D-OPTIMAL POPULATION DESIGNS FOR THE SIMPLE LINEAR RANDOM COEFFICIENTS REGRESSION MODEL

Legesse Kassa Debusho

Department of Statistics,
University of Pretoria, IT Building,
Pretoria 0002, South Africa.

Linda M. Haines

Department of Statistical Sciences,
University of Cape Town,
Rondebosch 7700, South Africa.

Abstract: In this paper D-optimal population designs for the simple linear random coefficients regression model with values of the explanatory variable taken from a set of equally spaced, non-repeated time points are considered. The D-optimal designs depend on the values of the variance components and locally optimal designs are therefore considered. It is shown that, if the time points are linearly transformed, then the D-optimal population designs for both the fixed effects and the variance components do not necessarily map onto one another. This result is illustrated numerically by means of a simple example.

Keywords: Linear regression; Random coefficients; Information matrix; Population designs; D-optimality

1. PROBLEM

The present paper is broadly concerned with optimal experimental design for linear mixed models fitted to longitudinal data when the fixed effects and variance component parameters are of particular interest. The essential problem is that of choosing the numbers of individuals to be allocated to various groups or cohorts and of choosing the times for taking measurements on the individuals within each group.

The construction of optimal population designs for longitudinal models has been extensively studied in the design literature in various contexts. However, there are relatively few results on design spaces comprising a finite number of time points. Thus Abt, Gaffke, Liski and Sinha (1998) investigated optimal designs for the precise estimation of the linear and quadratic regression coefficients and for growth prediction in the quadratic regression model with a random intercept numerically. However, they only considered a limited number of individual designs on which to base the population designs. Further studies, which are also of a highly computational nature, are presented in the papers by Ouwens, Tan and Berger (2002) and Berger and Tan (2004) on maximin and robust designs for linear mixed models. Nie (2007) provided optimal designs for estimation of both fixed effects and variance components for linear mixed models but his results apply only for designs with an even number of time points per individual over the design space [-1, 1]. Most recently Tekle, Tan and Berger (2008) considered constructing D-optimal cohort designs for linear mixed models numerically using a finite number of time points, while Debusho and Haines (2008) constructed optimal designs for the simple linear regression model with a random intercept term and with values of the explanatory variable taken from a set of equally spaced time points.

The aim of this paper is to compute optimal designs for the simple linear random coefficients regression model with values of the explanatory variable taken from a set of equally spaced time points. The model, some basic ideas and notation, an appropriate equivalence theorem and the dependence of the designs on a linear transformation of time points are introduced in Section 2. The nature and the numerical construction of D-optimal population designs for the model of interest, based on individual designs for which the time points are not repeated, is discussed in Section 3 and some concluding remarks are given in Section 4.

Legesse Kassa Debusho is the corresponding author, e-mail: legesse.debusho@up.ac.za.

2. PRELIMINARIES

2.1. Model and information matrices

Consider a longitudinal experiment with K individuals such that each of these individuals provides measurements at d_i time points t_{ij} taken from the set $\{0, \ldots, k\}$ where k is an integer with $k \geq 1$, $j = 1, 2, \ldots, d_i$ and $i = 1, 2, \ldots, K$. Suppose that a simple linear random coefficient regression model provides an appropriate fit to the data. Then the model for the jth observation on the ith individual, y_{ij}, at the time point t_{ij} is given by

$$y_{ij} = \beta_0 + b_{0i} + \beta_1\, t_{ij} + b_{1i}\, t_{ij} + e_{ij}, \quad j = 1, 2, \ldots, d_i \text{ and } i = 1, 2, \ldots, K. \tag{1}$$

where the intercept β_0 and the slope β_1 are fixed effects, b_{0i} and b_{1i} are random effects particular to the ith individual and e_{ij} is an error term associated with the ijth observation. Furthermore it is assumed that $b_{0i} \sim \mathcal{N}(0, \sigma^2_{b_0})$, that $b_{1i} \sim \mathcal{N}(0, \sigma^2_{b_1})$, that $e_{ij} \sim \mathcal{N}(0, \sigma^2_e)$, that $Cov(e_{ij}, b_{0i}) = Cov(e_{ij}, b_{1i}) = Cov(e_{ij}, e_{ij'}) = 0$ and that $Cov(b_{0i}, b_{1i}) = \sigma_{b_0 b_1}$ where $i, j, j' = 1, \ldots, K$ with $j \neq j'$.

The linear mixed model (1) can be expressed succinctly in matrix form as

$$\mathbf{y}_i = \mathbf{X}_i\, \boldsymbol{\beta} + \mathbf{Z}_i\, \mathbf{b}_i + \mathbf{e}_i, \quad i = 1, \ldots, K$$

where $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{id_i})$, $\mathbf{X}_i = (\mathbf{1}_{d_i}\ \mathbf{t}_i)$ is the design matrix with $\mathbf{1}_{d_i}$ the $d_i \times 1$ vector of 1's and $\mathbf{t}_i$ the vector of time points $(t_{i1}, t_{i2}, \ldots, t_{id_i})$, $\boldsymbol{\beta} = (\beta_0, \beta_1)$, $\mathbf{Z}_i = \mathbf{X}_i$, $\mathbf{b}_i = (b_{0i}, b_{1i})$ and $\mathbf{e}_i = (e_{i1}, e_{i2}, \ldots, e_{id_i})$. It now follows immediately from the assumptions introduced earlier that the mean and the variance matrix of the observed vector $\mathbf{y}_i$ are given by

$$E(\mathbf{y}_i) = \mathbf{X}_i\, \boldsymbol{\beta} \text{ and } Var(\mathbf{y}_i) = \mathbf{V}_i = \mathbf{X}_i\, \mathbf{G}\, \mathbf{X}_i^T + \sigma^2_e\, \mathbf{I}_{d_i}$$

respectively, where

$$\mathbf{G} = \begin{pmatrix} \sigma^2_{b_0} & \sigma_{b_0 b_1} \\ \sigma_{b_0 b_1} & \sigma^2_{b_1} \end{pmatrix}$$

is the variance matrix of the random effects b_{0i} and b_{1i} and $i = 1, \ldots, K$. It is in fact convenient to regard σ^2_e as a nuisance parameter and to express $\mathbf{V}_i$ as

$$\mathbf{V}_i = \sigma^2_e \left(\mathbf{X}_i\, \mathbf{G}_\gamma\, \mathbf{X}_i^T + \mathbf{I}_{d_i} \right) \quad \text{with} \quad \mathbf{G}_\gamma = \frac{1}{\sigma^2_e} \mathbf{G} = \begin{pmatrix} \gamma_{b_0} & \gamma_{b_0 b_1} \\ \gamma_{b_0 b_1} & \gamma_{b_1} \end{pmatrix},$$

thus introducing $\gamma_{b_0} = \dfrac{\sigma^2_{b_0}}{\sigma^2_e}$, $\gamma_{b_1} = \dfrac{\sigma^2_{b_1}}{\sigma^2_e}$ and $\gamma_{b_0 b_1} = \dfrac{\sigma_{b_0 b_1}}{\sigma^2_e}$ as ratios of variance components.

Following Debusho (2004), the information matrix for the fixed effects $\boldsymbol{\beta}$ at the vector of time points $\mathbf{t}_i$ can immediately be derived as

$$\mathbf{I}_\beta(\mathbf{t}_i) = \mathbf{X}_i^T \mathbf{V}_i^{-1} \mathbf{X}_i = \begin{pmatrix} \mathbf{1}_{d_i}^T\, \mathbf{V}_i^{-1}\, \mathbf{1}_{d_i} & \mathbf{1}_{d_i}^T\, \mathbf{V}_i^{-1}\, \mathbf{t}_i \\ \mathbf{t}_i^T\, \mathbf{V}_i^{-1}\, \mathbf{1}_{d_i} & \mathbf{t}_i^T\, \mathbf{V}_i^{-1}\, \mathbf{t}_i \end{pmatrix}$$

and furthermore the information matrix for the ratios of variance components, written $\boldsymbol{\theta} = (\gamma_{b_0}, \gamma_{b_1}, \gamma_{b_0 b_1})$, can be expressed as

$$\mathbf{I}_\theta(\mathbf{t}_i) = \begin{pmatrix} h_{(\gamma_{b_0}, \gamma_{b_0})} & h_{(\gamma_{b_0}, \gamma_{b_1})} & h_{(\gamma_{b_0}, \gamma_{b_0 b_1})} \\ h_{(\gamma_{b_0}, \gamma_{b_1})} & h_{(\gamma_{b_1}, \gamma_{b_1})} & h_{(\gamma_{b_1}, \gamma_{b_0 b_1})} \\ h_{(\gamma_{b_0}, \gamma_{b_0 b_1})} & h_{(\gamma_{b_1}, \gamma_{b_0 b_1})} & h_{(\gamma_{b_0 b_1}, \gamma_{b_0 b_1})} \end{pmatrix}$$

where the elements of this matrix are functions of the elements of $\mathbf{I}_\beta(\mathbf{t}_i)$. Specifically
$h_{(\gamma_{b_0}, \gamma_{b_0})} = \frac{1}{2}\, [\mathbf{I}_\beta(\mathbf{t}_i)]^2_{11}$, $h_{(\gamma_{b_0}, \gamma_{b_1})} = \frac{1}{2}\, [\mathbf{I}_\beta(\mathbf{t}_i)]^2_{12}$, $h_{(\gamma_{b_0}, \gamma_{b_0 b_1})} = [\mathbf{I}_\beta(\mathbf{t}_i)]_{11}\, [\mathbf{I}_\beta(\mathbf{t}_i)]_{12}$,
$h_{(\gamma_{b_1}, \gamma_{b_1})} = \frac{1}{2}\, [\mathbf{I}_\beta(\mathbf{t}_i)]^2_{22}$, $h_{(\gamma_{b_1}, \gamma_{b_0 b_1})} = [\mathbf{I}_\beta(\mathbf{t}_i)]_{22}\, [\mathbf{I}_\beta(\mathbf{t}_i)]_{12}$, and
$h_{(\gamma_{b_0 b_1}, \gamma_{b_0 b_1})} = [\mathbf{I}_\beta(\mathbf{t}_i)]^2_{12} + [\mathbf{I}_\beta(\mathbf{t}_i)]_{11}\, [\mathbf{I}_\beta(\mathbf{t}_i)]_{22}$ for $i = 1, \ldots, K$.

The information matrices for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ over all K individuals with associated sets of time points $\mathbf{t}_i, i = 1, \ldots, K$, are thus given by $\sum_{i=1}^{K} \mathbf{I}_\beta(\mathbf{t}_i)$ and $\sum_{i=1}^{K} \mathbf{I}_\theta(\mathbf{t}_i)$ respectively and correspond, at least approximately, to the inverses of the variance matrices of the maximum likelihood (ML) and the restricted maximum likelihood (REML) estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ (Verbeke and Molenberghs, 2000, p. 64).

2.2. Population designs and the D-optimality criterion

Consider an individual design for model (1) which comprises non-repeated time points. Then the d-point design $t = (t_1, \ldots, t_d)$, with $t_j \in \{0, 1, \ldots, k\}$ and $0 \leq t_1 < t_2 < \ldots < t_d \leq k$, which puts equal weight on each point is termed a d-point individual design. The space of all such designs can thus be defined as the set

$$S_{d,k} = \{t : t = (t_1, t_2, \ldots, t_d), \, t_j \in \{0, 1, \ldots, k\}, \; j = 1, \ldots, d, \; 0 \leq t_1 < t_2 < \ldots < t_d \leq k\}$$

and comprises $N_d = \binom{k+1}{d}$ designs.

Consider now a population design comprising r distinct individual designs with n_i individuals allocated to the design with d_i time points $\mathbf{t}_i = (t_{i1}, \ldots, t_{id_i})$ for $i = 1, \ldots, r$. Suppose further that the cost incurred in taking a single observation is constant and that no extra costs are incurred on recruiting the $\sum_{i=1}^{r} n_i$ individuals to the study. Then the information matrix for the parameter α at this population design on a per observation basis, where α denotes either β or θ, is given by

$$\frac{1}{N} \sum_{i=1}^{r} n_i \mathbf{I}_\alpha(\mathbf{t}_i) = \sum_{i=1}^{r} \frac{n_i d_i}{N} \mathbf{M}_\alpha(\mathbf{t}_i)$$

where $N = \sum_{i=1}^{r} n_i d_i$ is the total number of observations taken and $\mathbf{M}_\alpha(\mathbf{t}_i) = \frac{1}{d_i} \mathbf{I}_\alpha(\mathbf{t}_i)$ is the standardized information matrix at the individual design $\mathbf{t}_i, i = 1, \ldots, r$. Now consider relaxing the condition that n_i be an integer and introducing the approximate population design

$$\xi = \left\{ \begin{array}{ccc} t_1, & \ldots, & t_r \\ w_1, & \ldots, & w_r \end{array} \right.$$

with w_i replacing $\frac{n_i d_i}{N}$ and thus with $0 < w_i < 1$ and $\sum_{i=1}^{r} w_i = 1$. Then the weight w_i represents the proportion of the total number of observations taken at the individual design $\mathbf{t}_i$ and the information matrix for the parameter α at the population design ξ is given by $\mathbf{M}_\alpha(\xi) = \sum_{i=1}^{r} w_i \mathbf{M}_\alpha(\mathbf{t}_i)$. Note that if the individual designs within the population design comprise the same number of time points, that is $d_i = d$, then the proportion of individuals allocated to the design $\mathbf{t}_i$ is equal to the weight w_i and that otherwise this proportion can immediately be recovered as $v_i = \frac{w_i / d_i}{\sum_{i=1}^{r} w_i / d_i}$ for $i = 1, \ldots, r$.

Interest in the present paper centres in particular on the construction of D-optimal population designs for the fixed effects $\boldsymbol{\beta}$ and the variance components $\boldsymbol{\theta}$ in model (1). Specifically the D-optimal criterion is defined in the usual way as

$$\Psi_D(\xi) = -\ln |\mathbf{M}_\alpha(\xi)| = -\ln \left| \sum_{i=1}^{r} w_i \mathbf{M}_\alpha(\mathbf{t}_i) \right|$$

and clearly depends on the variance components $\sigma_{b_0}^2$, $\sigma_{b_1}^2$ and σ_e^2 through their ratios $\gamma_{b_0} = \frac{\sigma_{b_0}^2}{\sigma_e^2}$, $\gamma_{b_1} = \frac{\sigma_{b_1}^2}{\sigma_e^2}$ and $\gamma_{b_0 b_1} = \frac{\sigma_{b_0 b_1}}{\sigma_e^2}$. In order to accommodate this dependence, designs which are locally optimal in the sense of Chernoff (1953) are considered, with a best guess for the unknown variance components being adopted. The General Equivalence Theorem relating to approximate D-optimal population designs follows from the results presented in Debusho and Haines (2008) and, more fundamentally, is a special case of the Equivalence Theorem for multivariate design settings given in Fedorov (1972, p.212). The theorem is based on the fact that the directional derivative of $\Psi_D(\xi) = -\ln |\mathbf{M}_\alpha(\xi)|$ at ξ in the direction of an individual design $\mathbf{t}$ is given by

$$\phi(\mathbf{t}, \xi) = p - tr\{\mathbf{M}_\alpha(\xi)^{-1} \, \mathbf{M}_\alpha(\mathbf{t})\}$$

where p is the number of parameters in α and is stated without proof as follows:

Theorem 2.1: *For the random intercept model (1) and individual designs* $\mathbf{t}$ *taken from a space of designs* T*, the following three conditions on the* D*-optimal population design* ξ^* *are equivalent:*

1. *The design* ξ^* *minimizes* $-\ln|\mathbf{M}_\alpha(\xi)|$.
2. *The design* ξ^* *minimizes* $\max_{\mathbf{t}\in T}\ tr\{\mathbf{M}_\alpha^{-1}(\xi)\,\mathbf{M}_\alpha(\mathbf{t})\}$.
3. *The directional derivative* $\phi(\mathbf{t},\xi^*)$ *attains its minimum at the support designs of* ξ^* *and* $max_{\mathbf{t}\in T} tr\{\mathbf{M}_\alpha^{-1}(\xi^*)\mathbf{M}_\beta(\mathbf{t})\} = p$.

The theorem is important in that it can be invoked to confirm the global optimality or otherwise of candidate D-optimal population designs and also in that it forms the basis for algorithmic design construction. Finally note that the D-efficiency of a design ξ_1 relative to a design ξ_2 is given by $\left\{\frac{|\mathbf{M}_\alpha(\xi_1)|}{|\mathbf{M}_\alpha(\xi_2)|}\right\}^{\frac{1}{p}}$ (Atkinson, Donev and Tobias, 2007, p. 151). The ratio is raised to the power $\frac{1}{p}$ so that the D-efficiency can be interpreted in terms of the sample size. For example, if the D-efficiency of a design ξ_1 relative to a design ξ_2 is 3, then 3 times the number of observations in design ξ_2 are needed for ξ_2 to be as efficient as design ξ_1.

2.3. Linear transformation of the time points

Suppose that the time points comprising the generic vector $\mathbf{t} = (t_1, \ldots, t_d)$ are linearly transformed as

$$t_j^* = u + v\,t_j,\ \ j = 1, \ldots, d$$

where u and v are constants. Then the design matrix for the transformed points is given by $\mathbf{X}^* = \mathbf{X}\,\mathbf{A}$ where $\mathbf{X} = (\mathbf{1}_d\ \mathbf{t})$ and $\mathbf{A} = \begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix}$. The transformed model for the ith individual can therefore be written as

$$\mathbf{y}_i = \mathbf{X}_i^*\,\boldsymbol{\beta}^* + \mathbf{Z}_i^*\,\mathbf{b}_i^* + \mathbf{e}_i,$$

where $\boldsymbol{\beta}^* = \mathbf{A}^{-1}\,\boldsymbol{\beta}$, $\mathbf{Z}_i^* = \mathbf{X}_i^*$ and $\mathbf{b}_i^* = \mathbf{A}^{-1}\,\mathbf{b}_i$ with $\mathbf{b}_i^* \sim \mathcal{N}(\mathbf{0}, \mathbf{G}^*)$ and $\mathbf{G}^* = \mathbf{A}^{-1}\,\mathbf{G}\,(\mathbf{A}^{-1})^T$. Thus a linear transformation of the time points induces the transformation $\boldsymbol{\beta}^* = \mathbf{A}^{-1}\,\boldsymbol{\beta}$ in the fixed effects and $\mathbf{b}_i^* = \mathbf{A}^{-1}\,\mathbf{b}_i$ in the random effects. More particularly the variance matrix of the random effects after transformation, namely $\mathbf{G}^*$, depends on the matrix $\mathbf{A}$ and the structure of the original variance matrix $\mathbf{G}$, and hence of $\mathbf{G}_\gamma$, may well not be preserved (Longford 1993, pp. 93-98). Specifically, since $\mathbf{A}^{-1} = \begin{pmatrix} 1 & -\frac{u}{v} \\ 0 & \frac{1}{v} \end{pmatrix}$, it follows that the variance matrix for the random effects in the transformed model corresponds to $\mathbf{G}^* = \sigma_e^2\,\mathbf{G}_\gamma^*$ where

$$\mathbf{G}_\gamma^* = \mathbf{A}^{-1}\,\mathbf{G}_\gamma\,(\mathbf{A}^{-1})^T = \begin{pmatrix} \gamma_{b_0} - 2\,\frac{u}{v}\,\gamma_{b_0b_1} + \frac{u^2}{v}\gamma_{b_1} & \frac{\gamma_{b_0b_1}}{v} - \frac{u}{v^2}\,\gamma_{b_1} \\ \frac{\gamma_{b_0b_1}}{v} - \frac{u}{v^2}\,\gamma_{b_1} & \frac{\gamma_{b_1}}{v^2} \end{pmatrix}. \tag{2}$$

Note however that the variance matrix of $\mathbf{y}_i$ does not change with a linear transformation of the time points, that is $\mathbf{V}_i^* = \mathbf{X}_i^*\,\mathbf{G}^*\,(\mathbf{X}^*)^T + \sigma_e^2\,\mathbf{I} = \mathbf{V}_i$, and thus that the associated information matrices are related in a straightforward manner as $\mathbf{I}_{\beta^*}(\mathbf{t}_i^*) = \mathbf{A}^T\,\mathbf{I}_\beta(\mathbf{t}_i)\,\mathbf{A}$.

3. RESULTS

3.1. Random intercept model

Consider the random intercept model, that is model (1) with the random effect b_{1i} omitted and thus with $\theta = \gamma_{b_0}$, $\mathbf{Z}_i = \mathbf{1}_d$ and $\mathbf{G}_\gamma = \begin{pmatrix} \gamma_{b_0} & 0 \\ 0 & 0 \end{pmatrix}$. Then $V_i = \sigma_e^2(I + \gamma_{b_0} J)$ where J is a $d_i \times d_i$ matrix of 1's and does not depend on $\mathbf{t}_i$, $i = 1, \ldots, K$. It thus follows immediately that the D-optimal population designs for the fixed effects $\boldsymbol{\beta}$ and for the variance component γ_{b_0} in this model setting are invariant to linear transformations of the time points, in accord with the detailed findings of Debusho and Haines (2008).

3.2. Random slope model

Consider now the random slope model, that is model (1) with the random effect b_{0i} omitted and thus with $\theta = \gamma_{b_1}$, $\mathbf{Z}_i = \mathbf{t}_i$ and $V_i = \sigma_e^2(I + \gamma_{b_1}\mathbf{t}_i\mathbf{t}_i^T)$, $i = 1, \ldots, K$. Then, for a linear transformation of the time points of the form $t_j^* = u + vt_j, j = 1, \ldots, d$, it follows that

$$\mathbf{G}_\gamma = \begin{pmatrix} 0 & 0 \\ 0 & \gamma_{b_1} \end{pmatrix} \quad \text{but} \quad \mathbf{G}_\gamma^* = \frac{\gamma_{b_1}}{v^2} \begin{pmatrix} u^2 & -u \\ -u & 1 \end{pmatrix}$$

and thus that the variance matrix for the random effects has a totally different structure unless $u = 0$. This means that if the time points comprising a design which is optimal for a random slope model are linearly transformed, then the resultant design is not necessarily optimal for the associated transformed random slope model.

3.3. Numerical example

Suppose that a longitudinal study is to be planned with model (1) fitted to the data and with the model parameters $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ estimated as precisely as possible. Suppose further that observations for the response variable can only be taken at the time points 0, 1, 2, 3 and 4, i.e. $k = 4$, and that the best guess for the variance matrix of the random effects is given by $\mathbf{G}_\gamma = \begin{pmatrix} 1 & -0.05 \\ -0.05 & 0.25 \end{pmatrix}$. with σ_e^2 known and equal to 1. The D-optimal population designs for this model setting based on d-point individual designs were constructed numerically and are summarized, together with values of the optimality criterion, in Table 1. It is interesting to note that not all of the D-optimal population designs for $\boldsymbol{\beta}$ are optimal for $\boldsymbol{\theta}$. Thus the D-optimal population design for $\boldsymbol{\beta}$ based on single-point individual designs puts equal weight on the designs at the extreme time points 0 and 4, whereas the D-optimal population design for $\boldsymbol{\theta}$ puts weights on the individual designs at 0 and at 4 and also at the internal points 1 and 2.

Suppose now that the time points $0, 1, 2, 3$ and 4 are linearly transformed according to the relation $t_j^* = t_j - 2$ to give new time points $-2, -1, 0, 1$ and 2. The D-optimal population designs for model (1) based on these transformed time points were computed, again numerically, and are presented, together with values of the optimality criterion, in Table 2. Observe that not all of these new designs are linearly transformed versions of the designs in Table 1. Thus the D-optimal population design with $d = 3$ for the precise estimation of $\boldsymbol{\theta}$ is based on one individual design, namely $(0, 1, 4)$, when taken from the time points $0, 1, 2, 3$ and 4 but on two individual designs, namely $(-2, -1, 2)$ and $(-2, 1, 2)$, when taken from the time points $-2, -1, 0, 1$ and 2. In fact, in this particular case,

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{G}_\gamma^* = \mathbf{A}^{-1}\,\mathbf{G}_\gamma\,(\mathbf{A}^{-1})^T = \begin{pmatrix} 1.8 & 0.45 \\ 0.45 & 0.25 \end{pmatrix}$$

and thus the population designs of Table 2 are D-optimal for model (1) with time points $-2, -1, 0, 1$ and 2 and with the variance matrix for the random effects given by $\mathbf{G}_\gamma^*$ and not by $\mathbf{G}_\gamma$.

Finally it should be noted that D-optimal population designs for different best guesses of the variance matrix of the random effects $\mathbf{G}_\gamma$ can clearly differ. For example, the D-optimal population design for the precise estimation of $\boldsymbol{\theta}$ in model (1) with

Table 1. D-optimal population designs for model (1) based on the time points $\{0, 1, 2, 3, 4\}$ with B indicating the best design over the set of all individual designs.

(a) D-optimal population designs for $\boldsymbol{\beta}$

d	Design, ξ_d^*	$\Psi_D(\xi_d^*)$
1	$\begin{matrix}(0) & (4) \\ \frac{1}{2} & \frac{1}{2}\end{matrix}$	1.030
2	$\begin{matrix}(0, 4) \\ 1\end{matrix}$	0.971
3	$\begin{matrix}(0,1,4) \\ 1\end{matrix}$	1.591
4	$\begin{matrix}(0,1,3,4) \\ 1\end{matrix}$	2.109
5	$\begin{matrix}(0,1,2,3,4) \\ 1\end{matrix}$	2.521
B	$\begin{matrix}(0, 4) \\ 1\end{matrix}$	0.971

(b) D-optimal population designs for $\boldsymbol{\theta}$

d	Design, ξ_d^*	$\Psi_D(\xi_d^*)$
1	$\begin{matrix}(0) & (1) & (2) & (4) \\ 0.323 & 0.144 & 0.209 & 0.324\end{matrix}$	5.297
2	$\begin{matrix}(0, 4) \\ 1\end{matrix}$	2.219
3	$\begin{matrix}(0,1,4) \\ 1\end{matrix}$	2.864
4	$\begin{matrix}(0,1,3,4) \\ 1\end{matrix}$	3.554
5	$\begin{matrix}(0,1,2,3,4) \\ 1\end{matrix}$	4.120
B	$\begin{matrix}(0, 4) \\ 1\end{matrix}$	2.219

Table 2. D-optimal population designs for model (1) based on the linearly transformed time points $\{-2, -1, 0, 1, 2\}$ with B indicating the best design over the set of all individual designs and w a weight such that $0 < w < \frac{1}{2}$.

(a) D-optimal population designs for $\boldsymbol{\beta}$

d	Design, ξ_d^*	$\Psi_D(\xi_d^*)$
1	$\begin{matrix} (-2) & (2) \\ \frac{1}{2} & \frac{1}{2} \end{matrix}$	0.806
2	$\begin{matrix} (-2, 2) \\ 1 \end{matrix}$	0.806
3	$\begin{matrix} (-2, 1, 2) \\ 1 \end{matrix}$	1.591
4	$\begin{matrix} (-2, -1, 1, 2) \\ 1 \end{matrix}$	1.940
5	$\begin{matrix} (-2, -1, 0, 1, 2) \\ 1 \end{matrix}$	2.345
B	$\begin{matrix} (-2) & (2) & (-2, 2) \\ w & w & (1-2w) \end{matrix}$	0.806

(b) D-optimal population designs for $\boldsymbol{\theta}$

d	Design, ξ_d^*	$\Psi_D(\xi_d^*)$
1	$\begin{matrix} (-2) & (0) & (2) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{matrix}$	4.216
2	$\begin{matrix} (-2, 2) \\ 1 \end{matrix}$	1.726
3	$\begin{matrix} (-2, -1, 2) & (-2, 1, 2) \\ 0.013 & 0.987 \end{matrix}$	2.505
4	$\begin{matrix} (-2, -1, 1, 2) \\ 1 \end{matrix}$	3.048
5	$\begin{matrix} (-2, -1, 0, 1, 2) \\ 1 \end{matrix}$	3.594
B	$\begin{matrix} (-2, 2) \\ 1 \end{matrix}$	1.726

$\mathbf{G}_\gamma = \begin{pmatrix} 0.25 & 0.05 \\ 0.05 & 2.00 \end{pmatrix}$ based on individual one-point designs taken from the set of time points $\{0, 1, 2, 3, 4\}$ puts equal weight on the designs at 0, at 1 and at 4 and is not the same as the corresponding population design for $\mathbf{G}_\gamma = \begin{pmatrix} 1 & -0.05 \\ -0.05 & 0.25 \end{pmatrix}$ given in Table 1.

4. CONCLUSIONS

In this study locally D-optimal population designs for the simple linear random coefficients regression model with values of the explanatory variable taken from a set of equally spaced, non-repeated time points are considered. A linear transformation of the time points t_i to the time points $t_i^* = u + vt_i^*, i = 1, \ldots, d$, is introduced. It is then shown that the D-optimal population designs for the random coefficients model with random effects variance matrix $\mathbf{G}$ and time points $t_i, i = 1, \ldots, d$, do not necessarily coincide with optimal designs for the same model setting and the same variance matrix $\mathbf{G}$ but with the linearly transformed time points $t_i^*, i = 1, \ldots, d$. Rather, these D-optimal population designs with variance matrix $\mathbf{G}$ and time points t_i coincide with the optimal designs for the random coefficients model with variance matrix $\mathbf{G}^* = \mathbf{AGA}^T$ where $\mathbf{A} = \begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix}$ and time points $t_i^*, i = 1, \ldots, d$. The results are illustrated algebraically by means of the random slope model and numerically by a simple example.

The results of this paper are based on the assumption that the degree to which observations are correlated is the same for every pair of observations within an individual. However it is common for observations measured on an individual at time points close to each other are more highly correlated than observations measured at time points which are well separated. This knowledge could well be used to develop optimal designs within a broader framework of linear mixed models than that presented here. Furthermore the models used in the present study were restricted to one explanatory variable, namely time. However linear mixed effects models can accommodate more variables. There is therefore scope for further research into the construction of optimal designs in such cases.

Acknowledgements
Legesse Debusho would like to thank the University of Pretoria, South Africa, for financial support through a research grant provided by the Research Development Programme. Linda Haines would like to thank the University of KwaZulu-Natal, the University of Cape Town and the National Research Foundation, South Africa, for financial support.

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.

Berger, M.P.F. and Tan, F.E.S. (2004). Robust designs for linear mixed models. *Applied Statistics*, **53**, 569-581.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics*, **24**, 586-602.

Debusho L.K. (2004). Optimal population designs for linear mixed models. Ph.D. Thesis, University of Kwazululu-Natal, South Africa.

Debusho L.K. and Haines, L.M. (2008). V- and D-optimal population designs for the simple linear regression model with a random intercept term. *Journal of Statistical Planning and Inference*, **138**, 1116-1130.

Nie, L. (2007). Optimal designs for linear mixed-effects models. *Journal of Mathematical Science*, **1**, 1-16.

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

Tekle, F.B., Tan, F.E.S. and Berger, M.P.F. (2008). D-optimal cohort designs for linear mixed-effects models. *Statistics in Medicine*, **27**, 2586-2600.