Ann. Inst. Statist. Math.
Vol. 44, No. 2, 229–238 (1992)

ON EXACT D-OPTIMAL DESIGNS FOR REGRESSION MODELS WITH CORRELATED OBSERVATIONS

WOLFGANG BISCHOFF

Institute of Mathematical Stochastics, Department of Mathematics, University of Karlsruhe, D-7500 Karlsruhe 1, Germany

(Received January 8, 1990; revised October 22, 1990)

Abstract. Let τ^* be an exact D-optimal design for a given regression model $Y_\tau = X_\tau\boldsymbol{\beta} + Z_\tau$. In this paper sufficient conditions are given for designing how the covariance matrix of Z_τ may be changed so that not only τ^* remains D-optimal but also that the best linear unbiased estimator (BLUE) of $\boldsymbol{\beta}$ stays fixed for the design τ^*, although the covariance matrix of Z_{τ^*} is changed. Hence under these conditions a best, according to D-optimality, BLUE of $\boldsymbol{\beta}$ is known for the model with the changed covariance matrix. The results may also be considered as determination of exact D-optimal designs for regression models with special correlated observations where the covariance matrices are not fully known. Various examples are given, especially for regression with intercept term, polynomial regression, and straight-line regression. A real example in electrocardiography is treated shortly.

Key words and phrases: D-optimality, exact designs, correlated observations, linear regression, robustness against disturbances.

1. Introduction

Designing problems for regression with correlated observations have, so far as the author knows, only been solved in an asymptotic-optimal way (cf. Sacks and Ylvisaker (1966, 1968, 1970), Bickel and Herzberg (1979), Bickel *et al.* (1981)). Recently Näther (1985) gave conditions under which the best linear unbiased continuous estimator of the regression parameters degenerates to a discrete best linear unbiased estimator. But since designing is mainly important if the costs or time of an observation cannot be neglected, the experimenter is interested in exact optimal designs.

Another hybrid approach is to choose a suitable design only from those which are optimal in the uncorrelated case (cf. Kiefer and Wynn (1981), Budde (1984)). Note, the covariance function must be known for designing in the cases mentioned above.

In this paper exact D-optimal designs for regression models with correlated observations are considered where the covariance matrices are not fully known.

In fact, let τ^* be an exact D-optimal design for a given regression model $Y_\tau = X_\tau\beta + Z_\tau$ and let $\tilde{Y}_\tau = X_\tau\beta + \tilde{Z}_\tau$ be a second regression model. Then sufficient conditions are given for designing so that not only τ^* is D-optimal for the second model, too, but also that the best linear unbiased estimator (BLUE) of β stays fixed for the design τ^*, although the covariance matrix is changed from $\mathrm{Cov}(Z_{\tau^*})$ to $\mathrm{Cov}(\tilde{Z}_{\tau^*})$. Hence under these conditions a best, with respect to D-optimality, BLUE of β is known for the second model. The results may also be considered as determination of exact D-optimal designs for regression models with special correlated observations where the covariance matrices are not fully known. In Section 2 the above problem is explained more exactly. For that purpose D-optimal-invariance will be introduced and discussed in Section 2. Also a real example in electrocardiography is treated.

In Section 3 sufficient conditions for D-optimal-invariance are given for a rather general linear model, see Theorem 3.1. Afterwards in Subsection 3.1 Theorem 3.1 is specialized in covariance matrices not depending on τ and an application to regression with intercept term is given. In Subsection 3.2 Theorem 3.1 is applied to special covariance matrices depending on τ and the result is illustrated by polynomial regression. Afterwards two straight-line regression models are considered. The proof of Theorem 3.1 is given in Section 4.

2. The regression model, the problem, and notations

Consider a linear regression $y(t) = \beta^T \cdot f(t)$, $t \in \mathcal{E}$, where t is the controlled variable from the experimental region $\mathcal{E}$, $f = (f_1, \ldots, f_m)^T$ is a given $\mathbb{R}^m$-valued function on $\mathcal{E}$ with linearly independent components, and $\beta \in \mathbb{R}^m$ is the unknown coefficient vector. Assume that under an exact design $\tau = (t_1, \ldots, t_n)^T$ of size n we have a linear model, where the expectation and covariance matrix of the observations are $X_\tau\beta$ and $\sigma^2 D_\tau$, respectively, and $X_\tau = (f_j(t_i))_{1\le i\le n, 1\le j\le m}$, D_τ is a known positive definite $(n \times n)$-matrix, and $\sigma^2 > 0$ known or unknown, but independent of τ. This linear model will be written $\mathrm{LM}(X_\tau, D_\tau)$. Denote by $\mathcal{E}_n$ the set of all exact designs of size n under which β is estimable, that is

$$\mathcal{E}_n = \{\tau = (t_1, \ldots, t_n)^T : t_i \in \mathcal{E} \text{ for all } i = 1, \ldots, n, \text{ and } \mathrm{rank}(X_\tau) = m\}.$$

So, for τ ranging over $\mathcal{E}_n$, we have a class of linear models which we denote by $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$. Now suppose that the covariance structure is changed from $\sigma^2 D_\tau$ to $\kappa^2 C_\tau$ for all $\tau \in \mathcal{E}_n$, where C_τ are given positive definite matrices and $\kappa^2 > 0$ known or unknown, but independent of τ. This leads to another class $\mathrm{LM}(X_\tau, C_\tau : \tau \in \mathcal{E}_n)$ of linear models. Suppose that $\tau^* \in \mathcal{E}_n$ is D-optimal (for estimating β) in $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$, that is

$$\det(X_{\tau^*}^T D_{\tau^*}^{-1} X_{\tau^*}) = \max_{\tau\in\mathcal{E}_n} \det(X_\tau^T D_\tau^{-1} X_\tau).$$

The question is, when τ^* remains D-optimal and its BLUE for β remains unchanged, when going to $\mathrm{LM}(X_\tau, C_\tau : \tau \in \mathcal{E}_n)$, that is

$$\det(X_{\tau^*}^T C_{\tau^*}^{-1} X_{\tau^*}) = \max_{\tau\in\mathcal{E}_n} \det(X_\tau^T C_\tau^{-1} X_\tau) \tag{2.1}$$

and

$$(X_{\tau^*}^T C_{\tau^*}^{-1} X_{\tau^*})^{-1} X_{\tau^*}^T C_{\tau^*}^{-1} = (X_{\tau^*}^T D_{\tau^*}^{-1} X_{\tau^*})^{-1} X_{\tau^*}^T D_{\tau^*}^{-1}. \tag{2.2}$$

If this is true, then we want to call τ^* *D-optimal-invariant* (for $\mathrm{LM}(X_\tau, C_\tau : \tau \in \mathcal{E}_n)$ and $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$).

Condition (2.2) is a generalization of the question when the ordinary least squares estimator and the BLUE of a given linear model are identical. There is a large literature on this subject, see, for example, Zyskind (1967), Kruskal (1968) and Haberman (1975).

Condition (2.1) means τ^* is D-optimal in $\mathrm{LM}(X_\tau, C_\tau : \tau \in \mathcal{E}_n)$. In a treatment design situation, most papers to date assume that the underlying correlation structure does not depend on the treatment allocation. More, exact D-optimal designs for linear regression models are mostly only known if the condition $D_\tau = I_n$ is fulfilled for all $\tau \in \mathcal{E}_n$ where I_n is the $(n \times n)$-unity matrix. Little is known on D-optimal designs of polynomial regression for uncorrelated observations but with variances depending on τ (see Karlin and Studden (1966), Bischoff (1988) and cf. Federov (1972)). Something is known about exact designs for special factorial linear models with correlated observations not depending on τ (see Budde (1984), Kunert and Martin (1987)). Literature on exact D-optimal designs for correlated observations depending on τ is not known by the author.

Now let τ^* be a known D-optimal design of $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$ then the meaning of D-optimal-invariance is that τ^* is not only a D-optimal design of an in general more complicated $\mathrm{LM}(X_\tau, C_\tau : \tau \in \mathcal{E}_n)$ but also the best, according to D-optimality, BLUE of β can be evaluated. This is not always possible, compare the following example.

Example 2.1. Consider the linear model of straight-line regression on $\mathcal{E} = [-1, 1]$, that is, X_τ is a design matrix of the regression functions $f_1(t) \equiv 1$, $f_2(t) = t$, $t \in \mathcal{E} = [-1, 1]$. For $\tau = (t_1, \ldots, t_n)^T \in \mathcal{E}_n$ let $D_\tau = I_n$ and

$$C_\tau(a, b) = \mathrm{diag}(at_1 + a + b, \ldots, at_n + a + b), \qquad a \in [0, \infty), \quad b \in (0, \infty).$$

Satz 6.5 of Bischoff (1988) implies $\tau_1^* = \begin{pmatrix} -1 & 1 \\ [n/2] & [(n+1)/2] \end{pmatrix}$ and $\tau_2^* = \begin{pmatrix} -1 & 1 \\ [(n+1)/2] & [n/2] \end{pmatrix}$ are the only D-optimal designs for $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$ as well as for $\mathrm{LM}(X_\tau, C_\tau(a, b) : \tau \in \mathcal{E}_n)$ where as usual $\begin{smallmatrix} t \\ k \end{smallmatrix}$ means k experiments are performed at t and $[s] = \max\{k \in \mathbb{N} \cup \{0\} \mid k \leq s\}$. (Note, even if we write designs as above we want to understand them as elements of $\mathcal{E}_n$.) But obviously the BLUEs of β in $\mathrm{LM}(X_{\tau_i^*}, D_{\tau_i^*})$ and $\mathrm{LM}(X_{\tau_i^*}, C_{\tau_i^*}(a, b))$ do not coincide neither for $i = 1$ nor for $i = 2$, if $a \neq 0$. Hence, although all D-optimal designs of $\mathrm{LM}(X_\tau, C_\tau(a, b) : \tau \in \mathcal{E}_n)$ are known no best, with respect to D-optimality, BLUE can be calculated for $\mathrm{LM}(X_\tau, C_\tau(a, b) : \tau \in \mathcal{E}_n)$ if a is unknown. This shows also that (2.2) is not superfluous in the definition of D-optimal-invariance in mathematical sense as well as from a practical point of view.

The results of this paper are applicable to many practical situations. For example, consider body surface potential mapping (BSPM), which is the representation of cardiac electric potentials as they occur on the body surface during the electric activity of the heart. If cylindrical regression (see Bischoff *et al.* (1987)) is used to evaluate the data by the BLUE, then the covariance matrix C of the observations is needed. If $C = \sigma^2 I_n$ then a D-optimal design τ^* is well-known (see Bischoff *et al.* (1987)) but if the data are obtained at the same point of time as usually, $C = \sigma^2 I_n$ is not suitable. Then $C = \kappa_1^2 I_n + \kappa_2^2 \mathbf{1}\mathbf{1}^T$ is more convenient (cf. Bischoff *et al.* (1987)) where $\mathbf{1} = (1, \dots, 1)^T \in \mathbb{R}^n$. The next example may be considered as one of the simplest models of BSPM.

Example 2.2. Consider the linear model of trigonometric regression on $\mathcal{E} = [0, 2\pi)$, that is, X_τ is a design matrix of the regression functions

$$f_1(t) \equiv 1, \quad f_{2k}(t) = \cos kt, \quad f_{2k+1}(t) = \sin kt,$$
$$t \in \mathcal{E} = [0, 2\pi), \quad k = 1, \dots, r \quad (r \in \mathbb{N} \text{ fixed})$$

(that is, $m = 2r + 1$) and for $\tau \in \mathcal{E}_n$ let $D_\tau = I_n$ and

$$C_\tau(\theta) = I_n + \theta \cdot \mathbf{1}\mathbf{1}^T, \qquad \theta \in (-1/n, \infty) \text{ known or unknown.}$$

It is well-known that the design $\tau^* = (0, 2\pi/n, 2 \cdot 2\pi/n, 3 \cdot 2\pi/n, \dots, (n-1) \cdot 2\pi/n)^T \in \mathcal{E}_n$ is D-optimal in the ordinary trigonometric regression model, that is, in $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$. Corollary 3.2 will show that τ^* is D-optimal-invariant for $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$ and $\mathrm{LM}(X_\tau, C_\tau(\theta) : \tau \in \mathcal{E}_n)$ for each θ. So, even if the true θ of $C_\tau(\theta)$ is unknown, the best, according to D-optimality, BLUE of β in $\mathrm{LM}(X_\tau, C_\tau(\theta) : \tau \in \mathcal{E}_n)$ is the ordinary least squares estimator $(X_{\tau^*}^T X_{\tau^*})^{-1} X_{\tau^*}^T$.

3. Statement of the main results and examples

In the following, no distinction is made between a matrix and the linear mapping induced by that matrix in respect of the standard basis of unit vectors. In this section let

$$V_\tau := \mathrm{range}(X_\tau), \quad U_\tau := \mathrm{nullspace}(X_\tau^T), \qquad \tau \in \mathcal{E}_n,$$

and the set of all real symmetric positive definite $(n \times n)$-matrices is denoted by $\mathcal{M}$. In the sequel for $D \in \mathcal{M}$ and a subset W of $\mathbb{R}^n$ the following subset of $\mathcal{M}$ is needed

$$\Gamma(D; W) := \left\{ D + \sum_{i=1}^{r} \alpha_i \cdot \boldsymbol{a}_i \boldsymbol{a}_i^T \in \mathcal{M} \mid r \in \mathbb{N}; \alpha_i \in \mathbb{R}, \boldsymbol{a}_i \in W \text{ for } i = 1, \dots, r \right\}.$$

THEOREM 3.1. *Let* $\tau^* \in \mathcal{E}_n$ *be* D*-optimal in* $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$. *If* $C_\tau \in \Gamma(D_\tau; V_\tau \cup D_\tau(U_\tau))$ *and*

$$\det(X_\tau^T D_\tau^{-1} X_\tau) \cdot \det(D_\tau) \cdot \det(C_{\tau^*}) \le \det(X_{\tau^*}^T D_{\tau^*}^{-1} X_{\tau^*}) \cdot \det(D_{\tau^*}) \cdot \det(C_\tau), \tag{3.1}$$

for all $\tau \in \mathcal{E}_n$, *then* $\det(X_\tau^T C_\tau^{-1} X_\tau) \leq \det(X_{\tau^*}^T C_{\tau^*}^{-1} X_{\tau^*})$ *and* $(X_\tau^T C_\tau^{-1} X_\tau)^{-1} \cdot X_\tau^T C_\tau^{-1} = (X_\tau^T D_\tau^{-1} X_\tau)^{-1} X_\tau^T D_\tau^{-1}$ *for all* $\tau \in \mathcal{E}_n$, *and hence* τ^* *is D-optimal-invariant.*

Remark. (a) Notice, range$(C_\tau - D_\tau) \subseteq V_\tau \cup D_\tau(U_\tau)$ is sufficient for $C_\tau \in \Gamma(D_\tau; V_\tau \cup D_\tau(U_\tau))$. (b) Since τ^* is D-optimal in $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$ the condition

$$\det D_\tau \cdot \det C_{\tau^*} \leq \det D_{\tau^*} \cdot \det C_\tau \tag{3.2}$$

is sufficient for (3.1). (c) Condition (3.1) may not be omitted as the following example shows.

Example 3.1. Consider the linear model $\mathrm{LM}(X_\tau, C_\tau : \tau \in \mathcal{E}_3)$ of straight-line regression on $\mathcal{E} = [-1, 1]$, that is, X_τ is a design matrix of the regression functions $f_1(t) \equiv 1$, $f_2(t) = t$, $t \in \mathcal{E} = [-1, 1]$ and let be

$$\forall \tau = (t_1, t_2, t_3)^T \in \mathcal{E}_3 : C_\tau = I_3 + |t_1 \cdot t_2 \cdot t_3| \cdot \mathbf{1}\mathbf{1}^T.$$

It is well-known that $\tau^* = (1, 1, -1)^T \in \mathcal{E}_3$ is D-optimal but $\tau_0 = (1, 0, -1)^T \in \mathcal{E}_3$ is not D-optimal in $\mathrm{LM}(X_\tau, I_3 : \tau \in \mathcal{E}_3)$. Put $D_\tau = I_3$ then range$(C_\tau - D_\tau) \subseteq$ range(X_τ) for all $\tau \in \mathcal{E}_3$. But condition (3.1) is not fulfilled for τ^* and $\tau = \tau_0$. Further we obtain $\det(X_{\tau^*}^T C_{\tau^*}^{-1} X_{\tau^*}) = 2 < 6 = \det(X_{\tau_0}^T C_{\tau_0}^{-1} X_{\tau_0})$, that is, τ^* is not D-optimal in $\mathrm{LM}(X_\tau, C_\tau : \tau \in \mathcal{E}_3)$.

The next two subsections concern with specializations of Theorem 3.1. In Subsection 3.1 the case is considered that $D_\tau \equiv D$ and $C_\tau \equiv C$ for all $\tau \in \mathcal{E}_n$, afterwards in Subsection 3.2 the matrix C_τ may depend on τ.

3.1 *The covariance matrices do not depend on* τ

Theorem 3.1 implies Corollary 3.1 and Corollary 3.2 is a consequence of Corollary 3.1.

COROLLARY 3.1. *If* $\tau^* \in \mathcal{E}_n$ *is D-optimal in* $\mathrm{LM}(X_\tau, D : \tau \in \mathcal{E}_n)$ *and if* $C \in \Gamma(D; V_\tau \cup D(U_\tau))$ *for all* $\tau \in \mathcal{E}_n$ *then* τ^* *is D-optimal-invariant for* $\mathrm{LM}(X_\tau, D : \tau \in \mathcal{E}_n)$ *and* $\mathrm{LM}(X_\tau, C : \tau \in \mathcal{E}_n)$.

COROLLARY 3.2. *If* $\tau^* \in \mathcal{E}_n$ *is D-optimal in* $\mathrm{LM}(X_\tau, D : \tau \in \mathcal{E}_n)$ *and if* $\mathbf{1} \in V_\tau \cup D(U_\tau)$ *for all* $\tau \in \mathcal{E}_n$ *then* τ^* *is D-optimal-invariant for* $\mathrm{LM}(X_\tau, D : \tau \in \mathcal{E}_n)$ *and* $\mathrm{LM}(X_\tau, D + \theta\mathbf{1}\mathbf{1}^T : \tau \in \mathcal{E}_n)$ *where* $\theta \in (-n^{-2} \cdot \mathbf{1}^T D \mathbf{1}, \infty)$ *is known or unknown.*

Example 3.2. If τ^* is a D-optimal design for an ordinary polynomial or trigonometric regression or in general for every linear model with intercept term then Corollary 3.2 implies that τ^* is D-optimal-invariant for $\mathrm{LM}(X_\tau, I_n : \tau \in \mathcal{E}_n)$ and $\mathrm{LM}(X_\tau, I_n + \theta\mathbf{1}\mathbf{1}^T : \tau \in \mathcal{E}_n)$ where $\theta \in (-1/n, \infty)$ is known or unknown. For ordinary polynomial regression of order k and special numbers n of observations as well as for ordinary trigonometric regression of each order exact D-optimal designs are well-known, cf. Federov (1972); see also Gaffke (1987).

3.2 *The covariance matrices depend on* τ

Consider the linear model $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$. In short, we only investigate the case $C_\tau \in \Gamma(D_\tau; V_\tau)$ in the following, although it would have no difficulties to consider $C_\tau \in \Gamma(D_\tau; V_\tau \cup D_\tau(U_\tau))$. Let $\boldsymbol{a}(\tau) \in V_\tau \backslash \{\mathbf{0}\}$ then for all $\theta \in [0, \infty)$ we define

$$C_\tau(\theta) := D_\tau + \theta \cdot [\boldsymbol{a}(\tau)^T D_\tau^{-1} \boldsymbol{a}(\tau)]^{-1} \cdot \boldsymbol{a}(\tau)\boldsymbol{a}(\tau)^T. \tag{3.3}$$

Obviously $\mathrm{range}(C_\tau(\theta) - D_\tau) \subseteq V_\tau$ for each $\tau \in \mathcal{E}_n$ and $\theta \in [0, \infty)$. Using a standard matrix result we get $\det C_\tau(\theta) = (\det D_\tau) \cdot (1 + \theta)$. So condition (3.2) is fulfilled and Theorem 3.1 can be applied. Hence the following result is proved.

COROLLARY 3.3. *If* $\tau^* \in \mathcal{E}_n$ *is* D*-optimal in* $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$ *then* τ^* *is* D*-optimal-invariant for* $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$ *and* $\mathrm{LM}(X_\tau, C_\tau(\theta) : \tau \in \mathcal{E}_n)$ *where* $C_\tau(\theta)$ *is defined in* (3.3) *and* $\theta \in [0, \infty)$ *is known or unknown.*

Example 3.3. Consider the linear model $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$ of polynomial regression of order k with

$$D_\tau := \mathrm{diag}(v(t_1), \ldots, v(t_n)), \qquad v : \mathcal{E} \to (0, \infty) \text{ arbitrary function}$$

where $\tau = (t_1, \ldots, t_n)^T \in \mathcal{E}_n$. For special variance functions v and for $n = k + 1$ Karlin and Studden (1966) determined exact D-optimal designs. By Corollary 3.3 these designs are also D-optimal-invariant for $\mathrm{LM}(X_\tau, D_\tau + \theta \cdot [\boldsymbol{a}(\tau)^T D_\tau^{-1} \boldsymbol{a}(\tau)]^{-1} \cdot \boldsymbol{a}(\tau)\boldsymbol{a}(\tau)^T : \tau \in \mathcal{E}_n)$ and $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$ where $\boldsymbol{a}(\tau) \in \mathrm{range}(X_\tau) \backslash \{\mathbf{0}\}$, $\theta \in [0, \infty)$ are known or unknown. For example, it may be chosen as $\boldsymbol{a}(\tau) = \mathbf{1}$ for all $\tau \in \mathcal{E}_n$.

Now we give two further examples where Theorem 3.1 may be applied.

Example 3.4. Consider the ordinary straight-line regression $\mathrm{LM}(X_\tau, I_n : \tau \in \mathcal{E}_n)$ on $\mathcal{E} = [-1, 1]$, that is, $X_\tau = (\mathbf{1}\ \tau)$ for $\tau \in \mathcal{E}_n$. It is well-known that τ_i^*, $i = 1, 2$, considered in Example 2.1 is D-optimal in $\mathrm{LM}(X_\tau, I_n : \tau \in \mathcal{E}_n)$. Let $c > 1$, $\theta \in (0, \infty)$ and for all $\tau \in \mathcal{E}_n$

$$C_\tau(\theta, c) := I_n + \theta(\tau^T\tau)^{-c} \cdot \tau\tau^T$$

then τ_i^* is D-optimal-invariant for $\mathrm{LM}(X_\tau, I_n : \tau \in \mathcal{E}_n)$ and $\mathrm{LM}(X_\tau, C_\tau(\theta, c) : \tau \in \mathcal{E}_n)$ where $c > 1$ and $\theta \in (0, \infty)$ are known or unknown; indeed, the assertion follows by Theorem 3.1 since

$$\forall \tau \in \mathcal{E}_n : \det C_\tau(\theta, c) = 1 + \theta \cdot (\tau^T\tau)^{1-c} \geq 1 + \theta \cdot (\tau_i^{*T}\tau_i^*)^{1-c} = \det C_{\tau_i^*}(\theta, c)$$

Example 3.5. Consider the linear model $\mathrm{LM}(X_\tau, D_\tau : \tau \in \mathcal{E}_n)$ of straight-line regression on $\mathcal{E} = [-1, 1]$ with

$$D_\tau(v) = \mathrm{diag}(v(t_1), \ldots, v(t_n)), \qquad v : \mathcal{E} \to (0, \infty) \text{ arbitrary function}$$

where $\tau = (t_1, \dots, t_n)^T \in \mathcal{E}_n$. Bischoff (1988), Satz 6.5, showed that τ_i^*, $i = 1, 2$, considered in Example 2.1 is D-optimal in $\mathrm{LM}(X_\tau, D_\tau(v) : \tau \in \mathcal{E}_n)$ iff $v(t) \geq [(t-1)^2 v(-1) + (t+1)^2 v(1)]/4$ for all $t \in \mathcal{E}$. If we assume $v(-1) = v(1)$, $v(t) \leq v(1)$ for all $t \in \mathcal{E}$ and for the sake of simplicity $v(1) = 1$ the inequality above is equivalent to:

$$\forall t \in \mathcal{E} : \frac{1}{2}(t^2 + 1) \leq v(t) \leq 1. \tag{3.4}$$

Now let $\theta \in (-1/(2n), \infty)$ and

$$C_\tau(v, \theta) := \operatorname{diag}(v(t_1), \dots, v(t_n)) + \theta \cdot \mathbf{1}\mathbf{1}^T$$

where $\tau = (t_1, \dots, t_n)^T \in \mathcal{E}_n$ and v fulfills (3.4). Obviously $\operatorname{range}(C_\tau(v, \theta) - D_\tau(v)) \subseteq \operatorname{range}(X_\tau)$. Since $\det C_\tau(v, \theta) = \prod_{i=1}^n v(t_i) \cdot (1 + \theta \cdot \sum_{i=1}^n v^{-1}(t_i))$ we get for each $\tau \in \mathcal{E}_n$:

$$\begin{aligned}
\det D_{\tau_i^*}(v) \cdot \det C_\tau(v, \theta) &= \det D_{\tau_i^*}(v) \cdot \det D_\tau(v) \cdot \left(1 + \theta \cdot \sum_{i=1}^n v^{-1}(t_i)\right) \\
&\geq \det D_\tau(v) \cdot \det D_{\tau_i^*}(v) \cdot (1 + \theta \cdot n) \\
&= \det D_\tau(v) \cdot \det C_{\tau_i^*}(v, \theta).
\end{aligned}$$

So by Theorem 3.1 follows that τ_1^* and τ_2^* are D-optimal-invariant for $\mathrm{LM}(X_\tau, D_\tau(v) : \tau \in \mathcal{E}_n)$ and $\mathrm{LM}(X_\tau, C_\tau(v, \theta) : \tau \in \mathcal{E}_n)$ where $\theta \in (-1/(2n), \infty)$ and v satisfying (3.4) are known or unknown.

4. Proof of Theorem 3.1

Let X be a real $(n \times m)$-matrix with $\operatorname{rank}(X) = m$, let $\boldsymbol{x}_i \in \mathbb{R}^n$ be the i-th columnvector of X, and let $D \in \mathcal{M}$; we define

$$V := \operatorname{range}(X), \qquad U := \operatorname{nullspace}(X^T), \qquad U_D := \operatorname{nullspace}(X^T D^{-1}),$$
$$\forall \boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^n : \langle \boldsymbol{x}, \boldsymbol{z} \rangle_D := \boldsymbol{x}^T D^{-1} \boldsymbol{z}.$$

Note, $U_D = D(U)$. Before Theorem 3.1 is proved we consider some lemmas.

LEMMA 4.1. *Let $\boldsymbol{a}_{m+1}, \dots, \boldsymbol{a}_n \in \mathbb{R}^n$ be chosen so that $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m, \boldsymbol{a}_{m+1}, \dots, \boldsymbol{a}_n$ is a basis of $\mathbb{R}^n$, define $A := (\boldsymbol{a}_{m+1} \mid \cdots \mid \boldsymbol{a}_n)$, and let $D \in \mathcal{M}$ then*

$$\begin{aligned}
&\det(X^T D^{-1} X) \cdot \det(A^T D^{-1} A - A^T D^{-1} X (X^T D^{-1} X)^{-1} X^T D^{-1} A) \\
&\quad = \det(X^T X) \cdot \det(A^T A - A^T X (X^T X)^{-1} X^T A) \cdot \det D^{-1}.
\end{aligned}$$

PROOF. Let B be an $(n \times n)$-matrix that is partitioned as follows $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$, where B_{ij} is a real $(n_i \times n_j)$-matrix, $i, j = 1, 2$, and where $n_1 + n_2 = n$. If B_{11} is a nonsingular matrix, the determinant of B can be written as $\det B = \det B_{11} \cdot \det(B_{22} - B_{21}B_{11}^{-1}B_{12})$ (cf. Graybill ((1983), Theorem 8.2.1, p. 183)). Hence, we obtain for the patterned matrix $(X \mid A)$:

$$\begin{aligned}
&\det((X \mid A)^T D^{-1}(X \mid A)) \\
&\quad = \det\begin{pmatrix} X^T D^{-1} X & X^T D^{-1} A \\ A^T D^{-1} X & A^T D^{-1} A \end{pmatrix} \\
&\quad = \det(X^T D^{-1} X) \cdot \det(A^T D^{-1} A - A^T D^{-1} X (X^T D^{-1} X)^{-1} X^T D^{-1} A).
\end{aligned}$$

On the other hand, since $(X \mid A)$ and D^{-1} are $(n \times n)$-matrices we obtain:

$$\begin{aligned}
&\det((X \mid A)^T D^{-1}(X \mid A)) \\
&\quad = \det((X \mid A)^T (X \mid A)) \cdot \det D^{-1} \\
&\quad = \det(X^T X) \cdot \det(A^T A - A^T X (X^T X)^{-1} X^T A) \cdot \det D^{-1}. \qquad \square
\end{aligned}$$

Let P denote the unique matrix representation of the projector on U along V in respect of the standard basis of unit vectors: $P = I_n - X(X^T X)^{-1} X^T$.

LEMMA 4.2. *Let $D \in \mathcal{M}$, let $\boldsymbol{a}_{m+1}, \ldots, \boldsymbol{a}_n$ be an orthonormal basis of $D(U)$ with respect to $\langle \cdot, \cdot \rangle_D$ and define $A := (\boldsymbol{a}_{m+1} \mid \cdots \mid \boldsymbol{a}_n)$ then*

$$\det(X^T D^{-1} X) = \det(X^T X) \cdot \det(A^T P A) \cdot \det D^{-1}.$$

PROOF. Since $A^T D^{-1} A = I_{n-m}$ and $X^T D^{-1} A = \mathbf{0}_{m,n-m}$, where $\mathbf{0}_{i,j}$ is the $(i \times j)$-null-matrix, Lemma 4.1 implies:

$$\begin{aligned}
\det(X^T D^{-1} X) &= \det(X^T X) \cdot \det(A^T (I_n - X(X^T X)^{-1} X^T) A) \cdot \det D^{-1} \\
&= \det(X^T X) \cdot \det(A^T P A) \cdot \det D^{-1}. \qquad \square
\end{aligned}$$

LEMMA 4.3. *Let $D, C \in \mathcal{M}$. Let $\boldsymbol{a}_{m+1}, \ldots, \boldsymbol{a}_n$ and $\boldsymbol{b}_{m+1}, \ldots, \boldsymbol{b}_n$ be orthonormal bases of $D(U)$ and $C(U)$ with respect to $\langle \cdot, \cdot \rangle_D$ and $\langle \cdot, \cdot \rangle_C$, respectively. Further define $A := (\boldsymbol{a}_{m+1} \mid \cdots \mid \boldsymbol{a}_n)$ and $B := (\boldsymbol{b}_{m+1} \mid \cdots \mid \boldsymbol{b}_n)$ then*

$$\begin{aligned}
&\det(X^T C^{-1} X) \\
&\quad = \det(X^T D^{-1} X) \cdot \det(B^T P B) \cdot \det(A^T P A)^{-1} \cdot \det D \cdot \det C^{-1}.
\end{aligned}$$

PROOF. Apply Lemma 4.2 to $\det(X^T C^{-1} X)$ and $\det(X^T D^{-1} X)$ then the result of Lemma 4.3 is a direct consequence. $\square$

LEMMA 4.4. *Let D, $C \in \mathcal{M}$ then*
(a) $\operatorname{range}(C - D) \subseteq V \Rightarrow \forall \boldsymbol{x}, \boldsymbol{z} \in D(U) : \langle \boldsymbol{x}, \boldsymbol{z} \rangle_D = \langle \boldsymbol{x}, \boldsymbol{z} \rangle_C$,
(b) $\operatorname{range}(C - D) \subseteq D(U) \Rightarrow \forall \boldsymbol{x}, \boldsymbol{z} \in V : \langle \boldsymbol{x}, \boldsymbol{z} \rangle_D = \langle \boldsymbol{x}, \boldsymbol{z} \rangle_C$.

PROOF. (a): Since $\operatorname{range}(C - D) \subseteq V$ implies $Du = Cu$ for all $u \in U$ the assertion follows. (b): Exchange V for $D(U)$ in (a) then we obtain (b). $\square$

For $C, D \in \mathcal{M}$ we define

$$g(C) = (X^T C^{-1} X)^{-1} X^T C^{-1}, \qquad g(D) = (X^T D^{-1} X)^{-1} X^T D^{-1}.$$

The next lemma gives a necessary and sufficient condition for $g(C) = g(D)$.

LEMMA 4.5. *Let C, $D \in \mathcal{M}$ then $g(C) = g(D) \Leftrightarrow C(U) = D(U)$.*

PROOF. "$\Rightarrow$" follows since nullspace($g(C)$) $= C(U)$ and nullspace($g(D)$) $= D(U)$. "$\Leftarrow$": Remember $V = \text{range}(X)$. Since

$$\begin{aligned}
&\forall \boldsymbol{v} \in V \exists \boldsymbol{y} \in \mathbb{R}^m : X\boldsymbol{y} = \boldsymbol{v} \Rightarrow \forall \boldsymbol{v} \in V : g(C)(\boldsymbol{v}) = \boldsymbol{y} = g(D)(\boldsymbol{v}),\\
&\forall \boldsymbol{w} \in C(U) = D(U) : C^{-1}\boldsymbol{w} \in U, D^{-1}\boldsymbol{w} \in U\\
&\qquad \Rightarrow \forall \boldsymbol{w} \in C(U) = D(U) : g(C)(\boldsymbol{w}) = \boldsymbol{0} = g(D)(\boldsymbol{w}),
\end{aligned}$$

and $V \cap D(U) = \{\boldsymbol{0}\}$ we obtain $g(C) = g(D)$. □

We recall the definitions $V_\tau = \text{range}(X_\tau)$ and $U_\tau = \text{nullspace}(X_\tau^T)$.

PROOF OF THEOREM 3.1. Because $C_\tau \in \Gamma(D_\tau; V_\tau \cup D_\tau(U_\tau))$ implies $C_\tau(U_\tau) = D_\tau(U_\tau)$ we have $(X_\tau^T C_\tau^{-1} X_\tau)^{-1} X_\tau^T C_\tau^{-1} = (X_\tau^T D_\tau^{-1} X_\tau)^{-1} X_\tau^T D_\tau^{-1}$ for all $\tau \in \mathcal{E}_n$ by Lemma 4.5.

Since $C_\tau \in \Gamma(D_\tau; V_\tau \cup D_\tau(U_\tau))$ a matrix $\tilde{C}_\tau \in \mathcal{M}$ exists with $\tilde{C}_\tau \in \Gamma(D_\tau; V_\tau)$ and $\tilde{C}_\tau \in \Gamma(C_\tau; D_\tau(U_\tau))$ for all $\tau \in \mathcal{E}_n$. By Lemma 4.4(b) we obtain $X_\tau^T C_\tau^{-1} X_\tau = X_\tau^T \tilde{C}_\tau^{-1} X_\tau$, hence we may assume range$(C_\tau - D_\tau) \subseteq V_\tau$ for all $\tau \in \mathcal{E}_n$ in the sequel. Taking into account Lemma 4.4(a) there exists an orthonormal basis $\boldsymbol{a}_{m+1}, \ldots, \boldsymbol{a}_n$ of $C_\tau(U_\tau) = D_\tau(U_\tau)$ with respect to $\langle\cdot,\cdot\rangle_{C_\tau}$ and $\langle\cdot,\cdot\rangle_{D_\tau}$. So Lemma 4.3 implies

$$\begin{aligned}
&\det(X_\tau^T C_\tau^{-1} X_\tau)\\
&\quad = \det(X_\tau^T D_\tau^{-1} X_\tau) \cdot \det(A^T P A) \cdot \det(A^T P A)^{-1} \cdot \det D_\tau \cdot \det C_\tau^{-1}\\
&\quad = \det(X_\tau^T D_\tau^{-1} X_\tau) \cdot \det D_\tau \cdot \det C_\tau^{-1}\\
&\quad \le \det(X_{\tau^*}^T D_{\tau^*}^{-1} X_{\tau^*}) \cdot \det D_{\tau^*} \cdot \det C_{\tau^*}^{-1},
\end{aligned}$$

where $A := (\boldsymbol{a}_{m+1} \mid \cdots \mid \boldsymbol{a}_n)$. Considering range$(C_{\tau^*} - D_{\tau^*}) \subseteq V_{\tau^*}$ and Lemma 4.4(a) there exists an orthonormal basis $\boldsymbol{b}_{m+1}, \ldots, \boldsymbol{b}_n$ of $C_{\tau^*}(U_{\tau^*}) = D_{\tau^*}(U_{\tau^*})$ with respect to $\langle\cdot,\cdot\rangle_{C_{\tau^*}}$ and $\langle\cdot,\cdot\rangle_{D_{\tau^*}}$. Again Lemma 4.3 implies:

$$\begin{aligned}
&\det(X_{\tau^*}^T D_{\tau^*}^{-1} X_{\tau^*}) \cdot \det D_{\tau^*} \cdot \det C_{\tau^*}^{-1}\\
&\quad = \det(X_{\tau^*}^T D_{\tau^*}^{-1} X_{\tau^*}) \cdot \det(B^T P B) \cdot \det(B^T P B)^{-1} \cdot \det D_{\tau^*} \cdot \det C_{\tau^*}^{-1}\\
&\quad = \det(X_{\tau^*}^T C_{\tau^*}^{-1} X_{\tau^*}),
\end{aligned}$$

where $B := (\boldsymbol{b}_{m+1} \mid \cdots \mid \boldsymbol{b}_n)$. Hence τ^* is D-optimal in LM$(X_\tau, C_\tau : \tau \in \mathcal{E}_n)$. □

Acknowledgements

I am very grateful to the referees for their helpful remarks in improving the readability considerably, especially to the referee who detected an error in the first draft.

REFERENCES

Bickel, P. J. and Herzberg, A. M. (1979). Robustness of design against autocorrelation in time I: asymptotic theory, optimality for location and linear regression, *Ann. Statist.*, **7**, 77–95.

Bickel, P. J., Herzberg, A. M. and Schilling, M. (1981). Robustness of design against autocorrelation in time II: optimality, theoretical and numerical results for the first-order-autoregressive process, *J. Amer. Statist. Assoc.*, **76**, 870–877.

Bischoff, W. (1988). Über konkrete D-optimale Versuchspläne für die Geradenregression, Docteral Thesis, University of Karlsruhe.

Bischoff, W., Cremers, H. and Fieger, W. (1987). Optimal regression models in electrocardiography, *Innovation et Technologie en Biologie et Medcine*, **8**, 24–34.

Budde, M. (1984). Optimale Zweifachblockpläne bei seriell korrelierten Fehlern, *Metrika*, **31**, 203–213.

Cramér, H. (1963). *Mathematical Methods of Statistics*, Princeton University Press, New Jersey.

Federov, V. V. (1972). *Theory of Optimal Experiments*, Academic Press, New York.

Gaffke, N. (1987). On D-optimality of exact linear regression designs with minimum support, *J. Statist. Plann. Inference*, **15**, 189–204.

Graybill, F. A. (1983). *Matrices with Applications in Statistics*, Wadsworth, Belmont, California.

Haberman, S. J. (1975). How much do Gauss-Markov and least squares estimates differ? A coordinate-free approach, *Ann. Statist.*, **3**, 982–990.

Karlin, S. and Studden, W. J. (1966). Optimal experimental designs, *Ann. Math. Statist.*, **37**, 783–815.

Kiefer, J. and Wynn, H. P. (1981). Optimum balanced block and Latin square designs for correlated observations, *Ann. Statist.*, **9**, 737–757.

Kruskal, W. (1968). When are Gauss-Markov and least squares estimators identical? A coordinate-free approach, *Ann. Math. Statist.*, **39**, 70–75.

Kunert, J. and Martin, R. J. (1987). On the optimality of finite Williams II(a) designs, *Ann. Statist.*, **15**, 1604–1628.

Näther, W. (1985). Exact designs for regression models with correlated errors, *Statistics*, **16**, 479–484.

Sacks, J. and Ylvisaker, D. (1966). Design for regression problems with correlated errors, *Ann. Math. Statist.*, **37**, 66–89.

Sacks, J. and Ylvisaker, D. (1968). Design for regression problems with correlated errors: many parameters, *Ann. Math. Statist.*, **39**, 49–69.

Sacks, J. and Ylvisaker, D. (1970). Designs for regression problems with correlated errors III, *Ann. Math. Statist.*, **41**, 2057–2074.

Zyskind, G. (1967). On canonical forms non-negative covariance matrices and best and simple least squares linear estimators in linear models, *Ann. Math. Statist.*, **38**, 1092–1109.