A novel trace test for the mean parameters in a multivariate growth curve model

Jemila S. Hamid, Joseph Beyene, and Dietrich von Rosen

Research Report
Centre of Biostochastics

Swedish University of Agricultural Sciences

Report 2009:04
ISSN 1651-8543

A novel trace test for the mean parameters in a multivariate growth curve model

JEMILA S. HAMID[1] and JOSEPH BEYENE

Biostatistics Methodology Unit, Child Health Evaluative Sciences, Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Canada

DIETRICH VON ROSEN

Centre of Biostochastics
Swedish University of Agricultural Sciences
P.O. Box 7032, SE-750 07 Uppsala, Sweden

Abstract

A trace test for the mean parameters of the Growth Curve Model is proposed. It is constructed using the restricted maximum likelihood followed by an estimated likelihood ratio approach. The statistic reduces to the Lawley-Hotelling's trace test for the classical multivariate analysis of variance model. Our test statistic is, therefore, a natural extension of the classical trace test. We show that the distribution of the test under the null hypothesis does not depend on the unknown covariance matrix Σ. We also show that the distributions under the null and alternative hypotheses can be represented as sums of weighted central and non-central chi-square random variables, respectively. Under the null hypothesis, the Satterthwaite approximation is used to get an approximate critical point. A novel Satterthwaite type approximation is proposed to obtain an approximate power. A numerical example is provided as an illustration in which the data consists of two groups where measurements at four time points are taken from each individual.

Keywords: Estimated likelihood, growth curve model, Lawley-Hotelling trace test, restricted likelihood, Satterthwaite approximation, sums of weighted chi-squares.

[1]E-mail address to the correspondence author: jemila@utstat.toronto.edu

1 Introduction

The Growth Curve Model (GCM) is a Generalized Multivariate Analysis of Variance (GMANOVA) model which is especially useful for growth curve applications (Potthoff and Roy, 1964). It plays important roles in the study of repeated measurements and longitudinal data. It is particularly useful when we have short to moderate time series where one can not apply standard time series approaches. GCM has many applications and arises in many situations. Its principal application, among other things, is in analyzing trends or growth curves which are extensively applied in biostatistics, epidemiology and medical research (Pan and Fang, 2002; von Rosen, 1989). The model also arises when we have a linearly constrained mean and hence it is a natural generalization of the MANOVA model (Potthoff and Roy, 1964; Khatri, 1966; Kollo and von Rosen, 2005).

Suppose, for example, that we have k groups where q repeated observations are taken on a given experimental unit in each group. If these observations can be associated with some continuous variable, such as time, temperature or concentration, then they form a response curve which is assumed to be a polynomial in time of degree p-1. The expected value of the i^{th} group can, therefore, be described as (Pan and Fang, 2002)

$$b_{0,i} + b_{1,i}t + ... + b_{p-1,i}t^{p-1} + \epsilon_i, \quad i = 1, 2, ..., k. \tag{1}$$

This can formally be modeled using the following multivariate bilinear setup, which is referred to as the GCM (Pan and Fang, 2002; Kollo and von Rosen, 2005)

$$\mathbf{X} = \mathbf{ABC} + \mathbf{E}, \tag{2}$$

where $\mathbf{X} : p \times n$ and $\mathbf{B} : q \times k$ are the observation and parameter matrices, respectively; $\mathbf{A} : p \times q$ and $\mathbf{C} : k \times n$ are the within and between individual design matrices, respectively. The columns of $\mathbf{E}$ are assumed to be independently distributed as a p-variate normal distribution with mean zero and an unknown positive definite covariance matrix $\mathbf{\Sigma}$.

It is important to note here that $\mathbf{B}$ in (2) is an unknown parameter matrix consisting of the coefficients of the polynomials described in (1) whereas $\mathbf{A}$ and $\mathbf{C}$ are known design matrices. Moreover, the between individual design matrix $\mathbf{C}$ is precisely the same design matrix as used in the theory of univariate and multivariate linear models which includes univariate analysis of variance and regression models (Kollo and von Rosen, 2005). If $\mathbf{A}$=$\mathbf{I}$, GCM reduces to the ordinary MANOVA model. Note also that the mean structure in (2) is bilinear contrary to the MANOVA model which has a linear structure. Consequently, the maximum likelihood estimator of its mean parameters is a non linear random expression which causes many difficulties when considering inference (Kollo and von Rosen, 2005).

GCM was first introduced by Potthoff and Roy (1964) although similar growth curve situations had been considered earlier. The first paper considering growth curves appeared in Wishart (1938). Since then different aspects of the model have been studied, among many others, by Rao (1965), Khatri (1966), Kollo and von Rosen (2005). We refer to von Rosen (1991) and Srivastava and von Rosen (1999) for general overview of the model. Kshirsagar and Smith (1995) provided a simple introduction about the model from an applied perspective. Statistical diagnostics about GCMs is discussed in Pan and Fang (2002). The latter also gives an excellent background with illustrations using several practical examples. More advanced theory about the model can be found in Kollo and von Rosen (2005).

Hypothesis tests in the MANOVA model have been considered by many and several test statistics have been proposed (Anderson, 2003). Among them are Wilks Lamda, Roy's maximum root and Lawely-Hotelling trace tests. To our knowledge, except in special cases, there is no exact test for testing the mean structure in the growth curve model. Approximate tests based on the likelihood ratio are available and some authors have proposed tests applicable under some special cases (Chi and Weerahandi, 1998; Lin and Lee, 2003).

In this paper, we propose an intuitive, simple and practical alternative to the likelihood ratio test. The test statistics is constructed using a restricted likelihood approach followed by an estimated likelihood. We demonstrate that the proposed test statistics reduces to the Lawley-Hotelling trace test when A=I (that is, for the MANOVA model). Our test can, therefore, be considered as an extension of the Lawely-Hotelling test to GCM. Moreover, we show that our test statistic is a function of von Rosen's (1995) residuals, defined by taking the bilinear structure in the model into account. This desirable property allows interpretability and easy understanding in practical situations. Residuals in univariate models have been used as diagnostic tools for validating estimated models. Their application, however, is limited in multivariate models in general and the growth curve model in particular. Diagnostics for GCM has been considered by many (Liski, 1991; Pan and Fang, 2002; von Rosen, 1995; Hamid and von Rosen, 2006). However, there has been no studies connecting dispersion estimation with the residuals in the model.

We show that the distribution of our proposed statistic under the null hypothesis is independent of the unknown covariance matrix Σ and provide its expected value. We also show that the exact distributions of the test statistic under the null and alternative hypotheses can be represented as weighted sums of central and non-central chi-square random variables. We use a conditional approach where we condition on a natural ancillary statistic to derive an approximate distribution both under the null and alternative hypotheses.

2 A trace test for the mean in GCM

Suppose that GCM given in (2) has been fitted to data and we would like to know if the assumed model fits the data well. In this section we formulate the hypothesis and propose suitable statistic for testing this hypothesis. We use a restricted maximum likelihood approach followed by an estimated likelihood to construct test. See Searle et al. (1992) for more details about the procedure. First we write the likelihood as a product of two terms and maximize the second part of the likelihood to get an estimator for the covariance matrix $\mathbf{\Sigma}$. We the replace the unknown covariance matrix by its estimator to get the estimated likelihood which is then maximized under H_o and $H_o \cup H_1$, where H_o and H_1 are the null and alternative hypotheses, respectively. We discuss why the test appears to be natural, reasonable and easy to apply in practice. We also try to interpret the statistic in connection with the corresponding interpretation for the residuals given in the growth curve model.

Suppose that we have fitted a GCM and we would like to test the hypothesis that,

$$H_o : \mathbf{B} = \mathbf{0},$$
$$H_1 : \mathbf{B} \neq \mathbf{0}.$$

Now consider the likelihood function for GCM which is given by

$$L = \alpha |\mathbf{\Sigma}|^{-\frac{n}{2}} e^{-\frac{1}{2} tr\{\mathbf{\Sigma}^{-1}(\mathbf{X}-\mathbf{ABC})(\mathbf{X}-\mathbf{ABC})'\}},$$

where $\alpha = (2\pi)^{-\frac{1}{2}np}$. We can rewrite the above likelihood function as a product of two terms

$$L = \alpha e^{-\frac{1}{2} tr\{\mathbf{\Sigma}^{-1}(\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{C}-\mathbf{ABC})(\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{C}-\mathbf{ABC})'\}} |\mathbf{\Sigma}|^{-\frac{n}{2}} e^{-\frac{1}{2} tr\{\mathbf{\Sigma}^{-1}\mathbf{S}\}},$$

where $\mathbf{S} = \mathbf{X}(\mathbf{I} - \mathbf{C}'(\mathbf{CC}')^{-}\mathbf{C})\mathbf{X}'$ and the superscript "-" represents the generalized inverse.

Let us proceed by taking the second part of the likelihood which is given by

$$|\mathbf{\Sigma}|^{-\frac{n}{2}} e^{-\frac{1}{2} tr\{\mathbf{\Sigma}^{-1}\mathbf{S}\}}.$$

Maximize the above expression to get an estimator for the covariance matrix $\mathbf{\Sigma}$ which is given by (Srivastava and Khatri, 1979; Kollo and von Rosen, 2005)

$$\hat{\mathbf{\Sigma}} = \frac{1}{n}\mathbf{S}.$$

Once again, consider the likelihood function but this time use the estimator of the covariance matrix instead of the covariance matrix itself. The likelihood reduces to the following expression

$$EL = \alpha_1|\mathbf{S}|^{-\frac{n}{2}}e^{-\frac{1}{2}ntr\{\mathbf{S}^{-1}(\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{C}-\mathbf{ABC})(\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{C}-\mathbf{ABC})'\}}, \quad (3)$$

where $\alpha_1 = n^{\frac{n}{2}}(2\pi)^{-\frac{1}{2}np}e^{-\frac{1}{2}pn}$ and EL stands for estimated likelihood. Next we maximize the expression in (4) under H_o and $H_o \cup H_1$. Under H_o, i.e., when $\mathbf{B} = \mathbf{0}$, and the maximum of the estimated likelihood equals

$$\alpha_1|\mathbf{S}|^{-\frac{n}{2}}e^{-\frac{1}{2}ntr\{\mathbf{S}^{-1}\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{CX}'\}}. \quad (4)$$

Under $H_o \cup H_1$, the maximum of the estimated likelihood can be obtained by replacing the observed mean structure $\mathbf{ABC}$ by its estimated maximum likelihood estimator which is given by

$$\mathbf{A\hat{B}C} = \mathbf{A}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-}\mathbf{A}'\mathbf{S}^{-1}\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{C}.$$

The maximum of (4) under the alternative, therefore, becomes

$$\alpha_1|\mathbf{S}|^{-\frac{n}{2}}e^{-\frac{1}{2}ntr\{\mathbf{S}^{-1}(\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{C}-\mathbf{A\hat{B}C})(\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{C}-\mathbf{A\hat{B}C})'\}}. \quad (5)$$

where $\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{C} - \mathbf{A\hat{B}C}$ is von Rosen's (1995) residual, which we will denote by $\mathbf{R}$. Note that this residual is the difference between the observed and estimated mean structures. Now define a test statistic by taking the ratio between (5) and (6), which can be written as

$$\frac{e^{-\frac{1}{2}ntr\{\mathbf{S}^{-1}\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{CX}'\}}}{e^{-\frac{1}{2}ntr\{\mathbf{S}^{-1}\mathbf{RR}'\}}}, \quad (6)$$

and the hypothesis is rejected for small values of the ratio. It is possible to show that the above ratio takes values between zero and one. Equivalently, one can use the logarithm of the ratio, which can be re-written as

$$tr\{\mathbf{S}^{-1}\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{CX}'\} - tr\{\mathbf{S}^{-1}\mathbf{RR}'\}, \quad (7)$$

where the null hypothesis is now rejected for large values of (8).

Let us now consider the residual, $\mathbf{R}$, mentioned above. It can be re-written as (Hamid and von Rosen, 2006)

$$\mathbf{R} = \mathbf{SA}^o(\mathbf{A}^{o'}\mathbf{SA}^o)^{-}\mathbf{A}^{o'}\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{C}, \tag{8}$$

where $\mathbf{A}^o$ is a matrix of full rank spanning the orthogonal complement to the column space of $\mathbf{A}$. The desired test which is given in the next proposition can then be obtained from (8) by using the expression in (9) for $\mathbf{R}$ and the fact that tr($\mathbf{AB}$)=tr($\mathbf{BA}$) for any two matrices $\mathbf{A}$ and $\mathbf{B}$ of proper sizes. Although we focus on the hypothesis in (3) in this paper, we would like to note that a test statistics for H_o : $\mathbf{GB=0}$ can easily be obtained, where $\mathbf{G}$ is a known matrix. Our method, therefore, can be used to test more general hypotheses such as group differences.

Proposition 1. *Consider the GCM given in (2). Suppose also that we are interested in testing the hypothesis provided in (3). A test statistic is given by*

$$\phi(\boldsymbol{X}) = tr\{\boldsymbol{S}^{-1}\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{XC}'(\boldsymbol{CC}')^{-}\boldsymbol{CX}'\}. \tag{9}$$

The null hypothesis is rejected when $\phi(X) > c$, where c is calculated such that $P_{H_o}(\phi(X) > c) = \alpha$, where α is the desired level for the test.

Observe that the test given in Proposition 1 is always greater or equal to zero. Moreover, the test is equivalent to the ratio given in (7) where the numerator is a function of the observed mean structure, $\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{C}$, and the denominator is a function of the residual, $\mathbf{R}$, which is obtained by subtracting the estimated mean structure from the observed mean. This means our that proposed test compares the observed mean and the residuals. In other words, the test compares the observed and estimated means and rejects the hypothesis when they are "close" to each other, i.e., when the residuals are very "small". This characteristic of the test statistic we believe is very desirable and what makes the test natural, since it is a well known fact that comparing the observed and estimated values is the proper way to evaluate an estimated model.

Recall that the classical multivariate analysis of variance model is a special case of GCM when $\mathbf{A}$=$\mathbf{I}$. Our test statistic, when $\mathbf{A}$=$\mathbf{I}$, reduces to

$$tr\{\mathbf{S}^{-1}\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{CX}'\},$$

which is the same as the well known Lawley-Hotelling trace test, sometimes called the generalized Hotelling's test. This interesting finding suggests that our test is a natural extension of the Lawely-Hotelling trace test to GCM.

Under the null hypothesis the distribution of $(\phi(\mathbf{X}))$ is independent of $\mathbf{\Sigma}$ which is not obvious to see. This fact is shown in the following theorem. An important consequence of the theorem is that under the null hypothesis one can, without loss of generality, assume that $\mathbf{\Sigma} = \mathbf{I}$. As a result, the critical point is free of any unknown parameter. On the other hand, the distribution of the test under the alternative hypothesis depends on $\mathbf{\Sigma}$. That means the power of the test depends on both $\mathbf{\Sigma}$ and $\mathbf{B}$. However, one can use an estimate of $\mathbf{\Sigma}$ to get an estimate of the power of the test which could be used as a measure of performance.

Theorem 1. *Consider the hypothesis in (3). Under the null hypothesis, the distribution of the test given in (10) is independent of the unknown covariance matrix $\mathbf{\Sigma}$.*

Proof. Let $\mathbf{A}^o$ be a matrix of full rank spanning the orthogonal complement to the space generated by the columns of $\mathbf{A}$. We can write the test $\phi(\mathbf{X})$ as

$$\phi(\mathbf{X}) = tr\{\mathbf{XC}'(\mathbf{CC}')^-\mathbf{CX}'\mathbf{S}^{-1}\} - tr\{\mathbf{XC}'(\mathbf{CC}')^-\mathbf{CX}'\mathbf{A}^o(\mathbf{A}^{o'}\mathbf{SA}^o)^{-1}\mathbf{A}^{o'}\}.$$

The first term in the above expression is invariant under the transformation $\mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{X}$. It is therefore possible to replace $\mathbf{X}$ by $\mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{X}$ which shows that the distribution of the first term is independent of $\mathbf{\Sigma}$. For the second term, we can rewrite it as

$$tr\{\mathbf{C}'(\mathbf{CC}')^-\mathbf{CX}'\mathbf{A}^o(\mathbf{A}^{o'}\mathbf{SA}^o)^{-1}\mathbf{A}^{o'}\mathbf{X}\}.$$

Now, let us write $\mathbf{A}^{o'}\mathbf{X}$ as

$$(\mathbf{A}^{o'}\mathbf{\Sigma}\mathbf{A}^o)^{\frac{1}{2}}(\mathbf{A}^{o'}\mathbf{\Sigma}\mathbf{A}^o)^{-\frac{1}{2}}\mathbf{A}^{o'}\mathbf{X}.$$

Observe that we can rewrite $(\mathbf{A}^{o'}\mathbf{\Sigma}\mathbf{A}^o)^{\frac{1}{2}}(\mathbf{A}^{o'}\mathbf{SA}^o)^{-1}(\mathbf{A}^{o'}\mathbf{\Sigma}\mathbf{A}^o)^{\frac{1}{2}}$ as

$$((\mathbf{A}^{o'}\mathbf{\Sigma}\mathbf{A}^o)^{-\frac{1}{2}}\mathbf{A}^{o'}\mathbf{X}(\mathbf{I} - \mathbf{C}'(\mathbf{CC}')^-\mathbf{C})\mathbf{X}'\mathbf{A}^o(\mathbf{A}^{o'}\mathbf{\Sigma}\mathbf{A}^o)^{-\frac{1}{2}})^{-1}.$$

Consequently, it remains to show that the distribution of $(\mathbf{A}^{o'}\mathbf{\Sigma}\mathbf{A}^o)^{-\frac{1}{2}}\mathbf{A}^{o'}\mathbf{X}$ is independent of $\mathbf{\Sigma}$. However, the expression is a linear function of a multivariate normal random variable, and as a result, it is enough to show that the mean and dispersion matrices are independent of $\mathbf{\Sigma}$.

Under the null hypothesis $\boldsymbol{E}[\mathbf{X}] = \mathbf{ABC} = \mathbf{0}$ which implies

$$\boldsymbol{E}[(\mathbf{A}^{o'}\mathbf{\Sigma}\mathbf{A}^o)^{-\frac{1}{2}}\mathbf{A}^{o'}\mathbf{X}] = \mathbf{0}.$$

Moreover,

$$\boldsymbol{D}[(\mathbf{A}^{o'}\mathbf{\Sigma}\mathbf{A}^o)^{-\frac{1}{2}}\mathbf{A}^{o'}\mathbf{X}] = (\mathbf{A}^{o'}\mathbf{\Sigma}\mathbf{A}^o)^{-\frac{1}{2}}\mathbf{A}^{o'}\mathbf{\Sigma}\mathbf{A}^o(\mathbf{A}^{o'}\mathbf{\Sigma}\mathbf{A}^o)^{-\frac{1}{2}} = \mathbf{I}.$$

□

The expected value of $\phi(\mathbf{X})$ is presented in the following theorem. One can see from the theorem that the expression for the expectation consists of two parts: one part which is independent of $\mathbf{B}$, which in fact is the expected value of the test under the null hypothesis (see Corollary 1). The other part is an "increasing" function of $\mathbf{B}$. That means the "more" $\mathbf{B}$ differs from $\mathbf{0}$, the more likely the hypothesis is to be rejected. In other words, it suggests that the power of the test increases as $\mathbf{B}$ "increases". The results from the numerical example also support this argument, however, $\mathbf{B}$ is a matrix and it is not easy to quantify the difference between $\mathbf{B}$ and $\mathbf{0}$. One way of looking at it is through the non-centrality parameter which will be discussed in the next section.

Theorem 2. *Let $\phi(\boldsymbol{X})$ be as given in Proposition 1, then*

$$\boldsymbol{E}[\phi(\boldsymbol{X})] = \beta\rho(\boldsymbol{C})\rho(\boldsymbol{A}) + \tfrac{1}{n-p-1} tr\{\boldsymbol{ABC}(\boldsymbol{ABC})'\boldsymbol{\Sigma}^{-1}\},$$

where $\beta = \frac{n-1}{(n-p-1)(n-p+\rho(\boldsymbol{A})-1)}$ *and $\rho(.)$ is the rank of a matrix.*

Proof. The expression inside the trace function in (10) is the product of two independent terms. We can therefore write the expectation as

$$\boldsymbol{E}[\phi] tr\{\boldsymbol{E}[\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{CX}']\boldsymbol{E}[\mathbf{S}^{-1}\mathbf{A}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-}\mathbf{A}'\mathbf{S}^{-1}]\}.$$

Observe that the first expectation on the right hand side of the above expression is the expectation of a non-central Wishart random variable. Therefore,

$$\boldsymbol{E}[\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{CX}'] = \rho(\mathbf{C})\boldsymbol{\Sigma} + \mathbf{ABC}(\mathbf{ABC})'.$$

For the second expectation, we write the expression in its canonical form to get

$$\begin{aligned}\boldsymbol{E}[\mathbf{S}^{-1}\mathbf{A}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-}\mathbf{A}'\mathbf{S}^{-1}] &= \tfrac{\rho(\mathbf{A})}{(n-p-1)(n-p+\rho(\mathbf{A})-1)}\boldsymbol{\Sigma}^{-1} \\ &+ \tfrac{1}{n-p+\rho(\mathbf{A})-1}\boldsymbol{\Sigma}^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A})^{-}\mathbf{A}'\boldsymbol{\Sigma}^{-1}.\end{aligned}$$

More on the canonical representation can be found in Hamid (2001). The desired result follows because

$$tr\{\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A})^{-}\mathbf{A}'\boldsymbol{\Sigma}^{-1}\} = \rho(\mathbf{A})$$

and

$$\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A})^{-}\mathbf{A}' = \mathbf{A}'$$

which completes the proof. □

Corollary 1. *Consider the hypothesis given in (3). Under the null hypothesis, the expectations in Theorem 2 reduces to*

$$\boldsymbol{E}[\phi(\boldsymbol{X})] = \beta\rho(\boldsymbol{C})\rho(\boldsymbol{A})$$

where β is as given in Theorem 2.

3 Distribution of the trace test

Unfortunately, the exact distribution for the test statistics proposed in Proposition 1 is difficult to obtain. As a result, in practical situations one needs to implement alternative approaches to calculate the critical point. For instance, one can approximate the distribution using the first two moments based on Edgeworth expansion. One can also use permutation or bootstrapping approaches to obtain an empirical null distribution. In this paper, however, we use a conditional approach where we condition on a natural ancillary statistic. That is, we calculate the critical point for a given $\mathbf{S}$, where $\mathbf{S} = \mathbf{X}(\mathbf{I} - \mathbf{C}'(\mathbf{C}\mathbf{C}')^{-}\mathbf{C})\mathbf{X}'$ is an ancillary statistic for the parameter of interest, i.e. $\mathbf{B}$. Apart from a great simplification provided by conditioning, conditioning like sufficiency and invariance, leads to a reduction of the data (Lehmann, 1986). When the problem involves ancillary statistics conditioning is appropriate since it makes the inference more relevant to the situation at hand. For GCM, $\mathbf{S} = \mathbf{X}(\mathbf{I} - \mathbf{C}'(\mathbf{C}\mathbf{C}')^{-}\mathbf{C})$ has a Wishart distribution with parameters $\mathbf{\Sigma}$ and $n - \rho(\mathbf{C})$. Its distribution, therefore, is independent of the parameter of interest, i.e. $\mathbf{B}$, although it depends on the covariance matrix $\mathbf{\Sigma}$. This shows that $\mathbf{S}$ is ancillary for B. What we shall do in this section is to find a critical point for the proposed test by conditioning on the ancillary statistic $\mathbf{S}$. Before proceeding with the conditional approach, however, we will first show that the exact distributions of the proposed test statistics under the null and alternative hypotheses can be written as a linear combinations of independent chi-square random variables. We will then use a conditional approach that allows us to use existing approximations.

Theorem 3. *Consider the test statistic provided in Proposition 1 for testing the hypothesis given in (3). Under the null hypothesis, the distribution of the test $\phi(\boldsymbol{X})$ can be represented as*

$$\phi(\boldsymbol{X}) \equiv \sum \boldsymbol{W}_{ii}\boldsymbol{\Lambda}_{ii}, \tag{10}$$

where $\boldsymbol{W}_{ii'}s$ are independently distributed as chi-square random variables with $\rho(\boldsymbol{C})$ degrees of freedom and $\boldsymbol{\Lambda}_{ii}s$ are non negative constants which are functions of S. The "$\equiv$" in equation (11) represents equivalence in distribution.

Proof. Consider the test statistic provided in Proposition 1,

$$\phi(\mathbf{X}) = tr\{\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{CX}'\mathbf{S}^{-1}\mathbf{A}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-}\mathbf{A}'\mathbf{S}^{-1}\},$$

and assume, without loss of generality, that $\mathbf{\Sigma} = \mathbf{I}$. This is possible due to the fact that the null distributions is independent of $\mathbf{\Sigma}$.

Therefore, under the null hypothesis, i.e., when B=0, we have

$$\mathbf{W} = \mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{CX}' \sim \mathbf{W}(\mathbf{I}, \rho(\mathbf{C})),$$

where $\mathbf{W}(\mathbf{I}, \rho(\mathbf{C}))$ represents a Wishart distribution with parameters $\mathbf{I}$ and $\rho(\mathbf{C})$. We may therefore rewrite $\phi(\mathbf{X})$ as

$$\phi(\mathbf{X}) = tr\{\mathbf{WP}\},$$

where $\mathbf{P} = \mathbf{S}^{-1}\mathbf{A}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-}\mathbf{A}'\mathbf{S}^{-1}$, is independent of $\mathbf{W}$. P is a symmetric positive semi-definite matrix, as a result we could decompose it as

$$\mathbf{P} = \mathbf{\Gamma\Lambda\Gamma}',$$

where $\mathbf{\Gamma}$ is an orthogonal matrix, $\mathbf{\Lambda}$ is a diagonal matrix where the diagonal elements $\mathbf{\Lambda}_{ii}$ are the ith eigenvalues of $\mathbf{P}$. On the other hand $\mathbf{W}$ can be written as the sum of $\rho(\mathbf{C})$ independent random matrices as,

$$\mathbf{W} = \sum_{i=1}^{\rho(\mathbf{C})} w_i w_{i'}, \tag{11}$$

where $w_i \sim N_p(\mathbf{0}, \mathbf{I})$ (Kollo & von Rosen (2005). Consequently,

$$\phi(\mathbf{X}) \equiv tr\{\mathbf{W\Gamma\Lambda\Gamma}'\} \equiv tr\{\mathbf{W\Lambda}\},$$

where the last statement was possible since tr$\{\mathbf{AB}\}$=tr$\{\mathbf{BA}\}$ for any two matrices, the Wishart distribution is rotation invariant and $\mathbf{\Gamma}$ is an orthogonal matrix which is independent of $\mathbf{W}$. Now using the property of the trace function we get

$$\phi(\mathbf{X}) \equiv \sum \mathbf{W}_{ii}\mathbf{\Lambda}_{ii},$$

where the $\mathbf{W}_{ii'}s$ are the diagonal elements of $\mathbf{W}$. Moreover, using the representation in (12), it is possible to show that they are independently distributed as a chi-square distribution with $\rho(\mathbf{C})$ degrees of freedom. $\square$

Theorem 4. *Let* $\mu = \sqrt{n}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{ABC}$*. The distribution of* $\phi(\boldsymbol{X})$ *under the alternative can be written as*

$$\phi(\boldsymbol{X}) \equiv \sum \boldsymbol{W}_{ii}\boldsymbol{\Lambda}_{ii}; \tag{12}$$

where $\boldsymbol{\Lambda}_{ii}$ *are positive constants,* $\boldsymbol{W}_{ii'}s$ *are independently distributed as a non-central chi-square random variable with* $\rho(\boldsymbol{C})$ *degrees of freedom and* $\lambda_i = \sum_{j=1}^{\rho(\boldsymbol{C})} \mu_{ij}^2$ *is the non-centrality parameter with* μ_{ij} *as the* $(i,j)^{th}$ *element of* μ.

Proof. Once again, consider the test statistic provided in Proposition 1. It can be rewritten as

$$\phi(\mathbf{X}) = tr\{\{\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{CX}'\Sigma^{-\frac{1}{2}}\}\{\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{S}^{-1}\mathbf{A}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-}\mathbf{A}'\mathbf{S}^{-1}\boldsymbol{\Sigma}^{\frac{1}{2}}\}\}.$$

Now consider the first part of the above expression. It is possible to show that (see, for example, Kollo and von Rosen, 2005)

$$\mathbf{W} = \boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{CX}'\boldsymbol{\Sigma}^{-\frac{1}{2}} \sim \mathbf{W}_p(\mathbf{I}, \rho(\mathbf{C}), \boldsymbol{\Delta}),$$

where $\boldsymbol{\Delta} = \mu\mu'$. Let

$$\mathbf{P} = \boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{S}^{-1}\mathbf{A}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-}\mathbf{A}'\mathbf{S}^{-1}\boldsymbol{\Sigma}^{\frac{1}{2}}.$$

$\mathbf{P}$ is a symmetric positive semi-definite matrix. As a result we can decompose it as

$$\mathbf{P} = \boldsymbol{\Gamma\Lambda\Gamma}',$$

where $\boldsymbol{\Gamma}$ is an orthogonal matrix and $\boldsymbol{\Lambda}$ is a diagonal matrix where its diagonal elements are the eigen values of $\mathbf{P}$. Under the alternative hypothesis, $\phi_{H_1}(\mathbf{X})$ can therefore be written as

$$\phi(\mathbf{X}) = tr\{\mathbf{WP}\} = tr\{\mathbf{W}\boldsymbol{\Lambda}\} = \sum \mathbf{W}_{ii}\boldsymbol{\Lambda}_{ii}.$$

However, using a representation similar to (12) for $\mathbf{W}$, it is possible to show that $\mathbf{W}_{ii'}s$ are independently distributed as non-central chi-square with $\rho(\mathbf{C})$ degrees of freedom and non-centrality parameter $\lambda_i = \sum_{j=1}^{\rho(\mathbf{C})} \mu_{ij}^2$. $\square$

As shown in Theorems 3 and 4, the distribution of the proposed test statistics can be represented as a weighted sum of central and non-central chi-square random variables, respectively. However, the weights $\boldsymbol{\Lambda}_{ii'}s$ are functions of $\mathbf{S}$ and hence random. We, therefore, consider a conditional approach where the expressions given in (11) and (13) are considered for given $\mathbf{S}$ (equivalently, on

$\mathbf{\Lambda}_{ii}$). Consequently, the $\mathbf{\Lambda}_{ii'}s$ are no longer random and can be considered as constants.

A weighted sum of independent chi-square random variables arise very frequently in practical situations, see for example Johnson & Kotz (1968), Mathai (1982) and Moschopoulos (1985). Exact distribution for the sum has been given as an infinite series form in Kotz et al. (1967), Mathai (1982) and Moschopoulos (1985) where in the latter the applications in different areas such as queue type problems and engineering were given. It was mentioned that their representation is computationally convenient since the coefficients are easily computed by simple recursive relations. However, the exact distribution is too complicated to be applied in practical situations which brings a need for a good and reasonable approximation. Several approximations have been proposed, see for example Moschopoulos (1985), where a series representation of the exact distribution is given and it was suggested that one can use a truncated version of the series where the truncation error is readily obtainable.

In this paper, we use the well known Satterthwaite approximation which will be presented briefly below. For more details about the approximation and its extension of the approximation to linear combination of independent Wishart random variables see Statterthwaite (1949) and Tan & Gupta (1983), respectively. In the latter paper, some Monte Carlo results were given to demonstrate the closeness of the approximation and the studies indicate that the approximation in general is fairly good.

Suppose there are p experimental groups where observations in each of the groups follow a normal distribution with mean zero and variance σ_i^2, the weighted sum of squares of observations can be represented as

$$Z = \sum_{i=1}^{p} a_i \sigma_i^2 \chi^2_{f_i},$$

where the $\chi^2_{f_i'}s$ are independent chi-square random variables and the $a_{i'}s$ positive constants. The well known approximation of Z is given by (see Tan and Gupta, 1983):

$$Z \sim a\chi^2_f, \tag{13}$$

where

$$a = \frac{\sum_{i=1}^{p} a_i^2 f_i \sigma_i^4}{\sum_{i=1}^{p} a_i f_i \sigma_i^2}, \quad f = \frac{(\sum_{i=1}^{p} a_i f_i \sigma_i^2)^2}{\sum_{i=1}^{p} a_i^2 f_i \sigma_i^4}.$$

The f and a in the above two equations are obtained by equating the first two moments of both sides of equation (14). However, in practical situations the $\sigma_i^{2'}s$ are unknown. In this case, a and f are estimated by replacing the

$\sigma_i^{2'} s$ by their estimates. This approximation is known as the Satterthwaite approximation (Satterthwaite, 1946).

The distribution of $\sum \mathbf{W}_{ii}\mathbf{\Lambda}_{ii}$ in Theorem 3 is, therefore, approximated by $a\chi^2_f$, where χ^2_f is a chi-square random variable with f degrees of freedom and a is a positive constant. The unknown parameters a and f are then obtained by equating the first two moments of $\sum \mathbf{W}_{ii}\mathbf{\Lambda}_{ii}$ and $a\chi^2_f$. We shall present this result in the following theorem.

Theorem 5. *Consider the test statistic given in Proposition 1. Its null distribution for a given S can be approximated by that of $a\chi^2_f$ where χ^2_f is a chi-square random variable with f degrees of freedom and a is positive constant. The constants a and f are given by*

$$a = \frac{\sum_{i=1}^{p} \mathbf{\Lambda}_{ii}^2}{\sum_{i=1}^{p} \mathbf{\Lambda}_{ii}}, \quad f = \frac{\rho(\boldsymbol{C})(\sum_{i=1}^{p} \mathbf{\Lambda}_{ii})^2}{\sum_{i=1}^{p} \mathbf{\Lambda}_{ii}^2},$$

where $\mathbf{\Lambda}_{ii}$ are the eigen values, and $\rho(\boldsymbol{C})$ is the degrees of freedom of the chi-square random variables in Theorem 3.

The critical point c($\mathbf{S}$) can then be calculated from

$$P_{H_o}(a\chi^2_f > c(\mathbf{S})) = \alpha.$$

For the distribution under the alternative hypothesis, the representation provided in Theorem 4 enables us to use existing results for such sums. However, contrary to the distribution under the null hypothesis, under the alternative it depends on the covariance matrix $\mathbf{\Sigma}$, which in many practical situations is unknown. Consequently, an estimator is needed to get an estimated power. A brief discussion about two alternative estimators is given in the next section. Theorem 4 also shows that the distribution of the test depends on the parameter B. It is important to note that this dependence is through μ and hence through the non-centrality parameters λ_i's.

Linear combinations of non-central chi-square random variables were considered among others by Press (1966) where they provided an expression for exact distribution function. It was also mentioned that this kind of distribution arises in classifying an unknown vector into one of two multivariate normal populations with unequal means and covariance matrices. However, the distribution is too complicated to be used in practical applications although there exist several algorithms to numerically solve the series, see for example Imhof (1961). Satterthwaite's approximation, used for sums of chi-square random variables, can not be used directly for sums of non-central chi-square random variables either.

Here we propose a novel approach that can be used to extend Satterthwaite's approximation to weighted sums of non-central chi-square random variables. Suppose

$$Z = \sum_{i=1}^{p} a_i \chi^2_{f_i,\lambda_i},$$

where $\chi^2_{f_i,\lambda_i}$ is distributed as non-central chi-square with f_i degrees of freedom and non-centrality parameter λ_i. Our approach involves the well-known decomposition of a non-central chi-square random variable into two independent components. One part which is distributed as a non-central chi-square distribution with one degree of freedom and non-centrality parameter λ. The second component is distributed as a central chi-square distribution with f-1 degrees of freedom.

Let $x_i \sim N(\mu_i, 1)$, i=1,2,...,f. Then it is well known that

$$\sum x_i^2 \sim \chi^2_{f,\lambda},$$

where $\lambda^2 = \sum \mu_i^2$. We can make an orthogonal transformation form $x_1, x_2, ..., x_f$ to $y_1, y_2, ..., y_f$ such that $y_1, y_2, ..., y_f$ are independently normally distributed with a unit variance and $\boldsymbol{E}(y_1) = \lambda$, and $\boldsymbol{E}(y_i) = 0$ for i=2,3,...,f (Rao, 1973). As a result,

$$\sum_{i=1}^{f} x_i^2 = \sum_{i=1}^{f} y_i^2 = y_1^2 + \sum_{i=2}^{f} y_i^2. \tag{14}$$

Observe that y_1^2 is a non-central chi-square random variable with 1 degree of freedom and non-centrality parameter λ. Whereas, $\sum_{i=2}^{f} y_i^2$ is distributed as central chi-square with f-1 degrees of freedom. Moreover, it is important to note that the two terms in (15) are independent. We can, therefore, write $\chi^2_{f,\lambda}$ as

$$\chi^2_{f,\lambda} = \chi^2_{1,\lambda} + \chi^2_{f-1}.$$

Consequently, $a\chi^2_{f,\lambda}$ and $\sum_{i=1}^{p} a_i \chi^2_{f_i,\lambda_i}$ can be written as

$$a\chi^2_{f,\lambda} = a\chi^2_{1,\lambda} + a\chi^2_{f-1},$$

$$\sum_{i=1}^{p} a_i \chi^2_{f_i,\lambda_i} = \sum_{i=1}^{p} a_i \chi^2_{1,\lambda_i} + \sum_{i=1}^{p} a_i \chi^2_{f_i-1}.$$

An approximate distribution of Z can, therefore, be given by

$$Z \sim a\chi^2_{f,\lambda},$$

where a and f in are obtained by equating the first two moments of $a\chi^2_{f-1}$ and $\sum_{i=1}^{p} a_i\chi^2_{f_i-1}$. The non-centrality parameter λ is calculated by equating the first moments of $a\chi^2_{1,\lambda}$ and $\sum_{i=1}^{p} a_i\chi^2_{1,\lambda_i}$, where a is replaced by its estimator. The estimators for a, f and λ are presented in the following theorem.

Theorem 6. *The distribution of $\phi(\boldsymbol{X}|\boldsymbol{S})$ under the alternative hypothesis can be approximated by $a\chi^2_{f,\lambda}$, where a, f and λ are given by*

$$\begin{aligned}
a &= \frac{\sum_{i=1}^{p} \mathbf{\Lambda}_{ii}^2}{\sum_{i=1}^{p} \mathbf{\Lambda}_{ii}}, \\
f &= \frac{(\rho(\boldsymbol{C})-1)[\sum_{i=1}^{p} \mathbf{\Lambda}_{ii}]^2}{\sum_{i=1}^{p} \mathbf{\Lambda}_{ii}^2} + 1, \\
\lambda &= \frac{\{\sum_{i=1}^{p} \mathbf{\Lambda}_{ii}(1+\lambda_i)\}\{\sum_{i=1}^{p} \mathbf{\Lambda}_{ii}\}}{\sum_{i=1}^{p} \mathbf{\Lambda}_{ii}^2} - 1,
\end{aligned}$$

where the $\mathbf{\Lambda}_{ii'}s$ and $\lambda_{i'}s$ are as given in Theorem 4, and $\rho(\boldsymbol{C})$ is the degrees of freedom of the non-central chi-square random variables, $\boldsymbol{W}_{ii'}s$.

The power of the test, under the alternative hypothesis, is then calculated as

$$P_{H_1}(a\chi^2_{f,\lambda} > c(\mathbf{S})).$$

Observe that the test depends on $\mathbf{B}$ only through the non-central parameter λ, and this non-centrality parameter increases the more B differs from zero. We have shown this in the numerical example given in the next section. It is also interesting to see that the power is a monotone function of λ.

4 Estimation of $\mathbf{\Sigma}$ for power calculations

Recall that the distribution of the conditional test under the alternative hypothesis depends on the unknown covariance matrix $\mathbf{\Sigma}$. As a result, the approximate distribution also depends on $\mathbf{\Sigma}$, see Theorem 6. In practical situations, one needs to find a reasonable estimator for $\mathbf{\Sigma}$ to obtain an estimator for the approximate power.

One possible estimator for $\mathbf{\Sigma}$ is (1/n)$\mathbf{S}$. This estimator was obtained by maximizing the part of the likelihood, and was used to get the estimated likelihood when defining the test. Moreover, it is possible to show that $\mathbf{S}$ provides sufficient information in absence of knowledge about the parameter $\mathbf{B}$. See a paper by Sprott (1975) on marginal and conditional sufficiency. One can also use the unbiased estimator $(1/(n-\rho(C)))\mathbf{S}$. Furthermore, the two estimators mentioned above are functions of the ancillary statistic and are considered as constants. However, for a completely specified alternative, $\mathbf{S}$ does not provide sufficient information about $\mathbf{\Sigma}$. Another estimator that might be used is

$$\hat{\mathbf{\Sigma}} = \frac{1}{n}\{\mathbf{S} + (\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{C} - \widehat{\mathbf{ABC}})(\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{C} - \widehat{\mathbf{ABC}})'\}. \tag{15}$$

This estimator gives more information than S, provided that there is some knowledge about $\mathbf{B}$. This is particularly true when $\mathbf{B}$ is known. However, using (16) as an estimator for $\mathbf{\Sigma}$ brings complications to the conditional approach which will not be discussed in this paper. Nevertheless, one can use the estimate after the data has been obtained to get an estimate for the power. This will be shown in the next section using a numerical example.

5 Numerical example

In this section we give a numerical example to illustrate the results presented in the previous sections. We consider the Potthoff and Roy (1964) data. This data was also considered by von Rosen (1995) to illustrate how one can use the residuals in the growth curve model. The data consist of dental measurements on eleven girls and sixteen boys at four ages (8, 10, 12 and 14). Each measurement is the distance, in millimeters, from the center of pituitary to pteryo-maxillary fissure. Suppose the growth curve model has been fitted to the data where the mean for both the girls and boys are assumed to follow linear growth. The design and parameter matrices are given as follows:

$$\mathbf{B} = \begin{pmatrix} b_{01} & b_{02} \\ b_{11} & b_{12} \end{pmatrix}, \qquad \mathbf{A} = \begin{pmatrix} 1 & 8 \\ 1 & 10 \\ 1 & 12 \\ 1 & 14 \end{pmatrix} \qquad \text{and} \qquad \mathbf{C}_{2\times 27} = \begin{pmatrix} \mathbf{1}_{11} & \mathbf{0}_{16} \\ \mathbf{0}_{11} & \mathbf{1}_{16} \end{pmatrix}.$$

Suppose we are interested to test the following hypothesis:

$$H_o : \mathbf{B} = \mathbf{0}$$
$$H_1 : \mathbf{B} \neq 0.$$

The proposed test is given by

$$\phi(\mathbf{X}) = tr\{\mathbf{XC}'(\mathbf{CC}')^{-}\mathbf{CX}'\mathbf{S}^{-1}\mathbf{A}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-}\mathbf{A}'\mathbf{S}^{-1}\},$$

where the hypothesis is rejected when $\phi(\mathbf{X}|\mathbf{S}) > c(\mathbf{S})$, where $c(\mathbf{S})$ is obtained from

$$P_{H_o}(a\chi^2_f > c(\mathbf{S})) = \alpha,$$

where a and f are as given in Theorem 5. The observed value of the test for the above data is $\phi(\mathbf{X}|\mathbf{S}) = 175.12$. Therefore, we reject the hypothesis if c($\mathbf{S}$) is greater than 175.12. The calculated values of a and f are 0.02 and 3, respectively. Suppose the level of significance $\alpha = 0.05$, c($\mathbf{S}$) is then obtained from

$$P_{H_o}(0.02\chi^2_3 > c(\mathbf{S})) = 0.05,$$

and the value of the c($\mathbf{S}$) obtained is 0.16, which is much smaller than the observed value for the test indicating that the data gives strong evidence towards rejecting the hypothesis that $\mathbf{B} = \mathbf{0}$. Recall from Section 2 that our test indicates the difference between the observed and estimated mean values. Our result, therefore, indicates that these two values are close to each other and hence linear growth curves seem appropriate to describe the mean structure for both the girls and the boys. Similar statement was made by Potthoff & Roy when they analyzed the data for the first time. von Rosen (1995) also showed that residuals, which are obtained as a difference between the observed and estimated means, are very small which lead to the conclusion that the assumed linear curves seem to fit the data well.

Let us now calculate the estimated power for the above test. First, we shall replace the unknown covariance matrix $\mathbf{\Sigma}$ by $\frac{1}{n}\mathbf{S}$. That is the matrices μ and $\mathbf{P}$ given in Theorem 4 become

$$\mu = \sqrt{n}\mathbf{S}^{-\frac{1}{2}}\mathbf{ABC}$$
$$P = \frac{1}{n}\mathbf{S}^{-\frac{1}{2}}\mathbf{A}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-}\mathbf{A}'\mathbf{S}^{-\frac{1}{2}}$$

Recall that both the exact and approximate powers depend on $\mathbf{B}$. That means we need to specify the value of $\mathbf{B}$ under the alternative hypothesis. Suppose that we are testing the hypothesis that $\mathbf{B} = \mathbf{0}$ against the alternative

$$\mathbf{B} = \begin{pmatrix} 7.43 & 5.84 \\ 0.48 & 0.83 \end{pmatrix}.$$

The calculated values of a and f for the data under the alternative hypothesis are 0.04 and 3, respectively. Note that these values do not depend on the value of $\mathbf{B}$ and and hence remain the same for all $\mathbf{B}$. However, the value of the non-centrality parameter depends on the choice of $\mathbf{B}$. For our data set and the above specified $\mathbf{B}$, the value of the non-centrality parameter obtained is $\lambda = 10.07$. Consequently, the estimated power of the test is calculated as

$$P = P_{\mathbf{B}}(0.04\chi^2_{3,10.07} > c(\mathbf{S})),$$

where $c(\mathbf{S}) = 0.16$. The estimated power obtained is 0.94 which is very large. Now consider another alternative, say

$$\mathbf{B} = \begin{pmatrix} 10.71 & 9.32 \\ 0.91 & 0.87 \end{pmatrix}.$$

As mentioned above the values of a and f remain the same, i.e., $a = 0.04$ and $f = 3$. However, we get $\lambda = 21.11$ which is larger than the value obtained for the previous alternative. The estimated power for this alternative is 0.998 which, as expected, is larger than that of the previous alternative. Suppose now that

$$\mathbf{B} = \begin{pmatrix} 2.91 & 1.32 \\ 0.091 & 0.087 \end{pmatrix}.$$

For this alternative, we get $\lambda = 1.89$ and the estimated power is 0.5, smaller than the above two values. We have also tried $\mathbf{B}$ values very close to zero and the power of the test becomes smaller. However, in all the alternatives we observed that the power is larger than the level of significance which suggests that the test might be unbiased. Further investigation is, however, required to show that this is always the case. We would like to note here that the above values are estimated values after $\mathbf{\Sigma}$ has been replaced by $\frac{1}{n}\mathbf{S}$. However the fact that $\mathbf{S}$ gives all information about $\mathbf{\Sigma}$ is no longer true when $\mathbf{B}$ has a specified value. Consequently, it is possible to show that the above estimator underestimates the approximated power and perhaps it is also true for the exact power as well. The estimator given (16) gives more information about $\mathbf{\Sigma}$ under the alternative hypothesis. Unfortunately, this estimator is no longer a function of only the ancillary statistic which brings complications to the problem. However, the estimate may be used at the last stage. That is, after conditioning has been done and after the data is observed. We have tried to calculate the estimated power using (16) as an estimator for $\mathbf{\Sigma}$ instead of $\hat{\mathbf{\Sigma}} = \frac{1}{n}\mathbf{S}$, and we obtained a relatively large power estimates. For instance, for the last alternative, i.e., for

$$\mathbf{B} = \begin{pmatrix} 2.91 & 1.32 \\ 0.091 & 0.087 \end{pmatrix},$$

we obtained $a = 5.01$, $f = 2$ and $\lambda = 1.02$. The estimated power obtained is 0.99 which is much larger than the previous value, which was 0.5. It is also important to note that the estimated power now depends on B not only through λ but also through a and f. As a result, the three parameters are different for different $\mathbf{B}$ values specified in the alternative.

6 Discussion

A trace test for the mean parameters in the Growth Curve Model is proposed. The test is constructed using a restricted followed by an estimated likelihood approaches. The covariance matrix $\mathbf{\Sigma}$ is estimated from one part of the likelihood function, then $\mathbf{\Sigma}$ is replaced by its estimator to get the estimated likelihood. Our proposed test statistic reduces to the Lawley-Hotelling's trace test for the classical MANOVA model indicating that our test can be regarded as the Lawley-Hotelling trace test for the GMANOVA model. Moreover, we have demonstrated that our proposed test is equivalent to a ratio of the observed and estimated growth curves. The test compares the observed and the estimated mean growth curves and the hypothesis is rejected when the differences are large. This intuitive interpretation of our test statistic provides easy understanding and we hope will encourage researchers to adapt our method in practical applications.

We showed that the distributions of the test statistic under the null and alternative hypotheses can be represented as weighted sums of central and non-central chi-square random variables, respectively. We have also provided appropriate methods for calculating approximate critical values as well as estimate the power of the test.

Results from the numerical example indicate that the performance of our proposed statistic is fairly good. Moreover, mathematical as well as numerical results suggest the test is unbiased and that the power of the proposed statistic is a monotone function of the non-centrality parameter which is a function of the mean parameter $\mathbf{B}$. However, further studies including simulation studies are required to study the power and assess the performance of our proposed statistic. We plan to investigate this issues in the future.

We would also like to note that the hypothesis given in (3) can be formulated in its general form, i.e.

$$H_o : \mathbf{GBF} = \mathbf{0},$$
$$H_1 : \mathbf{GBF} \neq \mathbf{0},$$

where $\mathbf{G}$ and $\mathbf{F}$ are any two matrices. This kind of general formulation can for example be used if one is interested in comparing two growth curves which could be done by choosing suitable elements for the matrices $\mathbf{F}$ and $\mathbf{G}$. We are also currently working on extending the test statistic to the Extended Growth Curve Model.

References

[1] Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Hoboken, NJ: John Wiley & Sons.

[2] Banken, L. (1984). *Eine Verallgemeinerung des GMANOVA Modells.* Dissertation, University of Trier, Trier, Germany.

[3] Basu, D. (1964). Recovery of ancillary information. *Sankhyā Ser. A* **26**, 3-16.

[4] Basu, D. (1975). Statistical information and likelihood. *Sankhyā Ser. A* **160**, 268-282.

[5] Basu, D. (1977). On the elimination of nuisance parameters. *J. Amer. Statist. Assoc.* **72**, 355-366.

[6] Basu, D. (1978). On partial sufficiency: A review. *J. Stat. Plann. and Inf.* **2**, 1-13.

[7] Fisher (1956). *Statistical Methods and Scientific Inference*, Oliver & Boyed, London, UK.

[8] Fujikoshi, Y. (1974). Asymptotic expansions of the non-null distributions of three statistics in GMANOVA. *Ann. Inst. Statist. Math.* **26**, 289-297.

[9] Hamid, J. (2001). Residuals in the extended growth curve model. U.U.D.M. Report 2001:29, ISSN 1101-3591, Department of Mathematics, Uppsala university, Uppsala, Sweden.

[10] Hamid, J.S., von Rosen D. (2006). Residuals in the extended growth curve model. *Scand. Journal of Statistics* **33**, 121-138.

[11] Imhof, J.P. (1961), Computing the distribution of quadratic forms in normal variables. *Biometrika* **48**, 419-426.

[12] Johnson, N.L., Kotz, S. (1968). Tables of distributions of positive definite quadratic functions in normal vectors. *Ann. Math. Statist.* **37**, 468-479.

[13] Kariya, T. (1978). The General MANOVA problem. *Ann. Statist.* **6**, 1, 200-214.

[14] Khatri, C.G. (1966). A note on a MANOVA model applied to the problems in growth curve. *Ann. Inst. Statist. Math.* **18**, 75-86.

[15] Kollo, T., von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices.* Springer, New York.

[16] Kshirsagar, A.M., Smith, W.B. (1995). *Growth Curves*, Marcel Dekker, New York.

[17] Kotz, S., Johnson, N.L., Boyd, D.W. (1967). Series representations of distributions of quadratic forms in normal variables. I. central case, *Ann. Math. Statist.* **38**, 823-837.

[18] Lehmann, E.L. (1986). *Testing Statistical Hypotheses*, Springer-Verlag, New York.

[19] Liski, E.P. (1991). Detecting influential measurements in a growth curve model. *Biometrics* **47**, 659-668.

[20] Mathai, A. M. (1982). Storage capacity of a dam with gamma type inputs. *Ann. Inst. Statist. Math.* **34**, 3, A, 591-597.
[21] Moschopoulos, P.G. (1985). The didtribution of the sum of independent gamma random variables. *Ann. Inst. Statist. Math.* **37**, A, 541-544.
[22] Pan, J.X., Fang, K.T. (1995). Multiple outlier detection in growth curve model with unstructured covariance matrix. *Ann. Inst. Statist. Math.* **47**, 137-153.
[23] Pan, J.X., Fang, K.T. (1996). Influential observations in the growth curve model with unstructured covariance matrix. *Comput. Statist. Data Anal.* **22**, 71-87.
[24] Pan, J.X., Fang, K.T. (2002). *Growth Curve Model and Statistical Diagnostics.* Springer, New York.
[25] Patterson, H.D., Thompson, R. (1971). Recovery of interblock information when block sizes are unequal. *Biometrika* **58**, 545-554.
[26] Potthoff, R.F., Roy, S.N. (1964). A generalized multivariate analysis of variance model useful specially for growth curves. *Biometrika* **51**, 313-326.
[27] Press, S.J. (1966) Linear combination of non-central chi-square random variates, *Ann. Math. Statist.* **37**, 480-487.
[28] Rao, C.R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika* **52**, 447-458.
[29] Rao, C.R. (1973). *Linear Statistical Inference and its Applications*, John Wiley & Sons, New York.
[30] von Rosen, D. (1989). Maximum likelihood estimators in multivariate linear normal models. *J. Multivariate Anal.* **31**, 187-200.
[31] von Rosen. D. (1990). Moments of a multivariate linear model with an application to the growth curve model. *J. Multivariate Anal.* **35**, 243-259.
[32] von Rosen, D. (1991). The growth curve model: A review. *Comm. Statist. Theory Methods* **20**, 2791-1822.
[33] von Rosen, D. (1995a). Influential observations in multivariate linear models. *Scand. J. Statist.* **22**, 207-222.
[34] von Rosen, D. (1995b). Residuals in the growth curve model. *Ann. Inst. Statist. Math.* **47**, 129-136.
[35] Satterthwaite, F.E. (1946). An approximate distribution of estimates of variance components. *Biometrics* **2**, 110-114.
[36] Searle, S.R. et al. (1992). *Variance Components.* Wiley, New York.
[37] Sen, A., Srivastava, M.S. (1990). *Regression Analysis, Theory and Applications.* Springer, New York.
[38] Sprott, D.A. (1975). Marginal and conditional sufficiency. *Biometrika* **62**, 3, 599-605.

[39] Srivastava, M.S., Carter, E.M. (1983). *An Introduction to Applied Multivariate Statistics.* North Holland, New York.

[40] Srivastava, M.S., Khatri, C.G. (1979). *An Introduction to Multivariate Statistics.* North Holland, New York.

[41] Tan, W.Y., Gupta, R. P. (1983). On approximating a linear combination of central Wishart matrices with possitive coefficients. *Commun. Statist.-Theor. Meth.* **12**, 22, 2589-2600.

[42] Yanagihara, H. (2001). Asymptotic expansions of the null distributions of three statistics in a non normal GMANOVA model. *Hiroshima Math. J.* **31**, 213-262.