1.000	1.000	1.000
	50	30	0.990	0.983	0.983	0.981	0.983	0.977
	50	50	1.000	1.000	1.000	1.000	1.000	1.000
	50	100	1.000	1.000	1.000	1.000	1.000	1.000
(II)	10	30	0.121	0.164	0.123	0.155	0.121	0.145
	10	50	0.204	0.246	0.220	0.252	0.235	0.267
	10	100	0.423	0.499	0.474	0.536	0.478	0.515
	30	30	0.519	0.482	0.539	0.513	0.541	0.525
	30	50	0.888	0.896	0.864	0.855	0.848	0.855
	30	100	0.998	0.999	0.997	0.997	0.998	0.997
	50	30	0.663	0.683	0.681	0.683	0.622	0.640
	50	50	0.952	0.956	0.935	0.954	0.936	0.948
	50	100	1.000	1.000	0.999	1.000	1.000	1.000
(III)	10	30	0.288	0.290	0.298	0.328	0.314	0.353
	10	50	0.553	0.602	0.562	0.603	0.565	0.606
	10	100	0.958	0.968	0.935	0.965	0.944	0.966
	30	30	0.955	0.949	0.940	0.942	0.946	0.933
	30	50	1.000	1.000	0.999	0.999	1.000	0.999
	30	100	1.000	1.000	1.000	1.000	1.000	1.000
	50	30	0.994	0.992	0.988	0.991	0.989	0.984
	50	50	1.000	1.000	1.000	1.000	1.000	1.000
	50	100	1.000	1.000	1.000	1.000	1.000	1.000

Table : Power, $q = 10$

- The MLE for the mean for a Growth Curve model is a weighted estimator with the inverse of the sample covariance matrix, which is very unstable for p close to N and singular for N less than p. This fact makes the MLE not suitable for *'large p and small N'*.
- We have modified the MLE to an unbiased and unweighted estimator, just by removing the inverse of the sample covariance matrix.
- We have proposed three new test statistic, T_2 which is based on the unweighted estimator and T_3 and T_4 which can handle the high-dimensional case, when $p > N$.
- The test statistic T_3 is invariant under the group of orthogonal transformations, but not invariant under a diagonal transformation, which is an undesirable feature.
- Test statistic T_4 has the benefit of being invariant under diagonal transformations.

- We compare the LRT T_1, based on the MLE, with the three new tests proposed in this paper.
- For *'large p and small N'* the attained significance levels (ASL) are better and controlled for test statistic T_4.
- We have also shown that the power for T_4 seems to be similar or greater than the other tests, even for *'small p and large N'*.

- Kollo, T. and von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices*. Springer, Dordrecht, The Netherlands.
- Potthoff, R. F., and Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curves. *Biometrika* 51:313–326.
- Rao, C. R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Vol I (L. M. LeCam and J. Neyman, eds.). University of California Press, Berkeley, pages 355–372.
- Srivastava, M. S. (2005). Some tests concerning the covariance matrix in high dimensional data. *Journal of the Japan Statistical Society* 35(2):251– 272.
- Srivastava, M. S. (2007). Multivariate theory for analyzing high dimensional data. *Journal of the Japan Statistical Society* 37(1):53–86.

- Srivastava, M. S., and Du, M. (2008). A test for the mean vector with fewer observations than the dimension. *Journal of Multivariate Analysis* 99:386–402.
- Srivastava, M. S. and Fujikoshi, Y. (2006). Multivariate analysis of variance with fewer observations than the dimension. *Journal of Multivariate Analysis* 97(9):1927–1940.
- Srivastava, M. S. and Khatri, C. G. (1979). *An Introduction to Multivariate Statistics*. Elsvier North Holland, New York, USA.
- Srivastava, M. S. and Singull, M. (2014). Test for the mean in a Growth Curve model for high dimensions. Submitted to *Journal of Multivariate Analysis*.
- Yamada, T., and Srivastava, M. S. (2012). A test for multi-variate analysis of variance in high dimension. *Communication in Statistics – Theory and Methods* 41:2602–2615.

For more references see

- Srivastava, M. S. (2013). On testing the equality of mean vectors in high dimension. *Acta et Commentationes Universitatis Tartuensis de Mathematica* 17(1):31–56.