Biostatistics (2013), **0**, 0, *pp.* 1–14
doi:10.1093/biostatistics/ms˙V4

Analysis of two-factor unreplicated experiments with one factor random

Kai Wang*

Department of Biostatistics, College of Public Health, University of Iowa, Iowa City, IA 52242, USA

kai-wang@uiowa.edu

Summary

This work concerns two-factor unreplicated experiments in which one factor is random. The model is presented in a columns-linear form. After applying identifiability conditions, model parameters are estimated by the maximum likelihood and the restricted/residual maximum likelihood. For either likelihood, likelihood ratio tests and score tests for nonadditivity of the two factors as well as the noncentrality parameters are presented. The performance of these tests is compared to Mandel's test (which is designed for the same null) using simulation studies. Application of these tests are illustrated by an example.

Key words: two-way ANOVA; additivity; random effect; Tukey test; Mandel test

1. Introduction

Research in studies of additivity in two-factor experiments that have no replication has a long history. The most famous one is Tukey's test of additivity (Tukey, 1949). This 1-df test is designed for a specific non-additivity structure. A more general test is the Mendel's test (Mandel, 1961). Recent developments include Mandel (1971), Johnson and Graybill (1972), Boik (1993), Tusell (1990), and Franck *and others* (2013). A recent review of this research on this topic is presented in Alin and Kurt (2006).

We focus on a non-additivity structure that is columns-linear. This structure is equivalent to the rows-linear structure considered by Mandel (1961). In particular, we allow the effect of the row factor to be random. For instance, in studies of comparing different methods of measurement, the effect of subjects can be regarded as random. However, tests for mixed effects in this context seem to be non-existant. Rasch *and others* (2009) simply applied the tests designed for fixed effects and studied their performance in mixed effects setting.

This research is motivated by studies comparing methods of measurement. So another focus is parameter estimation. The estimated parameters are useful in calibrating different methods. This is in contrast to existing studies on testing non-additivity.

This paper is organized as follows. We first define the model with special attention given to

* Kai Wang, Department of Biostatistics, N322 CPHB, College of Public Health, University of Iowa, Iowa City, IA 52242, USA. E-mail: `kai-wang@uiowa.edu`, phone: (319) 384-1594, fax: (319) 384-1591.

identifiability of model parameters. Model parameters are estimated by the maximum likelihood and then by the restricted maximum likelihood. Likelihood ratio tests and score tests are introduced for testing non-additivity. These tests and the Mandel's test are compared in simulation studies and an empirical study.

2. Methods

Consider a two-way unreplicated experiment with factor A and factor B. There are n levels for factor A and m levels for factor B. The response y_{ij} corresponding to level i of A and level j of level B is assumed to follow the following model

$$y_{ij} = \alpha_j + \beta_j u_i + e_{ij}, \quad e_{ij} \sim N(0, \sigma^2),\ i = 1, \ldots, n,\ j = 1, \ldots, m, \tag{2.1}$$

where α_j, β_j, u_i, and σ^2 are model parameters. In this notation, i is the row index and j the column index. This model assumes the response is linear in each column with column-specific intercept α_j and slope β_j. It is equivalent to the rows-linear model considered in Mandel (1961). However, it appears convenient to lay out data this way in studies comparing methods of measurements: the columns correspond to different methods while the rows correspond to subjects. Model (2.1) is popular in studies comparing methods of measurements (Carstensen, 2010). It extends the traditional Bland-Altman method (Bland and Altman, 1986).

Let $u. = n^{-1} \sum_i u_i$. Equations in (2.1) can be written

$$y_{ij} = (\alpha_j + \beta_j u.) + \beta_j (u_i - u.) + e_{ij}.$$

That is, without loss of generality, it can be assumed that $\sum_i u_i = 0$. α_j represents the mean response of column j on an "average" subject whose value is $u.$. Furthermore, the term $\beta_j u_i$ remains the same when β_j is multiplied by a factor while u_i is divided by the same factor and so does when β_j and u_i switch their signs. So β_j is normalized such that $\sum_i u_i^2 = 1$ and $\beta_1 \geqslant 0$. Overall, the total number of parameters in (2.1) is $2m + n - 1$.

Taking each row of matrix $\{y_{ij}\}$ as a response from a subject, it is natural to treat the row effect as random. From now on it is assumed that u_i follows a normal distribution with mean 0 and variance 1: $u_i \sim N(0, 1)$. The variance 1 reflects the scale normalization on u_i. Let $\mathbf{y}_i$ be a column vector consisting of the ith row of matrix $\{y_{ij}\}$ and

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}.$$

With these vector notations, model (2.1) assumes the following more concise form:

$$\mathbf{y}_i | u_i \sim MVN(\boldsymbol{\alpha} + u_i \boldsymbol{\beta}, \sigma^2 \mathbf{I}), \quad u_i \sim N(0, 1).$$

This is a special case of mixed effects model. It is special because the variance of the random effect u is fixed at 1. It can be shown that this model is equivalent to (details omitted)

$$\mathbf{y}_i \sim MVN(\boldsymbol{\alpha}, \boldsymbol{\Sigma}), \quad \text{where } \boldsymbol{\Sigma} = \boldsymbol{\beta}\boldsymbol{\beta}^t + \sigma^2 \mathbf{I}. \tag{2.2}$$

The null hypothesis of interest is

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_j = \beta$$

for a common β. Under this hypothesis, the effect of the row factor and that of the column factor are additive. There is no interaction between these two factors. The alternative is that H_0 does not hold.

2.1 *MLE of model parameters*

The log-likelihood function for model (2.2) is

$$l(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2) = -\frac{nm}{2}\log(2\pi) - \frac{n}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}tr\left[\boldsymbol{\Sigma}^{-1}\sum_i(\mathbf{y}_i - \boldsymbol{\alpha})(\mathbf{y}_i - \boldsymbol{\alpha})^t\right].$$

Here $tr(\mathbf{A})$ means the sum of the diagonal elements of matrix $\mathbf{A}$, i.e., the trace of $\mathbf{A}$. The first-order derivatives are (Appendix)

$$\begin{aligned}
\dot{l}_{\boldsymbol{\alpha}} &= \boldsymbol{\Sigma}^{-1}\sum_i(\mathbf{y}_i - \boldsymbol{\alpha}), \\
\dot{l}_{\boldsymbol{\beta}} &= -n\boldsymbol{\Sigma}^{-1}\left[\boldsymbol{\beta} - n^{-1}\sum_i(\mathbf{y}_i - \boldsymbol{\alpha})(\mathbf{y}_i - \boldsymbol{\alpha})^t\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}\right], \\
\dot{l}_{\sigma^2} &= -\frac{n}{2}\left[tr(\boldsymbol{\Sigma}^{-1}) - tr\left(\boldsymbol{\Sigma}^{-2}\cdot n^{-1}\sum_i(\mathbf{y}_i - \boldsymbol{\alpha})(\mathbf{y}_i - \boldsymbol{\alpha})^t\right)\right].
\end{aligned}$$

Define

$$\mathbf{S} = n^{-1}\sum_i(\mathbf{y}_i - \mathbf{y}_\cdot)(\mathbf{y}_i - \mathbf{y}_\cdot)^t$$

and $\mathbf{y}_\cdot = n^{-1}\sum_i \mathbf{y}_i$. Setting each of the three first-order derivatives to 0 and solving them simultaneously for $\boldsymbol{\alpha}, \boldsymbol{\beta}$, and σ^2, we obtain the following relationships their maximum likelihood estimates (MLE) obey (Appendix):

$$\begin{aligned}
\hat{\boldsymbol{\alpha}} &= \mathbf{y}_\cdot \\
\hat{\boldsymbol{\beta}} &= (\hat{\boldsymbol{\beta}}^t\hat{\boldsymbol{\beta}} + \hat{\sigma}^2)^{-1}\mathbf{S}\hat{\boldsymbol{\beta}}, \\
\hat{\sigma}^2 &= m^{-1}(tr(\mathbf{S}) - \hat{\boldsymbol{\beta}}^t\hat{\boldsymbol{\beta}}).
\end{aligned}$$

There is an explicit solution for $\hat{\boldsymbol{\alpha}}$ but not for $\hat{\sigma}^2$ and $\hat{\boldsymbol{\beta}}$. However, the last two expressions can be used iteratively to find approximate solutions for $\hat{\sigma}^2$ and $\hat{\boldsymbol{\beta}}$.

Under the null hypothesis H_0, the MLEs of β and σ^2 are directly available. The MLE of β is equal to the square root of $\hat{\beta}^2$, where

$$\hat{\beta}^2 = \frac{\mathbf{1}^t\mathbf{S}\mathbf{1} - tr(\mathbf{S})}{m(m-1)}$$

is the average of the off-diagonal elements of $\mathbf{S}$. The MLE of σ^2 is equal to

$$\hat{\sigma}_0^2 = \frac{tr(\mathbf{S})}{m} - \hat{\beta}^2,$$

which is the difference between the averages of the diagonal elements of $\mathbf{S}$ and its off-diagonal elements. These results are easily interpretable.

Having the MLEs of the model parameters under both the null and the alternative, the likelihood ratio test can be formed:

$$\Lambda = 2[\log l(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}, \hat{\sigma}^2) - \log l(\hat{\boldsymbol{\alpha}}, \hat{\beta}\mathbf{1}, \hat{\sigma}_0^2)].$$

Note that the MLE of $\boldsymbol{\alpha}$ is the same under the null as under the alternative. When H_0 holds, Λ asymptotically follows a chi-square distribution with degrees of freedom equal to $m-1$.

The Fisher information matrix is block-diagonal (Appendix):

$$\mathbf{F} = \begin{pmatrix} n\boldsymbol{\Sigma}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_1 \end{pmatrix}, \tag{2.3}$$

where

$$\mathbf{F}_1 = n \begin{pmatrix} (\boldsymbol{\beta}^t\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta})\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}\boldsymbol{\beta}^t\boldsymbol{\Sigma}^{-1} & \boldsymbol{\Sigma}^{-2}\boldsymbol{\beta} \\ \boldsymbol{\beta}^t\boldsymbol{\Sigma}^{-2} & 0.5tr(\boldsymbol{\Sigma}^{-2}) \end{pmatrix}.$$

According to standard asymptotic theory, a score statistic is defined by

$$T = (\dot{l}_{\boldsymbol{\alpha}^t}, \dot{l}_{\boldsymbol{\beta}^t}, \dot{l}_{\sigma^2})\mathbf{F}^{-1}(\dot{l}_{\boldsymbol{\alpha}^t}, \dot{l}_{\boldsymbol{\beta}^t}, \dot{l}_{\sigma^2})^t \text{ evaluated at the null.}$$

It is easy to verify that $\dot{l}_{\boldsymbol{\alpha}} = \mathbf{0}$, therefore

$$T = (\dot{l}_{\boldsymbol{\beta}^t}, \dot{l}_{\sigma^2})\mathbf{F}_1^{-1}(\dot{l}_{\boldsymbol{\beta}^t}, \dot{l}_{\sigma^2})^t \text{ evaluated at the null.}$$

It is also straightforward to verify that $\dot{l}_{\sigma^2} = 0$ under the null. Using block-matrix inverse formula,

$$T = n^{-1}\dot{l}_{\boldsymbol{\beta}^t}\left[(\boldsymbol{\beta}^t\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta})\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}\boldsymbol{\beta}^t\boldsymbol{\Sigma}^{-1} - \frac{2}{tr(\boldsymbol{\Sigma}^{-2})}\boldsymbol{\Sigma}^{-2}\boldsymbol{\beta}\boldsymbol{\beta}^t\boldsymbol{\Sigma}^{-2}\right]^{-1}\dot{l}_{\boldsymbol{\beta}}.$$

Since

$$\begin{aligned}\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta} &= \frac{1}{\sigma^2}\left(\mathbf{I} - \frac{1}{\boldsymbol{\beta}^t\boldsymbol{\beta} + \sigma^2}\boldsymbol{\beta}\boldsymbol{\beta}^t\right)\boldsymbol{\beta} \\ &= \frac{1}{\boldsymbol{\beta}^t\boldsymbol{\beta} + \sigma^2}\boldsymbol{\beta},\end{aligned}$$

each of the last two terms between the brackets is proportional to $\boldsymbol{\beta}\boldsymbol{\beta}^t$. Using the Sherman-Morrison formula,

$$T = n^{-1}\frac{\dot{l}_{\boldsymbol{\beta}^t}\boldsymbol{\Sigma}\dot{l}_{\boldsymbol{\beta}}}{(\boldsymbol{\beta}^t\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta})} + \text{ terms depending on } (\boldsymbol{\beta}^t\dot{l}_{\boldsymbol{\beta}})^2.$$

The terms depending on $(\boldsymbol{\beta}^t\dot{l}_{\boldsymbol{\beta}})^2$ are 0 because $\boldsymbol{\beta} = \beta\mathbf{1}$ under the H_0 and

$$\begin{aligned}\boldsymbol{\beta}^t\dot{l}_{\boldsymbol{\beta}} &\propto \mathbf{1}^t(\mathbf{1} - \hat{\boldsymbol{\Sigma}}^{-1}\mathbf{S}\mathbf{1}) \\ &= 0.\end{aligned}$$

Plugging in the MLE estimates of β and σ_0^2 under H_0 and simplifying (details omitted), the score statistic for H_0 is

$$\begin{aligned}T &= n^{-1}\frac{\boldsymbol{\beta}^t\boldsymbol{\beta} + \sigma^2}{\boldsymbol{\beta}^t\boldsymbol{\beta}}\dot{l}_{\boldsymbol{\beta}^t}\boldsymbol{\Sigma}\dot{l}_{\boldsymbol{\beta}} \\ &= \frac{n}{\hat{\sigma}_0^2 \cdot \mathbf{1}^t\mathbf{S}\mathbf{1}}[\mathbf{1}^t\mathbf{S}^2\mathbf{1} - m^{-1}(\mathbf{1}^t\mathbf{S}\mathbf{1})^2]\end{aligned}$$

Since $\mathbf{1}^t\mathbf{S}^2\mathbf{1} = (\mathbf{S1})^t(\mathbf{S1})$, statistic T depends on the variation of in the row sums of $\mathbf{S}$. If H_0 holds, no variation is expected.

The non-centrality parameter (NCP) of T is

$$\lim_{n\to\infty} n^{-1}T = \frac{\mathbf{1}^t\boldsymbol{\Sigma}^2\mathbf{1} - m^{-1}(\mathbf{1}^t\boldsymbol{\Sigma}\mathbf{1})^2}{\sigma^2(\mathbf{1}^t\boldsymbol{\Sigma}\mathbf{1})}$$
$$= \frac{m^2[Ave(\boldsymbol{\beta})]^2}{\sigma^2(m[Ave(\boldsymbol{\beta})]^2+\sigma^2)} \cdot D(\boldsymbol{\beta}),$$

where

$$Ave(\boldsymbol{\beta}) = m^{-1}\sum_j \beta_j$$

is the average of the β_js and

$$D(\boldsymbol{\beta}) = m^{-1}\sum_j \beta_j^2 - [Ave(\boldsymbol{\beta})]^2$$

measures the divergence in β_js. The power of T at significance level α is

$$\Pr(X > \chi^2_{1-\alpha,m-1})$$

where X follows a chi-square distribution with $df = m-1$ and non-centrality parameter NCP and $\chi^2_{1-\alpha,m-1}$ is the critical value from a chi-square distribution with $df = m-1$ and non-centrality parameter 0.

Since the likelihood ratio statistic Λ is asymptotically equivalent to the score statistic T, they share the same NCP.

2.2 *Restricted maximum likelihood estimation*

The maximum likelihood method is known to generate biased estimate of variance components such as σ^2. A restricted maximum likelihood is useful in reducing the bias. We consider $n-m$ linear combinations of the responses y_{ij} such that the distribution of these linear combinations are free of the parameters in the mean structure (i.e., the $\boldsymbol{\alpha}$ parameter). Define

$$\mathbf{y}^* = \begin{pmatrix} \mathbf{y}_1^* \\ \mathbf{y}_2^* \\ \vdots \\ \mathbf{y}_{n-1}^* \end{pmatrix}_{(n-1)m\times 1}$$

where $\mathbf{y}_i^* = \mathbf{y}_i - \mathbf{y}_n, i = 1, 2, \ldots, n-1$. The distribution of $\mathbf{y}^*$ is multivariate normal with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Omega}$ equal to

$$\boldsymbol{\Omega} = \begin{pmatrix} 2\boldsymbol{\Sigma} & \boldsymbol{\Sigma} & \ldots & \boldsymbol{\Sigma} \\ \boldsymbol{\Sigma} & 2\boldsymbol{\Sigma} & \ldots & \boldsymbol{\Sigma} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\Sigma} & \boldsymbol{\Sigma} & \ldots & 2\boldsymbol{\Sigma} \end{pmatrix}$$
$$= (\mathbf{I} + \mathbf{11}^t)_{(n-1)\times(n-1)} \otimes \boldsymbol{\Sigma}_{m\times m}.$$

Here "⊗" denotes the Kronecker product of two matrices. In this formulation, subject n is used as the reference. It turns out that the choice of the reference subject does not affect the restricted likelihood. The likelihood for $\mathbf{y}^*$ is the restricted likelihood for the untransformed, or the original, data. The logarithm of this likelihood is

$$l^*(\boldsymbol{\beta}, \sigma^2) = -\frac{(n-1)m}{2}\log(2\pi) - \frac{1}{2}\log|\boldsymbol{\Omega}| - \frac{1}{2}tr\left[\boldsymbol{\Omega}^{-1}\mathbf{y}^*(\mathbf{y}^*)^t\right].$$

Since

$$\begin{aligned}|\boldsymbol{\Omega}| &= |\mathbf{I} + \mathbf{1}\mathbf{1}^t|^m|\boldsymbol{\Sigma}|^{n-1} = n^m|\boldsymbol{\Sigma}|^{n-1},\\ \boldsymbol{\Omega}^{-1} &= (\mathbf{I} + \mathbf{1}\mathbf{1}^t)^{-1} \otimes \boldsymbol{\Sigma}^{-1}\\ &= (\mathbf{I} - n^{-1}\mathbf{1}\mathbf{1}^t) \otimes \boldsymbol{\Sigma}^{-1},\end{aligned}$$

and

$$\begin{aligned}tr\left[\boldsymbol{\Omega}^{-1}\mathbf{y}^*(\mathbf{y}^*)^t\right] &= tr\left[\boldsymbol{\Sigma}^{-1}\left(\sum_{i=1}^{n-1}\mathbf{y}_i^*(\mathbf{y}_i^*)^t - n^{-1}\sum_{i=1}^{n-1}\mathbf{y}_i^*\sum_{i=1}^{n-1}(\mathbf{y}_i^*)^t\right)\right]\\ &= tr\left[\boldsymbol{\Sigma}^{-1}\left(\sum_{i=1}^{n}\mathbf{y}_i^*(\mathbf{y}_i^*)^t - n^{-1}\sum_{i=1}^{n}\mathbf{y}_i^*\sum_{i=1}^{n}(\mathbf{y}_i^*)^t\right)\right] \qquad (\text{define } \mathbf{y}_n^* = \mathbf{y}_n - \mathbf{y}_n = \mathbf{0})\\ &= tr\left[\boldsymbol{\Sigma}^{-1}\left(\sum_{i=1}^{n}(\mathbf{y}_i - \mathbf{y}_n)(\mathbf{y}_i - \mathbf{y}_n)^t - n^{-1}\sum_{i=1}^{n}(\mathbf{y}_i - \mathbf{y}_n)\sum_{i=1}^{n}(\mathbf{y}_i - \mathbf{y}_n)^t\right)\right]\\ &= tr\left[\boldsymbol{\Sigma}^{-1}\left(\sum_{i=1}^{n}\mathbf{y}_i\mathbf{y}_i^t - n^{-1}\sum_{i=1}^{n}\mathbf{y}_i\sum_{i=1}\mathbf{y}_i^t\right)\right]\\ &= ntr\left(\boldsymbol{\Sigma}^{-1}\mathbf{S}\right),\end{aligned}$$

the likelihood function $l^*(\boldsymbol{\beta}, \sigma^2)$ can be written

$$l^*(\boldsymbol{\beta}, \sigma^2) = -\frac{(n-1)m}{2}\log(2\pi) - \frac{m}{2}\log(n) - \frac{n-1}{2}\log|\boldsymbol{\Sigma}| - \frac{n-1}{2}tr\left(\boldsymbol{\Sigma}^{-1}\mathbf{S}^*\right)$$

with $\mathbf{S}^* = \frac{n}{n-1}\mathbf{S}$. This likelihood does not depend on $\boldsymbol{\alpha}$. Similar to likelihood $l(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2)$, the restricted maximum likelihood estimates (REMLEs) of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ satisfy the following relationship:

$$\begin{aligned}\tilde{\boldsymbol{\beta}} &= (\tilde{\boldsymbol{\beta}}^t\tilde{\boldsymbol{\beta}} + \tilde{\sigma}^2)^{-1}\mathbf{S}^*\tilde{\boldsymbol{\beta}},\\ \tilde{\sigma}^2 &= m^{-1}(tr(\mathbf{S}^*) - \tilde{\boldsymbol{\beta}}^t\tilde{\boldsymbol{\beta}}).\end{aligned}$$

Under H_0, the REMLE of β and σ^2 can be solved explicitly:

$$\begin{aligned}\tilde{\beta}^2 &= \frac{\mathbf{1}^t\mathbf{S}^*\mathbf{1} - tr(\mathbf{S}^*)}{m(m-1)},\\ \tilde{\sigma}^2 &= \frac{tr(\mathbf{S}^*)}{m} - \tilde{\beta}^2.\end{aligned}$$

These REMLEs relate to those from $l(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2)$ in the following way:

$$\tilde{\boldsymbol{\beta}} = \sqrt{\frac{n}{n-1}}\hat{\boldsymbol{\beta}}, \quad \tilde{\sigma}^2 = \frac{n}{n-1}\sigma^2, \quad \tilde{\beta} = \sqrt{\frac{n}{n-1}}\hat{\beta}, \quad \tilde{\sigma}_0^2 = \frac{n}{n-1}\sigma_0^2. \tag{2.4}$$

The likelihood ratio statistic is

$$\Lambda^* = (n-1) \cdot [\max_{\boldsymbol{\beta},\sigma^2}(-\log|\boldsymbol{\Sigma}| - tr(\boldsymbol{\Sigma}^{-1}\mathbf{S}^*) - \max_{\beta,\sigma^2}(-\log|\boldsymbol{\Sigma}| - tr(\boldsymbol{\Sigma}^{-1}\mathbf{S}^*)].$$

It is straightforward to show that

$$\Lambda^* = \frac{n-1}{n}\Lambda.$$

It is also straightforward to verify that

$$T^* = \frac{n-1}{n}T.$$

2.3 *A Prediction problem*

A prediction problem in studies of methods of measurement is how to predict the measurement of method j, denoted by y_j, given the measurements of the others, denoted by $\mathbf{y}_{-j}$, on a subject. Let u be the unknown "true" value of the subject. The joint distribution of $(\mathbf{y}^t_{-j}, u)^t$ is

$$\begin{pmatrix} \mathbf{y}_{-j} \\ u \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{\alpha}_{-j} \\ 0 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}^{-1}_{-j,-j} & \boldsymbol{\beta}_{-j} \\ \boldsymbol{\beta}^t_{-j} & 1 \end{pmatrix}\right).$$

Here a subscript $-j$ means the jth row (or the jth column for $\boldsymbol{\Sigma}^{-1}$) is removed. We have

$$\begin{aligned} E(u|\mathbf{y}_{-j}) &= \boldsymbol{\beta}^t_{-j}\boldsymbol{\Sigma}^{-1}_{-j,-j}(\mathbf{y}_{-j} - \boldsymbol{\alpha}_{-j}) \\ &= \frac{\boldsymbol{\beta}^t_{-j}(\mathbf{y}_{-j} - \boldsymbol{\alpha}_{-j})}{\boldsymbol{\beta}^t_{-j}\boldsymbol{\beta}_{-j} + \sigma^2}, \\ Var(u|\mathbf{y}_{-j}) &= 1 - \boldsymbol{\beta}^t_{-j}\boldsymbol{\Sigma}^{-1}_{-j,-j}\boldsymbol{\beta}_{-j} \\ &= \frac{\sigma^2}{\boldsymbol{\beta}^t_{-j}\boldsymbol{\beta}_{-j} + \sigma^2}. \end{aligned}$$

A prediction of the measurement by method j is

$$\begin{aligned} y^p_j &= \alpha_j + \beta_j E(u|\mathbf{y}_{-j}) \\ &= \left(\alpha_j - \frac{\beta_j \cdot \boldsymbol{\beta}^t_{-j}\boldsymbol{\alpha}_{-j}}{\boldsymbol{\beta}^t_{-j}\boldsymbol{\beta}_{-j} + \sigma^2}\right) + \frac{\beta_j}{\boldsymbol{\beta}^t_{-j}\boldsymbol{\beta}_{-j} + \sigma^2} \cdot \boldsymbol{\beta}^t_{-j}\mathbf{y}_{-j}. \end{aligned} \tag{2.5}$$

This is a linear function in $\mathbf{y}_{-j}$. The prediction variance is

$$\begin{aligned} Var(y^p_j) &= Var(u|\mathbf{y}_{-j}) + \sigma^2 \\ &= \sigma^2 + \frac{\sigma^2}{\boldsymbol{\beta}^t_{-j}\boldsymbol{\beta}_{-j} + \sigma^2}. \end{aligned} \tag{2.6}$$

A 95% prediction limits for y_j is

$$y^p_j \pm 1.96\sqrt{Var(y^p_j)}.$$

In this calculation, the unknown parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and σ^2 are to be substituted by their respective MLEs or REMLEs.

Table 1. Simulated rejection rate under the null and the alternative. The first two values for $\boldsymbol{\beta}$ correspond to the null and the the last two values correspond to the alternative.

$\boldsymbol{\beta}$	Nominal Level	Statistic Λ	Λ^*	T	T^*	Mendel
$(1,1)^t$	0.10	0.098	0.097	0.097	0.096	0.096
	0.05	0.044	0.044	0.044	0.044	0.044
	0.01	0.015	0.015	0.013	0.013	0.013
$(1,1,1,1)^t$	0.10	0.103	0.099	0.101	0.098	0.098
	0.05	0.053	0.050	0.049	0.046	0.047
	0.01	0.008	0.007	0.007	0.006	0.007
$(1.1,1.2,1.3,1.4)^t$	0.10	0.505	0.498	0.497	0.493	0.493
	0.05	0.394	0.389	0.389	0.381	0.388
	0.01	0.198	0.194	0.185	0.177	0.185
$(1.1,1.2,1.3,1.4,1.5,1.6)^t$	0.10	0.926	0.924	0.925	0.924	0.925
	0.05	0.885	0.879	0.879	0.876	0.876
	0.01	0.717	0.709	0.705	0.701	0.705

3. Simulation Studies

The goal of this simulation study is to investigate the performance of the likelihood ratio statistics and score statistics as well as the Mandel's test (Mandel, 1961). The Mandel's test statistic is

$$Mandel = \frac{\sum_j (b_j - 1)^2 \sum_i (y_{i\cdot} - y_{\cdot\cdot})^2}{\sum_i \sum_j [(y_{ij} - y_{\cdot j}) - b_j(y_{i\cdot} - y_{\cdot\cdot})]^2} \cdot (n-2), \tag{3.7}$$

where $y_{i\cdot} = m^{-1}\sum_i y_{ij}$, $y_{\cdot j} = n^{-1}\sum_i y_{ij}$, $y_{\cdot\cdot} = (nm)^{-1}\sum_{i,j} y_{ij}$, and

$$b_j = \frac{\sum_i y_{ij}(y_{i\cdot} - y_{\cdot\cdot})}{\sum_i (y_{i\cdot} - y_{\cdot\cdot})^2}.$$

Under H_0, Mandel follows an F distribution with $m-1$ and $(m-1)(n-2)$ degrees of freedom.

Data were generated from model (2.2) with $\boldsymbol{\alpha} = \mathbf{0}$ and $\sigma^2 = 1$. For the study of type I error rate, two situations with $\boldsymbol{\beta} = (1,1)^t$ and $\boldsymbol{\beta} = (1,1,1,1)^t$, respectively, are considered. Another two $\boldsymbol{\beta}$ vectors were considered for the study of power: $\boldsymbol{\beta} = (1.1,1.2,1.3,1.4)^t$ and $\boldsymbol{\beta} = (1.1,1.2,1.3,1.4,1.5,1.6)^t$. The number of subjects n is fixed at 100. The simulated rejection rates over 1000 replicates are reported in table 1. All the tests have satisfactory type I error rate. Their power are also similar to sach other.

4. An Example

The data frame named `hba1c` in the R package `MethComp` contains measurements of HbA1c (glycosylated haemoglobin) on blood samples from 38 individuals. A venous blood sample and a capillary blood sample were obtained from each individual. Three analyzers were then used to determine the level of HbA1c in each sample. So the total number of methods is $m = 2 \times 3 = 6$. Each sample was analyzed on five different days. This data were analyzed in Carstensen (2004).

Fig. 1. Plot of the HbA1c measurement averaged over 5 different sample analysis days. One line corresponds to one measurement method.

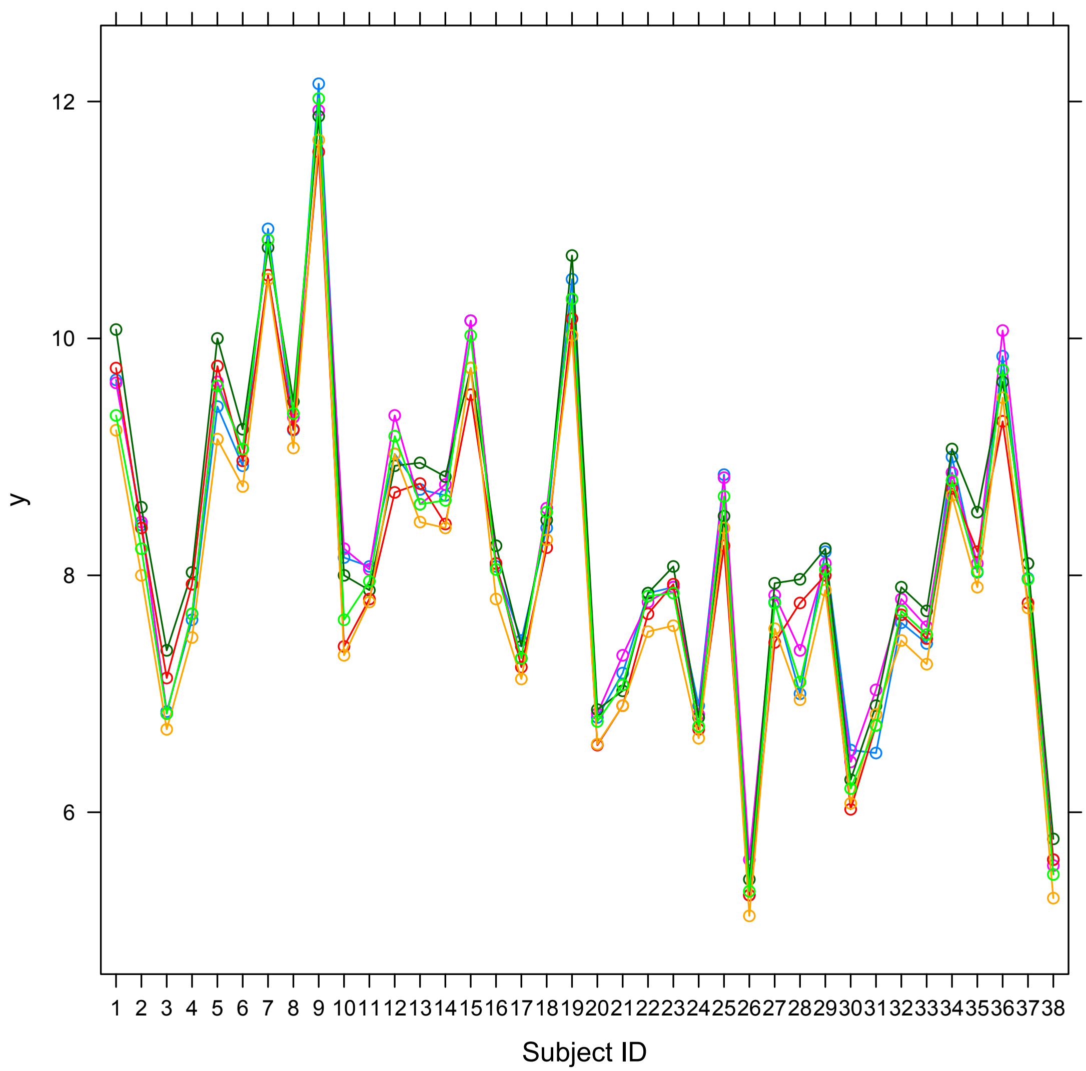

Table 2. MLEs of the model parameter for the HbA1c data. Each method consists of two parts: analyzer (`BR.V2`, `BR.VC`, or `Tosoh`) and type of blood sample (`Cap` or `Ven`). REMLEs can be obtained easily by using formulae in (2.4).

	Method					
Parameter	`BR.V2.Cap`	`BR.VC.Cap`	`Tosoh.Cap`	`BR.V2.Ven`	`BR.VC.Ven`	`Tosoh.Ven`
$\boldsymbol{\alpha}$	8.228289	8.344518	7.955263	8.281798	8.096053	8.171053
$\boldsymbol{\beta}$	1.382097	1.328687	1.329025	1.339047	1.290518	1.375324
σ^2	0.02196932					

Table 3. P-value of various tests for $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6$.

Λ	Λ^*	T	T^*	Mandel
0.08759307	0.09619505	0.09643245	0.10556074	0.1034894

In our analysis, measurements from each method over these 5 days are averaged so there is no replicated measurement per method per individual. The data is graphically presented in figure 1.

The MLEs of the model parameters are presented in table 2. There is no significant difference among the 6 β coefficients (table 3). Based on the results in table 2, prediction equations for each method given measurements from all others can be computed using equations (2.5) and (2.6). These prediction equations are shown in table 4. For instance, the prediction equation for `BR.V2.Cap` is

$$\begin{aligned}\texttt{BR.V2.Cap} = & -0.2216 + 0.2062\texttt{BR.VC.Cap} + 0.2063\texttt{Tosoh.Cap} \\ & +0.2079\texttt{BR.V2.Ven} + 0.2003\texttt{BR.VC.Ven} \\ & +0.2135\texttt{Tosoh.Ven}.\end{aligned}$$

A salient feature of these prediction equations is that they use all other methods simultaneously. In contrast, the equations reported in Carstensen (2004) are for pair-wise prediction — only one predicting method is used in each equation. The prediction standard error is about 10 times as large as reported here.

5. Discussion

We have presented analysis methods for two-factor unreplicated experiments where one factor is random. As this research is motivated by studies of comparing methods of measurement, its foci include parameter estimation, tests of additivity, and prediction of one method given measurements of other methods. In this context, treating one factor (i.e., subject) as random is more reasonable than treating it as fixed effect. In addition, it reduces the number of nuisance parameters and is expected to result in more powerful tests (although the simulated power of the proposed tests are similar to that of the Mandel's test).

The programming of the method proposed in this work is straightforward. The author's R code used for the analysis in this paper is available upon request.

Table 4.

Predicting Method	Predicted method					
	`BR.V2.Cap`	`BR.VC.Cap`	`Tosoh.Cap`	`BR.V2.Ven`	`BR.VC.Ven`	`Tosoh.Ven`
Intercept	−0.2216	0.3093	−0.1584	0.1595	0.2868	−0.2406
`BR.V2.Cap`	0	0.2029	0.2030	0.2052	0.1950	0.2130
`BR.VC.Cap`	0.2062	0	0.1952	0.1972	0.1874	0.2048
`Tosoh.Cap`	0.2063	0.1952	0	0.1973	0.1875	0.2049
`BR.V2.Ven`	0.2079	0.1966	0.1967	0	0.1889	0.2064
`BR.VC.Ven`	0.2003	0.1895	0.1896	0.1916	0	0.1989
`Tosoh.Ven`	0.2135	0.2020	0.2020	0.2042	0.1940	0
Prediction Standard Error	0.02444	0.02440	0.02440	0.02440	0.02437	0.02443

References

Alin, Aylin and Kurt, S. (2006). Testing non-additivity (interaction) in two-way anova tables with no replication. *Statistical Methods in Medical Research* **15**, 63–85.

Bland, J. M. and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **i**, 307–310.

Boik, Robert J. (1993). Testing additivity in two-way classifications with no replications:the locally best invariant test. *Journal of Applied Statistics* **20**(1), 41–55.

Carstensen, Bendix. (2004). Comparing and predicting between several methods of measurement. *Biostatistics* **5**(3), 399–413.

Carstensen, Bendix. (2010). Comparing methods of measurement: Extending the LoA by regression. *Stat Med* **29**(3), 401–410.

Franck, Christopher T., Nielsen, Dahlia M. and Osborne, Jason A. (2013). A method for detecting hidden additivity in two-factor unreplicated experiments. *Computational Statistics and Data Analysis* **67**, 95–104.

Johnson, Dallas E. and Graybill, Franklin A. (1972). An analysis of a two-way model with interaction and no replication. *Journal of The American Statistical Association* **67**(340), 862–868.

Mandel, John. (1961). Non-additivity in two-way analysis of variance. *Journal of the American Statistical Association* **56**(296), 878–888.

Mandel, John. (1971). A new analysis of variance model for non-additive data. *Technometrics* **13**(1), 1–18.

Rasch, Dieter, Rusch, Thomas, Šimečková, Marie, Kubinger, Klaus D., Moder, Karl and Šimeček, Petr. (2009). Tests of additivity in mixed and fixed effect two-way anova models with single sub-class numbers. *Stat Papers* **50**, 905–916.

Tukey, John. (1949). One degree of freedom for non-additivity. *Biometrics* **5**(3), 232–242.

Tusell, Fernando. (1990). Testing for interaction in two-way anova tables with no replication. *Computational Statistics and Data Analysis* **10**(1), 29–45.

Appendix Derivations related to maximum likelihood estimates

Because of the general relationships that

$$\frac{\partial \log|\mathbf{\Sigma}|}{\partial\theta} = \mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta}\right)$$

and

$$\frac{\partial\mathbf{\Sigma}^{-1}}{\partial\theta} = -\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta}\mathbf{\Sigma}^{-1},$$

it is straightforward to compute the first-order derivatives.

$$\dot{l}_{\boldsymbol{\alpha}} = \mathbf{\Sigma}^{-1}\sum_i(\mathbf{y}_i - \boldsymbol{\alpha}), \tag{5.8}$$

$$\begin{aligned}
\dot{l}_{\beta_m} &= -\frac{n}{2}\frac{\partial\log|\mathbf{\Sigma}|}{\partial\beta_m} - \frac{1}{2}tr\left[\frac{\partial\mathbf{\Sigma}^{-1}}{\partial\beta_m}\sum_i(\mathbf{y}_i-\boldsymbol{\alpha})(\mathbf{y}_i-\boldsymbol{\alpha})^t\right]\\
&= -\frac{n}{2}tr(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\beta_m}) + \frac{1}{2}tr\left[\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\beta_m}\mathbf{\Sigma}^{-1}\sum_i(\mathbf{y}_i-\boldsymbol{\alpha})(\mathbf{y}_i-\boldsymbol{\alpha})^t\right]\\
&= -\frac{n}{2}tr(\mathbf{\Sigma}^{-1}(\boldsymbol{\beta}\mathbf{1}_m^t + \mathbf{1}_m\boldsymbol{\beta}^t)) + \frac{1}{2}tr\left[\mathbf{\Sigma}^{-1}(\boldsymbol{\beta}\mathbf{1}_m^t + \mathbf{1}_m\boldsymbol{\beta}^t)\mathbf{\Sigma}^{-1}\sum_i(\mathbf{y}_i-\boldsymbol{\alpha})(\mathbf{y}_i-\boldsymbol{\alpha})^t\right]\\
&= -\frac{n}{2}tr(\mathbf{\Sigma}^{-1}(\boldsymbol{\beta}\mathbf{1}_m^t + \mathbf{1}_m\boldsymbol{\beta}^t)) + \frac{1}{2}\sum_i(\mathbf{y}_i-\boldsymbol{\alpha})^t\mathbf{\Sigma}^{-1}(\boldsymbol{\beta}\mathbf{1}_m^t + \mathbf{1}_m\boldsymbol{\beta}^t)\mathbf{\Sigma}^{-1}(\mathbf{y}_i-\boldsymbol{\alpha})\\
&= \mathbf{1}_m^t[-n\mathbf{\Sigma}^{-1}\boldsymbol{\beta} + \sum_i\mathbf{\Sigma}^{-1}(\mathbf{y}_i-\boldsymbol{\alpha})(\mathbf{y}_i-\boldsymbol{\alpha})^t\mathbf{\Sigma}^{-1}\boldsymbol{\beta}],
\end{aligned} \tag{5.9}$$

$$\dot{l}_{\boldsymbol{\beta}} = -n\mathbf{\Sigma}^{-1}\left[\boldsymbol{\beta} + \mathbf{\Sigma}^{-1}\cdot n^{-1}\sum_i(\mathbf{y}_i-\boldsymbol{\alpha})(\mathbf{y}_i-\boldsymbol{\alpha})^t\cdot\mathbf{\Sigma}^{-1}\boldsymbol{\beta}\right],$$

$$\begin{aligned}
\dot{l}_{\sigma^2} &= -\frac{n}{2}\frac{\partial\log|\mathbf{\Sigma}|}{\partial\sigma^2} - \frac{1}{2}tr\left[\frac{\partial\mathbf{\Sigma}^{-1}}{\partial\sigma^2}\sum_i(\mathbf{y}_i-\boldsymbol{\alpha})(\mathbf{y}_i-\boldsymbol{\alpha})^t\right]\\
&= -\frac{n}{2}tr(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\sigma^2}) + \frac{1}{2}tr\left[\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\sigma^2}\mathbf{\Sigma}^{-1}\sum_i(\mathbf{y}_i-\boldsymbol{\alpha})(\mathbf{y}_i-\boldsymbol{\alpha})^t\right]\\
&= -\frac{n}{2}tr(\mathbf{\Sigma}^{-1}) + \frac{n}{2}tr\left[\mathbf{\Sigma}^{-2}\cdot n^{-1}\sum_i(\mathbf{y}_i-\boldsymbol{\alpha})(\mathbf{y}_i-\boldsymbol{\alpha})^t\right].
\end{aligned} \tag{5.10}$$

From (5.8), the $\boldsymbol{\alpha}$ is equal to

$$\hat{\boldsymbol{\alpha}} = \mathbf{y}_{\cdot\cdot}$$

Substituting $\hat{\boldsymbol{\alpha}}$ into (5.9) and (5.10), which is equivalent to replacing $n^{-1}\sum_i(\mathbf{y}_1-\boldsymbol{\alpha})(\mathbf{y}_i-\boldsymbol{\alpha})^t$ by $\mathbf{S}$,

$$\begin{aligned}
\boldsymbol{\beta} &= \mathbf{S}\mathbf{\Sigma}^{-1}\boldsymbol{\beta}\\
&= \frac{1}{\boldsymbol{\beta}^t\boldsymbol{\beta}+\sigma^2}\mathbf{S}\boldsymbol{\beta},
\end{aligned} \tag{5.11}$$

$$tr(\mathbf{\Sigma}^{-1}) = tr(\mathbf{\Sigma}^{-2}\mathbf{S}). \tag{5.12}$$

From (5.11),

$$\boldsymbol{\beta}^t \mathbf{S} \boldsymbol{\beta} = \boldsymbol{\beta}^t \boldsymbol{\beta}(\boldsymbol{\beta}^t \boldsymbol{\beta} + \sigma^2) \tag{5.13}$$

we have from (5.12) and (5.13)

$$\sigma^2 = \frac{tr(\mathbf{S}) - \boldsymbol{\beta}^t \boldsymbol{\beta}}{m}. \tag{5.14}$$

The Fisher information is computed using standard formulae for matrix expectation.

$$\begin{aligned}
E(\dot{l}_{\boldsymbol{\alpha}} \dot{l}_{\boldsymbol{\alpha}^t}) &= n\boldsymbol{\Sigma}^{-1}, \\
E(\dot{l}_{\boldsymbol{\alpha}} \dot{l}_{\boldsymbol{\beta}^t}) &= E([\boldsymbol{\Sigma}^{-1} \sum_i (\mathbf{y}_i - \boldsymbol{\alpha})] \cdot [-n\boldsymbol{\beta}^t \boldsymbol{\Sigma}^{-1} + \sum_i \boldsymbol{\beta}^t \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\alpha})(\mathbf{y}_i - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-1}] \\
&= n\boldsymbol{\Sigma}^{-1} E[(\mathbf{y} - \boldsymbol{\alpha})\boldsymbol{\beta}^t \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\alpha})(\mathbf{y} - \boldsymbol{\alpha})^t] \boldsymbol{\Sigma}^{-1} \\
&= \mathbf{0}, \\
E(\dot{l}_{\boldsymbol{\alpha}} \dot{l}_{\sigma^2}) &= \frac{n}{2} E[\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\alpha}) \cdot (\mathbf{y} - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\alpha})] \\
&= \mathbf{0}, \\
E(\dot{l}_{\boldsymbol{\beta}} \dot{l}_{\boldsymbol{\beta}^t}) &= E(-n\boldsymbol{\Sigma}^{-1} + \sum_i \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\alpha})(\mathbf{y}_i - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-1}) \boldsymbol{\beta}\boldsymbol{\beta}^t (-n\boldsymbol{\Sigma}^{-1} + \sum_i \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\alpha})(\mathbf{y}_i - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-1}) \\
&= -n\boldsymbol{\Sigma}^{-1} \boldsymbol{\beta}\boldsymbol{\beta}^t \boldsymbol{\Sigma}^{-1} + n\boldsymbol{\Sigma}^{-1} E[(\mathbf{y} - \boldsymbol{\alpha})(\mathbf{y} - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta}\boldsymbol{\beta}^t \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\alpha})(\mathbf{y} - \boldsymbol{\alpha})^t] \boldsymbol{\Sigma}^{-1} \\
&= n\boldsymbol{\Sigma}^{-1} \boldsymbol{\beta}\boldsymbol{\beta}^t \boldsymbol{\Sigma}^{-1} + n(\boldsymbol{\beta}^t \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta}) \boldsymbol{\Sigma}^{-1} \\
E(\dot{l}_{\boldsymbol{\beta}} \dot{l}_{\sigma^2}) &= \frac{1}{2} E[-ntr(\boldsymbol{\Sigma}^{-1}) + \sum_i (\mathbf{y}_i - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\alpha})] \cdot [-n\boldsymbol{\Sigma}^{-1} + \sum_i \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\alpha})(\mathbf{y}_i - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-1}] \boldsymbol{\beta} \\
&= \frac{1}{2} E[-ntr(\boldsymbol{\Sigma}^{-1}) + \sum_i (\mathbf{y}_i - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\alpha})] \cdot [\sum_i \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\alpha})(\mathbf{y}_i - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-1}] \boldsymbol{\beta} \\
&= \frac{1}{2} E[-n^2 tr(\boldsymbol{\Sigma}^{-1}) \boldsymbol{\Sigma}^{-1} + n(n-1) tr(\boldsymbol{\Sigma}^{-1}) \boldsymbol{\Sigma}^{-1} + n(\mathbf{y} - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-2} (\mathbf{y} - \boldsymbol{\alpha}) \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\alpha})(\mathbf{y} - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-1}] \boldsymbol{\beta} \\
&= \frac{n}{2} E[-tr(\boldsymbol{\Sigma}^{-1}) + (\mathbf{y} - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-2} (\mathbf{y} - \boldsymbol{\alpha}) \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\alpha})(\mathbf{y} - \boldsymbol{\alpha})^t] \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta} \\
&= -\frac{n}{2} tr(\boldsymbol{\Sigma}^{-1}) \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta} + \frac{n}{2} \boldsymbol{\Sigma}^{-1} E[(\mathbf{y} - \boldsymbol{\alpha})(\mathbf{y} - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-2} (\mathbf{y} - \boldsymbol{\alpha})(\mathbf{y} - \boldsymbol{\alpha})^t] \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta} \\
&= -\frac{n}{2} tr(\boldsymbol{\Sigma}^{-1}) \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta} + \frac{n}{2} \boldsymbol{\Sigma}^{-1} [2\mathbf{I} + tr(\boldsymbol{\Sigma}^{-1}) \boldsymbol{\Sigma}] \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta} \\
&= n\boldsymbol{\Sigma}^{-2} \boldsymbol{\beta} \\
E(\dot{l}_{\sigma^2})^2 &= \frac{1}{4}(-n^2 [tr(\boldsymbol{\Sigma}^{-1})]^2 + n(n-1)[tr(\boldsymbol{\Sigma}^{-1})]^2 + nE[(\mathbf{y} - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-2} (\mathbf{y} - \boldsymbol{\alpha})]^2) \\
&= \frac{1}{4}(-n[tr(\boldsymbol{\Sigma}^{-1})]^2 + nE[(\mathbf{y} - \boldsymbol{\alpha})^t \boldsymbol{\Sigma}^{-2} (\mathbf{y} - \boldsymbol{\alpha})]^2) \\
&= \frac{n}{2} tr(\boldsymbol{\Sigma}^{-2})
\end{aligned}$$

[Received August 1, 2010; revised October 1, 2010; accepted for publication November 1, 2010]