Contents

Abstract

This article introduces a class of unbalanced experimental designs, called split-factorial nested designs, which allow for the estmation of both response surface effects (fixed effects of crossed factors) and variance components arising from nested random effects. An iterated least squared (ITLS) method using sufficient statistics is given for calculating maximum likelihood estimates (ML) of the parameters in a mixed model. Simulation results show that advantages for the unbalanced designs are greatest when error variance is small.

Key Words: central composite design; iterated least squares; variance components.

1 Introduction

In many experimental settings, the measured response is affected not only by the fixed effects of crossed factors, but also by the random effects (usually nested) of sampling and measurement procedures. The fixed effect estimates can be used to optimize the process, and knowing which variance source (variance component) is largest could help to focus quality improvement of the process. Such estimation is necessiated by the need to indentify various sources of variations, which is required to be controlled to improve the quality of the final product. For example, in an experiment to study certain critical dimension of molded part, machine settings such as mold zone temperatures, or screw speed could be the crossed factors of interest while shift-to-shift variation, part-to-part variation, and measurement-to-measurement variation might be the random effect. Balanced nested designs is usually used for this purpose, owing to its simplicity for statistical analysis. However, balanced nested designs have a defect in yielding more information on the lower level factors than on the leading factors, which are higher in hierarchy. To eliminate this defect, we consider designs that are mixtures of structures $(Sq)_2$, $(Sq)_1$ and $(Sq)_{2/1}$ shown as tree diagrams in Figure 1. A nested design with q second-stage units will be denoted as $(Hq)_{x,w,z}$, where x, w, and z are the numbers of $(Sq)_2$, $(Sq)_1$ and $(Sq)_{2/1}$ structures, respectively. Note that $(Sq)_2$ and $(Sq)_1$ structures are balanced with two replicates and one replicate per second-stage unit, respectively. Notation for those designs which include only structures of $(Sq)_2$ and $(Sq)_1$ will be $(Hq)_{x,w}$, the omitted subscript z being understood to be zero. Also, $(H2)_{0,0,z}$ and $(H2)_{x,w}$ are three-stage Bainbridge (1963) and Anderson (1961) designs, respectively. For the estimation of response surface effects, the fixed effects of crossed factors are often studied with a 2^k, 3^k factorial design or central composite design (ccd, see Myers and Montgomery (1995)), where k is the number of crossed factors (denoted as $A_1, \ldots, A_k$). Both crossed

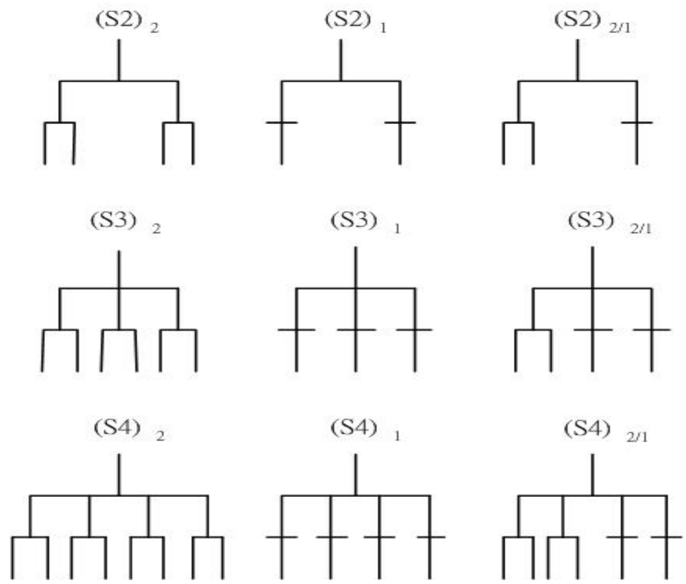

Figure 1: Digrams of $(Hq)_{x,w,z}$ designs with $q = 2, 3, 4$.

factor effects and variance components could be estimated by performing a $(Hq)_{x,w,z}$ nested design at each design point in a factorial design or ccd. However, this would require many obervations, which often is not feasible or economical. For example, for ccd, the crossed factor design points are the 2^k factorial points augmented by $2k$ axial and n_c central points of the cube, which would require $N = (q(2x + w + z) + [q/2]z) \times (2^k + 2k + n_c)$ observations, where $[m]$ is the greatest integer $\leq m$. To reduce the number of observations, we construct a split-factorial nested design by taking a $(Hq)_{x,0}$, $(Hq)_{0,w}$ or $(Hq)_{0,0.z}$ design at each crossed factor design point. Our approach is motivated by the designs proposed by Bruce et al. (2002). In their article, a methodology for designing a split factorial experiment is introduced for a 2^k factorial design and $q = 2^d$ (where d is an integer) variance components associated with nested random effects. The design points are split into q sub-experiments by d blocking generators such that the sub-experiments gathers information on only one of the q variance components, i.e., in the i^{th} $(i = 1, \ldots, q)$ sub-experiment, a nested structure that branches only at the i^{th} level (into n branches, say) will be run at each of 2^{k-d} design points. Under this designs, they showed that the ordinary least squared (OLS) estimator for fixed effects are BLUE and the variance component estimators from the mean squared errors on the ANOVA table are also minimum variance among unbiased quadratic estimators. In this note, similar approaches would be implemented to ccd. For example, for ccd, a blocking generator can be used to split the 2^k factorial points into 2 sub-experiments, each with a $(Hq)_{x,0}$ or $(Hq)_{0,w}$ nested design. Similarly, we can use an even number of n_c and split the axial and central points into 2 sub-experiments. When $w = 2x$, the total observation N is reduced by half. Table 1 shows a split-factorial nested design (denoted by $\frac{1}{2}(Hq)_{x,0} + \frac{1}{2}(Hq)_{0,w}$) with crossed factor design points from ccd ($k = 3$, $n_c = 4$ and blocking generator $B = A_1A_2A_3$). Note that design points 1 through 8 denote 2^k factorial points, points 9 through 14 denote the axial points and points 15 through 18 denote the central points.

A response model for a split-factorial nested design is

$$\mathbf{Y} = \mathbf{X}\mathbf{b} + \mathbf{u}, \tag{1.1}$$

where $\mathbf{Y}$ is an $N \times 1$ vector of observations, $\mathbf{X}$ is an $N \times p$ matrix of estimable response surface contrasts including a constant column, $\mathbf{b}$ is a vector of p unknown coefficient parameter; $\mathbf{u} = \mathbf{Z}_1\mathbf{u}_1 + \ldots + \mathbf{Z}_c\mathbf{u}_c + \mathbf{e}$, $\mathbf{u}_i \sim N(\mathbf{0}, \sigma_i^2\mathbf{I}_{m_i})$ for $i = 1, \ldots, c$, $\mathbf{e} \sim N(\mathbf{0}, \sigma_e^2\mathbf{I}_N)$; $\mathbf{u}_i$ for $i = 1, \ldots, c$ and $\mathbf{e}$ are all independent of one another. Thus, $\mathbf{u} \sim N(\mathbf{0}, \mathbf{V})$, with $\mathbf{V} = \sum_{i=1}^{c} \sigma_i^2\mathbf{G}_i + \sigma_e^2\mathbf{I}_N$ where $\mathbf{G}_i = \mathbf{Z_i}\mathbf{Z_i}^T$ for $i = 1, \ldots, c$. Note that $\mathbf{Z}_i$ ($i = 1, \ldots, c$) is an $N \times m_i$ indicator matrix associated with the i^{th} variance component and $\mathbf{u_i}$ is a $m_i \times 1$ vector consisting of normally distributed independent random effect parameters associated with the i^{th} variance components. Note that for a three-stage nested design c is equal to 2.

Now, let $\theta^T = (\mathbf{b^T}, \theta_2^T)$ where $\theta_2^T = (\sigma_1^2, \ldots, \sigma_c^2, \sigma_e^2)$. In Section 2, an iterated least squared method is proposed for calculating maximum likelihood (ML) estimation of θ in model (1.1).

Table 1: A split-factorial nested design $\frac{1}{2}(Hq)_{x,0} + \frac{1}{2}(Hq)_{0,w}$ with crossed factor design points from ccd ($k = 3$, $n_c = 4$ and $B = A_1A_2A_3$)

TRT	Design Point	A_1	A_2	A_3	B	sub-exp.
1	1	-1	-1	-1	-1	$(Hq)_{0,w}$
2	2	-1	-1	1	1	$(Hq)_{x,0}$
3	3	-1	1	1	-1	$(Hq)_{0,w}$
4	4	-1	1	-1	1	$(Hq)_{x,0}$
5	5	1	-1	1	-1	$(Hq)_{0,w}$
6	6	1	-1	-1	1	$(Hq)_{x,0}$
7	7	1	1	-1	-1	$(Hq)_{0,w}$
8	8	1	1	1	1	$(Hq)_{x,0}$
9	9	-1.732	0	0	-	$(Hq)_{0,w}$
10	10	1.732	0	0	-	$(Hq)_{x,0}$
11	11	0	-1.732	0	-	$(Hq)_{0,w}$
12	12	0	1.732	0	-	$(Hq)_{x,0}$
13	13	0	0	-1.732	-	$(Hq)_{0,w}$
14	14	0	0	1.732	-	$(Hq)_{x,0}$
15	15	0	0	0	-	$(Hq)_{0,w}$
15	16	0	0	0	-	$(Hq)_{x,0}$
15	17	0	0	0	-	$(Hq)_{0,w}$
15	18	0	0	0	-	$(Hq)_{x,0}$

2 Maximum likelihood estimation using ITLS

The non-linear estimation procedure used to obtain ML estimates in a split-factorial nested desing is known as iterated least squares, denoted by ITLS (see Anderson 1961). What follows is a general discussion of the ITLS procedure. Briefly, the method consists of obtaining from the data a set of linear statistics $\mathbf{S}_1$ needed to estimate $\mathbf{b}$ and a set of quadratic "statistics" $\mathbf{S}_2$ (i.e., "pivotal quantities", some of which may depend on $\mathbf{b}$) needed to estimate the variance components θ_2. Let $\mathbf{S} = (\mathbf{S}_1^T, \mathbf{S}_2^T)^T$ and $Var(\mathbf{S}) = \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}$.

We shall choose $\mathbf{S}_1$ and $\mathbf{S}_2$ such that $\mathbf{\Sigma}_{12} = \mathbf{\Sigma}_{21} = \mathbf{0}$. Note that this can always be achieved because we could choose $\mathbf{S}_1 = \mathbf{Y}$ and $\mathbf{S}_2 = vech(\mathbf{Q})$ to be a vector which contains the elements from the upper triangle of $\mathbf{Q} = (\mathbf{Y} - \mathbf{Xb})(\mathbf{Y} - \mathbf{Xb})^T$. Next, let $\mathbf{H}$ be a matrix of constants such that $E(\mathbf{S}) = \mathbf{H}\theta$. The ITLS estimator of θ (denoted by $\hat{\theta}$) proposed by Anderson (1961) is the solution to the equation $\mathbf{H^T\Sigma^{-1}H}\theta = \mathbf{H^T\Sigma^{-1}S}$. It is proved by Goldstein (1986) that the ITLS solution is ML if $\mathbf{S}_1$ and $\mathbf{S}_2$ are as we have just described (provided that the ITLS solution is contained within the parameter space). However, the ITLS method is most useful when we can use an orthogonal transformation to reduce the number of elements in $\mathbf{S}_1$ and $\mathbf{S}_2$ to a smaller set of sufficient statistics. It is claimed by Jennrich and Moore (1975) that if $\mathbf{S}$ contains the sufficient statistics for $\mathbf{b}$, given the variance components, and the sufficient statistics for the variance components, given $\mathbf{b}$, then the ITLS method will yield ML estimates. This is proved in Appendix under normality assumptions.

For a split-factorial $p(Hq)_{x,0}+(1-p)(Hq)_{0,w}$ $(0 \leq p \leq 1)$ design, $\mathbf{S}$ can be chosen such that $Var(\mathbf{S})$ is diagonal. For example, consider the design given in Table 1. Let n_1 and n_2 be the number of crossed factor points with a $(Hq)_{0,w}$ and $(Hq)_{x,0}$ design, respectively. Let n_{c1} and n_{c2} be the number of central points with a $(Hq)_{0,w}$ and $(Hq)_{x,0}$ design, respectively; Consider the example given in Table 1. For $i = 1, \ldots, 14$, let $\bar{y}_i$ denote the mean of observations from the i^{th} design point of Table 1. Let $\bar{y}_{c_1}$ denote the mean of observations from the 15^{th} and 17^{th} design point of Table 1 (i.e., the central points with a $(Hq)_{0,w}$ design). Similarly, let $\bar{y}_{c_2}$ denote the mean of observations from the 16^{th} and 18^{th} design point of Table 1 (i.e., the central points with a $(Hq)_{x,0}$ design). Let $\mu_i(\mathbf{b})$ $(i = 1, \ldots, 18)$ denotes the mean of the i^{th} design point, which is is a function of $\mathbf{b}$. Note that $\mu_{15}(\mathbf{b}) = \cdots = \mu_{\mathbf{18}}(\mathbf{b})$, which are the mean of central points. Let SSA_1 and SSB_1 denote the sum of squares for first-stage and second-stage, repectively from the sub-experiment with $(Hq)_{0,w}$ designs. Let SSA_2 and SSB_2 and SSE_2 denote the sum of squares for first-stage, second-stage and error, respectively, from the sub-experiment with $(Hq)_{x,0}$ designs. Let $E_1 = q\sigma_1^2 + \sigma_2^2 + \sigma_e^2$, $E_2 = 2q\sigma_1^2 + 2\sigma_2^2 + \sigma_e^2$, $E_3 = \sigma_2^2 + \sigma_e^2$, and $E_4 = 2\sigma_2^2 + \sigma_e^2$. For a $\frac{1}{2}(Hq)_{x,0} + \frac{1}{2}(Hq)_{0,w}$ design described in Table 1, the linear statistics for estimation of $\mathbf{b}$ and the quadratic statistics for estimation of variance components $\theta_{\mathbf{2}}$ along with their degrees of freedom, expectations and variances are shown in Table 2. Thus, the statistics $\mathbf{S} = (\mathbf{S_1^T}, \mathbf{S_2^T})^{\mathbf{T}}$ is given by $\mathbf{S}_1 = (\bar{y}_1, \ldots, \bar{y}_{c_1}, \bar{y}_{c_2})^T$, and $\mathbf{S}_2 = (qw(\bar{y}_1 - \mu_1)^2, \ldots, 2n_{c2}qx(\bar{y}_{c_2} - \mu_{15})^2, SSA_1, SSA_2, \ldots, SSE_2)^T$. Suppose that the response data is fitted using a second-order model with $\mathbf{b} = (b_0, b_1, b_2, b_3, b_{11}, b_{22}, b_{33}, b_{12}, b_{13}, b_{23})^T$. Thus, $E[\mathbf{S_1}] = \mathbf{H_1 b}$, $E[\mathbf{S_2}] = \mathbf{H_2}\theta_{\mathbf{2}}$, and

$$\mathbf{H} = \begin{pmatrix} \mathbf{H_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{H_2} \end{pmatrix}, \text{ where}$$

$$\mathbf{H_1} = \begin{pmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

and

$$\mathbf{H_2} = \begin{pmatrix} q & 1 & 1 \\ 2q & 2 & 1 \\ . & . & . \\ . & . & . \\ q & 1 & 1 \\ 2q & 2 & 1 \\ n_1(w-1)q & n_1(w-1) & n_1(w-1) \\ 2n_2(x-1)q & 2n_2(x-1) & n_2(x-1) \\ 0 & n_1w(q-1) & n_1w(q-1) \\ 0 & 2n_2x(q-1) & n_2x(q-1) \\ 0 & 0 & n_2xq \end{pmatrix}.$$

Also, $Var(\mathbf{S}) = \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma_{11}} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma_{22}} \end{pmatrix}$ where $\mathbf{\Sigma_{11}} = Diag(v_1, v_2, \ldots, v_1, v_2, v_3, v_4)$ and $\mathbf{\Sigma_{22}} = Diag(v_5, v_6, \ldots, v_5, v_6, v_7, v_8, v_9, v_{10}, v_{11})$.

Table 2: Table 2. Linear and quadratic statistics for a $\frac{1}{2}(Hq)_{x,0} + \frac{1}{2}(Hq)_{0,w}$ design

Statistic	*df*	*Expectation*	*Variance*
$\bar{y}_1$	-	$\mu_1(\mathbf{b})$	$E_1/wq = v_1$
$\bar{y}_2$	-	$\mu_2(\mathbf{b})$	$E_2/2xq = v_2$
.	.	.	.
.	.	.	.
$\bar{y}_{13}$	-	$\mu_{13}(\mathbf{b})$	$E_1/wq = v_1$
$\bar{y}_{14}$	-	$\mu_{14}(\mathbf{b})$	$E_2/2xq = v_2$
$\bar{y}_{c_1}$	-	$\mu_{15}(\mathbf{b})$	$E_1/n_{c1}wq = v_3$
$\bar{y}_{c_2}$	-	$\mu_{15}(\mathbf{b})$	$E_2/2n_{c2}xq = v_4$
$qw(\bar{y}_1 - \mu_1(\mathbf{b}))^{\mathbf{2}}$	1	E_1	$2E_1^2 = v_5$
$2qx(\bar{y}_2 - \mu_2(\mathbf{b}))^{\mathbf{2}}$	1	E_2	$2E_2^2 = v_6$
.	.	.	.
.	.	.	.
$qw(\bar{y}_{13} - \mu_{13}(\mathbf{b}))^{\mathbf{2}}$	1	E_1	$2E_1^2 = v_5$
$2qx(\bar{y}_{14} - \mu_{14}(\mathbf{b}))^{\mathbf{2}}$	1	E_2	$2E_2^2 = v_6$
$n_{c1}qw(\bar{y}_{c_1} - \mu_{15}(\mathbf{b}))^{\mathbf{2}}$	1	E_1	$2E_1^2 = v_5$
$2n_{c2}qx(\bar{y}_{c_2} - \mu_{15}(\mathbf{b}))^{\mathbf{2}}$	1	E_2	$2E_2^2 = v_6$
SSA_1	$n_1(w-1)$	$n_1(w-1)E_1$	$2n_1(w-1)E_1^2 = v_7$
SSA_2	$n_2(x-1)$	$n_2(x-1)E_2$	$2n_2(x-1)E_2^2 = v_8$
SSB_1	$n_1w(q-1)$	$n_1w(q-1)E_3$	$2n_1w(q-1)E_3^2 = v_9$
SSB_2	$n_2x(q-1)$	$n_2x(q-1)E_4$	$2n_2x(q-1)E_4^2 = v_{10}$
SSE_2	n_2xq	$n_2xq\sigma_e^2$	$2n_2xq\sigma_e^4 = v_{11}$

Similar to Table 1, a split-factorial $\frac{1}{2}(Hq)_{0,0,z} + \frac{1}{2}(Hq)_{0,w}$ design can be constructed by replacing $(Hq)_{x,0}$ of Table 1 with $(Hq)_{0,0,z}$. For this case, $\mathbf{S}$ can not be chosen such that $Var(\mathbf{S})$ is diagonal. To illustrate how to expediently handle a case with $z \neq 0$, consider a $(H4)_{0,0,z}$ design. Let y_{ijkl} denote the l^{th} measurement of the k^{th} second stage unit from the j^{th} first stage unit in the i^{th} design point. Allocate subscripts $k = 1, 2, 3, 4$ to the second stage units left to right as they appear in the $(S4)_{2/1}$ structure (Figure 1). For the i^{th} $(i = 2, 4, \ldots, 18)$ design point of Table 1, define $y^*_{ij1} = y_{ij11} + y_{ij12} + y_{ij21} + y_{ij22}$, $y^*_{ij2} = y_{ij31} + y_{ij41}$, and a 2×2 matrix $\mathbf{T_i} = \begin{pmatrix} T_{i11} & T_{i12} \\ T_{i12} & T_{i22} \end{pmatrix}$, where $T_{i11} = \sum_j [y^*_{ij1} - 4\mu_i(\mathbf{b})]^2$, $T_{i12} = \sum_j [(y^*_{ij1} - 4\mu_i(\mathbf{b}))(y^*_{ij2} - 2\mu_i(\mathbf{b}))]$, $T_{i22} = \sum_j [y^*_{ij2} - 2\mu_i(\mathbf{b})]^2$. Then

$$E[\mathbf{T_i}] = \mathbf{E_T} = z \begin{pmatrix} E_{11} & E_{12} \\ E_{12} & E_{22} \end{pmatrix},$$

where $E_{11} = 16\sigma_1^2 + 8\sigma_2^2 + 4\sigma_e^2$, $E_{12} = 8\sigma_1^2$, and $E_{22} = 4\sigma_1^2 + 2\sigma_2^2 + 2\sigma_e^2$. Since $\mathbf{T_i}$ $(i = 2, 4, \ldots, 18)$ is a Wishart matrix with z degrees of freedom, we have

$$Var(\mathbf{T_i}) = \mathbf{\Sigma_T} = \begin{pmatrix} 2zE_{11}^2 & 2zE_{11}E_{12} & 2zE_{12}^2 \\ 2zE_{11}E_{12} & z(E_{11}E_{22} + E_{12}^2) & 2zE_{12}E_{22} \\ 2zE_{12}^2 & 2zE_{12}E_{22} & 2zE_{22}^2 \end{pmatrix}.$$

Define $SSB_{2(2/1)} = \sum_i \sum_j [y_{ij11} + y_{ij12} - y_{ij21} - y_{ij22}]^2/4$, $SSB_{1(2/1)} = \sum_i \sum_j [y_{ij31} - y_{ij41}]^2/2$, and $SSE_{2(2/1)} = \sum_i \sum_j [(y_{ij11} - y_{ij12})^2 + (y_{ij21} - y_{ij22})^2]/4$. Note that $SSB_{2(2/1)}$ and $SSB_{1(2/1)}$ are sum of squares due to second stage units for sets with one replicate and two replicates, respectively. Clearly, $SSB_{s(2/1)}$, $s = 1, 2$, is distributed as a chi-squared random variable with $n_2 z$ degrees of freedom multiplied by $s\sigma_2^2 + \sigma_e^2$ and $SSE_{2(2/1)}$ is distributed as a chi-squared random variable with $2n_2 z$ degrees of freedom multiplied by σ_e^2. Let $\bar{y}^*_{is}$ $(i = 2, 4, \ldots, 14; s = 1, 2)$ denote the mean of observations from y^*_{ijs}; $\bar{y}^*_{c_{2s}}$ $(s = 1, 2)$ denote the mean of observations from y^*_{16js} and y^*_{18js} (i.e., central design points), and $\mathbf{T_{c2}} = \mathbf{T_{16}} + \mathbf{T_{18}}$.

For a $\frac{1}{2}(H4)_{0,0,z} + \frac{1}{2}(Hq)_{0,w}$ design described above, the linear and quadratic statistics are shown in Table 3. Based on Table 3, the statistics $\mathbf{S}$, and its corresponding matrix $\mathbf{H}$ and covariance matrix $\mathbf{\Sigma}$ can be obtained.

Table 3. Linear and quadratic statistics for a $\frac{1}{2}(Hq)_{0,0,z} + \frac{1}{2}(Hq)_{0,w}$ design

Statistic	*df*	*Expectation*	*Variance*
$\bar{y}_1$	-	$\mu_1(\mathbf{b})$	E_1/wq
$\bar{y}^*_{21}$	-	$\mu_2(\mathbf{b})$	$E_{11}/16$
$\bar{y}^*_{22}$	-	$\mu_2(\mathbf{b})$	$E_{22}/4$
.	.	.	.
.	.	.	.
$\bar{y}_{13}$	-	$\mu_{13}(\mathbf{b})$	E_1/wq
$\bar{y}^*_{141}$	-	$\mu_{14}(\mathbf{b})$	$E_{11}/16$
$\bar{y}^*_{142}$	-	$\mu_{14}(\mathbf{b})$	$E_{22}/4$
$\bar{y}_{c_1}$	-	$\mu_{15}(\mathbf{b})$	$E_1/n_{c1}wq$
$\bar{y}^*_{c_{21}}$	-	$\mu_{15}(\mathbf{b})$	$E_{11}/16n_{c2}$
$\bar{y}^*_{c_{22}}$	-	$\mu_{15}(\mathbf{b})$	$E_{22}/4n_{c2}$
$qw(\bar{y}_1 - \mu_1(\mathbf{b}))^{\mathbf{2}}$	1	E_1	$2E_1^2$
$\mathbf{T_2}$	z	$\mathbf{E_T}$	$\mathbf{\Sigma_T}$
.	.	.	.
.	.	.	.
$qw(\bar{y}_{13} - \mu_{13}(\mathbf{b}))^{\mathbf{2}}$	1	E_1	$2E_1^2$
$\mathbf{T_{14}}$	z	$\mathbf{E_T}$	$\mathbf{\Sigma_T}$
$n_{c1}qw(\bar{y}_{c_1} - \mu_{15}(\mathbf{b}))^{\mathbf{2}}$	1	E_1	$2E_1^2$
$\mathbf{T_{c2}}$	$n_{c2}z$	$n_{c2}\mathbf{E_T}$	$n_{c2}\mathbf{\Sigma_T}$
SSA_1	$n_1(w-1)$	$n_1(w-1)E_1$	$2n_1(w-1)E_1^2$
SSB_1	$n_1w(q-1)$	$n_1w(q-1)E_3$	$2n_1w(q-1)E_3^2$
$SSB_{2(2/1)}$	n_2z	n_2zE_4	$2n_2zE_4^2$
$SSB_{1(2/1)}$	n_2z	n_2zE_3	$2n_2zE_3^2$
$SSE_{2(1/1)}$	$2n_2z$	$2n_2z\sigma_e^2$	$4n_2z\sigma_e^4$

3 Simulation Results

Using the ITLS procedure outlined above, simulation results were obtained for the split-factorial nested designs $p(Hq)_{x,0}+(1-p)(Hq)_{0,w}$ and $p(Hq)_{0,0,z}+(1-p)(Hq)_{0,w}$ as described in Table 2 and Table 3, respectively. Several SAS IML (SAS Institute, Inc. 2003) programs were constructed to compute 5000 sets of ML estimates $\hat{\theta} = (\hat{\mathbf{b}}^T, \hat{\sigma}_1^2, \hat{\sigma}_2^2, \hat{\sigma}_e^2)^T$. Our computed results were for $k = 3$, $n_c = 4$ ($2^k + 2k + n_c = 18$) and $N = 72, 108, 144$. For $N = 72$ (i.e., four observations per design point), the designs studied include $\frac{1}{2}(H2)_{1,0} + \frac{1}{2}(H2)_{0,2}$ and $(H2)_{1,0}$ designs. For $N = 108$ (i.e., six observations per design point), the designs studied include $\frac{1}{2}(H4)_{0,0,1} + \frac{1}{2}(H3)_{0,2}$, $\frac{1}{2}(H3)_{1,0} + \frac{1}{2}(H3)_{0,2}$, and $(H3)_{1,0}$ designs. For $N = 144$ (i.e., eight observations per design point) the designs studied include $\frac{1}{2}(H2)_{2,0} + \frac{1}{2}(H2)_{0,4}$, $\frac{1}{2}(H4)_{1,0} + \frac{1}{2}(H4)_{0,2}$, $(H2)_{2,0}$ and $(H4)_{1,0}$ designs. The values of parameter $\mathbf{b}$ were chosen based on an application dealing with effects on cracking of titanium alloy (see Exercise 6.5 of Myers and Montgomery (1995), page 269). The three factors are pouring temperature (A_1), titanium content (A_2) and amount of grain refiner (A_3). The design of crossed factor effects are the same as listed in Table 1. The response data is fitted using a second-order model. The estimated coefficients are then used as parameter values of $\mathbf{b}$, which are $b_0 = 1.27003$, $b_1 = -0.167164$, $b_2 = -0.077662$, $b_3 = 0.117$, $b_{11} = 0.060953$, $b_{12} = 0.0425$, $b_{22} = 0.059286$, $b_{13} = -0.06$, $b_{23} = -0.0575$, and $b_{33} = 0.042619$. The values of variance components were chosen to give six sets of combination with $(\sigma_1^2, \sigma_2^2, \sigma_e^2) = (0.2, 0.4, 0.8)$, $(0.2, 0.8, 0.4), \ldots, (0.8, 0.4, 0.2)$. Table 3 lists the simulated mean squared error (MSE) of $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$ and $\hat{\sigma}_e^2$, (denoted by $M(\hat{\sigma}_1^2)$, $M(\hat{\sigma}_2^2)$ and $M(\hat{\sigma}_e^2)$, respectively), and sum of the three MSEs (denoted by $M(\hat{\theta}_2)$). Table 3 also lists the sum of the simulated MSE of $\hat{b}_0, \hat{b}_1, \ldots, \hat{b}_{23}$ (denoted by $M(\mathbf{b})$). Based on the simulation results of Table 3, we conclude that:

(i) Given N, for most of cases studied, the $\frac{1}{2}(Hq)_{x,0} + \frac{1}{2}(Hq)_{0,w}$ designs perform best for estimation of **b**.

(ii) Given $N = 72, 144$, if $\sigma_e^2/\sigma_i^2 > 1$ $(i = 1, 2)$, then the balanced designs $(H2)_{x,0}$ perform best for estimation of θ_2. However, if $\sigma_e^2/\sigma_i^2 < 1$ $(i = 1, 2)$, then the $\frac{1}{2}(H2)_{x,0} + \frac{1}{2}(H2)_{0,w}$ designs perform best. Increased efficiency for use of $\frac{1}{2}(H2)_{x,0} + \frac{1}{2}(H2)_{0,w}$ designs is largest when σ_1^2 is the largest.

(iii) Given $N = 108$, if $\sigma_e^2/\sigma_i^2 > 1$ $(i = 1, 2)$, then the balanced design $(H3)_{1,0}$ performs best for estimation of θ_2. However, if $\sigma_e^2/\sigma_i^2 < 1$ $(i = 1, 2)$, then the $\frac{1}{2}(H3)_{1,0} + \frac{1}{2}(H3)_{0,2}$ design performs best. In this case, we found no advantage for the $\frac{1}{2}(H4)_{0,0,1} + \frac{1}{2}(H3)_{0,2}$ design.

Table 4. Simulated mean squared error of the ML estimate of $\hat{\theta}$

σ_1^2	σ_2^2	σ_e^2	N	Design	$M(\hat{\sigma}_1^2)$	$M(\hat{\sigma}_2^2)$	$M(\hat{\sigma}_e^2)$	$M(\hat{\theta}_2)$	$M(\hat{\mathbf{b}})$
0.2	0.4	0.8	72	$\frac{1}{2}(H2)_{1,0}+\frac{1}{2}(H2)_{0,2}$	0.065	0.179	0.297	0.541	0.699
0.2	0.4	0.8	72	$(H2)_{1,0}$(balanced)	0.039	0.062	0.035	0.135	0.721
0.2	0.8	0.4	72	$\frac{1}{2}(H2)_{1,0}+\frac{1}{2}(H2)_{0,2}$	0.037	0.108	0.046	0.190	1.186
0.2	0.8	0.4	72	$(H2)_{1,0}$(balanced)	0.038	0.093	0.009	0.140	1.165
0.4	0.2	0.8	72	$\frac{1}{2}(H2)_{1,0}+\frac{1}{2}(H2)_{0,2}$	0.147	0.121	0.246	0.514	0.720
0.4	0.2	0.8	72	$(H2)_{1,0}$(balanced)	0.138	0.030	0.033	0.202	0.747
0.4	0.8	0.2	72	$\frac{1}{2}(H2)_{1,0}+\frac{1}{2}(H2)_{0,2}$	0.109	0.058	0.005	0.171	1.188
0.4	0.8	0.2	72	$(H2)_{1,0}$(balanced)	0.139	0.071	0.002	0.212	1.370
0.8	0.2	0.4	72	$\frac{1}{2}(H2)_{1,0}+\frac{1}{2}(H2)_{0,2}$	0.254	0.026	0.022	0.303	0.807
0.8	0.2	0.4	72	$(H2)_{1,0}$(balanced)	0.416	0.017	0.008	0.441	1.048
0.8	0.4	0.2	72	$\frac{1}{2}(H2)_{1,0}+\frac{1}{2}(H2)_{0,2}$	0.252	0.026	0.005	0.283	0.937
0.8	0.4	0.2	72	$(H2)_{1,0}$(balanced)	0.437	0.025	0.002	0.464	1.211
0.2	0.4	0.8	108	$\frac{1}{2}(H4)_{0,0,1}+\frac{1}{2}(H3)_{0,2}$	0.123	0.095	0.070	0.288	0.607
0.2	0.4	0.8	108	$\frac{1}{2}(H3)_{1,0}+\frac{1}{2}(H3)_{0,2}$	0.042	0.089	0.126	0.257	0.561
0.2	0.4	0.8	108	$(H3)_{1,0}$(balanced)	0.037	0.037	0.030	0.104	0.547
0.2	0.8	0.4	108	$\frac{1}{2}(H4)_{0,0,1}+\frac{1}{2}(H3)_{0,2}$	0.135	0.060	0.016	0.212	1.103
0.2	0.8	0.4	108	$\frac{1}{2}(H3)_{1,0}+\frac{1}{2}(H3)_{0,2}$	0.031	0.048	0.014	0.093	0.969
0.2	0.8	0.4	108	$(H3)_{1,0}$(balanced)	0.038	0.051	0.008	0.097	0.992
0.4	0.2	0.8	108	$\frac{1}{2}(H4)_{0,0,1}+\frac{1}{2}(H3)_{0,2}$	0.145	0.092	0.064	0.302	0.553
0.4	0.2	0.8	108	$\frac{1}{2}(H3)_{1,0}+\frac{1}{2}(H3)_{0,2}$	0.103	0.053	0.090	0.246	0.561
0.4	0.2	0.8	108	$(H3)_{1,0}$(balanced)	0.125	0.022	0.022	0.169	0.667
0.4	0.8	0.2	108	$\frac{1}{2}(H4)_{0,0,1}+\frac{1}{2}(H3)_{0,2}$	0.164	0.051	0.004	0.218	1.239
0.4	0.8	0.2	108	$\frac{1}{2}(H3)_{1,0}+\frac{1}{2}(H3)_{0,2}$	0.091	0.032	0.003	0.126	1.150
0.4	0.8	0.2	108	$(H3)_{1,0}$(balanced)	0.135	0.038	0.002	0.175	1.188
0.8	0.2	0.4	108	$\frac{1}{2}(H4)_{0,0,1}+\frac{1}{2}(H3)_{0,2}$	0.221	0.090	0.015	0.326	0.726
0.8	0.2	0.4	108	$\frac{1}{2}(H3)_{1,0}+\frac{1}{2}(H3)_{0,2}$	0.209	0.014	0.011	0.234	0.740
0.8	0.2	0.4	108	$(H3)_{1,0}$(balanced)	0.395	0.010	0.007	0.413	0.889
0.8	0.4	0.2	108	$\frac{1}{2}(H4)_{0,0,1}+\frac{1}{2}(H3)_{0,2}$	0.227	0.072	0.004	0.304	0.909
0.8	0.4	0.2	108	$\frac{1}{2}(H3)_{1,0}+\frac{1}{2}(H3)_{0,2}$	0.210	0.014	0.003	0.226	0.872
0.8	0.4	0.2	108	$(H3)_{1,0}$(balanced)	0.414	0.014	0.002	0.430	1.032

Table 4. (continued)

σ_1^2	σ_2^2	σ_e^2	N	Design	$M(\hat{\sigma}_1^2)$	$M(\hat{\sigma}_2^2)$	$M(\hat{\sigma}_e^2)$	$M(\hat{\theta}_2)$	$M(\hat{\mathbf{b}})$
0.2	0.4	0.8	144	$\frac{1}{2}(H2)_{2,0}+\frac{1}{2}(H2)_{0,4}$	0.025	0.047	0.034	0.106	0.410
0.2	0.4	0.8	144	$(H2)_{2,0}$(balanced)	0.025	0.032	0.018	0.075	0.464
0.2	0.4	0.8	144	$\frac{1}{2}(H4)_{1,0}+\frac{1}{2}(H4)_{0,2}$	0.033	0.056	0.069	0.158	0.473
0.2	0.4	0.8	144	$(H4)_{1,0}$(balanced)	0.035	0.026	0.017	0.077	0.553
0.2	0.8	0.4	144	$\frac{1}{2}(H2)_{2,0}+\frac{1}{2}(H2)_{0,4}$	0.025	0.042	0.009	0.077	0.860
0.2	0.8	0.4	144	$(H2)_{2,0}$(balanced)	0.029	0.043	0.004	0.077	0.959
0.2	0.8	0.4	144	$\frac{1}{2}(H4)_{1,0}+\frac{1}{2}(H4)_{0,2}$	0.029	0.035	0.009	0.073	0.921
0.2	0.8	0.4	144	$(H4)_{1,0}$(balanced)	0.036	0.034	0.004	0.075	1.040
0.4	0.2	0.8	144	$\frac{1}{2}(H2)_{2,0}+\frac{1}{2}(H2)_{0,4}$	0.050	0.030	0.028	0.108	0.344
0.4	0.2	0.8	144	$(H2)_{2,0}$(balanced)	0.048	0.020	0.017	0.085	0.403
0.4	0.2	0.8	144	$\frac{1}{2}(H4)_{1,0}+\frac{1}{2}(H4)_{0,2}$	0.082	0.032	0.045	0.158	0.477
0.4	0.2	0.8	144	$(H4)_{1,0}$(balanced)	0.115	0.015	0.016	0.146	0.599
0.4	0.8	0.2	144	$\frac{1}{2}(H2)_{2,0}+\frac{1}{2}(H2)_{0,4}$	0.054	0.033	0.002	0.089	0.955
0.4	0.8	0.2	144	$(H2)_{2,0}$(balanced)	0.066	0.039	0.001	0.106	1.077
0.4	0.8	0.2	144	$\frac{1}{2}(H4)_{1,0}+\frac{1}{2}(H4)_{0,2}$	0.079	0.023	0.002	0.104	1.080
0.4	0.8	0.2	144	$(H4)_{1,0}$(balanced)	0.126	0.027	0.001	0.154	1.249
0.8	0.2	0.4	144	$\frac{1}{2}(H2)_{2,0}+\frac{1}{2}(H2)_{0,4}$	0.079	0.013	0.008	0.099	0.413
0.8	0.2	0.4	144	$(H2)_{2,0}$(balanced)	0.097	0.009	0.004	0.111	0.540
0.8	0.2	0.4	144	$\frac{1}{2}(H4)_{1,0}+\frac{1}{2}(H4)_{0,2}$	0.183	0.010	0.008	0.202	0.683
0.8	0.2	0.4	144	$(H4)_{1,0}$(balanced)	0.355	0.007	0.004	0.366	0.947
0.8	0.4	0.2	144	$\frac{1}{2}(H2)_{2,0}+\frac{1}{2}(H2)_{0,4}$	0.081	0.013	0.002	0.096	0.539
0.8	0.4	0.2	144	$(H2)_{2,0}$(balanced)	0.107	0.013	0.001	0.122	0.688
0.8	0.4	0.2	144	$\frac{1}{2}(H4)_{1,0}+\frac{1}{2}(H4)_{0,2}$	0.184	0.009	0.002	0.196	0.809
0.8	0.4	0.2	144	$(H4)_{1,0}$(balanced)	0.371	0.009	0.001	0.381	1.093

4 Discussion and Conclusion

However, the testing for lack-of-fit requires further investigation. In practice, it is not necessary to split the factorial points into half as descirbed in Table 2 and Table 3. One can choose different combination of the three structures, i.e., $p_1(Hq)_{x,0}+p_2(Hq)_{0,w}+(1-p_1-p_2)(Hq)_{0,0,z}$, where $0 \leq p_i \leq 1$. The optimal choices of q, p_i x, w and z depend on the relative values of the variance components and total sample size. Often there will be some prior data, or at least some intuition, available concerning relative magnitudes of the variance components. Then, based on the best prior information available, one can arrive at an optimal choice by simulated MSEs of all possible designs.

Appendix

Proof of equivalence of ITLS estimates and ML estimates

It is now shown that if $\mathbf{S}$ contains the sufficient statistics for $\mathbf{b}$ given the variance components, and the sufficient statistics for the variance component given $\mathbf{b}$, then the ITLS method will yield maximum likelihood estimates under normality assumptions. Let the model be as in (1.1) and let $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, ..., \mathbf{Z}_c]$. The log likelihood function for the multivariate normal model is

$$2 \log L = -tr\big(\mathbf{V}^{-1}\mathbf{Q}\big) - \log |\mathbf{V}| + constant,$$

where $\mathbf{Q} = (\mathbf{Y} - \mathbf{Xb})(\mathbf{Y} - \mathbf{Xb})^T$.

Let $\mathbf{P}_a$ be a matrix of orthonormal eigenvectors of $\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^-\mathbf{Z}^T$ associated with the eigenvalue one which has mutiplicity h and $\mathbf{P}_b$ be a matrix of orthonormal eignevectors of $\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^-\mathbf{Z}^T$ associated with the eigenvalue zero which has multiplicity $N - h$. Let $\mathbf{P} = [\mathbf{P}_a \ \mathbf{P}_b]$ and $\mathbf{Y}^* = \mathbf{P}^T(\mathbf{Y} - E(\mathbf{Y}))$. Then

$$\mathbf{V}^* = \mathbf{P}^T\mathbf{V}\mathbf{P} = Var(\mathbf{Y}^*) = \begin{pmatrix} \mathbf{D}_a & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_b \end{pmatrix}$$

, where $\mathbf{D}_b = \sigma^2_{c+1}\mathbf{I}_{N-h}$. Further simplification may not be possible for an arbitrary design, but with a suitable choice of design it will be possible to choose $\mathbf{P}_a$ not dependent on θ_2, but such that $\mathbf{D}_a$ itself has the block diagonal structure

$$\begin{pmatrix} \mathbf{D_1} & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & \mathbf{D_l} & & & & & & & \\ & & & D_{l+1} & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & \mathbf{D}_{l+r_1} & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & \mathbf{D}_{l+r_1+r_2} & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \mathbf{D}_{l+r_1+\ldots+r_a} \end{pmatrix}$$

where for some integers $l,\ r_1,\ r_2, \ldots, r_a$, $\mathbf{D}_i = \mathbf{P}_i^T \mathbf{V} \mathbf{P}_i$ for $i = 1, \ldots, l + r_1 + \ldots + r_a = k$ (say), the $\mathbf{D}_i$ for $i > l$ are equal in groups, i.e.,

$$\mathbf{D}_{l+1} = \ldots = \mathbf{D}_{l+r_1},$$

$$\mathbf{D}_{l+r_1+1} = \ldots = \mathbf{D}_{l+r_1+r_2},$$

$$\begin{matrix} . & & . \\ . & & . \\ . & & . \end{matrix}$$

$$\mathbf{D}_{l+r_1+\ldots+r_{a-1}+1} = \ldots = \mathbf{D}_{l+r_1+\ldots+r_a}$$

and $E(\mathbf{P}_i^T \mathbf{Y}) = E(\mathbf{Y}_i^*) = \mathbf{0}$ for $i > l$.

Now let

$$\mathbf{Q}^* = \begin{pmatrix} \mathbf{Q}_a^* & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_b^* \end{pmatrix}$$

where

$$\mathbf{Q}_a^* = \begin{pmatrix} \mathbf{Q}_{11}^* & \dots & \mathbf{Q}_{1k}^* \\ \vdots & \dots & \vdots \\ \mathbf{Q}_{k1}^* & \dots & \mathbf{Q}_{kk}^* \end{pmatrix}$$

$$\mathbf{Q}_b^* = \mathbf{P}_b^T[\mathbf{Y} - E(\mathbf{Y})][\mathbf{Y} - E(\mathbf{Y})]^T\mathbf{P}_b$$

In (3) $\mathbf{Q}_{ij}^* = \mathbf{P}_i^T[\mathbf{Y} - E(\mathbf{Y})][\mathbf{Y} - E(\mathbf{Y})]^T\mathbf{P}_j$. Also put $\mathbf{S} = [\mathbf{S}_1^T\ \mathbf{S}_2^T]^T$ where

$$\mathbf{S}_1 = \begin{pmatrix} \mathbf{P}_1^T\mathbf{Y} \\ \vdots \\ \mathbf{P}_l^T\mathbf{Y} \end{pmatrix}, \mathbf{S}_2 = \begin{pmatrix} Vech(\mathbf{Q}_{11}^*) \\ \vdots \\ Vech(\mathbf{Q}_{ll}^*) \\ Vech(\mathbf{S}_{r_1}) \\ \vdots \\ Vech(\mathbf{S}_{r_a}) \\ tr(\mathbf{Q}_b^*) \end{pmatrix},$$

$$\mathbf{S}_{r_1} = \sum_{i=l+1}^{l+r_1} \mathbf{P}_i^T\mathbf{Y}\mathbf{Y}^T\mathbf{P}_i = \sum_{i=l+1}^{i=l+r_1} \mathbf{Q}_{ii}^*,$$

$$\begin{matrix} . & & . \\ . & & . \end{matrix}$$

$$\mathbf{S}_{r_a} = \sum_{i=l+r_1+\ldots+r_{a-1}+1}^{i=l+r_1\ldots+r_a} \mathbf{P}_i^T\mathbf{Y}\mathbf{Y}^T\mathbf{P}_i = \sum_{i=l+r_1+\ldots+r_{a-1}+1}^{i=l+r_1+\ldots+r_a=k} \mathbf{Q}_{ii}^*.$$

and $\mathbf{S}_{ri}$ is a Wishart matrix with r_i degrees of freedom (or a multiple of chi-square if $\mathbf{S}_{r_i}$ is 1×1) and $tr(\mathbf{Q}_b^*)$ is distributed as $\sigma_e^2\chi_{N-h}^2$.

Then $\mathbf{S}_1, \mathbf{S}_{r_1}, \ldots, \mathbf{S}_{r_a}$ and $tr(\mathbf{Q}_b^*)$ are sufficient statistics for $(\mathbf{b}^\mathbf{T}, \theta_2^T)$. To see this, note that the density function

$$
\begin{aligned}
f(\mathbf{Y}) &= (2\pi)^{0.5N}|\mathbf{V}|^{-0.5}\exp[-0.5(\mathbf{Y}-\mathbf{Xb})^T\mathbf{V}^{-1}(\mathbf{Y}-\mathbf{Xb})] \\
&= (2\pi)^{-0.5N}|\mathbf{V}|^{-0.5}\exp[-0.5(\mathbf{Y}-\mathbf{Xb})^T\mathbf{P}\mathbf{V}^{*-1}\mathbf{P}^T(\mathbf{Y}-\mathbf{Xb})] \\
&= (2\pi)^{-0.5N}|\mathbf{V}|^{-0.5}\exp\{-0.5[\sum_{i=1}^{k}(\mathbf{Y}-\mathbf{Xb})^T\mathbf{P}_i\mathbf{D}_i^{-1}\mathbf{P}_i^T(\mathbf{Y}-\mathbf{Xb})+ \\
&\quad \sigma_e^{-2}(\mathbf{Y}-\mathbf{Xb})^T\mathbf{P}_b\mathbf{P}_b^T(\mathbf{Y}-\mathbf{Xb})]\} \\
&= (2\pi)^{-0.5N}|\mathbf{V}|^{-0.5}\exp\{-0.5[\sum_{i=1}^{l}(\mathbf{Y}-\mathbf{Xb})^T\mathbf{P}_i\mathbf{D}_i^{-1}\mathbf{P}_i^T(\mathbf{Y}-\mathbf{Xb})+ \\
&\quad \sum_{i=l+1}^{k}\mathbf{Y}^T\mathbf{P}_i\mathbf{D}_i^{-1}\mathbf{P}_i^T\mathbf{Y}+\sigma_{c+1}^{-2}(\mathbf{Y}-\mathbf{Xb})^T\mathbf{P}_b\mathbf{P}_b^T(\mathbf{Y}-\mathbf{Xb})]\} \\
&= (2\pi)^{-0.5N}|\mathbf{V}|^{-0.5}\exp\{-0.5[\sum_{i=1}^{l}tr(\mathbf{D}_i^{-1}\mathbf{P}_i^T(\mathbf{Y}-\mathbf{Xb})(\mathbf{Y}-\mathbf{Xb})^T\mathbf{P}_i)+ \\
&\quad \sum_{i=l+1}^{k}tr(\mathbf{D}_i^{-1}\mathbf{P}_i^T\mathbf{Y}\mathbf{Y}^T\mathbf{P}_i)+\sigma_{c+1}^{-2}tr(\mathbf{Q}_b^*)]\} \\
&= (2\pi)^{-0.5N}|\mathbf{V}|^{-0.5}\exp\{-0.5[\sum_{i=1}^{l}tr(\mathbf{D}_i^{-1}\mathbf{P}_i^T(\mathbf{Y}-\mathbf{Xb})(\mathbf{Y}-\mathbf{Xb})^T\mathbf{P}_i)+ \\
&\quad tr(\mathbf{D}_{l+r_1}^{-1}\mathbf{S}_{r_1})+tr(\mathbf{D}_{l+r_1+r_2}^{-1}\mathbf{S}_{r_2})+\ldots+tr(\mathbf{D}_{l+r_1+\ldots+r_a}^{-1}\mathbf{S}_{r_a})+\sigma_e^{-2}tr(\mathbf{Q}_b^*)]\}
\end{aligned}
\tag{1}
$$

Thus, the density function depends on $\mathbf{Y}$ only through the statistics in $\mathbf{S}_1, \mathbf{S}_{r_i}$, for $i = 1, \ldots, a$, and $tr(\mathbf{Q}_b^*)$. By the factorization criterion, $\mathbf{S}_1, \mathbf{S}_{r_1}, \ldots, \mathbf{S}_{r_a}$, and $tr(\mathbf{Q}_b^*)$ are joint sufficient statistics for θ_2 and $\mathbf{b}$. Clearly, if $\mathbf{b}$ is known, the statistics $\mathbf{Q}_{ii}^* = \mathbf{P}_i^T(\mathbf{Y}-\mathbf{Xb})(\mathbf{Y}-\mathbf{Xb})^T\mathbf{P}_i$ (for $i = 1, \ldots, l$) and $\mathbf{S}_{r_1}, \ldots, \mathbf{S}_{r_a}$,

$tr(\mathbf{Q}_b^*)$ are sufficient for estimating the variance components (θ_2). On the other hand, if the variance components are known then the $\mathbf{P}_i^T\mathbf{Y}$ (for $i = 1, \ldots, l$) are sufficient for $\mathbf{b}$. Note that the pivotal quantities, $vech(\mathbf{Q}_{ii}^*)$, for $i = 1, \ldots, l$, contained in $\mathbf{S}_2$ depend on $\mathbf{Y}$ only through $\mathbf{P}_i^T\mathbf{Y}$. The ITLS solution has to satisfy the following two equations.

$$\hat{\mathbf{b}} = (\mathbf{H}_1^T\mathbf{\Sigma}_1^{-1}\mathbf{H}_1)^{-1}\mathbf{H}_1^T\mathbf{\Sigma}_1^{-1}\mathbf{S}_1 \tag{A.1}$$

$$\hat{\theta}_2 = (\mathbf{H}_2^T\mathbf{\Sigma}_2^{-1}\mathbf{H}_2)^{-1}\mathbf{H}_2^T\mathbf{\Sigma}_2^{-1}\mathbf{S}_2, \tag{A.2}$$

where $\mathbf{\Sigma}_1 = Var(\mathbf{S}_1),\ E(\mathbf{S}_1) = \mathbf{H}_1\mathbf{b}, \mathbf{\Sigma}_2 = Var(\mathbf{S}_2)$ and

$$E(\mathbf{S}_2) = \mathbf{H}_2\theta_2 = \begin{pmatrix} vech(\mathbf{D}_1) \\ \vdots \\ vech(\mathbf{D}_l) \\ vech(r_1\mathbf{D}_{l+r_1}) \\ \vdots \\ vech(r_a\mathbf{D}_{l+r_1+\ldots+r_a}) \\ (N-h)\sigma_e^2 \end{pmatrix}$$

Now $\log L = tr(\mathbf{V}^{-1}\mathbf{Q}) - log|\mathbf{V}| + constant = tr(\mathbf{V}^{*-1}\mathbf{Q}^*) - log|\mathbf{V}^*| + constant.$

The estimation equations for MLEs of $\mathbf{b}$ and θ_2 are

$$\frac{\partial L}{\partial \mathbf{b}} = -\frac{\partial}{\partial \mathbf{b}} tr(\mathbf{V}^{*-1}\mathbf{Q}^*) = \mathbf{0} \tag{A.3}$$

$$\frac{\partial L}{\partial \sigma_i^2} = -\frac{\partial}{\partial \sigma_i^2} tr(\mathbf{V}^{*-1}\mathbf{Q}^*) + tr(\mathbf{V}^*\frac{\partial \mathbf{V}^{*-1}}{\partial \sigma_i^2}) = 0 \tag{A.4}$$

for $i = 1, \ldots, c+1$

Equation $(A.3)$ implies $\quad -\frac{\partial}{\partial \mathbf{b}} tr(\mathbf{V}^{*-1}\mathbf{Y}^*\mathbf{Y}^{*T}) = \mathbf{0}$

$$\Rightarrow -\frac{\partial}{\partial \mathbf{b}} tr(\mathbf{Y}^{*T}\mathbf{V}^{*-1}\mathbf{Y}^*) = \mathbf{0}$$

$$\Rightarrow -\frac{\partial}{\partial \mathbf{b}} \mathbf{Y}^{*T}\mathbf{V}^{*-1}\mathbf{Y}^* = \mathbf{0}$$

$$\Rightarrow -\frac{\partial}{\partial \mathbf{b}} (\mathbf{S}_1 - \mathbf{H}_1\mathbf{b})^T\mathbf{\Sigma}_1^{-1}(\mathbf{S}_1 - \mathbf{H}_1\mathbf{b}) = \mathbf{0}$$

$$\Rightarrow \hat{\mathbf{b}} = (\mathbf{H}_1\mathbf{\Sigma}_1^{-1}\mathbf{H}_1)^{-1}\mathbf{H}_1^T\mathbf{\Sigma}_1^{-1}\mathbf{S}_1$$

which is the same as $(A.1)$. Equation $(A.4)$ implies

$$-\frac{\partial}{\partial\sigma_i^2} tr(\mathbf{V}^{*-1}\mathbf{Q}_D^*) + tr(\mathbf{V}^*\frac{\partial\mathbf{V}^{*-1}}{\partial\sigma_i^2}) = 0 \text{ for } i = 1, \ldots, c, \qquad (A.5)$$

where

$$\mathbf{Q}_D^* = \begin{pmatrix} \mathbf{Q}_{aD}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_b^* \end{pmatrix}, \quad \mathbf{Q}_{aD}^* = \begin{pmatrix} \mathbf{Q}_{11}^* & & & & & \mathbf{0} \\ & \mathbf{Q}_{22}^* & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ \mathbf{0} & & & & & \mathbf{Q}_{kk}^* \end{pmatrix},$$

Based on Goldstein's (1986) result, solving $(A.5)$ is equivalent to minimizing

$$\mathbf{G}^* = vech(\mathbf{Q}_D^* - \mathbf{V}^*)^T \mathbf{\Sigma}_2^{*-1} vech(\mathbf{Q}_D^* - \mathbf{V}^*)$$

where $\mathbf{\Sigma}_2^* = var[vech(\mathbf{Q}_D^*)]$ Let $E[vech(\mathbf{Q}_D^*)] = vech(\mathbf{V}^*) = \mathbf{H}_2^*\theta_2$. Now,

$$\frac{\partial\mathbf{H}^*}{\partial\theta_2} = \mathbf{0}$$

$$\Rightarrow \mathbf{H}_2^{*T}\mathbf{\Sigma}_2^{*-1}\mathbf{H}_2^*\theta_2 = \mathbf{H}_2^{*T}\mathbf{\Sigma}_2^{*-1} vech(\mathbf{Q}_D^*)$$

$$\Rightarrow \mathbf{H}_2^T\mathbf{\Sigma}_2^{-1}\mathbf{H}_2\theta_2 = \mathbf{H}_2^T\mathbf{\Sigma}_2^{-1}\mathbf{S}_2$$

$$\Rightarrow \hat{\theta}_2 = (\mathbf{H}_2^T\mathbf{\Sigma}_2^{-1}\mathbf{H}_2)^{-1}\mathbf{H}_2^T\mathbf{\Sigma}_2^{-1}\mathbf{S}_2, \text{ which is the same as } (A.2).$$

References

Anderson, R. L. (1961). Designs for estimating variance components. Proc. Seven Conf. on the Design of Experiments in Army Research Development and Testing. Fort Monmouth, New Jersey, October 18-20, Inst. Stat. Memo Ser. 310, p. 781-823.

Bainbridge, T. R. (1963). Staggered nested designs for estimating variance components. *ASQC Annual Conference Transctions*, 93-103.

Burce, E. A. , Hui, L. , Alan, F. K. and Jeffrey, D. P. (2002). A class of experimental designs for estimating a response surface and variance components. *Tehnometrics* 44, 45-54.

Goldstein, H. (1986). Multilevel mixed linear analysis using iterative generalized least squares. *Biometrika* 73, 43-56.

Jennrich, R. I. and Moore, R. H. (1975). Maximum likelihood estimation by means of nonlinear least squares. *ASA Proceedings of the Statistical Computing Section*, 57-65.

Montgomery, D. C. and Myers, R. H., (1995), Response Surface Methodology, New York: *John Wiley*.