Metrika (2010) 72:281–294
DOI 10.1007/s00184-009-0253-4

Optimal prediction designs in finite discrete spectrum linear regression models

Radoslav Harman · František Štulajter

Received: 26 September 2008 / Published online: 7 April 2009

Abstract In the paper, we solve the n-point optimal prediction design problem for the simplest nontrivial finite discrete spectrum linear regression models with correlated observations. We show that for all the models in consideration, there exists an optimal prediction design supported on at most three distinct points, which can be computed using one-dimensional optimization. In some cases, an optimal prediction design can be found explicitly.

Keywords Optimal design · Random coefficient regression · Best linear unbiased predictor · Mean square error · Correlated observations · c-optimality

1 Introduction

Assume that we intend to perform an experiment consisting of n trials, and the characteristics of the random result of every trial can be influenced by the choice of a design point from the experimental domain $\mathcal{T}$. The vector of design points $\tau = (t_1, t_2, \ldots, t_n)' \in \mathcal{T}^n$ will be called an n-point design on $\mathcal{T}$, and the set $\mathrm{supp}(\tau)$ of all distinct design points will be called the support of τ.

Depending on the objectives of the experiment, the quality of designs on $\mathcal{T}$ is measured by a criterion of optimality, i.e., a function $\Phi : \mathcal{T}^n \to \mathbb{R}$, such that "better" designs yield smaller values of Φ. An n-point design τ_n^* on $\mathcal{T}$ is an optimal n-point design, if it minimizes Φ on $\mathcal{T}^n$. If the assumed model contains unknown parameters, the criterion of optimality usually measures the "size" of the information matrix for the

R. Harman (✉) · F. Štulajter
Department of Applied Mathematics and Statistics, Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovakia
e-mail: harman@fmph.uniba.sk

F. Štulajter
e-mail: stulajter@fmph.uniba.sk

estimation of the parameter. For instance, maximizing determinant of the information matrix corresponds to the best-known criterion of D-optimality.

Optimal design of experiments is studied most often for linear regression with uncorrelated errors, see, e.g., the monographs Pázman (1986) and Pukelsheim (1993). The problem of experimental design for the models with correlated observations is in general much more difficult, requiring complex numeric techniques, cf. Brimkulov et al. (1980), Müller and Pázman (1999) and Müller and Pázman (2003), usually without theoretical results for the exact nonasymptotic optimal design (cf. the recent papers Dette et al. 2008; Kiseľák and Stehlík 2008).

In the paper, we focus on the problem of optimal design for a model in which the observations are correlated. As we show, the optimal designs for the simplest finite discrete spectrum linear regression models (FDSLRMs) are given by explicit formulas or can be computed using basic numeric procedures, such as one-dimensional optimization. The model that we assume is a random coefficient regression model (i.e., a mixed effect model) which we view as a generalization of the FDSLRM for time series introduced in Štulajter (2003). An application of the model for real time series data is given in Štulajter and Witkovský (2004).

The objective of the experiment is assumed to be prediction of the value of a future observation. In contrast to the approach where the criterion is the integrated mean square error, or the minimax criterion (see, e.g., Sacks et al. 1989; Mukherjee 2003; Liski et al. 2002, Section 3.4), we assume that the target time point of prediction is known in advance. Therefore our optimality criterion is directly the mean square error (MSE) of the best linear unbiased predictor (BLUP) of the observation in the target point (cf. Goldberger 1962; Christensen 1991; Štulajter 2002).

The main difficulty in deriving the optimal prediction design is that the MSE of the BLUP is a function of not only the information matrix of regression parameters, but also of the inverse of the covariance matrix of observations, and both these matrices depend on the chosen design. Nevertheless, for some basic FDSLRMs it is possible to derive a suitable expression for the inverse of the covariance matrix and, consequently, find exact optimal prediction designs.

The optimal prediction designs do not depend on regression parameters, but do depend on the covariance parameter, that is, on the vector of variances of random coefficients and the variance of errors. We assume that the covariance parameter is arbitrary, but known, for instance from previous observations of the random process. In the case of population studies we can estimate the covariance parameter from the past data and the estimated value can be subsequently used for constructing an optimal prediction design of future measurements on any single individual.

General results on estimation of the covariance parameter based on time series data can be found in Rao and Kleffe (1988), Harville (2008) and, in the context of FDSLRMs in Štulajter and Witkovský (2004). The BLUPs with estimated covariance parameter are called empirical BLUPs (EBLUPs). It is very difficult to find explicit expressions for the MSEs of the EBLUPs and only some approximations are known, see, e.g., Das et al. (2004), Harville (2008) and Štulajter (2007). Therefore the construction of the optimal prediction design that takes the randomness in estimation of the covariance parameter into account is very complicated, and we leave this problem open for future research.

If all the points of a design $\tau \in \mathcal{T}^n$ are distinct, then we will say that τ is a design without repetitions, otherwise we say that τ is a design with repetitions. For some practical situations, the designs with repetitions are only an auxiliary mathematical tool. For instance, if t represents time, it is often necessary to consider only designs for which $t_{i+1} - t_i \geq \Delta$, for all $i = 0, 1, \ldots, n-1$, where Δ is a minimum delay between consecutive measurements. Nevertheless, if it is possible to use several independent measurement devices with the same operational characteristics, there is no reason to exclude designs with repetitions. Moreover, an optimal design with repetitions can be useful as an approximation of an efficient realizable design without repetitions that can be obtained, e.g., by replacing the repeated measurements by clusters of similar, but not identical points.

2 Prediction designs in finite discrete spectrum linear regression models

Let $f_1, \ldots, f_k, v_1, \ldots, v_l : [0, \infty) \to \mathbb{R}$ be given continuous functions. By the symbol $(f_1, \ldots, f_k | v_1, \ldots, v_l)$-FDSLRM or, briefly, by $(f_1, \ldots, f_k | v_1, \ldots, v_l)$, we will denote the FDSLRM in which, under the experimental design $\tau = (t_1, \ldots, t_n)' \in \mathcal{T}^n$, the observations $X_1, \ldots, X_n$ satisfy

$$X_m = \sum_{i=1}^{k} \beta_i f_i(t_m) + \sum_{j=1}^{l} Y_j v_j(t_m) + w_m, \quad m = 1, \ldots, n. \tag{1}$$

In the model (1) the coefficients $\beta_1, \ldots, \beta_k \in \mathbb{R}$ are unknown regression parameters, and the variables $Y_1, \ldots, Y_l$ represent the random coefficients. We will assume that the random vector $Y = (Y_1, \ldots, Y_l)'$ has the zero mean value and a diagonal covariance matrix $\mathrm{Cov}(Y) = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_l^2)$, $\sigma_j^2 > 0$ for all j, i.e., the random coefficients are uncorrelated. We will also assume that the vector $w = (w_1, \ldots, w_n)'$ of errors satisfies $E(w) = 0$, $\mathrm{Cov}(w) = \sigma^2 \mathbf{I}_n$, $\sigma^2 > 0$, and that the vectors Y and w are mutually uncorrelated. By $\nu = (\sigma^2, \sigma_1^2, \ldots, \sigma_l^2)'$ we denote the covariance parameter of the model (1). For simplicity, we will assume that the experimental domain is the interval $\mathcal{T} = [0, T]$, where T is a given positive constant, but the theory can be extended to cover also more general compact experimental domains.

The model (1) is equivalent to the finite discrete linear regression model for time series given in Štulajter (2003) if the design τ is without repetitions and $w_m = w(t_m)$, and thus $X_m = X(t_m)$.

Remarks on notation Let $\tau = (t_1, \ldots, t_n)'$ be a design on $[0, T]$. For any function $h : [0, T] \to \mathbb{R}$, we will use the notation $h(\tau)$ to denote the vector $(h(t_1), \ldots, h(t_n))'$. Similarly, for functions $g_1, \ldots, g_k : [0, \infty) \to \mathbb{R}$ and for a number $t \in \mathbb{R}$, we will use $g(t)$ to denote the vector $(g_1(t), \ldots, g_k(t))'$.

Let τ be an n-point design on $[0, T]$ and let $\mathbf{F}(\tau)$ be an $n \times k$ matrix defined by

$$\mathbf{F}(\tau) = (f_1(\tau), f_2(\tau), \ldots, f_k(\tau)).$$

Then the vector of observations $X(\tau) = (X_1, X_2, \ldots, X_n)'$ satisfies the linear regression model

$$X(\tau) = \mathbf{F}(\tau)\beta + \epsilon(\tau),$$

where $\beta = (\beta_1, \ldots, \beta_k)'$ and the random vector of errors

$$\epsilon(\tau) = (v_1(\tau), v_2(\tau), \ldots, v_l(\tau))Y + (w_1, \ldots, w_n)'$$

has the covariance matrix

$$\Sigma_\nu(\tau) = \sigma^2 \mathbf{I}_n + \sum_{j=1}^{l} \sigma_j^2 v_j(\tau) v_j'(\tau).$$

Let $d > 0$ and let $T_d = T + d$ be the target point of prediction. Let

$$r_{\nu d}(\tau) = \sum_{j=1}^{l} \sigma_j^2 v_j(T_d) v_j(\tau)$$

be the vector of covariances between the observations $X(\tau)$ and $X(T_d)$. Then the BLUP, $X_\nu^*(\tau, T_d)$, of $X(T_d)$ based on the observations $X(\tau)$ is (see Štulajter 2002)

$$X_\nu^*(\tau, T_d) = f(T_d)'\beta_\nu^*(\tau) + r_{\nu d}'(\tau)\Sigma_\nu^{-1}(\tau)(X(\tau) - \mathbf{F}(\tau)\beta_\nu^*(\tau)), \tag{2}$$

where

$$\beta_\nu^*(\tau) = (\mathbf{F}'(\tau)\Sigma_\nu^{-1}(\tau)\mathbf{F}(\tau))^{-1}\mathbf{F}'(\tau)\Sigma_\nu^{-1}(\tau)X(\tau).$$

The MSE of the BLUP is given by

$$\begin{aligned} \mathrm{MSE}_\nu \left[X_\nu^*(\tau, T_d)\right] = D_\nu\left[X(T_d)\right] &- r_{\nu d}'(\tau)\Sigma_\nu^{-1}(\tau) r_{\nu d}(\tau) \\ &+ (f(T_d) - \mathbf{F}'(\tau)\Sigma_\nu^{-1}(\tau) r_{\nu d}(\tau))'(\mathbf{F}'(\tau)\Sigma_\nu^{-1}(\tau)\mathbf{F}(\tau))^{-1} \\ &\times (f(T_d) - \mathbf{F}'(\tau)\Sigma_\nu^{-1}(\tau) r_{\nu d}(\tau)). \end{aligned} \tag{3}$$

The covariance parameter ν of the model (1) is assumed to be known and fixed. Therefore, we will omit ν from all subsequent notations.

In the paper, we consider the following problem: for given n, T, d, and ν find a design $\tau_n^* \in \mathcal{T}^n$, such that

$$\tau_n^* \in \underset{\tau_n \in \mathcal{T}^n}{\operatorname{argmin}}\, \mathrm{MSE}[X^*(\tau_n, T_d)].$$

Any such design τ_n^* will be called the optimal n-point design for prediction of $X(T_d)$ or, briefly, optimal n-point prediction design. Note that the optimal prediction design does not depend on the regression parameter β but, as we shall see later, it depends on the ratios σ^2/σ_j^2 for $j = 1, \ldots, l$ (unless $l = 0$ as explained in the next paragraph).

By $(f_1, \ldots, f_k|\emptyset)$-FDSLRM we will denote the model (1) with $l = 0$ in which the observations $X_1, \ldots, X_n$ satisfy

$$X_m = \sum_{i=1}^{k} \beta_i f_i(t_m) + w_m; \quad m = 1, \ldots, n. \tag{4}$$

Obviously (4) is the standard linear regression model with uncorrelated errors in which

$$X^*(\tau, T_d) = f(T_d)'(\mathbf{F}'(\tau)\mathbf{F}(\tau))^{-1}\mathbf{F}'(\tau)X(\tau),$$
$$\text{MSE}[X^*(\tau, T_d)] = \sigma^2(1 + f(T_d)'(\mathbf{F}'(\tau)\mathbf{F}(\tau))^{-1} f(T_d)).$$

That is, the problem of optimal prediction design in the model $(f_1, \ldots, f_k|\emptyset)$ is equivalent to the problem of exact c-optimal design in the theory for uncorrelated observations with the vector of coefficients $c = f(T_d) = (f_1(T_d), \ldots, f_k(T_d))'$, see, e.g., Pázman (1986, Chap. III).

By $(\emptyset|v_1, \ldots, v_l)$-FDSLRM we will denote the model (1) with $k = 0$ in which the observations $X_1, \ldots, X_n$ satisfy

$$X_m = \sum_{j=1}^{l} Y_j v_j(t_m) + w_m; \quad m = 1, \ldots, n. \tag{5}$$

In this case

$$X^*(\tau, T_d) = r_d'(\tau)\Sigma^{-1}(\tau)X(\tau),$$
$$\text{MSE}\left[X^*(\tau, T_d)\right] = D\left[X(T_d)\right] - r_d'(\tau)\Sigma^{-1}(\tau)r_d(\tau).$$

The models $(h|\emptyset)$ and $(\emptyset|h)$, where h is a continuous function, are trivial special cases of (4) and (5). It is straightforward to show that in both these models a design τ_n^* is prediction optimal if and only if

$$\text{supp}(\tau_n^*) \subseteq \left\{ t \in [0, T] : |h(t)| = \max_{0 \le s \le T} |h(s)| \right\}.$$

In the following sections we solve the problem of optimal prediction design for the simplest nontrivial FDSLRMs, i.e., $(\emptyset|1, h)$, $(1|h)$, $(h|1)$, and $(1, h|\emptyset)$, where 1 denotes the function identically equal to $1 \in \mathbb{R}$ and h is any function continuous on $[0, \infty)$.

2.1 Optimal prediction designs for $(\emptyset|1, h)$-FDSLRM

Assume that the observations $X_1, \ldots, X_n$ under the design $\tau = (t_1, \ldots, t_n)'$ are modeled by

$$X_m = A + Bh(t_m) + w_m; \quad m = 1, \ldots, n \tag{6}$$

with uncorrelated random coefficients A and B satisfying $E(A) = E(B) = 0, D(A) = \sigma_a^2 \in (0, \infty)$, and $D(B) = \sigma_b^2 \in (0, \infty)$. The random variables A and B are assumed to be uncorrelated with errors $w_1, \ldots, w_m$ of zero mean value and constant variance $\sigma^2 > 0$. Clearly, (6) corresponds to the model $(\emptyset|1, h)$.

Define

$$n_a = n + \sigma^2/\sigma_a^2, \tag{7}$$

$$n_b(\tau) = \sum_{i=1}^{n} h^2(t_i) + \sigma^2/\sigma_b^2. \tag{8}$$

Using the results from the paper Štulajter (2003) and the notations

$$\widetilde{h}(\tau) = \frac{1}{n_a} \sum_{i=1}^{n} h(t_i), \tag{9}$$

$$\widetilde{h^2}(\tau) = \frac{1}{n_a} \sum_{i=1}^{n} h^2(t_i), \tag{10}$$

$$\widetilde{X}(\tau) = \frac{1}{n_a} \sum_{i=1}^{n} X(t_i), \tag{11}$$

$$\widetilde{hX}(\tau) = \frac{1}{n_a} \sum_{i=1}^{n} h(t_i) X(t_i), \tag{12}$$

we obtain the following form of the BLUP and the MSE suitable for solving the problem of optimal prediction design:

$$X^*(\tau, T_d) = \frac{\left[n_b(\tau) n_a^{-1} - h(T_d)\widetilde{h}(\tau)\right] \widetilde{X}(\tau) + \left[h(T_d) - \widetilde{h}(\tau)\right] \widetilde{hX}(\tau)}{n_b(\tau) n_a^{-1} - \left(\widetilde{h}(\tau)\right)^2}, \tag{13}$$

and

$$\mathrm{MSE}\left[X^*(\tau, T_d)\right] = \sigma^2 \left\{1 + \frac{1}{n_a}\left(1 + \frac{\left[h(T_d) - \widetilde{h}(\tau)\right]^2}{n_b(\tau) n_a^{-1} - \left(\widetilde{h}(\tau)\right)^2}\right)\right\}. \tag{14}$$

Since $n_b(\tau) n_a^{-1} - (\widetilde{h}(\tau))^2 > 0$, the $\mathrm{MSE}[X^*(\tau, T_d)]$ attains its smallest possible value $\sigma^2(1 + 1/n_a)$ for any design $\tau_n^* \in \mathcal{T}^n$ satisfying

$$\widetilde{h}(\tau_n^*) = h(T_d). \tag{15}$$

We thus obtain the following theorem.

Theorem 1 *Let $\tau_n^* \in \mathcal{T}^n$ satisfy (15). Then τ_n^* is an optimal n-point design on* $[0, T]$ *for prediction of $X(T_d)$ in the model* $(\emptyset|1, h)$ *and*

$$X^*(\tau_n^*, T_d) = \widetilde{X}(\tau_n^*),$$
$$\text{MSE}[X^*(\tau_n^*, T_d)] = \sigma^2(1 + 1/n_a).$$

An n-point design satisfying condition (15) exists if and only if

$$S_n^* = \left\{t \in [0, T] : h(t) = \frac{n_a}{n} h(T_d)\right\} \neq \emptyset,$$

in which case all designs τ_n^* supported on S_n^* (i.e., $\text{supp}(\tau_n^*) \subseteq S_n^*$) are optimal. Therefore, we can find an n-point optimal prediction design without repetitions for all $n \leq \#S_n^*$. This happens for instance if h is a periodic function such as $h(t) = \cos(\lambda t)$ with a sufficiently large λ or T. The trigonometric periodic functions are often used to model cyclic phenomena in random processes in the case of the FDSLRMs for time series.

To give a simple illustrative example, take $T = 4$, $d = 0.25$, $\sigma^2/\sigma_a^2 = 9$, $\sigma^2/\sigma_b^2 = 1$ and $h(t) = \cos(3.2\pi t)$. Clearly, the function h is periodic with the period $5/8 = 0.625$. Since $h(T_d) = 0.309$, Theorem 1 implies that a design $\tau_n^* = (t_1^*, \ldots, t_n^*)'$ is optimal once

$$\frac{1}{n+9} \sum_{i=1}^{n} h(t_i^*) = 0.309. \tag{16}$$

Let, for example, $n = 6$. Then the optimality condition (16) is satisfied for infinitely many designs, and in particular for all 6-point designs supported on the set

$$\begin{aligned} S_6^* &= \{t \in [0, T] : h(t) = 0.7725\} \\ &= \{0.0684; 0.625i \pm 0.0684; \quad i = 1, 2, \ldots, 6\}. \end{aligned}$$

For $n \to \infty$, the optimal support set S_n^* tends to the set

$$\begin{aligned} S_\infty^* &= \{t \in [0, T] : h(t) = 0.309\} \\ &= \{0.125; 0.625i \pm 0.125; \quad i = 1, 2, \ldots, 6\}. \end{aligned}$$

The situation is illustrated in Fig. 1.

Clearly, in many interesting situations we have $S_n^* = \emptyset$, and Theorem 1 cannot be applied. Nevertheless, using the following lemma, we can easily construct an optimal prediction design for every model $(\emptyset|1, h)$.

Lemma 1 *Let $\Phi : \mathcal{T}^n \to \mathbb{R}$ be a criterion of design optimality and let $h : \mathcal{T} \to \mathbb{R}$ be any continuous function. Assume that for all designs $\tau_1 = (t_1^{(1)}, \ldots, t_n^{(1)})' \in \mathcal{T}^n$ and $\tau_2 = (t_1^{(2)}, \ldots, t_n^{(2)})' \in \mathcal{T}^n$ satisfying $\sum_{i=1}^n h(t_i^{(1)}) = \sum_{i=1}^n h(t_i^{(2)})$ and*

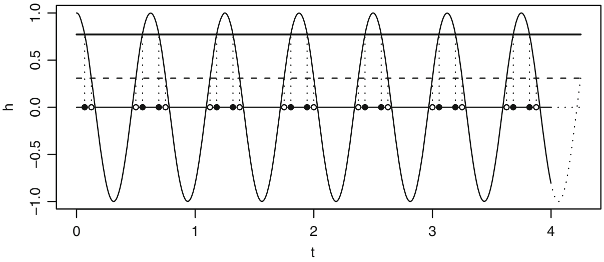

Fig. 1 The set S_6^* that determines a support of a 6-point optimal prediction design (*solid circles*) and the set S_∞^* that determines a support of a prediction optimal design for the size $n \to \infty$ (*empty circles*)

$\sum_{i=1}^n h^2(t_i^{(1)}) > \sum_{i=1}^n h^2(t_i^{(2)})$ *we have* $\Phi(\tau_1) \leq \Phi(\tau_2)$. *Then there exists an* n*-point* Φ*-optimal design* $\tau^* = (t_1^*, \ldots, t_n^*)'$ *satisfying*

$$h(t_1^*) = \cdots = h(t_m^*) = h_* := \min_{t \in \mathcal{T}} h(t),$$
$$h(t_{m+2}^*) = \cdots = h(t_n^*) = h^* := \max_{t \in \mathcal{T}} h(t)$$

for some $m \in \{0, \ldots, n-1\}$.

Proof For any design $\tau = (t_1, \ldots, t_n)'$ on $\mathcal{T}$ let

$$c(\tau) = \#\{i \in \{1, \ldots, n\} : h(t_i) = h_* \quad \text{or } h(t_i) = h^*\}.$$

In words, $c(\tau)$ denotes the number of support points of τ that are mapped on the boundary of the set $h(\mathcal{T})$. To prove the lemma, it is enough to show that, given any n-point design τ for which $c(\tau) < n - 1$, there exists τ^+, such that $c(\tau^+) > c(\tau)$ and $\Phi(\tau^+) \leq \Phi(\tau)$.

Let $\tau = (t_1, \ldots, t_n)'$ be an n-point design such that $c(\tau) < n - 1$. Then there exist two points t_i and t_j such that $h_* < h(t_i) < h(t_j) < h^*$. Since h is a continuous function, there exist $t_i^*, t_j^* \in \mathcal{T}$, such that $h(t_i) + h(t_j) = h(t_i^*) + h(t_j^*)$ and either $h(t_i^*) = h_*$ or $h(t_j^*) = h^*$. Let τ^+ be the design that replaces t_i, t_j by the couple t_i^*, t_j^* in τ, that is $\tau^+ = (t_1^+, \ldots, t_n^+)'$ satisfies $t_i^+ = t_i^*$, $t_j^+ = t_j^*$, and $t_k^+ = t_k$ for all $k \notin \{i, j\}$.

Clearly $\sum_{i=1}^n h(t_i^+) = \sum_{i=1}^n h(t_i)$ and $c(\tau^+) > c(\tau)$. In the same time, we have $h(t_i^*) + \delta = h(t_i) < h(t_j) = h(t_j^*) - \delta$ for some $\delta > 0$, which means

$$h^2(t_i^*) + h^2(t_j^*) = h^2(t_i) + h^2(t_j) + 2\delta(h(t_j) - h(t_i)) + 2\delta^2 > h^2(t_i) + h^2(t_j)$$

and consequently $\sum_{i=1}^n h^2(t_i^+) > \sum_{i=1}^n h^2(t_i)$. Using the assumption of the lemma, we obtain $\Phi(\tau^+) \leq \Phi(\tau)$. □

The statement of Lemma 1 can be reformulated as follows: assume the uniform probability on $\Omega = \{1, \ldots, n\}$. For a design $\tau = (t_1, \ldots, t_n)'$, let $Z_\tau : \Omega \to \mathbb{R}$ be the random variable defined $Z_\tau(i) = h(t_i)$ for all $i = 1, \dots, n$. Assume that for any couple of n-point designs τ_1 and τ_2 satisfying $E(Z_{\tau_1}) = E(Z_{\tau_2})$ and $D(Z_{\tau_1}) > D(Z_{\tau_2})$ we have $\Phi(\tau_1) \le \Phi(\tau_2)$. Then there exists an n-point Φ-optimal design τ^*, such that m support points of τ^* are mapped on h_*, $n - m - 1$ of the support points are mapped on h^*, and one remaining support point is in some unspecified position in $\mathcal{T}$. We will denote the class of designs of this form by $\mathcal{T}^*_{n,h}$.

Since $n_b(\tau)$ increases with $\sum_{i=1}^n h^2(t_i)$, the criterion $\Phi(\tau) = \mathrm{MSE}[X^*(\tau, T_d)]$, where the MSE is given by (14), satisfies assumptions of Lemma 1. Consequently, we obtain:

Theorem 2 *The class $\mathcal{T}^*_{n,h}$ contains an optimal n-point prediction design for the model* $(\emptyset|1, h)$.

Thus, to construct an optimal n-point prediction design on $[0, T]$, we do not need to solve an n-dimensional optimization problem on $[0, T]^n$; it is enough to solve n one-dimensional optimization problems on $[0, T]$. Firstly, we find the point $t_{\min}$ minimizing $h(t)$ on $[0, T]$ and the point $t_{\max}$ maximizing $h(t)$ on $[0, T]$. Second, for all $k \in \{0, \dots, n-1\}$ we set $t_i = t_{\min}$, $i = 1, \dots, k$; $t_i = t_{\max}$, $i = k+1, \dots, n-1$, and use a one-dimensional optimization routine to find the point $t_n \in [0, T]$ such that the design $\tau_n^{(k)} = (t_1, \dots, t_{n-1}, t_n)'$ has the minimal MSE. By Theorem 2 we know that the set $\tau_n^{(0)}, \dots, \tau_n^{(n-1)}$ contains an optimal n-point prediction design which we can select by simply comparing the corresponding MSEs. Using this method, we shall explore optimal designs for the model $(\emptyset|1, t)$ in Sect. 3. Notice also that Theorem 2 implies that for the model $(\emptyset|1, h)$ there always exists an optimal prediction design supported on at most three distinct points.

It turns out that the model $(\emptyset|1, h)$ encompasses the models $(1|h)$, $(h|1)$ and $(1, h|\emptyset)$ as "limit cases". It is an intuitively plausible fact which is also straightforward to verify rigorously from the general Eqs. (2) and (3), that the BLUP and MSE of the model $(\emptyset|1, h)$ turn to the corresponding equations for the model

- $(1|h)$ if $\sigma_a^2 \to \infty$, i.e., $\sigma^2/\sigma_a^2 \to 0$, which amounts to replacing n_a by n;
- $(h|1)$ if $\sigma_b^2 \to \infty$, i.e., $\sigma^2/\sigma_b^2 \to 0$, which amounts to replacing $n_b(\tau)$ by $\sum_{i=1}^n h^2(t_i)$;
- $(1, h|\emptyset)$ if both $\sigma_a^2 \to \infty$ and $\sigma_b^2 \to \infty$, i.e., $\sigma^2/\sigma_a^2 \to 0$ and $\sigma^2/\sigma_b^2 \to 0$, which amounts to replacing n_a by n, and $n_b(\tau)$ by $\sum_{i=1}^n h^2(t_i)$.

In the following subsections we give the resulting formulas and theorems for the models $(1|h)$, $(h|1)$ and $(1, h|\emptyset)$. Notice that replacing n_a by n means that the symbol $\widetilde{(\cdot)}$ defined in Eqs. (9)–(12) represents taking the ordinary average, i.e., it can be written by the usual symbol $\overline{(\cdot)}$.

2.2 Optimal prediction designs for $(1|h)$-FDSLRM

Assume that, under the design $\tau = (t_1, \dots, t_n)'$, the observations $X_1, \dots, X_n$ satisfy

$$X_m = a + Bh(t_m) + w_m; \quad m = 1, \dots, n, \tag{17}$$

with an unknown parameter $a \in \mathbb{R}$, and random coefficient B satisfying $E(B) = 0$, $D(B) = \sigma_b^2 \in (0, \infty)$, which is assumed to be uncorrelated with the vector $w = (w_1, \ldots, w_n)'$, and $\mathrm{Cov}(w) = \sigma^2 \mathbf{I}_n$, $\sigma^2 > 0$. Therefore, (17) corresponds to $(1|h)$-FDSLRM. Using the notation (8) we can write:

$$X^*(\tau, T_d) = \frac{\left[n_b(\tau)n^{-1} - h(T_d)\overline{h}(\tau)\right]\overline{X}(\tau) + \left[h(T_d) - \overline{h}(\tau)\right]\overline{hX}(\tau)}{n_b(\tau)n^{-1} - \left(\overline{h}(\tau)\right)^2}$$

and

$$\mathrm{MSE}\left[X^*(\tau, T_d)\right] = \sigma^2 \left\{1 + \frac{1}{n}\left(1 + \frac{\left[h(T_d) - \overline{h}(\tau)\right]^2}{n_b(\tau)n^{-1} - \left(\overline{h}(\tau)\right)^2}\right)\right\}. \tag{18}$$

The expression (18) and Lemma 1 imply the following theorem:

Theorem 3 *The class $\mathcal{T}^*_{n,h}$ contains an n-point optimal prediction design for $(1|h)$-FDSLRM. Moreover, if some design $\tau_n^* = (t_1^*, \ldots, t_n^*)' \in \mathcal{T}^n$ satisfies $\overline{h}(\tau_n^*) = h(T_d)$, then τ_n^* is an optimal n-point design on $[0, T]$ for prediction of $X(T_d)$ with*

$$X^*(\tau_n^*, T_d) = \overline{X}(\tau_n^*),$$
$$\mathrm{MSE}\left[X^*(\tau_n^*, T_d)\right] = \sigma^2\left(1 + 1/n\right).$$

Note that even in the case of the simplest optimal design $\tau_n^* = (t_1^*, \ldots, t_n^*)'$, i.e., if $h(t_i^*) = h(T_d)$ for all $i = 1, \ldots, n$, we have $\mathrm{Cov}(X_i, X_j) = \sigma_b^2 h^2(T_d)$ for all $i, j = 1, 2, \ldots, n$, $i \neq j$. Thus the components of the random vector $X(\tau_n^*)$ are correlated, unless $h(T_d) = 0$.

2.3 Optimal prediction designs for $(h|1)$-FDSLRM

Assume that the observations $X_1, \ldots, X_n$ under the design $\tau = (t_1, \ldots, t_n)'$ satisfy

$$X_m = bh(t_m) + A + w_m, \quad m = 1, \ldots, n, \tag{19}$$

with an unknown parameter $b \in \mathbb{R}$, and random coefficient A satisfying $E(A) = 0$, $D(A) = \sigma_a^2 \in (0, \infty)$, which is assumed to be uncorrelated with the vector $w = (w_1, \ldots, w_n)'$, and $\mathrm{Cov}(w) = \sigma^2 \mathbf{I}_n$, $\sigma^2 > 0$. The model (19) corresponds to $(h|1)$-FDSLRM. Using the notations (7) and (9)–(12) we can write:

$$X^*(\tau, T_d) = \frac{\left[\widetilde{h^2}(\tau) - h(T_d)\widetilde{h}(\tau)\right]\widetilde{X}(\tau) + \left[h(T_d) - \widetilde{h}(\tau)\right]\widetilde{hX}(\tau)}{\widetilde{h^2}(\tau) - \left(\widetilde{h}(\tau)\right)^2},$$

and

$$\text{MSE}[X^*(\tau, T_d)] = \sigma^2 \left\{ 1 + \frac{1}{n_a} \left(1 + \frac{\left[h(T_d) - \widetilde{h}(\tau)\right]^2}{\widetilde{h^2}(\tau) - \left(\widetilde{h}(\tau)\right)^2} \right) \right\}. \tag{20}$$

The expression (20) and Lemma 1 directly imply the following theorem:

Theorem 4 *The class $\mathcal{T}^*_{n,h}$ contains an n-point optimal prediction design for $(h|1)$-FDSLRM. Moreover, if some design $\tau^*_n = (t^*_1, \ldots, t^*_n)' \in \mathcal{T}^n$ satisfies $\widetilde{h}(\tau^*_n) = h(T_d)$, then τ^*_n is an optimal n-point design on $[0, T]$ for prediction of $X(T_d)$ with*

$$X^*(\tau^*_n, T_d) = \widetilde{X}(\tau^*_n),$$
$$\text{MSE}\left[X^*(\tau^*_n, T_d)\right] = \sigma^2(1 + 1/n_a).$$

Note that in the model $(1|h)$-FDSLRM, under any design $\tau = (t_1, \ldots, t_n)' \in \mathcal{T}^n$, the observations $X_1, \ldots, X_n$ are positively correlated: $\text{Cov}(X_i, X_j) = \sigma_a^2 > 0$.

2.4 Optimal prediction designs for $(1, h|\emptyset)$-FDSLRM

Consider the standard linear regression model with uncorrelated errors in which the observations $X_1, \ldots, X_n$ under the design $\tau = (t_1, \ldots, t_n)'$ satisfy:

$$X_m = a + bh(t_m) + w_m; \quad m = 1, \ldots, n, \tag{21}$$

where $a, b \in \mathbb{R}$ are parameters and $w_1, \ldots, w_m$ are uncorrelated errors with zero mean and variance σ^2. Clearly, (21) corresponds to $(1, h|\emptyset)$-FDSLRM.

If $h(t_i) \neq h(t_j)$ for some $i \neq j$, we have

$$X^*(\tau, T_d) = \frac{\left[\overline{h^2}(\tau) - h(T_d)\overline{h}(\tau)\right]\overline{X}(\tau) + \left[h(T_d) - \overline{h}(\tau)\right]\overline{hX}(\tau)}{\overline{h^2}(\tau) - \left(\overline{h}(\tau)\right)^2},$$

and

$$\text{MSE}[X^*(\tau, T_d)] = \sigma^2 \left\{ 1 + \frac{1}{n} \left(1 + \frac{\left[h(T_d) - \overline{h}(\tau)\right]^2}{\overline{h^2}(\tau) - \left(\overline{h}(\tau)\right)^2} \right) \right\}. \tag{22}$$

Using the expression (22) and Lemma 1 we obtain the following theorem.

Theorem 5 *The class $\mathcal{T}^*_{n,h}$ contains an n-point optimal prediction design for $(1, h|\emptyset)$-FDSLRM. Moreover, if some design $\tau^*_n = (t^*_1, \ldots, t^*_n)' \in \mathcal{T}^n$ satisfies $\overline{h}(\tau^*_n) = h(T_d)$, then τ^*_n is an optimal n-point design on $[0, T]$ for prediction of $X(T_d)$ with*

$$X^*(\tau^*_n, T_d) = \overline{X}(\tau^*_n),$$
$$\text{MSE}\left[X^*(\tau^*_n, T_d)\right] = \sigma^2(1 + 1/n).$$

The corresponding approximate c-optimal design with $c = (1, h(T_d))'$ can always be very efficiently calculated numerically (see, e.g., López-Fidalgo and Rodríguez-Díaz 2004 or Harman and Jurík 2008) or, in special cases, analytically. For instance if $h(t) = h(T_d)$ for some $t \in [0, T]$, then the Elfving theorem (see, e.g., Pázman 1986, Chapter III) directly implies that any approximate design ξ, such that the barycenter of $h(\xi)$ is equal to $h(T_d)$, is c-optimal. This is in accord with Theorem 5. However, using Theorem 5 we can calculate not only an approximate c-optimal design, but also an exact n-point c-optimal design for $c = (1, h(T_d))'$.

Note also that the information matrix for the estimation of the parameter $(a, b)'$ is

$$\mathbf{M}(\tau) = \mathbf{F}'(\tau)\mathbf{F}(\tau) = n \begin{pmatrix} 1 & \overline{h}(\tau) \\ \overline{h}(\tau) & \overline{h^2}(\tau) \end{pmatrix}.$$

Therefore, for any two designs τ_1 and τ_2 satisfying $\overline{h}(\tau_1) = \overline{h}(\tau_2)$ and $\overline{h^2}(\tau_1) \geq \overline{h^2}(\tau_2)$ we have $\mathbf{M}(\tau_1) \geq \mathbf{M}(\tau_2)$ in the Loewner ordering. That is, for the model $(1, h|\emptyset)$ the assumptions of Lemma 1 are satisfied for any Loewner isotonic criterion Φ, which covers all reasonable criteria of optimality (cf. Pukelsheim 1993). That is, for all standard criteria, such as D-, A-, E-optimality, the class $\mathcal{T}^*_{n,h}$ contains an exact n-point optimal design. To our best knowledge, this is a new contribution to the theory of exact optimal designs in the standard uncorrelated case.

3 Example

Consider the model $(\emptyset|1, t)$ on the experimental domain $[0, T]$, i.e., under the design $\tau = (t_1, \ldots, t_n)'$ the observations $X_1, \ldots, X_n$ satisfy:

$$X_m = A + Bt_m + w_m, \quad 0 \leq t_m \leq T, \; m = 1, \ldots, n, \tag{23}$$

where the uncorrelated random coefficients A and B satisfy $E(A) = E(B) = 0$, $D(A) = \sigma_a^2 \in (0, \infty)$, and $D(B) = \sigma_b^2 \in (0, \infty)$. The i.i.d. errors $w_1, \ldots, w_m$ with zero mean and variance σ^2 are uncorrelated with both A and B. In this section we study the optimal prediction designs for the model (23) for various values of T, d, $r_a = \sigma^2/\sigma_a^2$ and $r_b = \sigma^2/\sigma_b^2$. We will analyze the problem numerically using the fact that the class $\mathcal{T}^*_{n,t}$ always contains at least one optimal n-point prediction design; see Theorem 2.

In general it is not possible to exclude that all optimal n-point prediction designs for a model $(\emptyset|1, h)$ are supported on at least three points. Nevertheless, in our numerical experiments with $h(t) = t$ we were always able to find an optimal design supported only on two points, namely 0 and T. Thus, to fully describe the n-point optimal prediction designs resulting from our computations, is it enough to give the optimal number k of measurements to be taken in 0, with the understanding that $n - k$ observations are to be taken in T.

In Fig. 2, Graphs **b** (for $n = 10$) and **d** (for $n = 100$) exhibit the number k for the length of the experimental domain and the prediction delay set to $T = 1$ and $d = 1$. The horizontal and vertical axes correspond to values r_a and r_b varying between 0

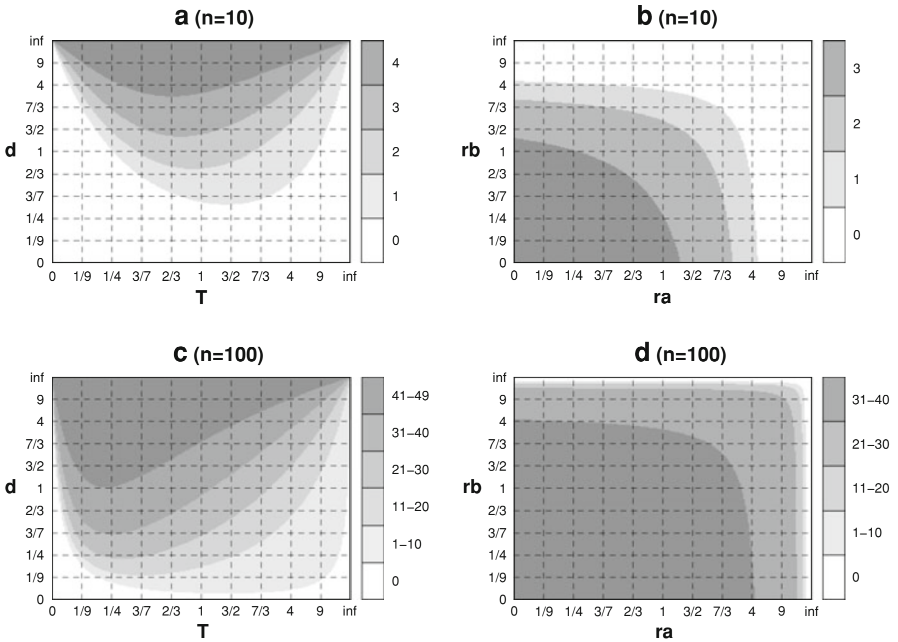

Fig. 2 Number k of measurements to be taken in 0 corresponding to the optimal n-point prediction designs for the model $(\emptyset|1, t)$. Graphs **a** and **c** correspond to fixed $r_a = r_b = 1$ and graphs **b** and **d** correspond to fixed $T = d = 1$. See the main text for details

and ∞. (Note that the axes are nonlinearly transformed such that the finite intervals on the graphs are mapped on $(0, \infty)$ by the functions $r_a/(r_a + 1)$ and $r_b/(r_b + 1)$.) Analogously, Graphs **a** (for $n = 10$) and **c** ($n = 100$) exhibit the optimal number k for the fixed coefficients $r_a = 1$ and $r_b = 1$, and the values T and d varying between 0 and ∞. (The transformations that enable us to display the limiting behaviors in 0 as well as in ∞ are $T/(T + 1)$ and $d/(d + 1)$.)

Notice also, that the optimal prediction designs for the model $(1|t)$, fixed $T = 1, d = 1$ and varying r_b are depicted on Graphs **b** (for $n = 10$) and **d** (for $n = 100$) on the line $r_a = 0$. Similarly, the optimal prediction designs for the model $(t|1)$, fixed $T = 1, d = 1$ and varying r_a are depicted on Graphs **b** (for $n = 10$) and **d** (for $n = 100$) on the line $r_b = 0$. Finally, the point with coordinates $r_a = r_b = 0$ on Graphs **b** and **d** gives the optimal number k of measurements in 0 for the simplest model $(1, t|\emptyset)$ with uncorrelated errors. Specifically, we see that for $n = 10$ the optimum number of measurements in 0 is $k = 3$, and for $n = 100$ the optimum k is some number between 31 and 40.

Remark If we adopt the conjecture that there always exists an optimal design for the model $(\emptyset|1, t)$ with k design points in 0 and $n - k$ design points in T, then we can find an approximate analytic formula for the optimum number k. Given the values T, d, r_a and r_b, the MSE depends only on the number $k \in \{0, 1, \ldots, n\}$. However, as is straightforward to verify, the continuous interpolation of the MSE as a function of $k \in \mathbb{R}$ has a unique minimum given by the formula

$$k_c = \frac{Tdn - T(T+d)r_a - 2r_b}{T(T+2d)}. \tag{24}$$

Hence, for the optimum number k we have $k \approx 0$ if $k_c < 0$, $k \approx k_c$ if $k_c \in [0, T]$ and $k \approx n$ in $k_c > n$. This approximation seems to perform well, at least for large values of n; e.g., it correctly approximates the behavior of the exactly computed optimal values k for the limiting cases of T, d, r_a and r_b. Nevertheless, a more rigorous analysis of this specific model is beyond the scope of the present paper.

Acknowledgments The authors would like to thank anonymous referees for useful comments on the first version of the paper. The research was supported by the Slovak VEGA-Grant No. 1/0077/09.

References

Brimkulov U, Krug G, Savanov V (1980) Numerical construction of exact experimental designs when the measurements are correlated. Zavodskaya Laboratoria 36:435–442 (in Russian)

Christensen R (1991) Linear models for multivariate, time series and spatial Data. Springer, New York

Das K, Jiang J, Rao JNK (2004) Mean squared error of empirical predictor. Ann Stat 32:818–840

Dette H, Kunert J, Pepelyshev A (2008) Exact optimal designs for weighted least squares analysis with correlated errors. Stat Sin 18:135–154

Goldberger AS (1962) Best linear unbiased prediction in the generalized linear regression model. J Am Stat Assoc 57:369–375

Harman R, Jurík T (2008) Computing c-optimal designs of experiments using the methods of linear programming. Comput Stat Data Anal 53(2):247–254

Harville DA (2008) Accounting for the estimation of variances and covariances in prediction under a general linear model: an overview. Tatra Mt Math Publ 39:1–15

Kiseľák J, Stehlík M (2008) Equidistant and D-optimal designs for parameters of Ornstein–Uhlenbeck process. Stat Probab Lett 78:1388–1396

Liski EP, Mandal NK, Shah KR, Sinha BK (2002) Topics in optimal design. Lecture notes in statistics 163. Springer, New York

López-Fidalgo J, Rodríguez-Díaz JM (2004) Elfving's method for m-dimensional models. Metrika 59:235–244

Mukherjee B (2003) Exactly Optimal Sampling Designs for Processes with a Product Covariance Structure. Can J Stat/La Revue Canadienne de Statistique 31:69–87

Müller WG, Pázman A (1999) An algorithm for the computation of optimum designs under a given covariance structure. Comput Stat 14:197–211

Müller WG, Pázman A (2003) Measures for designs in experiments with correlated errors. Biometrika 90:423–434

Pázman A (1986) Foundations of optimum experimental design. Reidel, Dordrecht

Pukelsheim F (1993) Optimal design of experiments. Wiley, New York

Rao CR, Kleffe J (1988) Estimation of variance components and applications. North-Holland, Amsterdam

Sacks J, Welch WJ, Mitchell TJ, Wynn HP (1989) Design and analysis of computer experiments. Stat Sci 4:409–423

Štulajter F (2002) Predictions in time series using regression models. Springer, New York

Štulajter F (2003) The MSE of the BLUP in a finite discrete spectrum LRM. Tatra Mt Math Publ 26:125–131

Štulajter F (2007) Mean squared error of the empirical best linear unbiased predictor in an orthogonal finite discrete spectrum linear regression model. Metrika 65:331–348

Štulajter F, Witkovský V (2004) Estimation of variances in orthogonal FDSLRMs. Metrika 60:105–118