Journal of
Mathematical Science ISSN 1934-7286
Volume 1, Issue 1, 2007

Optimal designs for linear mixed-effects models

Lei Nie, Department of Biostatistics, Bioinformatics, and Biomathematics,
Georgetown University, Washington, DC. ln54@georgetown.edu

Abstract

Optimal design theories of mixed-effects models are not well established, mainly because there are some substantial differences between them and fixed-effects models. The differences are “local optimality”, lack of orthogonality and independence, and two layout design structure. Despite all of these differences, optimal designs for mixed-effects models could be surprisingly similar to optimal designs for relevant fixed-effects models. These circumstances include some mixed-effects models for stability research for drug development, which initially motivated this research project. We construct optimal designs for many mixed-effects models. These optimal designs do not depend on unknown parameters and the optimality of these designs may extend to variance components estimation.

1 Introduction

As important alternatives of fixed-effects models, mixed-effects models have been very popular in practice. For one example, a random block design is an alternative model of a standard block design, since the block effect should be viewed as a random effect in many circumstances, see e.g. (Searle, Casella, and McCulloch 1992), (Cheng 1995), (Atkins and Cheng 1999), (Pinheiro and Bates 2000), and (Montgomery 2001), among others. For another example, linear models with random intercept and random slope are alternatives of simple regression models. These models were proposed in the stability research for drug development, by (Chow and Shao 1989) and (Chow and Shao 1991); a similar model was used for a longitudinal growth study, by (Snijders and Bosker 1999), (Ouwens, Tan, and Berger 2002), and references therein. Finally, more mixed-effects models are abundantly available in the literature. Readers are referred to some excellent books for details, e.g. (Christensen 1987), (Searle, Casella, and McCulloch 1992), (Pinheiro and Bates 2000), (Diggle, Heagetty, Liang, and Zeger 2002).

Designs for experiments using random-effects models and mixed-effects models attract more and more attentions. This study is motivated by the stability study for drug development, conducted in many pharmaceutical companies, through a random intercept and random slope longitudinal model, (Chow and Shao 1989), (Chow and Shao 1991), (Murphy and Weisman 1990), and (Chow and Wang 1994). In the stability study for drug development, the computation of shelf life of the drug is particularly important. The computation of shelf life depends on the parameters estimates obtained through some experiments. These experiments are usually very expensive. In order to be cost-efficient, experiments are expected to be carefully planned while the studies on optimal designs are important. Compared to approximate optimal designs, exact optimal designs are normally desirable, since both of the number of lots and the number of lots samples are usually small or moderate. Furthermore, computations of shelf life may depend on the estimation of both of the fixed effects and variance components parameters, namely the intercept, slope, and variances of random intercept and random slope. As a consequence, estimations for all parameters, including variance components, may be equally important.

Optimal design theories for linear fixed-effects regression models are extensively carried out in the past a few decades. However, there are relatively less results on optimal design theories for random or mixed-effects models. (Cheng 1995) and (Atkins and Cheng 1999) constructed optimal designs in the presence of random block effect. (Liski, Mandal, Sinha, and Abt 1997) and (Abt, Gaffke, Liski, and Sinha 1998) considered linear and quadratic growth mixed-effect models, with interclass correlation structure and autocorrelated structure; (Ouwens, Tan, and Berger 2002) discussed Maximin D-optimal designs for a random intercept and random slope longitudinal mixed-effects model. (Liski, Mandal, Shah, and Sinha 2002) considered optimal designs for a variety of models, including many random or mixed-effects models. (Hedayat, Stufken, and Yang 2006) discussed Optimal and efficient crossover designs when subject effects are random. Some elegant results were obtained in theses papers and references therein. However, many design problems remain unsolved, mainly because there are some substantial differences between fixed-effects models and mixed-effects models, in terms of theory and tools. For one of these differences, optimal designs in mixed-effects models may depend on unknown parameters: the variance components, see, e.g. (Cheng 1995), (Atkins and Cheng 1999), (Ouwens, Tan, and Berger 2002), and (Liski, Mandal, Shah, and Sinha 2002). For this reason, optimal designs are actually "local optimal designs", (Atkinson and Donev 1992). As we know, "local optimality" was a character of nonlinear designs. This character has made the major difference on optimal designs between linear and nonlinear models; it also has made people reluctant to adopt them in practice. Another difference is the lack of orthogonality and independence among estimations of effects parameters and variance parameters in the mixed-effects models. For linear fixed-effects models, the estimators of effect parameters and variance parameters are uncorrelated, which is termed as orthogonality in this paper, see formula (5) in Section 2. Furthermore, from formula (5) in Section 2, the asymptotic variance of the variance estimator is independent of the design point. Because of the orthogonality and independence, an optimal design for effects parameters is also an optimal design for the variance parameters. However, this is usually not true for the mixed-effects models, see Section 2 for details. Finally, designs for mixed-effects models have two layouts: the within-subject layout and the between-subject layout. For the within-subject layout, we need to determine the number (or weight) of lots samples and actual values of the covariate for each sample unit. For the between-subject subject, we need to determine the number (or weight) of lots. This two-layout structure could make the optimal designs very complicated for mixed-effects models, see Section 4 for some details.

Estimations for all parameters, including variance components, may be equally important. Optimal designs which are the best for estimating effects parameters under some suitable criteria, are not necessarily be optimal for estimating the variance components. Existing papers usually do not address this problem. In Section 3, we establish exact optimal designs theories for some mixed-effects models with some special structures. The optimality may extend to variance components parameters. The special structures allow us to avoid the complications due to the two-layout structure. The optimal designs we obtained do not depend on the variance components parameters, which therefore avoid the "local optimality". Since optimal designs considered here automatically handle both layouts, we shall not focus on these complications. However, in the discussion section, we shall briefly discuss the potential difficulties on optimal designs due to two layouts.

In Section 2, we introduce the model under our considerations. In Section 3, we construct exact optimal designs for some models, where the numbers of random effects components and the cluster size satisfy some conditions. Starting from the simplest case when the dimension of the random effect is 1, several cases have been considered. However, as indicated in Section 4, we only considered a small portion of models here. Further research is needed to address many other important differences between fixed-effects model and mixed-effects models, in terms of optimal design theories.

2 Models under considerations

The model considered is described as follows. Consider data partitioned into m clusters, where the ith cluster consists of n observations, $y_i = (y_{i1}, \ldots, y_{in})^T$, where y_{ij} is the observed value of the jth observation from the ith cluster. We consider the following linear mixed-effects model,

$$y_i = X_i\beta + Z_i b_i + \epsilon_i, \tag{1}$$

where β is a vector of u unknown fixed-effects parameters and b_i is a vector of w unobservable random effects. We also assume $\epsilon_i \sim N(0, \sigma^2 I_n)$, which means that $y_{i1}, ..., y_{in}$ are conditionally independent for given b_i. Here I_n is the $n \times n$ identity matrix. In the second stage, the unobservable random-effects vectors $b_1, ..., b_n$ are assumed to be a random sample from a normal distribution $N(0, \Psi)$. We assume that $\Psi = \mathrm{diag}\{\sigma_0^2, \sigma_1^2, \cdots, \sigma_{\mathrm{w}-1}^2\}$. Let Ω denote the parameter space for $\xi^T = (\beta^T, \sigma^2, \theta^T)$, where $\theta = (\sigma_0^2, \sigma_1^2, \cdots, \sigma_{w-1}^2)$. We assume that the first column of the design matrix X_i is $1_n = (1, \cdots, 1)^T$, therefore the model also assumes the existence of a general mean.

Let $V_i = \sigma^2 I_n + Z_i \Psi^{-1} Z_i^T$, following (Searle, Casella, and McCulloch 1992), page 239, the Fisher information matrix

$$I_\xi(d) = \begin{pmatrix} \sum_{i=1}^m X_i^T V_i^{-1} X_i & 0 & 0 \\ 0 & \frac{1}{2}\sum_{i=1}^m tr(V_i^{-2}) & I_{\sigma^2\theta} \\ 0 & I_{\sigma^2\theta} & I_{\theta\theta} \end{pmatrix}, \tag{2}$$

where d is the design with design matrices X_i and Z_i, $I_{\sigma^2\theta}$ is a $1 \times w$ matrix, and $I_{\theta\theta}$ is a $w \times w$ matrix,

$$\begin{aligned} I_{\sigma^2\theta}[1,k] &= \frac{1}{2}\sum_{i=1}^m tr(V_i^{-1} Z_{ik} Z_{ik}^T V_i^{-1}), & (3) \\ I_{\theta\theta}[k,l] &= \frac{1}{2}\sum_{i=1}^m tr(Z_{ik} Z_{ik}^T V_i^{-1} Z_{il} Z_{il}^T V_i^{-1}). & (4) \end{aligned}$$

Here Z_{ik} is the kth column vector of the matrix Z_i.

In this article, we mostly focus on the construction of the D-optimal design. By definition, (Atkinson and Donev 1992), a design d is called D-optimal if d maximizes the determination of the information matrix, i.e, maximizes $\det(I_\xi(d))$.

If $\dim(\theta) = w = 0$, i.e. the model (1) reduces to the following fixed-effects model,

$$y_i = X_i\beta + \epsilon_i,$$

and the information matrix becomes

$$I_\eta(d) = \begin{pmatrix} \sum_{i=1}^m X_i^T V^{-1} X_i & 0 \\ 0 & \frac{m}{2\sigma^4} \end{pmatrix}, \tag{5}$$

where $V = \sigma^2 I_n$ and $\eta^T = (\beta^T, \sigma^2)$. The variance of $\widehat{\sigma^2}$, the maximum likelihood estimator (MLE) of σ^2, are "independent" of the design matrix X_i and Z_i. Furthermore, the covariance between $\hat{\beta}$ and $\widehat{\sigma^2}$ is 0, where $\hat{\beta}$ denotes the MLE of β. In other words, for the fixed-effects model, estimation of β and σ^2 are "orthogonal" (the MLE of β and σ^2 are uncorrelated). As a consequence, an optimal design for estimating β is an optimal design for estimating σ^2. However this is certainly not true for mixed-effects models. In model (1), the parameter is $\xi = (\beta, \sigma^2, \theta)$. From the information matrix $I_\xi(d)$, it is clear that the variance of $\widehat{\sigma^2}$ and $\hat{\theta}$ both depend on the design matrix Z_i. Furthermore, the covariance between $\widehat{\sigma^2}$ and $\hat{\theta}$ is no longer 0, the "orthogonality" property is lost. As a consequence, for mixed-effects models, an optimal design for estimating β is not automatically optimal for estimating other parameters.

The differences between a fixed-effects model and a mixed-effects model suggest us to carefully study optimal designs for mixed-effects models. Before introducing the main results, we present some notations. Let $d = (d_1, \cdots, d_m)$ denote a design where d_i is the sub-design on the ith subject,

$$d_i = \begin{Bmatrix} a_{i1} & \cdots & a_{in_i} \\ w_{i1} & \cdots & w_{in_i} \end{Bmatrix} \tag{6}$$

where $a_{i1}, \cdots, a_{in_i}$ are distinct design points and $w_{i1}, \cdots, w_{in_i}$ are corresponding weights. Since we consider exact designs, we assume that nw_{ij} be an integer.

3 Main results

In this section, we shall consider several cases separately. We recall that $\dim(\theta) = w$ is the dimension of the random effect b_i.

3.1 Case 1: $\dim(\theta) = 1$ and $Z_i = 1_n$.

$Z_i = 1_n$ means there is only one random effect. In this case, b_i is usually called the random block effect, (Searle, Casella, and McCulloch 1992), (Cheng 1995), (Atkins and Cheng 1999), among others.

Proposition 1. *If $Z_i = 1_n$, then a D-optimal design for estimating β is also a D-optimal design for estimating ξ.*

Proof.

$$\det(I_\xi(d)) = \det\left(\sum_{i=1}^m X_i^T V_i^{-1} X_i\right) \det\left\{\begin{pmatrix} \frac{1}{2}\sum_{i=1}^m tr\left(V_i^{-2}\right) & I_{\sigma^2\theta} \\ I_{\sigma^2\theta} & I_{\theta\theta} \end{pmatrix}\right\}.$$

In this case, $V_i = \sigma^2 I_n + 1_n \Psi^{-1} 1_n^T$ does not depend on the design matrix X. It is therefore also clear, through (3) and (4), that $I_{\sigma^2\theta}$ and $I_{\theta\theta}$ do not depend on X since they only depend on $V_i's$. As a consequence, a D-optimal design for estimating ξ maximizes $\det(\sum_{i=1}^m X_i^T V_i^{-1} X_i)$, which is the information matrix for estimating β. $\square$

The proposition implies, D-optimal designs presented in Theorem 2.1 of (Atkins and Cheng 1999) are also D-optimal designs for estimating ξ, all parameters.

3.2 Case 2: $\dim(\theta) = 2$, $\dim(\beta) = 2$

In this case, we assume $Z_i = X_i$.

$$Z_i^T = X_i^T = \left\{ \begin{array}{c} 1, \cdots, 1 \\ x_{i1}, \cdots, x_{in} \end{array} \right\}_{2\times n}.$$

This is a simple random-effect linear model with random intercept b_{i0} and random slope b_{i1}. $b_{i0} \sim N(0, \sigma_0^2)$ and $b_{i1} \sim N(0, \sigma_1^2)$ are uncorrelated. This model has been used in many areas, e.g. (Chow and Shao 1989), (Chow and Shao 1991), (Murphy and Weisman 1990), (Chow and Wang 1994), (Ouwens, Tan, and Berger 2002), and (Liski, Mandal, Shah, and Sinha 2002).

By definition, $V_i = \sigma^2 I_n + \sigma_0^2 1_n 1_n^T + \sigma_1^2 x_i x_i^T$, with $x_i^T = (x_{i1}, \cdots, x_{in})$. It is easy to show that, from equations (3) and (4), the information for parameter θ for given σ^2 is

$$I_{\theta\theta} = \left(\begin{array}{cc} \sum_{i=1}^m (1_n^T V_i^{-1} 1_n)^2 & \sum_{i=1}^m (1_n^T V_i^{-1} x_i)^2 \\ \sum_{i=1}^m (x_i^T V_i^{-1} 1_n)^2 & \sum_{i=1}^m (x_i^T V_i^{-1} x_i)^2 \end{array} \right).$$

The information matrix for parameter $\gamma = (\beta, \sigma_0^2, \sigma_1^2)$, for given σ^2, is

$$I_\gamma(d) = \left(\begin{array}{cccc} \sum_{i=1}^m 1_n^T V_i^{-1} 1_n & \sum_{i=1}^m 1_n^T V_i^{-1} x_i & 0 & 0 \\ \sum_{i=1}^m x_i^T V_i^{-1} 1_n & \sum_{i=1}^m x_i^T V_i^{-1} x_i & 0 & 0 \\ 0 & 0 & \sum_{i=1}^m (1_n^T V_i^{-1} 1_n)^2 & \sum_{i=1}^m (1_n^T V_i^{-1} x_i)^2 \\ 0 & 0 & \sum_{i=1}^m (x_i^T V_i^{-1} 1_n)^2 & \sum_{i=1}^m (x_i^T V_i^{-1} x_i)^2 \end{array} \right). \quad (7)$$

In this model, the design point is x_{ij}, $i = 1, \cdots, m$, $j = 1, \cdots, n$.

Theorem 2. *Assume that n is an even number and the design space of x_{ij} is [-1,1], then the design $d = (d_1, \cdots, d_n)$, with*

$$d_i = \left(\begin{array}{cc} -1 & 1 \\ 0.5 & 0.5 \end{array} \right)$$

is the unique D-optimal design for estimating for estimating $\gamma = (\beta^T, \sigma_0^2, \sigma_1^2)$, for given σ^2, i.e. it maximizes $Det(I_\gamma)$.

The proof is given in Appendix 5.1.

3.3 Case 3: $\dim(\theta) = 2$, $\dim(\beta) = 4$

We assume two column vectors of Z_i are the first two columns of X_i, i.e. a mixed-effects linear model with random intercept b_{i0} and random slope b_{i1}. $b_{i0} \sim N(0, \sigma_0^2)$ and $b_{i1} \sim N(0, \sigma_1^2)$.

In this case, we assume $Z_i = X_i$.

$$X_i^T = \left\{ \begin{array}{c} 1, \cdots, 1 \\ x_{i1,1}, \cdots, x_{i1,n} \\ x_{i2,1}, \cdots, x_{i2,n} \\ x_{i3,1}, \cdots, x_{i3,n} \end{array} \right\}_{4 \times n},$$

and Z_i^T is the matrix formed by the first 2 rows of X_i^T. The design point $x_{ij} = (x_{i1,j}, x_{i2,j}, x_{i3,j})^T \in [-1,1]^3$.

Theorem 3. *Let n be a multiple of 4. Assume that the design space of x_{ij} is $[-1,1]^3$, then the design $d = (d_1, \cdots, d_n)$ with*

$$d_i = \left(\begin{array}{cccc} \{1,1,1\} & \{1,-1,-1\} & \{-1,1,-1\} & \{-1,-1,1\} \\ 0.25 & 0.25 & 0.25 & 0.25 \end{array} \right).$$

is a D-optimal design for estimating parameters β. i.e. the design maximizes $\det(\sum_{i=1}^m X_i^T V_i^{-1} X_i)$. Furthermore, this design is also a D-optimal design for estimating $\gamma = (\beta^T, \sigma_0^2, \sigma_1^2)$ for given σ^2, i.e. it maximizes $Det(I_\gamma)$.

The proof is given in Appendix 5.3.

Remark: Since the information for β and the information other parameters are "orthogonal", the D-optimal design for β in the theorem could be viewed as D_A or D_S-optimal design, (Silvey 1980).

3.4 Case 4: $\dim(\theta) = 2$, $\dim(\beta) = 3$

In this case, we assume two column vectors of Z_i are first two columns of X_i,

$$X_i^T = \left\{ \begin{array}{c} 1, \cdots, 1 \\ x_{i1,1}, \cdots, x_{i1,n} \\ x_{i2,1}, \cdots, x_{i2,n} \end{array} \right\}_{3 \times n},$$

and Z_i^T is the matrix formed by the first 2 rows of X_i^T. The design point $x_{ij} = (x_{i1,j}, x_{i2,j})^T \in [-1,1]^2$.

This is a mixed-effects linear model with random intercept b_{i0} and random slope b_{i1}. $b_{i0} \sim N(0, \sigma_0^2)$ and $b_{i1} \sim N(0, \sigma_1^2)$.

Theorem 4. *Let n be an even number. Assume that the design space of x_{ij} is $[-1,1]^2$, then the design $d = (d_1, \cdots, d_n)$, with*

$$d_i = \left(\begin{array}{cccc} \{1,1\} & \{1,-1\} & \{-1,1\} & \{-1,-1\} \\ 0.25 & 0.25 & 0.25 & 0.25 \end{array} \right)$$

is a D-optimal design for estimating parameters β. i.e. the design maximizes $\det(\sum_{i=1}^m X_i^T V_i^{-1} X_i)$. Furthermore, this design is also D-optimal design for estimating $\gamma = (\beta^T, \sigma_0^2, \sigma_1^2)$, for given σ^2, i.e. it maximizes $Det(I_\gamma)$.

The proof is given in Appendix 5.4.

3.5 Case 5: $\dim(\theta) = 4$, $\dim(\beta) = 4$

In this case, we assume $Z_i = X_i$,

$$X_i^T = Z_i^T = \left\{ \begin{array}{c} 1, \cdots, 1 \\ x_{i1,1}, \cdots, x_{i1,n} \\ x_{i2,1}, \cdots, x_{i2,n} \\ x_{i3,1}, \cdots, x_{i3,n} \end{array} \right\}_{4\times n},$$

i.e. a linear model with random intercept b_{i0} and 3 independent random slopes b_{i1}, b_{i2}, b_{i3}. $b_{il} \sim N(0, \sigma_l^2)$, $l = 0, 1, 2, 3$. In this section we assume n is a multiple of 4. The design point is $x_{ij} = (x_{i1,j}, x_{i2,j}, x_{i3,j})^T \in [-1, 1]^3$.

Theorem 5. *Let n be a multiple of 4. Assume that the design space of x_{ij} is $[-1, 1]^3$, then the design $d = (d_1, \cdots, d_n)$, with*

$$d_i = \begin{pmatrix} \{1,1,1\} & \{1,-1,-1\} & \{-1,1,-1\} & \{-1,-1,1\} \\ 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix}$$

is the unique D-optimal design for estimating parameters β. i.e. The design maximizes $\sum_{i=1}^m X_i^T V_i^{-1} X_i$.

The proof is given in Appendix 5.2. Again, the D-optimal design for β in the theorem could be viewed as D_A or D_S-optimal design, (Silvey 1980).

3.6 Case 6: $\dim(\theta) = 4l$, $\dim(\beta) = 4l$

In this case, we assume $Z_i = X_i$, i.e. a linear model with random intercept $b_{i,0}$ and $4l - 1$ independent random slopes $b_{i,1,}, \cdots, b_{i,4l-1}$. $b_{ih} \sim N(0, \sigma_h^2)$, $h = 0, \cdots, 4l - 1$. In this section we assume n is a multiple of $4l$. The design point $x_{ij} = (1, x_{i1,j}, \cdots, x_{i4l-1,j})^T \in [-1, 1]^{4l}$.

Theorem 6. *Let $w = 4l$ and n is a multiple of w. Furthermore, we assume that a $w \times w$ Hadamard matrix exists. Let $v_1 = (1, \cdots, 1)^T$. Assume that $v_1, \cdots, v_w$ are column vectors of the Haramard matrix, and the design space of x_{ij} is $[-1, 1]^{4l}$, then the design $d = (d_1, \cdots, d_n)$, with*

$$d_i = \begin{pmatrix} v_1, & \cdots, & v_{4l} \\ \frac{1}{4l}, & \cdots, & \frac{1}{4l} \end{pmatrix},$$

is a D-optimal design for estimating parameters β. i.e. The design maximizes $\det(\sum_{i=1}^m X_i^T V_i^{-1} X_i)$.

The proof can be done along with the lines of the proof for Theorem 5 and details are given in Section 6.2. The construction depends on the existence of the Hadamard matrix, for which readers are refer to (Hedayat, Sloane, and Stufken 1999), page 147 for details. The D-optimal design for β in the theorem could be viewed as D_A or D_S-optimal design, (Silvey 1980).

4 Discussion

We considered a number of cases where the exact optimal designs are simple and do not depend on the variance components. However, it is worth to mentioned that these models contribute to a small proportion of mixed-effect models. In general, the design could be much more complicated and difficult because of the following reasons. First, optimal designs depend on the variance and covariance structure of the random effects and random errors, (Cheng 1995), (Atkins and Cheng 1999), (Liski, Mandal, Shah, and Sinha 2002), (Ouwens, Tan, and Berger 2002). Therefore, it carried the properties of nonlinear fixed-effects model design, although it is a linear model. Second, the estimations of variances components are not "orthogonal and independent" to estimation of other parameters. Third, the design for mixed-effects model has two layouts. In order to establish the D-optimality, both layouts should be considered, particularly for exact designs when the cluster sizes are small. For the within-subject layout, the design criterion is similar to the D_A or D_s criteria, (Silvey 1980), (Atkinson and Donev 1992). However, for the between-subject layout, the design criterion is the standard D-optimal criteria. The two criteria made it difficult to use the general equivalence theory to simplify the questions. On the other hand, if we restricted on one layout, the problem may be easier. For example, we assume that $X_{ij} = X_{1j}$ and $Z_{ij} = Z_{1j}$ for all $j = 1, \cdots, n$, (Liski, Mandal, Shah, and Sinha 2002), then we need only to consider the D_s optimality; We may also fixed the number of n, and consider the between-subject layout, (Ouwens, Tan, and Berger 2002). However, many more tools are needed to develop before we can systematically solve the optimal design problem in linear mixed-effects models. As a matter of fact, the domain of design matrix is also a complicated issue to deal with. Readers are referred to (Liski, Mandal, Shah, and Sinha 2002) for details.

5 Appendix

5.1 Appendix: Proof of Theorem 2

Lemma 7. *Let $x^T = (x_1, \cdots, x_n)$, $1_n^T = (1, \cdots, 1)$ which is the n-vector with all elements being 1, $V = \sigma^2 I_n + \sigma_0^2 1_n 1_n^T + \sigma_1^2 x x^T$, where I_n is the $n \times n$ identity matrix.*

$$1_n^T V^{-1} 1_n = \frac{n}{\sigma^2 + n\sigma_0^2} \frac{1 + \frac{n\sigma_1^2}{\sigma^2}(\bar{x^2} - \bar{x}^2)}{1 + \frac{n\sigma_1^2}{\sigma^2}\left\{\bar{x^2} - \frac{n\bar{x}^2\sigma_0^2}{\sigma^2+n\sigma_0^2}\right\}}, \tag{8}$$

$$x^T V^{-1} 1_n = \frac{1}{\sigma^2 + n\sigma_0^2} \frac{n\bar{x}}{1 + \frac{\sigma_1^2}{\sigma^2}\left\{n\bar{x^2} - \frac{n^2\bar{x}^2\sigma_0^2}{\sigma^2+n\sigma_0^2}\right\}}, \tag{9}$$

$$x^T V^{-1} x = \frac{\frac{1}{\sigma^2}\left\{n\bar{x^2} - \frac{n^2\bar{x}^2\sigma_0^2}{\sigma^2+n\sigma_0^2}\right\}}{1 + \frac{\sigma_1^2}{\sigma^2}\left\{n\bar{x^2} - \frac{n^2\bar{x}^2\sigma_0^2}{\sigma^2+n\sigma_0^2}\right\}}, \tag{10}$$

where $\bar{x} = \sum_{i=1}^n x_i/n$ and $\bar{x^2} = \sum_{i=1}^n x_i^2/n$.

The proof of lemma 7 is given in Appendix 6.1.

Proof of Theorem 2

Proof.

$$I_\gamma = \begin{pmatrix} \sum_{i=1}^m 1_n^T V_i^{-1} 1_n & \sum_{i=1}^m 1_n^T V_i^{-1} x_i & 0 & 0 \\ \sum_{i=1}^m x_i^T V_i^{-1} 1_n & \sum_{i=1}^m x_i^T V_i^{-1} x_i & 0 & 0 \\ 0 & 0 & \sum_{i=1}^m (1_n^T V_i^{-1} 1_n)^2 & \sum_{i=1}^m (1_n^T V_i^{-1} x_i)^2 \\ 0 & 0 & \sum_{i=1}^m (x_i^T V_i^{-1} 1_n)^2 & \sum_{i=1}^m (x_i^T V_i^{-1} x_i)^2 \end{pmatrix},$$

where $V_i = \sigma^2 I_n + \sigma_0^2 1_n 1_n^T + \sigma_1^2 x_i x_i^T$ and $x_i^T = (x_{i1}, \cdots, x_{in})$. By Hadamard's determinant inequality, Marcus and Minc (1992), page 114, (see also Hedayat, Sloane, and Stufken (1999), page 166).

$$\begin{aligned} \det(I_\gamma) \quad \leq \quad & \left\{\sum_{i=1}^m 1_n^T V_i^{-1} 1_n\right\}\left\{\sum_{i=1}^m x_i^T V_i^{-1} x_i\right\} \\ & \left\{\sum_{i=1}^m (1_n^T V_i^{-1} 1_n)^2\right\}\left\{\sum_{i=1}^m (x_i^T V_i^{-1} x_i)^2\right\}, \end{aligned} \tag{11}$$

where the equality holds if and only if I_γ is a diagonal matrix. Then equation (9), Lemma 7 implies that the equality holds if and only if $\sum_{j=1}^n x_{ij} = 0$, $i = 1, \cdots, m$.

Using equation (8), we obtain

$$\sum_{i=1}^m 1_n^T V_i^{-1} 1_n \leq \frac{nm}{\sigma^2 + n\sigma_0^2}, \quad \sum_{i=1}^m (1_n^T V_i^{-1} 1_n)^2 \leq \frac{n^2 m}{(\sigma^2 + n\sigma_0^2)^2},$$

where the equalities hold if and only if $x_{ij}^2 = 1$. Finally, using equation (9), we obtain

$$\sum_{i=1}^m x_i^T V_i^{-1} x_i \leq \frac{mn}{\sigma^2 + \sigma_1^2 n}, \quad \sum_{i=1}^m (x_i^T V_i^{-1} x_i)^2 \leq \frac{mn^2}{(\sigma^2 + \sigma_1^2 n)^2},$$

where the equalities hold if and only if $x_{ij}^2 = 1$.

As a consequence, the inequality

$$\det(I_\gamma) \leq \frac{nm}{\sigma^2 + n\sigma_0^2} \frac{n^2 m}{(\sigma^2 + n\sigma_0^2)^2} \frac{mn}{\sigma^2 + \sigma_1^2 n} \frac{mn^2}{(\sigma^2 + \sigma_1^2 n)^2},$$

where the equality holds if and only if $\sum_{i=1}^n x_{ij} = 0$, $i = 1, \cdots, m$ and $x_{ij}^2 = 1$, $i = 1, \cdots, n$, $j = 1, \cdots, n$. The design given in Theorem 2 is the unique choice which satisfies the conditions $\sum_{i=1}^n x_{ij} = 0$, $i = 1, \cdots, m$ and $x_{ij}^2 = 1$, $i = 1, \cdots, m$, $j = 1, \cdots, n$. □

5.2 Appendix: Proof of Theorem 5

Before proving Theorem 5 for case 5 and Theorem 6 for case 6, we need the following two lemmas. Both lemmas are given based on case 6, which naturally apply to case 5.

Lemma 8. *Let* $A = (X_i^T X_i + \sigma^2 \Psi^{-1})^{-1}$ *and* a_{kk} *be the kth element on the diagonal, then*

$$a_{kk} \geq \frac{\sigma_{k-1}^2}{\sigma^2 + n\sigma_{k-1}^2}, \ k = 1, ..., 4l,$$

the equality holds if and only if $X_{ik}^T X_{ip} = 0$ *and* $X_{ik}^T X_{ik} = n$ *for* $k, p = 1, \cdots, 4l$.

Proof. We rewrite

$$A = \begin{pmatrix} n + \frac{\sigma^2}{\sigma_0^2} & 1_n^T B \\ B^T 1_n & D \end{pmatrix}^{-1},$$

where $1_n = (1, \cdots, 1)^T$ and B is obtained from X_i by dropping the first column. $D = B^T B + \sigma^2 \text{diag}\left\{\frac{1}{\sigma_1^2}, \cdots, \frac{1}{\sigma_{4l-1}^2}\right\}$. Note that,

$$\begin{aligned} a_{11} &= \left(n + \frac{\sigma^2}{\sigma_0^2} - 1_n^T B D^{-1} B^T 1_n\right)^{-1} \\ &\geq \left(n + \frac{\sigma^2}{\sigma_0^2}\right)^{-1} = \frac{\sigma_0^2}{\sigma^2 + n\sigma_0^2}, \end{aligned}$$

where the equality holds if and only if $X_{ik}^T 1_n$=0 for $k = 1, \cdots, 4l-1$. Here X_{ik} is the $k+1$th column of the matrix X_i. Let B_1 is the matrix obtained from X by dropping the second column X_{i1}, $D_1 = B_1^T B_1 + \sigma^2 \text{diag}\left\{\frac{1}{\sigma_0^2}, \frac{1}{\sigma_2^2}, \cdots, \frac{1}{\sigma_{4l-1}^2}\right\}$. Let X_{i1} be the second column vector of x_i. It can be show that

$$\begin{aligned} a_{22} &= \left(X_{i1}^T X_{i1} + \frac{\sigma^2}{\sigma_1^2} - X_{i1}^T B_1 D_1^{-1} B_1^T X_{i1}\right)^{-1} \\ &\geq \left(X_{i1}^T X_{i1} + \frac{\sigma^2}{\sigma_1^2}\right)^{-1} \\ &\geq \frac{\sigma_1^2}{\sigma^2 + n\sigma_1^2}, \end{aligned}$$

where the equality holds if and only if $X_{i2}^T X_{i2} = n$ and $X_{i2}^T X_{ik}$=0 for $k = 0, 2, \cdots, 4l-1$. Similarly we can show

$$a_{kk} \geq \frac{\sigma_{k-1}^2}{\sigma^2 + n\sigma_{k-1}^2},$$

where the equality holds if and only if $X_{ik}^T X_{ik} = n$ and $X_{ik}^T X_{ip}$=0 for $k \neq p$. □

Lemma 9. *Let* $C = X_i^T X_i (X_i^T X_i + \sigma^2 \Psi^{-1})^{-1}$ *then*

$$c_{kk} = 1 - \frac{a_{kk}\sigma^2}{\sigma_{k-1}^2}, \quad k = 1, \cdots, 4l-1,$$

where c_{kk} *and* a_{kk} *is the kth element on the diagonal of the matrices* C *and* $(X_i^T X_i + \sigma^2 \Psi^{-1})^{-1}$.

Proof. Let

$$\left(\begin{array}{c} d_1, ..., d_{4l} \end{array}\right)^T = (X_i^T X_i + \sigma^2 \Psi^{-1})^{-1} \left(\begin{array}{c} 1, 0, ..., 0 \end{array}\right)^T.$$

Note that d_1 is c_{11}. Then

$$X_i^T X_i \left(\begin{array}{c} d_1, ..., d_{4l} \end{array}\right)^T = \left(\begin{array}{c} 1 - \frac{d_1\sigma^2}{\sigma_0^2}, -\frac{d_1\sigma^2}{\sigma_1^2}, ..., -\frac{d_1\sigma^2}{\sigma_{4l-1}^2} \end{array}\right)^T.$$

Therefore,

$$\begin{aligned} c_{11} &= (1, 0, \cdots, 0) \; X_i^T X_i (X_i^T X_i + \sigma^2 \Psi^{-1})^{-1} \left(\begin{array}{c} 1, 0, ..., 0 \end{array}\right)^T \\ &= 1 - \frac{d_1 \sigma^2}{\sigma_0^2} = 1 - \frac{a_{11}\sigma^2}{\sigma_0^2} \end{aligned}$$

Similarly, we can prove $c_{kk} = 1 - \frac{a_{kk}\sigma^2}{\sigma_{k-1}^2}$, $k = 2, \cdots, 4l-1$. □

Proof of Theorem 5

Proof. The D-Optimal designs maximize

$$\det\left(\sum_{i=1}^m X_i^T V_i^{-1} X_i\right),$$

Note that,

$$\begin{aligned} & \sum_{i=1}^m X_i^T V_i^{-1} X_i \\ =\; & \frac{1}{\sigma^2} \sum_{i=1}^m \left[X_i^T X_i - X_i^T X_i (X_i^T X_i + \sigma^2 \Psi^{-1})^{-1} X_i^T X_i\right] \\ =\; & \sum_{i=1}^m \left[X_i^T X_i (X_i^T X_i + \sigma^2 \Psi^{-1})^{-1} \Psi^{-1}\right] \\ =\; & \sum_{i=1}^m \left[X_i^T X_i (X_i^T X_i + \Psi^{-1}\sigma^2)^{-1}\right] \Psi^{-1}. \end{aligned}$$

Now by Hadamard's determinant inequality, we obtain

$$\det\left\{\sum_{i=1}^m \left[X_i^T X_i (X_i^T X_i + \Psi^{-1}\sigma^2)^{-1}\right] \Psi^{-1}\right\} \le \prod_{k=1}^4 \frac{\lambda_{kk}}{\sigma_{k-1}^2},$$

where λ_{kk} is the kth diagonal element of the matrix $\sum_{i=1}^m \left[X_i^T X_i (X_i^T X_i + \Psi^{-1}\sigma^2)^{-1}\right]$. By Lemmas 9 and 8, we have $\lambda_{kk} \le mn\sigma_{k-1}^2/(\sigma^2 + n\sigma_{k-1}^2)$. Thus,

$$\det\left(\sum_{i=1}^m X_i^T V_i^{-1} X_i\right) \le \prod_{l=0}^3 \frac{mn}{\sigma^2 + n\sigma_l^2}.$$

The equality holds if and only if assumptions in Lemma 8 holds. for all $k, p = 1, 2, 3, 4$, i.e. X_i is the design matrix for design described in Theorem 5. $\square$

5.3 Appendix: Proof of Theorem 3

Proof. Let $x_{ik}^T = (x_{ik,1}, \cdots, x_{ik,n})$, $k = 1, 2, 3$.

$$I_\beta = \begin{pmatrix} \sum_{i=1}^m 1_n^T V_i^{-1} 1_n & \sum_{i=1}^m 1_n^T V_i^{-1} x_{i1} & \sum_{i=1}^m 1_n^T V_i^{-1} x_{i2} & \sum_{i=1}^m 1_n^T V_i^{-1} x_{i3} \\ \sum_{i=1}^m 1_n^T V_i^{-1} x_{i1} & \sum_{i=1}^m x_{i1}^T V_i^{-1} x_{i1} & \sum_{i=1}^m x_{i1}^T V_i^{-1} x_{i2} & \sum_{i=1}^m x_{i1}^T V_i^{-1} x_{i3} \\ \sum_{i=1}^m 1_n^T V_i^{-1} x_{i2} & \sum_{i=1}^m x_{i1}^T V_i^{-1} x_{i2} & \sum_{i=1}^m x_{i2}^T V_i^{-1} x_{i2} & \sum_{i=1}^m x_{i2}^T V_i^{-1} x_{i3} \\ \sum_{i=1}^m 1_n^T V_i^{-1} x_{i3} & \sum_{i=1}^m x_{i1}^T V_i^{-1} x_{i3} & \sum_{i=1}^m x_{i2}^T V_i^{-1} x_{i3} & \sum_{i=1}^m x_{i3}^T V_i^{-1} x_{i3} \end{pmatrix},$$

where $V_i = \sigma^2 I_n + \sigma_0^2 1_n 1_n^T + \sigma_1^2 x_{i1} x_{i1}^T$, and x_{il} is the $l + 1$th column vector of the matrix X_i.

By Hadamard's determinant inequality, we obtain,

$$\det(I_\beta) \leq \sum_{i=1}^m 1_n^T V_i^{-1} 1_n \prod_{l=1}^3 \left(\sum_{i=1}^m x_{il}^T V_i^{-1} x_{il} \right).$$

The equality holds when the matrix is diagonal. Using the Exercise 2.9, page 33, (Rao 1973), we obtain,

$$V_i^{-1} \quad = \quad U^{-1} - \sigma_1^2 U^{-1} x_{i1} (\sigma_1^2 x_{i1}^T U^{-1} x_{i1} + 1)^{-1} x_{i1}^T U^{-1},$$

where $U^{-1} = \sigma^{-2} I_n - \sigma^{-2} \sigma_0^2 1_n (\sigma^2 + n\sigma_0^2)^{-1} 1_n^T$. Thus

$$x_{i2}^T V_i^{-1} x_{i2} \quad \leq \quad x_{i2}^T U^{-1} x_{i2} = \sigma^{-2} x_{i2}^T x_{i2} - \sigma^{-2} \sigma_0^2 x_{i2}^T 1_n (\sigma^2 + n\sigma_0^2)^{-1} 1_n^T x_{i2} \leq n\sigma^{-2}.$$

Similarly, $x_{i3}^T V_i^{-1} x_{i3} \leq n\sigma^{-2}$. From the proof of Theorem 2, we have,

$$\sum_{i=1}^m 1_n^T V_i^{-1} 1_n \leq \frac{mn}{\sigma^2 + \sigma_0^2 n}, \quad \sum_{i=1}^m x_{i1}^T V_i^{-1} x_{i1} \leq \frac{mn}{\sigma^2 + \sigma_1^2 n},$$

thus,

$$\det(I_\beta) \leq \frac{m^2 n^2}{\sigma^4} \prod_{l=0}^1 \frac{mn}{\sigma^2 + \sigma_l^2 n}.$$

On the other hand, for the design given in Theorem 3, the following equality holds exactly,

$$\det(I_\beta) = \frac{m^2 n^2}{\sigma^4} \prod_{l=0}^1 \frac{mn}{\sigma^2 + \sigma_l^2 n}.$$

Along with lines of this proof, we can show the design is D-optimal for estimating $(\beta^T, \sigma_0^2, \sigma^2)$. $\square$

5.4 Appendix: Proof of Theorem 4

Proof.

$$I_\beta = \begin{pmatrix} \sum_{i=1}^m 1_n^T V_i^{-1} 1_n & \sum_{i=1}^m 1_n^T V_i^{-1} x_{i1} & \sum_{i=1}^m 1_n^T V_i^{-1} x_{i2} \\ \sum_{i=1}^m 1_n^T V_i^{-1} x_{i1} & \sum_{i=1}^m x_{i1}^T V_i^{-1} x_{i1} & \sum_{i=1}^m x_{i1}^T V_i^{-1} x_{i2} \\ \sum_{i=1}^m 1_n^T V_i^{-1} x_{i2} & \sum_{i=1}^m x_{i1}^T V_i^{-1} x_{i2} & \sum_{i=1}^m x_{i2}^T V_i^{-1} x_{i2} \end{pmatrix},$$

where $V_i = \sigma^2 I_n + \sigma_0^2 1_n 1_n^T + \sigma_1^2 x_{i1} x_{i1}^T$, and x_{il} is the $l+1$th column vector of the matrix X_i. By Hadamard's determinant inequality, we obtain,

$$\det(I_\beta) \leq \sum_{i=1}^m 1_n^T V_i^{-1} 1_n \prod_{l=1}^2 \left(\sum_{i=1}^m x_{il}^T V_i^{-1} x_{il} \right).$$

The equality holds when the matrix is diagonal. Using Lemma 7, similar to the proof of Theorem 3 we obtain,

$$x_{i2}^T V_i^{-1} x_{i2} \leq n\sigma^{-2}$$

and

$$\det(I_\beta) \leq \frac{mn}{\sigma^2} \prod_{l=0}^1 \frac{mn}{\sigma^2 + \sigma_l^2 n}.$$

On the other hand, we know that, for the design described in Theorem 3,

$$\det(I_\beta) = \frac{mn}{\sigma^2} \prod_{l=0}^1 \frac{mn}{\sigma^2 + \sigma_l^2 n}.$$

Similarly, we can show the design is D-optimal for estimating $(\beta^T, \sigma_0^2, \sigma^2)$. ☐

Acknowledgement: The author is grateful to all anonymous referees for many helpful suggestions and comments. The author also thanks my friend Dr. Min Yang for drawing my attention to this topic.

References

Abt, M., Gaffke, N., Liski, E. P., and Sinha, B. K. (1998). *Optimal designs in growth curve models – II. Correlated model for quadratic growth: Optimal designs for parameter estimation and growth prediction*. Journal of Statistical Planning and Inference *67*, 287–296.

Atkins, J. E. and Cheng, C.-S. (1999). *Optimal regression designs in the presence of random block effects*. Journal of Statistical Planning and Inference *77*, 321–335.

Atkinson, A. C. and Donev, A. N. (1992). *Optimum Experimental Designs*. New York: Oxford.

Cheng, C.-S. (1995). *Optimal regression designs under random block-effects models*. Statistica Sinica *5*, 485–497.

Chow, S.-C. and Shao, J. (1989). *Test for batch-to-batch variation in stability analysis*. Statistics in Medicine *8*, 883–890.

Chow, S.-C. and Shao, J. (1991). *Estimating drug shelf-life with random batches*. Biometrics *47*, 1071–1079.

Chow, S.-C. and Wang, S.-G. (1994). *On the estimation of variance components in stability analysis*. Communications in Statistics, Part A – Theory and Methods [Split from: @J(CommStat)] *23*, 289–303.

Christensen, R. (1987). *Plane answers to complex questions.* Springer,.

Diggle, P. J., Heagetty, P., Liang, K.-Y., and Zeger, S. L. (2002). *Analysis of Longitudinal Data.* Oxford: Oxford University Press.

Hedayat, A., Sloane, N. J. A., and Stufken, J. (1999). *Orthogonal arrays: theory and applications.* Springer-Verlag Inc.

Hedayat, S., Stufken, J., and Yang, M. (2006). *Optimal and efficient crossover designs when subject effects are random.* J. Amer. Statist. Assoc. *101*, 1031–1038.

Liski, E. P., Mandal, N. K., Shah, K., and Sinha, B. K. (2002). *Topics in Optimal designs.* Springer-Verlag Inc.

Liski, E. P., Mandal, N. K., Sinha, B. K., and Abt, M. (1997). *Optimal designs in growth curve models: Part I, Correlated model for linear growth: Optimal designs for slope parameter estimation and growth prediction*. Journal of Statistical Planning and Inference *64*, 141–150.

Marcus, M. and Minc, H. (1992). *A survey of matrix theory and matrix inequality.* Dover publications.

Montgomery, D. C. (2001). *Design and analysis of experiments.* John Wiley & Sons.

Murphy, J. R. and Weisman, D. (1990). *Using random slopes for estimating shelf life*. In *ASA Proceedings of the Biopharmaceutical Section*, pp. 196–203.

Ouwens, M. J. N. M., Tan, F. E. S., and Berger, M. P. F. (2002). *Maximin D-optimal designs for Longitudinal mixed effects models*. Biometrics *58*, 735–741.

Pinheiro, J. C. and Bates, D. M. (2000). *Mixed-Effect Models in S and S-Plus.* Statistics and Computing. New York: Springer Verlag.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications.* Wiley.

Searle, S. R., Casella, G., and McCulloch, C. E. (1992). *Variance components.* John Wiley & Sons.

Silvey, S. D. (1980). *Optimum Designs.* London: Chapman & Hall.

Snijders, T. A. B. and Bosker, R. J. (1999). *Multilevel analysis: an introduction to basic and advanced multilevel modeling.* Sage Publications Inc.

6 Technical report

6.1 Technical report: Proof of Lemma 7

Let $U = \sigma^2 I_n + \sigma_0^2 1_n 1_n^T$, Let

$$\bar{x^2} = \frac{\sum_{j=1}^n x_j^2}{n}, \quad , \bar{x} = \frac{\sum_{j=1}^n x_j}{n},$$

$$\begin{aligned} V^{-1} &= U^{-1} - \sigma_1^2 U^{-1} x (\sigma_1^2 x^T U^{-1} x + 1)^{-1} x^T U^{-1}, \\ U^{-1} &= \sigma^{-2} I_n - \sigma^{-2} \sigma_0^2 1_n (\sigma^2 + n\sigma_0^2)^{-1} 1_n^T. \end{aligned}$$

Simple computation leads to

$$U^{-1}1_n = \frac{1_n}{\sigma^2 + n\sigma_0^2},\ U^{-1}x = \frac{1}{\sigma^2}\left[x - \frac{n\sigma_0^2}{\sigma^2 + n\sigma_0^2}\bar{x}1_n\right].$$

So

$$\begin{aligned}
1_n^T U^{-1} 1_n &= \frac{n}{\sigma^2 + n\sigma_0^2}, \\
x^T U^{-1} 1_n &= \frac{n\bar{x}}{\sigma^2 + n\sigma_0^2}, \\
x^T U^{-1} x &= \frac{1}{\sigma^2}\left[n\bar{x^2} - \frac{n^2\bar{x}^2\sigma_0^2}{\sigma^2 + n\sigma_0^2}\right].
\end{aligned}$$

Thus

$$\sigma_1^2 x^T U^{-1} x + 1 = 1 + \frac{\sigma_1^2}{\sigma^2}\left[n\bar{x^2} - \frac{n^2\bar{x}^2\sigma_0^2}{\sigma^2 + n\sigma_0^2}\right].$$

Now,

$$\begin{aligned}
1_n^T V^{-1} 1_n &= 1_n^T U^{-1} 1_n - \sigma_1^2 1_n^T U^{-1} x(\sigma_1^2 x^T U^{-1} x + 1)^{-1} x^T U^{-1} 1_n \\
&= \frac{n}{\sigma^2 + n\sigma_0^2} - \sigma_1^2 \frac{n\bar{x}}{\sigma^2 + n\sigma_0^2}\left[1 + \frac{\sigma_1^2}{\sigma^2}\left\{n\bar{x^2} - \frac{n^2\bar{x}^2\sigma_0^2}{\sigma^2 + n\sigma_0^2}\right\}\right]^{-1}\frac{n\bar{x}}{\sigma^2 + n\sigma_0^2} \\
&= \frac{n}{\sigma^2 + n\sigma_0^2} - \sigma_1^2 \frac{n^2\bar{x}^2}{(\sigma^2 + n\sigma_0^2)^2}\left[1 + \frac{\sigma_1^2}{\sigma^2}\left\{n\bar{x^2} - \frac{n^2\bar{x}^2\sigma_0^2}{\sigma^2 + n\sigma_0^2}\right\}\right]^{-1} \\
&= \frac{1}{\sigma^2 + n\sigma_0^2}\frac{n + n^2\sigma_1^2\sigma^{-2}\bar{x^2} - n^3\sigma_1^2\sigma^{-2}\sigma_0^2/(\sigma^2 + n\sigma_0^2)\bar{x}^2 - n^2\sigma_1^2/(\sigma^2 + n\sigma_0^2)\bar{x}^2}{1 + \frac{\sigma_1^2}{\sigma^2}\left\{n\bar{x^2} - \frac{n^2\bar{x}^2\sigma_0^2}{\sigma^2 + n\sigma_0^2}\right\}} \\
&= \frac{1}{\sigma^2 + n\sigma_0^2}\frac{n + n^2\sigma_1^2\sigma^{-2}\bar{x^2} - n^2\sigma_1^2\sigma^{-2}\bar{x}^2}{1 + \frac{\sigma_1^2}{\sigma^2}\left\{n\bar{x^2} - \frac{n^2\bar{x}^2\sigma_0^2}{\sigma^2 + n\sigma_0^2}\right\}}.
\end{aligned}$$

Similarly

$$\begin{aligned}
x^T V^{-1} 1_n &= x^T U^{-1} 1_n - \sigma_1^2 x^T U^{-1} x(\sigma_1^2 x^T U^{-1} x + 1)^{-1} x^T U^{-1} 1_n \\
&= \frac{1}{\sigma^2 + n\sigma_0^2}\frac{n\bar{x}}{1 + \frac{\sigma_1^2}{\sigma^2}\left\{n\bar{x^2} - \frac{n^2\bar{x}^2\sigma_0^2}{\sigma^2 + n\sigma_0^2}\right\}}, \\
x^T V^{-1} x &= x^T U^{-1} x - \sigma_1^2 x^T U^{-1} x(\sigma_1^2 x^T U^{-1} x + 1)^{-1} x^T U^{-1} x \\
&= \frac{\frac{1}{\sigma^2}\left\{n\bar{x^2} - \frac{n^2\bar{x}^2\sigma_0^2}{\sigma^2 + n\sigma_0^2}\right\}}{1 + \frac{\sigma_1^2}{\sigma^2}\left\{n\bar{x^2} - \frac{n^2\bar{x}^2\sigma_0^2}{\sigma^2 + n\sigma_0^2}\right\}}.
\end{aligned}$$

6.2 Technical report: Proof of Theorem 6

Proof. As in the proof of Theorem 5,

$$\sum_{i=1}^{m} X_i^T V_i^{-1} X_i \quad = \quad \sum_{i=1}^{m} \left[X_i^T X_i (X_i^T X_i + \Psi^{-1}\sigma^2)^{-1}\right] \Psi^{-1}.$$

By Hadamard's determinant inequality, we obtain

$$\det\left\{\sum_{i=1}^{m} \left[X_i^T X_i (X_i^T X_i + \Psi^{-1}\sigma^2)^{-1}\right] \Psi^{-1}\right\} \leq \prod_{k=1}^{4l} \frac{\lambda_{kk}}{\sigma_{k-1}^2},$$

where λ_{kk} is the kth diagonal element of the matrix $\sum_{i=1}^{m} \left[X_i^T X_i (X_i^T X_i + \Psi^{-1}\sigma^2)^{-1}\right]$. By Lemmas 9 and 8, we have $\lambda_{kk} \leq mn\sigma_{k-1}^2/(\sigma^2 + n\sigma_{k-1}^2)$. Thus,

$$\det\left(\sum_{i=1}^{m} X_i^T V_i^{-1} X_i\right) \leq \prod_{l=0}^{4l-1} \frac{mn}{\sigma^2 + n\sigma_l^2}.$$

The equality holds if and only if assumptions in Lemma 8 holds. for all $k, p = 1, 2, 3, 4l$, i.e. X_i is the design matrix for design described in Theorem 5. □