Optimal designs for estimation and prediction in simple random-intercept models

YUE Rongxian[1], ZHOU Xiaodong[2]

(1. College of Mathematics and Sciences, Shanghai Normal University, Shanghai 200234, China; 2. School of Business Information Management, Shanghai University of Iuternational Business and Eemomics, Shanghai 201620, China)

Abstract: The paper is concerned with the optimal design problem of estimating linear combinations of the fixed and random effects, and predicting future observations of individual responses in a random intercept model. The variance components in the model are assumed to be known. The design criteria for the predictions are obtained from the mean squared error (MSE) of the estimator and the mean squared prediction error (MSPE) of the predictor. The exact n-point optimal designs and approximate optimal designs are discussed.

Key words: random intercept model; optimal designs; mean squared error

CLC number: O 212.6 **Document code:** A **Article ID:** 1000-5137(2013)05-0441-10

1 Introduction

Mixed models are widely used in various disciplines including agriculture, biology, medicine, physical sciences, education, and social and behavioral sciences. Mixed models incorporate both fixed effects and random effects. The fixed effects represent the mean values of the parameters in the population of individuals, and the random effects represent the individual deviations in the population. While the statistical analysis for mixed models has been well-developed, the optimal design problems of such models have become attractive in the recent two decades.

Optimal designs are considered in the presence of random block effects in [1-4]. In [5] and [6] optimal designs are considered for linear and quadratic growth mixed-effect models with interclass correlation structure and autocorrelated structure. In [7] and [8], the maximin D-optimal designs are discussed for a random intercept and random slope longitudinal mixed-effects models. The D-optimal designs are considered in [9] for linear regression models with a random intercept and first order autoregressive serial correlations. Schmelter[10-11] discussed the optimality of designs for single-group designs for certain mixed models, and then extended his results to group-wise designs for linear mixed models. Debusho and Haines[12] considered the V- and D-optimal population designs for the simple linear regression model with a random intercept term. More recently, Cheng et

Received date: 2013-05-20

Foundation item: This work is partially supported by the NSFC Grant (11071168); Special Funds for Doctoral Authorities of the Education Ministry (20103127110002); Innovation Program of Shanghai Municipal Education Commission (11zz116); E-Institutes of Shanghai Municipal Education Commission (E03004); Shanghai Leading Academic Discipline Project.

Biography: YUE Rongxian(1958-), male, professor, College of Mathematics and Sciences, Shanghai Normal University; ZHOU Xiaodong(1979-), male, associate professor, School of Business Information Management, Shanghai Universty of International Business and Economics.

al. [13] considered the optimal designs based on D-, G-, A-, I- and D_β-optimality criteria for random coefficient regression models with heteroscedastic errors.

Most research papers on optimal designs in mixed models are for estimating fixed effects and predicting the population mean response. However, the estimation of linear combinations of fixed and random effects and the prediction of individual responses can sometimes be of great interest. Such estimation and prediction problems often arise in many practical applications, such as the estimation of quality index, longitudinal studies, the selection index in quantitative genetics, plant varietal trials and small area estimation [14]. For example, in the area of medicine and health sciences, average growth curves should be studied, but estimating individual growth curves may also be very important because they may be useful in screening subjects for being at risk for some disease. Individual growth curves can also be used in predicting future observations of individuals. There are few studies on the design issues for estimation of linear combinations of fixed and random effects and prediction of individual responses in mixed models in literature.

In this paper we address the optimal design problem of estimating linear combinations of fixed and random effects, and predicting individual responses in a simple linear mixed model. The model that we assume is a simple random-intercept one which is the simplest mixed model. The random intercepts are assumed to have zero mean and non-zero mean, respectively. The variance components in the model are assumed to be known. The design criteria are obtained from the mean squared error (MSE) of the estimator and the mean squared prediction error (MSPE) of the predictor.

The rest of the paper is organized as follows: In Section 1 we specify the simple random intercept model, and consider the optimal design problem of estimating the linear combination of fixed and random effects. Section 2 considers the optimal design problem of predicting individual responses beyond the design region. Some concluding remarks are given in Section 3.

1 Model specification and optimal designs for estimation

1.1 Model specification

We consider the situation that we have K individuals with n observations each at the experimental settings $x_1, \cdots, x_n$ in the region $[0, h]$, and the individual effects have only influences on the overall level of the response. Let Y_{ij} be the jth observation of individual i at the setting x_j. We consider both a random-intercept model with and without a population mean intercept given by

$$y_{ij} = \beta_0 + \beta_1 x_j + b_i + \varepsilon_{ij}, i = 1, \cdots, K; j = 1, \cdots, n, \tag{1}$$

and

$$y_{ij} = \beta_1 x_j + b_i + \varepsilon_{ij}, i = 1, \cdots, K; j = 1, \cdots, n. \tag{2}$$

Here the term b_i denotes the individual random intercept distributed as $N(0, \sigma_b^2)$, the random observational errors ε_{ij} are assumed to be homoscedastic and distributed as $N(0, \sigma_\varepsilon^2)$, and all b_i and ε_{ij} are independent both within and between individuals. The intercept β_0 and the slope β_1 are fixed effects and the parameters σ_b^2 and σ_ε^2 comprise the variance components.

Both models (1) and (2) can be unified in matrix form as follows:

$$\boldsymbol{y}_i = \boldsymbol{X\beta} + b_i \mathbf{1}_n + \boldsymbol{\varepsilon}_i, i = 1, \cdots, K, \tag{3}$$

where $\boldsymbol{y}_i = (y_{i1}, \cdots, y_{in})^{\mathrm{T}}$ is the observation vector of individual i, $\boldsymbol{X} = [\mathbf{1}_n, \boldsymbol{x}]$ for model (1) and $\boldsymbol{X} = \boldsymbol{x}$ for model (2), where $\mathbf{1}_n$ the $n \times 1$ vector of 1's and $\boldsymbol{x} = (x_1, \cdots, x_n)^{\mathrm{T}}$. $\boldsymbol{\beta} = (\beta_0, \beta_1)^{\mathrm{T}}$ for model (1) and $\boldsymbol{\beta} = \beta_1$ for model (2). $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \cdots, \varepsilon_{in})^{\mathrm{T}}$. Furthermore, $b_i \sim N(0, \sigma_b^2)$, $\boldsymbol{\varepsilon}_i \sim N(0, \sigma_\varepsilon^2 \boldsymbol{I}_n)$, and b_i and $\boldsymbol{\varepsilon}_i$ are independent for $i=1,\cdots,K$. It is assumed that the variance components $\sigma_\varepsilon^2,\sigma_b^2$ are known. Then the population mean response and the variance of the response of individual i are given by

$$E(\boldsymbol{y}_i)=\boldsymbol{X\beta},\ \boldsymbol{\Sigma}=\mathrm{var}(\boldsymbol{y}_i)=\sigma_\varepsilon^2\boldsymbol{I}_n+\sigma_b^2\boldsymbol{J}_n,$$

for $i=1,\cdots,K$, where $\boldsymbol{J}_n=\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}$.

We further combine all the individual observations and error vectors, $\boldsymbol{y}_i$ and $\boldsymbol{\varepsilon}_i$, to establish the N-dimensional vectors $\boldsymbol{y}=(\boldsymbol{y}_1^{\mathrm{T}},\cdots,\boldsymbol{y}_K^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{\varepsilon}=(\boldsymbol{\varepsilon}_1^{\mathrm{T}},\cdots,\boldsymbol{\varepsilon}_K^{\mathrm{T}})^{\mathrm{T}}$, respectively. Here, $N=Kn$ is the total number of observations across the whole sample of individuals. The N-dimensional observation vector $\boldsymbol{y}$ can then be written as

$$\boldsymbol{y}=\boldsymbol{F\beta}+\boldsymbol{Zb}+\boldsymbol{e},\tag{4}$$

where $\boldsymbol{b}=(b_1,\cdots,b_K)^{\mathrm{T}}$, $\boldsymbol{Z}=\boldsymbol{I}_K\otimes\mathbf{1}_n$, $\boldsymbol{F}=1_K\otimes\boldsymbol{X}$. It follows that

$$E(\boldsymbol{y})=\boldsymbol{F\beta},\quad \boldsymbol{V}=\mathrm{cov}(\boldsymbol{y})=\boldsymbol{I}_K\otimes\boldsymbol{\Sigma},\quad \boldsymbol{G}=\mathrm{cov}(\boldsymbol{b})=\sigma_b^2\boldsymbol{I}_K.$$

1.2 Exact optimal design for estimation

We mainly consider the problem of estimating a general linear combination of the fixed and random effects of the following form

$$\mu=\boldsymbol{l}^{\mathrm{T}}\boldsymbol{\beta}+\boldsymbol{m}^{\mathrm{T}}\boldsymbol{b},$$

where $\boldsymbol{l}$ and $\boldsymbol{m}$ are known vectors which occur frequently in the paper. When the components of $\boldsymbol{\theta}=(\sigma_\varepsilon^2,\sigma_b^2)^{\mathrm{T}}$ is known, the best linear unbiased estimation (BLUE) of μ under model (4) is [15]

$$\hat{\mu}=\boldsymbol{l}^{\mathrm{T}}\hat{\boldsymbol{\beta}}+\boldsymbol{m}^{\mathrm{T}}\hat{\boldsymbol{b}},\tag{5}$$

where

$$\hat{\boldsymbol{\beta}}=(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{F})^{-1}\boldsymbol{F}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}=(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\bar{\boldsymbol{y}},\ \bar{\boldsymbol{y}}=\frac{1}{K}\sum_{i=1}^{K}\boldsymbol{y}_i,\tag{6}$$

and

$$\hat{\boldsymbol{b}}=\boldsymbol{GZ}^{\mathrm{T}}\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{F}\hat{\boldsymbol{\beta}}).\tag{7}$$

Let $\xi_n=\{x_1,x_2,\cdots,x_n\}$ be an exact n-point design on $[0,h]$, where $x_j\in[0,h]$, $j=1,\cdots,n$ with at least two distinct x_j's. We now derive the MSE of $\hat{\mu}(\boldsymbol{\theta})$ in (5). Note that the covariance matrix of $\hat{\boldsymbol{\beta}}$ in (6) is

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}})=(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{F})^{-1}=\frac{1}{K}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1},\tag{8}$$

and the MSE of $\hat{\boldsymbol{b}}$ in (7) is

$$\mathrm{MSE}(\hat{\boldsymbol{b}})=E(\hat{\boldsymbol{b}}-\boldsymbol{b})(\hat{\boldsymbol{b}}-\boldsymbol{b})^{\mathrm{T}}=\boldsymbol{G}-\boldsymbol{GZ}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{ZG}+\boldsymbol{G}^{\mathrm{T}}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{F}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{F})^{-1}\boldsymbol{F}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{ZG}.\tag{9}$$

Defining $\boldsymbol{S}=\boldsymbol{V}^{-1}\boldsymbol{ZGm}$, and noting that $\hat{\mu}$ in (5) is unbiased in the sense $E(\hat{\mu}-\mu)=0$, we see that the MSE of $\hat{\mu}$ is given by

$$\mathrm{MSE}(\hat{\mu},\xi_n)=E(\hat{\mu}-\mu)^2=g_1(\xi_n)+g_2(\xi_n),\tag{10}$$

where

$$g_1(\xi_n)=\boldsymbol{m}^{\mathrm{T}}(\boldsymbol{G}-\boldsymbol{GZ}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{ZG})\boldsymbol{m},$$

$$g_2(\xi_n)=\frac{1}{K}(\boldsymbol{l}-\boldsymbol{F}^{\mathrm{T}}\boldsymbol{S})^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}(\boldsymbol{l}-\boldsymbol{F}^{\mathrm{T}}\boldsymbol{S}).$$

For any design ξ_n on $[0,h]$, define

$$\tau=\frac{\sigma_b^2}{\sigma_\varepsilon^2},\quad \gamma=\frac{n\tau}{1+n\tau},\tag{11}$$

and

$$u_1 = \frac{1}{n}\sum_{j=1}^{n} x_j, u_2 = \frac{1}{n}\sum_{j=1}^{n} x_j^2. \tag{12}$$

It is easy to verify that

$$0 < u_1 < h, \quad u_1^2 < u_2 \leqslant h u_1. \tag{13}$$

The following theorem gives the expressions for g_1 and g_2.

Theorem 1 If the design matrix $\boldsymbol{X}$ in model (3) has full column rank, then $g_1(\xi_n)$ and $g_2(\xi_n)$ can be expressed as

$$g_1(\xi_n) = \frac{\sigma_\varepsilon^2 \gamma m_{**}}{n}, \tag{14}$$

and

$$g_2(\xi_n) = \begin{cases} \frac{\sigma_\varepsilon^2}{K}\left\{\frac{(l_1^2 u_2 - 2l_1 l_2 u_1 + l_2^2)}{n(u_2 - u_1^2)} + \tau l_1^2 - 2\tau l_1 m_* + \tau\gamma\right\} & \text{for model (1)} \\ \frac{\sigma_\varepsilon^2 (l - \gamma m_* u_1)^2}{N(u_2 - \gamma u_1^2)} & \text{for model (2)} \end{cases}, \tag{15}$$

where $m_* = \sum_{i=1}^{K} m_i$ and $m_{**} = \sum_{i=1}^{K} m_i^2$.

It follows that the MSE of $\hat{\mu}$ can be expressed as

$$\mathrm{MSE}(\hat{\mu}, \xi_n) = \begin{cases} \sigma_\varepsilon^2\left[\frac{(l_1^2 u_2 - 2l_1 l_2 u_1 + l_2^2)}{N(u_2 - u_1^2)} + q(\boldsymbol{l}, \boldsymbol{m}, \tau)\right] & \text{for model (1)} \\ \sigma_\varepsilon^2\left[\frac{(\gamma m_* u_1 - l)^2}{N(u_2 - \gamma u_1^2)} + \frac{\gamma m_{**}}{n}\right] & \text{for model (2)} \end{cases}, \tag{16}$$

where the term $q(\boldsymbol{l}, \boldsymbol{m}, \tau)$ is as follows

$$q(\boldsymbol{l}, \boldsymbol{m}, \tau) = \frac{\tau}{K}(l_1^2 - 2 l_1 m_* + \gamma m_*^2) + \frac{\gamma m_{**}}{n},$$

which does not depend on the design points x_i.

Therefore, in the case of known variance components $\boldsymbol{\theta}$ of the models (1) and (2), minimizing $\mathrm{MSE}(\hat{\mu}, \xi_n)$ in (16) over the possible n-point designs ξ_n is equivalent to minimizing the following criterion function $\Phi(\xi_n)$:

$$\Phi(\xi_n) = \begin{cases} \frac{l_1^2 u_2 - 2l_1 l_2 u_1 + l_2^2}{u_2 - u_1^2} & \text{for model (1)} \\ \frac{(l - \gamma m_* u_1)^2}{u_2 - \gamma u_1^2} & \text{for model (2)} \end{cases}. \tag{17}$$

A design, ξ_n^*, is said to be an MSE-optimal for predicting $\mu = \boldsymbol{l}^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{m}^{\mathrm{T}}\boldsymbol{b}$ if it minimizes $\Phi(\xi_n)$ in (17) over all the possible designs ξ_n on the region $[0, h]$.

Remark 1 For model (1), where an explicit population mean intercept β_0 is present, if the variance components $\boldsymbol{\theta}$ are known, then the MSE-optimal design for estimating $\boldsymbol{l}^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{m}^{\mathrm{T}}\boldsymbol{b}$ depends only on the coefficients $\boldsymbol{l}$, but not on $\boldsymbol{m}$. Furthermore, the MSE-optimality is just the c-optimality for the corresponding fixed effects model

$$y(x_j) = \beta_0 + \beta_1 x_j + \varepsilon_j, E(\varepsilon_j) = 0, \mathrm{var}(\varepsilon_j) = \sigma_\varepsilon^2, \quad j = 1, \cdots, n, \tag{18}$$

since

$$\Phi(\xi_n) = \frac{l_1^2 u_2 - 2l_1 l_2 u_1 + l_2^2}{u_2 - u_1^2} = \boldsymbol{l}^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{l}.$$

Hence, the c-optimal designs for model (18) remains MSE-optimal for model (1) with all values of γ. Note that the c-optimality for linear regression models including (18) has been well investigated in literature.

Remark 2 For model (2), where the population mean is through the origin, the MSE-optimal design for estimating $\boldsymbol{l}^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{m}^{\mathrm{T}}\boldsymbol{b}$ may depend not only on the coefficients $\boldsymbol{l}$ and $\boldsymbol{m}$, but also on the value of γ.

The following two cases for estimating $\mu = \boldsymbol{l}^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{m}^{\mathrm{T}}\boldsymbol{b}$ in the mixed model attract more attention in practice:

Case 1 For any $x \in [0,h]$, $\boldsymbol{l} = (1,x)^{\mathrm{T}}$ for model (1) and $\boldsymbol{l} = x$ for model (2), and $\boldsymbol{m} = 0$. Then $\mu = \boldsymbol{l}^{\mathrm{T}}\boldsymbol{\beta}$ is just the population mean response at $x \in [0,h]$ for all individuals. The MSE-criterion in (17) becomes

$$\Phi_{\mathrm{pop}}(\xi_n, x) = \begin{cases} \dfrac{u_2 - 2u_1 x + x^2}{u_2 - u_1^2} & \text{for model (1)} \\ \dfrac{x^2}{u_2 - \gamma u_1^2} & \text{for model (2)} \end{cases}. \tag{19}$$

Note that the designs minimizing (19) will depend on $x \in [0,h]$ for model (1), and not on x for model (2). One may consider the following designs:

$$\begin{cases} \text{The IMSE}_{\mathrm{pop}}\text{-optimal design which minimizes integral } \int_0^h \Phi_{\mathrm{pop}}(\xi_n, x)\,\mathrm{d}x \\ \text{The MMSE}_{\mathrm{pop}}\text{-optimal design which minimizes } \max\limits_{x \in [0,h]} \Phi_{\mathrm{pop}}(\xi_n, x) \end{cases}.$$

Case 2 For any $x \in [0,h]$, $\boldsymbol{l} = (1,x)^{\mathrm{T}}$ for model (1) and $\boldsymbol{l} = x$ for model (2), and $\boldsymbol{m} = (0,\cdots,1,\cdots,0)^{\mathrm{T}}$ with 1 in the ith cell and zero elsewhere. Then $\mu = \boldsymbol{l}^{\mathrm{T}}\boldsymbol{\beta} + b_i$ is just the ith individual response at $x \in [0,h]$. The MSE-criterion in (17) becomes

$$\Phi_{\mathrm{ind}}(\xi_n, x) = \begin{cases} \dfrac{u_2 - 2u_1 x + x^2}{u_2 - u_1^2} & \text{for model (1)} \\ \dfrac{(\gamma u_1 - x)^2}{u_2 - \gamma u_1^2} & \text{for model (2)} \end{cases}. \tag{20}$$

Note that the designs minimizing (20) will depend on $x \in [0,h]$, but not on i for both models (1) and (2). One may consider the following designs:

$$\begin{cases} \text{The IMSE}_{\mathrm{ind}}\text{-optimal design which minimizes integral } \int_0^h \Phi_{\mathrm{ind}}(\xi_n, x)\,\mathrm{d}x. \\ \text{The MMSE}_{\mathrm{ind}}\text{-optimal design which minimizes } \max\limits_{x \in [0,h]} \Phi_{\mathrm{ind}}(\xi_n, x). \end{cases}$$

To construct the exact n-point MSE-optimal designs, we introduce an algorithm in the following. Lemma 1 (18) is helpful for constructing the n-point optimal designs.

Lemma 1 Let Φ: $[0,h] \to \mathrm{R}$ be a criterion of design optimality. Assume that for all designs $\xi_n^{(1)} = (x_1^{(1)},\cdots,x_n^{(1)})^{\mathrm{T}} \in [0,h]^n$ and $\xi_n^{(2)} = (x_1^{(2)},\cdots,x_n^{(2)})^{\mathrm{T}} \in [0,h]^n$ satisfying $u_1^{(1)} = \frac{1}{n}\sum_{j=1}^n x_j^{(1)} = \frac{1}{n}\sum_{j=1}^n x_j^{(2)} = u_1^{(2)}$ and $u_2^{(1)} = \frac{1}{n}\sum_{j=1}^n (x_j^{(2)})^2 > \frac{1}{n}\sum_{j=1}^n (x_j^{(2)})^2 = u_2^{(2)}$ we have $\Phi(\xi_n^{(1)}) \leqslant \Phi(\xi_n^{(2)})$. Then there exists an n-point Φ-optimal design ξ_n^* satisfying

$$x_1 = \cdots = x_s = 0, \quad \text{and} \quad x_{s+2} = \cdots = x_n = h$$

for some $s \in \{0,1,\cdots,n-1\}$.

Note that for any given $\boldsymbol{l}$ and $\boldsymbol{m}$, the MSE-criterion function $\Phi(\xi_n)$ in (17) satisfies the condition in Lemma 1, since $\partial\Phi(\xi_n)/\partial u_2 \leqslant 0$.

Therefore, it is necessary to choose the n-point designs with only at most three distinct points in $\{0,t,h\}$ with some $t \in [0,h]$. So, the algorithm can be described as follows:

Step 1 For each $s \in \{0,1,\cdots,n-1\}$, set $x_j = 0$ for $j = 1,\cdots,s$, and $x_j = h$ for $j = s+1,\cdots,n-1$, and use a one-dimensional optimization routine to find the point $x_n \in [0,h]$ such that the design $\xi_n^s = \{x_1,\cdots,x_{n-1},x_n\}$ has the minimal value of Φ.

Step 2 Choose $s_* \in \{0,1,\cdots,n-1\}$ such that

$$\Phi(\xi_n^{s_*}) = \min\{\Phi(\xi_n^s), s = 0,1,\cdots,n-1\}.$$

Then $\xi_n^{s_*}$ is the MSE-optimal design.

1.3 Approximate optimal designs for the estimation

Due to Remark 1, we will only focus on the optimal designs for prediction in model (2) in the design region $[0,h]$ in this subsection. Note that the MSE-criterion function given in (17) for model (2) decreases as u_2 increases for any u_1 that remains constant. This means that if u_1 is fixed, then the design improves when the variance of $\boldsymbol{x} = (x_1,\cdots,x_n)^{\mathrm{T}}$ increases. To maximize the variance of $\boldsymbol{x}$ we have to place as many values x_j as possible close to 0 and h. On the other hand, for polynomial regression models one would like to choose the designs with minimum support points. So, we now consider approximate MSE-optimal designs with two support points 0 and h. That is, we consider approximate designs of the following form

$$\xi = \begin{Bmatrix} 0 & h \\ 1-w & w \end{Bmatrix}, \text{ where } 0 \leqslant w \leqslant 1. \tag{21}$$

The MSE-criterion function in (17) for model (2) at the approximate design ξ becomes

$$\Phi(\xi) = \frac{(l - \gamma h m_* w)^2}{h^2 w(1-\gamma w)}. \tag{22}$$

Minimizing (22) with respect to $w \in [0,1]$, one can obtain the optimal weight w^* in (21). The following theorem gives a result on optimal weights w^* for prediction of the population mean response and the ith individual response at $x \in [0,h]$ in model (2) under the IMSE- and MMSE-optimality, respectively.

Theorem 2 Consider model (2) with known variance components. The optimal weights of the approximate designs for predicting the population mean response $x\beta$ and the ith individual response $x\beta + b_i$, $x \in [0,h]$, under the IMSE- and MMSE-optimality criteria, respectively, are all the same as follows:

$$w^* = \begin{cases} \dfrac{1}{2\gamma} & \gamma > \dfrac{1}{2} \\ 1 & \gamma \leqslant \dfrac{1}{2} \end{cases},$$

where γ is as in (11).

2 Optimal designs for future responses of the individuals

Here, we consider the design problems of predicting a future observation in models (1) and (2). Suppose one wants to predict the response value y_i of individual i which belongs to the experimental subjects at a point $x \in (h,H]$. Note that

$$\mathrm{Cov}(\boldsymbol{y}_i, y_i(x)) = \sigma_b^2 \mathbf{1}_n.$$

When the components of $\boldsymbol{\theta}$ are known, the best linear unbiased predictor (BLUP) of $y_i(x)$, $x \in (h,H]$, is as follows:

$$\hat{y}_i(x) = \boldsymbol{x}_f^T\hat{\boldsymbol{\beta}} + \sigma_b^2\mathbf{1}_n^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i - \boldsymbol{X}\hat{\boldsymbol{\beta}}), \tag{23}$$

where $\boldsymbol{x}_f = (1, x)^{\mathrm{T}}$ for model (1) and $\boldsymbol{x}_f = x$ for model (2).

The mean squared prediction error (MSPE) of the predictor $\hat{y}_i(x)$ under the n-point design ξ_n is given by

$$\mathrm{MSPE}(\hat{y}_i(x), \xi_n) = \mathrm{E}[\hat{y}_i(x) - y_i(x)]^2 =$$

$$\mathrm{var}(y_i(x)) - \sigma_b^4\mathbf{1}_n^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{1}_n + \frac{1}{K}(\boldsymbol{x}_f - \sigma_b^2\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{1}_n)^{\mathrm{T}}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}(\boldsymbol{x}_f - \sigma_b^2\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{1}_n).$$

A straightforward calculation gives

$$\mathrm{MSPE}(\hat{y}_i(x), \xi_n) = \begin{cases} \sigma_\varepsilon^2\left[\dfrac{u_2 - 2u_1x + x^2}{N(u_2 - u_1^2)} + \dfrac{(K-1)\gamma}{K} + 1\right] & \text{for model (1)} \\ \sigma_\varepsilon^2\left[\dfrac{(x - \gamma u_1)^2}{N(u_2 - \gamma u_1^2)} + \dfrac{\gamma}{n} + 1\right] & \text{for model (2)} \end{cases}. \tag{24}$$

Note that this expression does not depend on i. Similar to the discussion in the previous subsection, we consider the approximate designs ξ of form (21), and find the expression for $\mathrm{MSPE}(\hat{y}_i(x), \xi)$. The corresponding MSPE-optimality criterion for predicting a future response at $x \in (h, H]$ is given by

$$\Psi(\xi, x) = \begin{cases} \dfrac{h^2w - 2hxw + x^2}{w - w^2} & \text{for model (1)} \\ \dfrac{(x - \gamma hw)^2}{w - \gamma w^2} & \text{for model (2)} \end{cases}. \tag{25}$$

Integrating $\Psi(\xi, x)$ with respect to x over $[h, H]$, we obtain the IMSPE-optimality criterion for predicting the response $y_i(x)$ over $(h, H]$ as follows:

$$\Psi_{IMSPE}(\xi) = \begin{cases} \dfrac{(H^2 + Hh + h^2)/3 - Hhw}{w - w^2} & \text{for model (1)} \\ \dfrac{(H^2 + Hh + h^2)/3 + \gamma^2h^2w - (H+h)\gamma hw}{w - \gamma w^2} & \text{for model (2)} \end{cases}, \tag{26}$$

Maximizing $\Psi(\xi, x)$ with respect to x over $[h, H]$, we obtain the MMSPE-optimality criterion for the prediction of the response $y_i(x)$ over $(h, H]$ as follows:

$$\Psi_{MMSPE}(\xi) = \begin{cases} \dfrac{h^2w - 2Hhw + H^2}{w - w^2} & \text{for model (1)} \\ \dfrac{(\gamma hw - H)^2}{w - \gamma w^2} & \text{for model (2)} \end{cases}, \tag{27}$$

Theorem 3 Consider models (1) and (2) with known variance components. For $H > h > 0$ let $\alpha = h/H$ and $A = (\alpha^2 + \alpha + 1)/3$. Then the weights of the approximate optimal designs for predicting the future response $y_i(x), x \in (h, H]$, under the IMSPE- and MMSPE-optimality criteria, respectively, are

$$w^*_{IMSPE} = \frac{A}{\alpha} - \sqrt{\frac{A^2}{\alpha^2} - \frac{A}{\alpha}}, w^*_{MMSPE} = \frac{1}{2 - \alpha}$$

for model (1), and

$$w^*_{IMSPE} = \begin{cases} \dfrac{A - \sqrt{A^2 - A[(1-\gamma)\alpha^2 + \alpha]}}{\gamma[\alpha^2(1-\gamma) + \alpha]} & \text{if } \gamma > \dfrac{2A-1}{\alpha^2}, \alpha^2\gamma^3 - (\alpha^2 + \alpha)\gamma^2 + 2A\gamma - A > 0 \\ 1 & \text{otherwise} \end{cases},$$

$$w^*_{MMSPE} = \begin{cases} \dfrac{1}{\gamma(2-\alpha)} & \text{if } \gamma > \dfrac{1}{2-\alpha} \\ 1 & \text{otherwise} \end{cases}$$

for model (2).

The plots of the optimal weights w^*_{IMSPE} and w^*_{MMSPE} vs the ratio $\alpha = h/H$ for models (1) and (2) are shown in Fig. 1 and Fig. 2, respectively. The following facts were observed:

(i) For the prediction of future responses on $(h, H]$ in models (1) and (2), the IMSPE- and MMSPE-optimal designs are seriously affected by the ratio $\alpha = h/H$.

(ii) The optimal weights at the setting h are all greater than 0.5 for all $\alpha > 0$.

(iii) For smaller values of α, the observations are evenly required at the two extreme settings of the design region $[0, h]$; For larger values of α, more observations are required at the extreme setting h.

(iv) In model (2), the IMSPE- and MMSPE-optimal designs are also seriously affected by the ratio $\gamma = n\sigma_b^2/(\sigma_\varepsilon^2 + n\sigma_b^2)$. For the same value of α, the smaller value of γ is, the more observations are required at the extreme setting h.

(v) There is not much difference between the two designs for the same α and same γ.

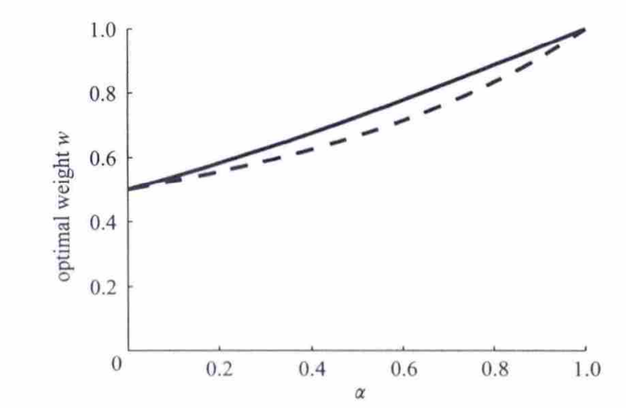

Figure 1 The optimal weights w^*_{IMSPE} (solid) and w^*_{MMSPE} (dashed) vs the ratio $\alpha = h/H$ for model (1)

3 Concluding remarks

The present paper studied the optimal design problems of estimation and prediction in simple random intercept models based on the MSE and MSPE optimality criteria on the design region $[0, h]$, respectively. For the estimation of the linear combination $\mu = \boldsymbol{l}^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{m}^{\mathrm{T}}\boldsymbol{b}$ in the random intercept model with a fixed intercept, the MSE-optimal designs depend on $\boldsymbol{l}$ but not on $\boldsymbol{m}$. For estimation of $\mu = \boldsymbol{l}^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{m}^{\mathrm{T}}\boldsymbol{b}$ in the random intercept model without a fixed intercept, the MSE-optimal designs depend not only on $\boldsymbol{l}$ and $\boldsymbol{m}$, but also on the ratio τ of the variance components. The IMSPE- and MMSPE-optimal designs in both models for predicting future observations of any individual response on $(h, H]$ depend on the ratio $\alpha = h/H$. Furthermore, the optimal designs in the model without a fixed intercept also depend on $\gamma = n\sigma_b^2/(\sigma_e{}^2 + n\sigma_b^2)$.

However, the variance components of the models are usually unknown in practice. It will be interesting to consider the effect of estimation for the variance components on the optimal designs. Moreover, it is also interesting to study the IMSPE- and MMSPE-optimal designs for more general random coefficient regression models.

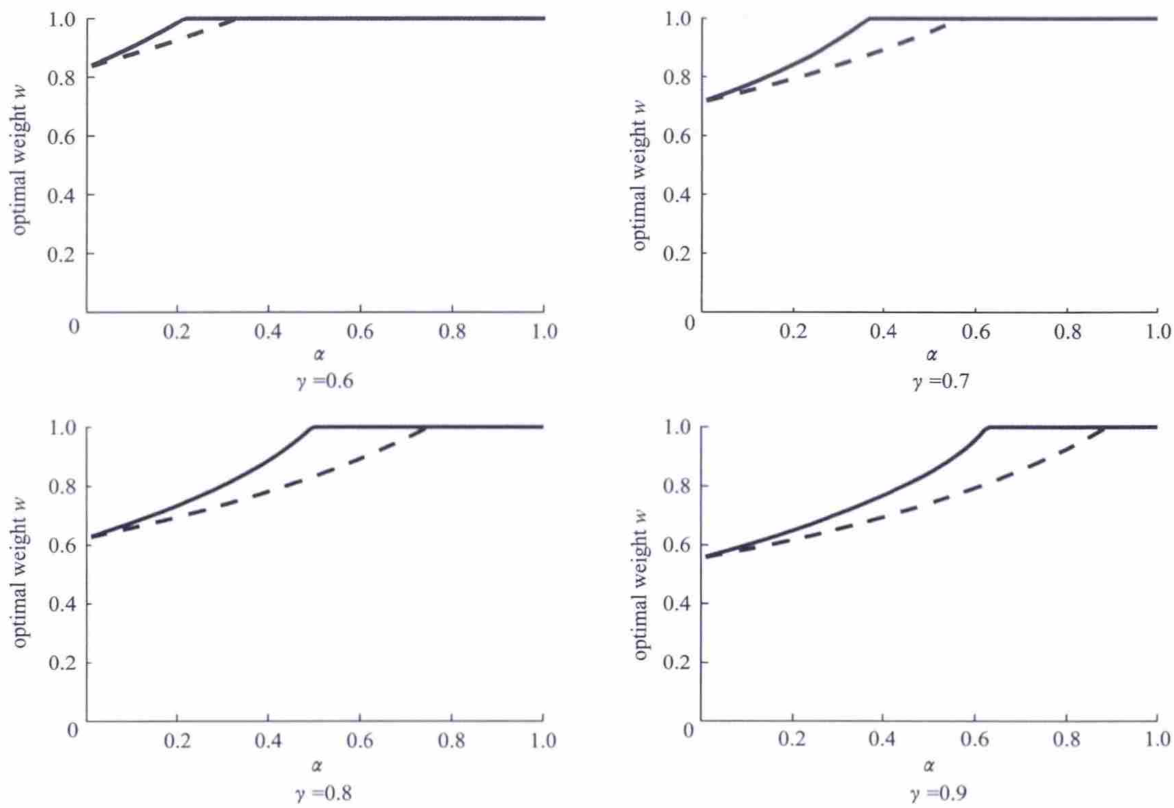

Figure 2　The optimal weights w^*_{IMSPE} (solid) and w^*_{MMSPE} (dashed) vs the ratio $\alpha = h/H$ for model (2) , where $\gamma = n\sigma_b^2/(\sigma_\varepsilon^2 + n\sigma_b^2)$

References:

[1] ATKINS J E, CHENG C S. Optimal regression designs in the presence of random block effects [J] . J Statist Plann Inference, 1999, 77: 321 – 335.

[2] CHENG C S. Optimal regression designs under random block-effects models [J] . Statist Sinica, 1995, 5, 485 – 497.

[3] GOOS P, VANDEBROEK M. D-optimal response surface designs in the presence of random block effects [J] . Comput Statist Data Anal, 2001, 37: 433 – 453.

[4] KHURI A. Response surface models with random block effects [J] . Technometrics, 1992, 34: 26 – 37.

[5] LISKI E P, MANDAL N K, SINHA B K, et al. Optimal designs in growth curve models: Part I, Correlated model for linear growth: Optimal designs for slope parameter estimation and growth prediction [J] . J Statist Plann Inference, 1997, 64: 141 – 150.

[6] ABT M, GAFFKE N, LISKI E P, et al. Optimal designs in growth curve models Ⅱ. correlated model for quadratic growth: optimal designs for parameter estimation and growth prediction [J] . J Statist Plann Inference, 1998, 67: 287 – 296.

[7] BERGER M P F, TAN F E S. Robust designs for linear mixed models [J] . Appl Statist, 2004, 53: 569 – 581.

[8] OUWENS M J N M, TAN F E S, BERGER M P F. Maximin D-optimal designs for longitudinal mixed effects models [J] . Biometrics, 2002, 58: 735 – 741.

[9] TAN F E S, BERGER M P F. Optimal allocation of time points for the random effects model [J] . Comm Statist-Simulations, 1999, 28: 517 – 540.

[10] SCHMELTER T. The optimality of single-group designs for certain mixed models [J] . Metrika, 2007, 65: 183 – 193.

[11] SCHMELTER T. Considerations on group-wise identical designs for linear mixed models [J] . J Statist Plann Inference, 2007, 137: 4003 – 4010.

[12] DEBUSHO L K, HAINES L M. V- and D-optimal population designs for the simple linear regression model with a random

intercept term [J]. J Statist Plann Inference,2008,138:1116 - 1130.

[13] CHENG J,YUE R-X,LIU X. Optimal Designs for Random Coefficient Regression Models with Heteroscedastic Errors [J]. Comm Statist-Theory and Methods,2012,42:2798 - 2809.

[14] ROBINSON G K. That BLUP is a good thing: The estimation of random effects (with discussion) [J]. Statist Sci,1991,6:15 - 51.

[15] HENDERSON C R. Best linear unbiased estimation and prediction under a selection model [J]. Biometrics,1975,31:423 - 447.

[16] HARMAN R,ŠTULAJTER F. Optimal prediction designs in finite discrete spectrum linear regression models [J]. Metrika,2010,72:281 - 294.

简单随机截距模型参数估计与响应预测的最优设计

岳荣先[1],周晓东[2]

(1. 上海师范大学 数理学院,上海 200234; 2. 上海对外经贸大学 商务信息管理学院,上海 201620)

摘 要: 研究简单随机截距模型中固定和随机效应线性组合的估计和个体未来观测值的预测的最优设计问题. 在模型的方差分量假定为已知的情况下,分别利用估计量的均方误差和预测量的预测均方误差建立设计准则,给出若干最优设计测度的解析解或精确设计数值算法.

关键词: 随机截距模型; 最优设计; 均方误差

(责任编辑: 冯珍珍)