ResearchGate

See discussions, stats, and author profiles for this publication at:
https://www.researchgate.net/publication/23629723

Computing -optimal experimental designs using the simplex method of linear programming

ARTICLE *in* COMPUTATIONAL STATISTICS & DATA ANALYSIS · DECEMBER 2008

Impact Factor: 1.4 · DOI: 10.1016/j.csda.2008.06.023 · Source: DBLP

CITATIONS
11

READS
50

2 AUTHORS, INCLUDING:

Radoslav Harman
Comenius University in Bratislava
30 PUBLICATIONS **125** CITATIONS

SEE PROFILE

Available from: Radoslav Harman
Retrieved on: 02 February 2016

Computational Statistics and Data Analysis 53 (2008) 247–254

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Computing c-optimal experimental designs using the simplex method of linear programming

Radoslav Harman [a,*], Tomáš Jurík [b]

[a] *Department of Applied Mathematics and Statistics, Faculty of Mathematics, Physics & Informatics, Comenius University, Mlynská Dolina, 84248 Bratislava, Slovakia*

[b] *Department of Mathematical Analysis and Numerical Mathematics, Faculty of Mathematics, Physics & Informatics, Comenius University, Mlynská Dolina, 84248 Bratislava, Slovakia*

ARTICLE INFO

Article history:
Received 2 June 2007
Received in revised form 24 June 2008
Accepted 30 June 2008
Available online 12 July 2008

ABSTRACT

An experimental design is said to be c-optimal if it minimizes the variance of the best linear unbiased estimator of $c^{\mathrm{T}}\beta$, where c is a given vector of coefficients, and β is an unknown vector parameter of the model in consideration. For a linear regression model with uncorrelated observations and a finite experimental domain, the problem of approximate c-optimality is equivalent to a specific linear programming problem. The most important consequence of the linear programming characterization is that it is possible to base the calculation of c-optimal designs on well-understood computational methods. In particular, the simplex algorithm of linear programming applied to the problem of c-optimality reduces to an exchange algorithm with different pivot rules corresponding to specific techniques of selecting design points for exchange. The algorithm can also be applied to "difficult" problems with singular c-optimal designs and relatively high dimension of β. Moreover, the algorithm facilitates identification of the set of all the points that can support some c-optimal design. As an example, optimal designs for estimating the individual parameters of the trigonometric regression on a partial circle are computed.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Consider the linear regression model $(\mathbf{f}, \mathcal{X})$ with uncorrelated homoscedastic observations y satisfying $E(y) = \mathbf{f}^{\mathrm{T}}(x)\beta$, where β is an unknown m-dimensional parameter and $\mathbf{f}$ is a known vector of regression functions linearly independent on the experimental domain $\mathcal{X} = \{x_1, \ldots, x_k\}$. Let $\mathbf{c}$ be a fixed nonzero m-dimensional vector.

The main purpose of this paper is to propose an algorithm based on the simplex method of linear programming for constructing $\mathbf{c}$-optimal approximate designs for the model $(\mathbf{f}, \mathcal{X})$. An experimental design $\xi^*_{\mathbf{c}}$ is said to be $\mathbf{c}$-optimal if, under $\xi^*_{\mathbf{c}}$, the best linear unbiased estimator of $\mathbf{c}^{\mathrm{T}}\beta$ has the minimal possible variance (see (2) for a more precise definition). For instance, if $\mathbf{c}$ is the i-th unit vector $\mathrm{e}_i \in \mathbb{R}^m$, then $\xi^*_{\mathbf{c}}$ is optimal for estimating individual parameters β_i, if $\mathbf{c} = \mathbf{f}(x)$, then $\xi^*_{\mathbf{c}}$ is optimal for estimating the mean value of response in x, and so on. Moreover, $\mathbf{c}$-optimal designs are closely related to several other criteria of design optimality, in particular the so-called standardized criteria (see Dette (1997)). We refer the reader to Pázman (1986) and Pukelsheim (1993) for details on general optimal design of experiments.

As is usual, by an approximate design we understand a probability ξ on $\mathcal{X}$ with the interpretation that $\xi(x)$ represents the proportion of the measurements to be taken in x. The set of all approximate designs on $\mathcal{X}$ will be denoted by Ξ. A design $\xi \in \Xi$ is said to be an exact design of size n, if and only if (iff) it can be realized by n measurements, i.e., iff $\xi(x_j) = n_j/n$ for all $j = 1, \ldots, k$ and some $n_j \in \mathbb{N} \cup \{0\}$.

* Corresponding author.
E-mail address: harman@fmph.uniba.sk (R. Harman).

0167-9473/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2008.06.023

Let $\Xi_{\mathbf{c}}$ be the set of all the designs $\xi \in \Xi$ that guarantee estimability of $\mathbf{c}^T\beta$, i.e., such that $\mathbf{c}$ belongs to the range of the information matrix

$$\mathbf{M}(\xi) = \sum_{x\in\mathcal{X}} \mathbf{f}(x)\mathbf{f}^{\mathrm{T}}(x)\xi(x). \tag{1}$$

If we perform n measurements according to an exact design $\xi \in \Xi_{\mathbf{c}}$ of size n, and if $\hat{\beta}_\xi$ is the least squares estimate of β, then the variance of the best linear unbiased estimator of the linear combination $\mathbf{c}^{\mathrm{T}}\beta$ is

$$\mathrm{Var}(\mathbf{c}^{\mathrm{T}}\hat{\beta}_\xi) = \sigma^2 n^{-1}\mathbf{c}^{\mathrm{T}}\mathbf{M}^{-}(\xi)\mathbf{c}, \tag{2}$$

where σ^2 is the variance of the individual observations y, and $\mathbf{M}^{-}(\xi)$ is a pseudoinverse of $\mathbf{M}(\xi)$. Note that for $\xi \in \Xi_{\mathbf{c}}$ the value $\mathbf{c}^{\mathrm{T}}\mathbf{M}^{-}(\xi)\mathbf{c}$ does not depend on the choice of the pseudoinverse.

Equality (2) motivates the following definition: Any design $\xi_{\mathbf{c}}^* \in \Xi_{\mathbf{c}}$ that minimizes $\mathbf{c}^{\mathrm{T}}\mathbf{M}^{-}(\xi)\mathbf{c}$ among all designs $\xi \in \Xi_{\mathbf{c}}$ is said to be $\mathbf{c}$-optimal for the model $(\mathbf{f}, \mathcal{X})$. We define the $\mathbf{c}$-optimal variance of the model $(\mathbf{f}, \mathcal{X})$ to be $\mathbf{c}^{\mathrm{T}}\mathbf{M}^{-}(\xi_{\mathbf{c}}^*)\mathbf{c}$, and the $\mathbf{c}$-optimal information matrix to be the information matrix of any $\mathbf{c}$-optimal design. It is possible to show that a $\mathbf{c}$-optimal design always exists, but it does not have to be unique. Moreover, for many standard models and choices of the vector $\mathbf{c}$, some or all of the $\mathbf{c}$-optimal information matrices are singular.

The Elfving set of the model $(\mathbf{f}, \mathcal{X})$ is defined by

$$\mathcal{E} = \mathrm{conv}\left(\mathbf{f}(\mathcal{X}) \cup -\mathbf{f}(\mathcal{X})\right),$$

where $\mathbf{f}(\mathcal{X}) = \{\mathbf{f}(x_1), \ldots, \mathbf{f}(x_k)\}$ and conv is the convex hull. By $\partial\mathcal{E}$ we will denote the boundary of $\mathcal{E}$. The Elfving theorem can be formulated as follows (Elfving (1952); see also Pázman (1986) or Pukelsheim (1993)).

Theorem 1. *Let $\xi \in \Xi$. Then the following statements are equivalent:*

(i) *The design ξ is $\mathbf{c}$-optimal for $(\mathbf{f}, \mathcal{X})$;*
(ii) *There exists $h > 0$ and a selection of signs $\epsilon(x) \in \{-1, 1\}$ for all $x \in \mathcal{X}$, such that*

$$h\mathbf{c} = \sum_{x\in\mathcal{X}} \epsilon(x)\mathbf{f}(x)\xi(x) \in \partial\mathcal{E}. \tag{3}$$

In such a case h^{-2} is the $\mathbf{c}$-optimal variance of the model $(\mathbf{f}, \mathcal{X})$.

In López-Fidalgo and Rodríguez-Díaz (2004), the Elfving theorem has been used to construct algorithms for $\mathbf{c}$-optimal designs (with a compact experimental domain, possibly infinite) requiring constrained nonlinear optimization routines and achieving a practically efficient performance for dimensions m up to 4. In the present paper, we use the Elfving theorem to reformulate the problem of constructing a $\mathbf{c}$-optimal design on a finite experimental domain as a specific problem of linear programming (LP). As a consequence, questions regarding $\mathbf{c}$-optimality on a finite experimental domain can be directly answered by the extensive theory of LP. Moreover, LP gives us algorithmic tools for calculating $\mathbf{c}$-optimal designs (such as the simplex method) that are rapid and reliable even in the case of relatively high dimensions of the parameter. With present computer hardware, the algorithm can be applied to problems with a very large experimental domain (e.g. thousands of design points). Thus, even in the case of an infinite experimental domain such as an interval, it is often possible to use a dense discretization with a negligible loss of efficiency.

2. Linear programming characterization of c-optimal designs

By the Elfving theorem, constructing a $\mathbf{c}$-optimal design is equivalent to finding the maximum scalar h, such that the vector $h\mathbf{c}$ belongs to the Elfving set. For the finite experimental domain $\mathcal{X}$, the Elfving set is a *polytope*, hence the $\mathbf{c}$-optimal design problem is equivalent to the specific problem of LP formulated below in Theorem 2. Our reference monographs for LP are Dantzig and Thapa (1997) and Vanderbei (2001). For simplicity, let $\mathbf{F} = (\mathbf{f}_1, \ldots, \mathbf{f}_{2k})$, where

$$\mathbf{f}_j = \begin{cases} \mathbf{f}(x_j), & j \in \{1, \ldots, k\} \\ -\mathbf{f}(x_{j-k}), & j \in \{k+1, \ldots, 2k\}. \end{cases} \tag{4}$$

Theorem 2. *Let $\xi \in \Xi$. Then the following statements are equivalent:*

(i) *The design ξ is $\mathbf{c}$-optimal for $(\mathbf{f}, \mathcal{X})$ and h^{-2} is the $\mathbf{c}$-optimal variance of the model $(\mathbf{f}, \mathcal{X})$;*
(ii) *$\xi(x_j) = \alpha_j + \alpha_{j+k}$ for all $j = 1, \ldots, k$ and for some solution $(\alpha^{\mathrm{T}}, h)^{\mathrm{T}} \in \mathbb{R}^{2k+1}$ of the LP problem:*

$$\max\left\{h \,\middle|\, \begin{pmatrix} \mathbf{F} & -\mathbf{c} \\ \mathbf{1}_{2k}^{\mathrm{T}} & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ h \end{pmatrix} = \begin{pmatrix} \mathbf{0}_m \\ 1 \end{pmatrix}, \alpha \geq \mathbf{0}_{2k}, h \geq 0\right\}. \tag{5}$$

Clearly, the set of all feasible solutions of the problem (5) is bounded and nonempty, which means that it has an optimal solution. Hence, the strong duality theorem of LP (e.g., Vanderbei (2001)) implies that the dual problem

$$\min\left\{g \,\middle|\, \begin{pmatrix} \mathbf{F}^{\mathrm{T}} & \mathbf{1}_{2k} \\ -\mathbf{c}^{\mathrm{T}} & 0 \end{pmatrix} \begin{pmatrix} \gamma \\ g \end{pmatrix} \geq \begin{pmatrix} \mathbf{0}_{2k} \\ 1 \end{pmatrix}, \gamma \in \mathbb{R}^m, g \in \mathbb{R}\right\} \tag{6}$$

also has an optimal solution. The weak duality theorem of LP implies that a feasible solution $(\alpha^{\mathrm{T}}, h)^{\mathrm{T}}$ of (5) and some feasible solution $(\gamma^{\mathrm{T}}, g)^{\mathrm{T}}$ of (6) are optimal for their respective problems if and only if $h = g$. In this case $\alpha_j(\mathbf{f}_j^{\mathrm{T}}\gamma + g) = 0$ for all $j = 1, 2, \ldots, 2k$, which is a consequence of the complementary slackness conditions. This directly gives yet another characterization of **c**-optimality on a finite experimental domain:

Theorem 3. *Let $\xi \in \Xi$. Then the following statements are equivalent:*

(i) *The design ξ is* **c***-optimal for $(\mathbf{f}, \mathcal{X})$ and h^{-2} is the* **c***-optimal variance of the model $(\mathbf{f}, \mathcal{X})$;*
(ii) *There exist $\alpha \geq \mathbf{0}_{2k}$ and $\gamma \in \mathbb{R}^m$ such that $\xi(x_j) = \alpha_j + \alpha_{j+k}$ for all $j = 1, \ldots, k$, $\mathbf{1}_{2k}^{\mathrm{T}}\alpha = 1$, $\mathbf{F}\alpha = h\mathbf{c}$, $\mathbf{F}^{\mathrm{T}}\gamma + h\mathbf{1}_{2k} \geq \mathbf{0}_{2k}$ and $-\mathbf{c}^{\mathrm{T}}\gamma \geq 1$.*

In this case, $\xi(x_j) > 0$ for some $j = 1, \ldots, k$ implies $|\mathbf{f}_j^{\mathrm{T}}\gamma| = h$.

In the usual terms of optimal design, Theorem 3 is an "equivalence theorem" for **c**-optimality on a finite experimental domain.

3. The simplex algorithm for computing c-optimal designs

The most useful consequence of Theorem 2 is that we can employ efficient linear optimization algorithms to numerically construct **c**-optimal designs. A straightforward possibility for solving linear programs is to use a general-purpose LP procedure implemented in an available software package. However, for extensive computations of **c**-optimal designs, it can be more efficient to use the simplex LP method adjusted to the special problems of type (5), as we describe in this section.

Assume the general LP problem in the following form:

$$\max\{\mathbf{z}^{\mathrm{T}}\mathbf{x} | \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}_{2k+1}\}. \tag{7}$$

In the problem (5) we have

$$\mathbf{z} = \begin{pmatrix} \mathbf{0}_{2k} \\ 1 \end{pmatrix}, \qquad \mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{2k+1}) = \begin{pmatrix} \mathbf{F} & -\mathbf{c} \\ \mathbf{1}_{2k}^{\mathrm{T}} & 0 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} \mathbf{0}_m \\ 1 \end{pmatrix}.$$

For a basis $B = (B_1, B_2, \ldots, B_{m+1})$, which is a sequence of $m + 1$ distinct indices from $\{1, 2, \ldots, 2k + 1\}$, we will denote $\mathbf{A}_B = (\mathbf{a}_{B_1}, \ldots, \mathbf{a}_{B_{m+1}})$ and $\mathbf{z}_B = (\mathbf{z}_{B_1}, \ldots, \mathbf{z}_{B_{m+1}})^{\mathrm{T}}$.

To use the simplex algorithm, we need to find an initial basis B such that $\mathbf{A}_B$ is regular and $\mathbf{A}_B^{-1}\mathbf{b}$ is a feasible basic solution, i.e., $\mathbf{A}_B^{-1}\mathbf{b} \geq \mathbf{0}_{m+1}$. For a general LP problem, the initial feasible basis is constructed by solving an auxiliary linear program. Nevertheless, in the specific problem (5), we are able to construct a feasible basis explicitly: Let $L = (l_1, l_2, \ldots, l_m)$ be a sequence of indices from $\{1, \ldots, k\}$, such that the vectors $\mathbf{f}_{l_1}, \mathbf{f}_{l_2}, \ldots, \mathbf{f}_{l_m}$ are linearly independent. (In **c**-optimal design problems, an m-tuple of such indices is known or simple to find.) Let $\mathbf{F}_L = (\mathbf{f}_{l_1}, \ldots, \mathbf{f}_{l_m})$ and define the signs $\epsilon_1, \ldots, \epsilon_m$ by

$$\epsilon_i = \begin{cases} +1 & \text{if } (\mathbf{F}_L^{-1}\mathbf{c})_i \geq 0, \\ -1 & \text{if } (\mathbf{F}_L^{-1}\mathbf{c})_i < 0, \end{cases} \tag{8}$$

for all components $i = 1, \ldots, m$. As is simple to verify, the signs ϵ_i are chosen such that the system of linear equalities

$$\begin{pmatrix} \epsilon_1\mathbf{f}_{l_1} & \cdots & \epsilon_m\mathbf{f}_{l_m} & -\mathbf{c} \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \tilde{\alpha} \\ h \end{pmatrix} = \begin{pmatrix} \mathbf{0}_m \\ 1 \end{pmatrix} \tag{9}$$

is regular. Let $(\tilde{\alpha}^{\mathrm{T}}, h)^{\mathrm{T}} \in \mathbb{R}^{2k+1}$ be the solution of (9). Obviously, equality (9) is equivalent to $\mathbf{F}_L(\epsilon_1\tilde{\alpha}_1, \ldots, \epsilon_m\tilde{\alpha}_m)^{\mathrm{T}} = h\mathbf{c}$ and $\sum_{i=1}^m \tilde{\alpha}_i = 1$. Thus, for any $i \in \{1, \ldots, m\}$ we have

$$\tilde{\alpha}_i = \epsilon_i^2\tilde{\alpha}_i = \epsilon_i(\epsilon_i\tilde{\alpha}_i) = \epsilon_i(h\mathbf{F}_L^{-1}\mathbf{c})_i, \tag{10}$$

$$h\sum_{i=1}^m \epsilon_i(\mathbf{F}_L^{-1}\mathbf{c})_i = 1. \tag{11}$$

Definition (8) guarantees that $\epsilon_i(\mathbf{F}_L^{-1}\mathbf{c})_i \geq 0$ for all indices i which together with (11) gives $h > 0$. From the positivity of h and equality (10) we see that $\tilde{\alpha}_i \geq 0$ for all indices i, which means that $(\tilde{\alpha}^{\mathrm{T}}, h)^{\mathrm{T}} \geq \mathbf{0}_{m+1}$. Let $B = (B_1, \ldots, B_m, 2k + 1)$ be

defined by $B_i = l_i$ if $\epsilon_i = +1$ and $B_i = l_i + k$ if $\epsilon_i = -1$ for all $i = 1, \ldots, m$; then $\mathbf{A}_B$ is the left-hand side matrix in (9), hence $\mathbf{A}_B^{-1}\mathbf{b} = (\tilde{\alpha}^{\mathrm{T}}, h)^{\mathrm{T}} \geq \mathbf{0}_{m+1}$ and therefore, B is a feasible initial basis.

Once we have a feasible basis B, the standard simplex algorithm proceeds as follows: As the first step, we need to calculate values $\bar{\mathbf{z}}_j = \mathbf{z}_j - \mathbf{z}_B^{\mathrm{T}}\mathbf{A}_B^{-1}\mathbf{a}_j$ for all $j \notin B$. Note that $\mathbf{F}_B = (\mathbf{f}_{B_1}, \ldots, \mathbf{f}_{B_m})$ is a regular $m \times m$ matrix, $\mathbf{0}_m \neq \mathbf{F}_B^{-1}\mathbf{c} \geq \mathbf{0}_m$ and denote $q = \mathbf{1}_m^{\mathrm{T}}\mathbf{F}_B^{-1}\mathbf{c} > 0$, $\mathbf{s}_j = \mathbf{F}_B^{-1}\mathbf{f}_j$ and $r_j = \mathbf{1}_m^{\mathrm{T}}\mathbf{s}_j - 1$ for all $j \notin B$. We have

$$\mathbf{A}_B^{-1} = q^{-1}\begin{pmatrix} q\mathbf{F}_B^{-1} - \mathbf{F}_B^{-1}\mathbf{c}\mathbf{1}_m^{\mathrm{T}}\mathbf{F}_B^{-1} & \mathbf{F}_B^{-1}\mathbf{c} \\ -\mathbf{1}_m^{\mathrm{T}}\mathbf{F}_B^{-1} & 1 \end{pmatrix}, \tag{12}$$

which implies for all $j \notin B$:

$$\bar{\mathbf{z}}_j = \mathbf{z}_j - \mathbf{z}_B^{\mathrm{T}}\mathbf{A}_B^{-1}\mathbf{a}_j = q^{-1}(\mathbf{1}_m^{\mathrm{T}}\mathbf{F}_B^{-1}\mathbf{f}_j - 1) = q^{-1}r_j. \tag{13}$$

The actual feasible solution has basis components

$$\begin{pmatrix} \tilde{\alpha} \\ h \end{pmatrix} = \begin{pmatrix} q^{-1}\mathbf{F}_B^{-1}\mathbf{c} \\ q^{-1} \end{pmatrix}, \tag{14}$$

and is optimal if and only if all $\bar{\mathbf{z}}_j \leq 0$, i.e., iff $\mathbf{1}_m^{\mathrm{T}}\mathbf{s}_j \leq 1$ for all $j \notin B$. Note also that the simplex algorithm keeps the matrix $\mathbf{A}_B$ regular, hence the indices i and $i + k$ cannot be simultaneously in B.

If the optimality condition is not satisfied, an index $j^* \notin B$ maximizing $\bar{\mathbf{z}}_j$ (or equivalently maximizing $\mathbf{1}_m^{\mathrm{T}}\mathbf{s}_j$) is chosen to be the new index entering the basis. (We use the so-called largest coefficient pivot rule.) The index B_{i^*} that leaves the basis is such that i^* minimizes

$$\left\{(\mathbf{A}_B^{-1}\mathbf{a}_{j^*})_i\right\}^{-1}\begin{pmatrix} \tilde{\alpha} \\ h \end{pmatrix}_i = \left\{\begin{pmatrix} \mathbf{s}_{j^*} - \tilde{\alpha}r_{j^*} \\ -q^{-1}r_{j^*} \end{pmatrix}_i\right\}^{-1}\begin{pmatrix} \tilde{\alpha} \\ h \end{pmatrix}_i \tag{15}$$

among all $i \in \{1, \ldots, m+1\}$ for which $(\mathbf{A}_B^{-1}\mathbf{a}_{j^*})_i > 0$. Clearly $(\mathbf{A}_B^{-1}\mathbf{a}_{j^*})_{m+1} = -q^{-1}r_{j^*} = -\bar{\mathbf{z}}_{j^*} < 0$, therefore $i^* \neq m+1$, which means that the index $B_{m+1} = 2k+1$ never leaves the basis. Note also that we have $\sum_{i=1}^m (\mathbf{A}_B^{-1}\mathbf{a}_{j^*})_i = \mathbf{1}_m^{\mathrm{T}}(\mathbf{s}_{j^*} - \tilde{\alpha}r_{j^*}) = 1$, that is, there does exist an index $i \in \{1, \ldots, m\}$ such that $(\mathbf{A}_B^{-1}\mathbf{a}_{j^*})_i > 0$. Finally, we update the basis B by replacing the index B_{i^*} by the index j^* and proceed with the next iteration.

Summarizing the ideas above we obtain the following algorithm for computing **c**-optimal designs on a finite experimental domain.

Simplex algorithm for c-optimality (SAC)

Input: The design points $x_1, \ldots, x_k$, the vector of regression functions $\mathbf{f}$ and a sequence $L = (l_1, \ldots, l_m)$ of indices from $\{1, \ldots, k\}$, such that $\mathbf{f}(x_{l_1}), \ldots, \mathbf{f}(x_{l_m})$ are independent vectors.

1. Set $\mathbf{F} \leftarrow (\mathbf{f}(x_1), \ldots, \mathbf{f}(x_k), -\mathbf{f}(x_1), \ldots, -\mathbf{f}(x_k))$, $B_{m+1} \leftarrow 2k+1$.
2. For all $i = 1, \ldots, m$ do: If $(\mathbf{F}_L^{-1}\mathbf{c})_i \geq 0$ then $B_i \leftarrow l_i$ else $B_i \leftarrow l_i + k$.
3. Set $h \leftarrow (1_m^{\mathrm{T}}\mathbf{F}_B^{-1}\mathbf{c})^{-1}$, $\tilde{\alpha} \leftarrow h\mathbf{F}_B^{-1}\mathbf{c}$, and $\mathbf{s}_j \leftarrow \mathbf{F}_B^{-1}\mathbf{F}_{(j)}$ for all $j \notin B$.
4. Set $j^* \leftarrow \min\left\{\mathrm{argmax}_{j \notin B}\mathbf{1}_m^{\mathrm{T}}\mathbf{s}_j\right\}$. If $\mathbf{1}_m^{\mathrm{T}}\mathbf{s}_{j^*} \leq 1$ go to Step 7.
5. Set $i^* \leftarrow \min\left\{\mathrm{argmin}_{i \in \{1,\ldots,m\}, \mathrm{d}_i > 0}\left(\frac{\tilde{\alpha}_i}{\mathrm{d}_i}\right)\right\}$, where $\mathrm{d} = \mathbf{s}_{j^*} - \tilde{\alpha}r_{j^*}$.
6. Update $B_{i^*} \leftarrow j^*$ and return to Step 3.
7. For all $i = 1, \ldots, m$ do: If $B_i \leq k$ then $x_i^* \leftarrow x_{B_i}$ else $x_i^* \leftarrow x_{B_i - k}$.

Output: The **c**-optimal design ξ assigning weights $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_m$ to design points $x_1^*, \ldots, x_m^*$, and **c**-optimal variance h^{-2}.

Thus, when applied to the problem of **c**-optimality, the simplex method reduces to an exchange algorithm in which the pivot rules (steps 4. and 5.) correspond to specific techniques of selecting design points for exchange. Moreover, at each step the algorithm emulates the known procedure of calculating optimal weights on independent support points. (That is, calculating **c**-optimal weights for the design supported on a given set $x_1^*, \ldots, x_m^*$ of points, where $\mathbf{f}(x_1^*), \ldots, \mathbf{f}(x_m^*)$ are independent; see Pukelsheim and Torsney (1991) for details and generalizations to other criteria.)

Notes on convergence of SAC. While it is possible to construct artificial examples of general LP problems that lead to a cycle of the simplex algorithm with the largest coefficient rule, cycling is very rare in practical problems. Moreover, we can avoid the theoretical possibility of cycling by replacing the step 4 of SAC by an "anticycling" pivot rule, such as the Bland's rule: $j^* \leftarrow \min\left\{j \notin B : \mathbf{1}_m^{\mathrm{T}}\mathbf{s}_j > 1\right\}$, which always terminates the simplex algorithm in a finite number of steps (see, e.g. Dantzig and Thapa (1997)). As for the rate of convergence of SAC, we observed very rapid convergence on numerous

concrete examples, but we cannot guarantee that the algorithm is polynomial. Nevertheless, there are polynomial methods for solving LP programs, such as the interior point method and the ellipsoid method (for details see, e.g. Roos et al. (2005) and Korte and Vygen (2002)). Hence **c**-optimality on a finite experimental domain is itself a polynomial problem.

Note also that the typical situation in **c**-optimal design problems is that the dimension m of the model is small (usually less than 10) which means that the calculation of the inverse of the $m \times m$ matrix $\mathbf{F}_B$ is rapid. On the other hand, the number of support points k is relatively very large, i.e. the choice of the entering variable is the most time consuming task. Thus, we can gain only a small increase of efficiency by using a numerically subtler approach that does not require a direct computation of $\mathbf{F}_B^{-1}$. (Note that we need to compute $\mathbf{F}_B^{-1}$ only once per iteration.)

Geometric interpretation of SAC. Let $\mathbf{f}(x_{l_1}), \ldots, \mathbf{f}(x_{l_m})$ be the m-tuple of independent vectors from the input of SAC and let $\mathcal{F}_B$ be the facet of the polytope $\mathrm{conv}\{\pm\mathbf{f}(x_{l_1}), \ldots, \pm\mathbf{f}(x_{l_m})\}$ satisfying $h\mathbf{c} \in \mathcal{F}_B$ for some $h > 0$. In Step 2, the basis B is initialized such that it includes either the index l_i (iff $\mathbf{f}(x_{l_i}) = \mathbf{f}_{l_i} \in \mathcal{F}_B$) or the index $l_i + k$ (iff $-\mathbf{f}(x_{l_i}) = \mathbf{f}_{l_i+k} \in \mathcal{F}_B$). Therefore, the vectors $\mathbf{f}_{B_1}, \ldots, \mathbf{f}_{B_m}$ belong to $\mathcal{F}_B$. In Step 3 we calculate the scalar h, and the vector $\tilde{\alpha}$ of the corresponding convex combination of vectors $\mathbf{f}_{B_1}, \ldots, \mathbf{f}_{B_m}$, such that $h\mathbf{c} = \sum_{i=1}^{m} \tilde{\alpha}_i \mathbf{f}_{B_i}$. For vectors $\mathbf{F}_{(j)} = \mathbf{f}_j, j \notin B$, we determine the coordinates $\mathbf{s}_j$ with respect to the basis $\mathbf{f}_{B_1}, \ldots, \mathbf{f}_{B_m}$. Note that $\mathbf{1}_m^{\mathrm{T}}\mathbf{s}_j$ is the scalar product of vectors $\mathbf{f}_j$ with a vector $\mathbf{n}_B = \mathbf{F}_B^{-T}\mathbf{1}_m$ normal to the hyperplane $\mathcal{P}_B = \mathrm{aff}\{\mathbf{f}_{B_1}, \ldots, \mathbf{f}_{B_m}\}$, where aff means the affine hull. In Step 4 we choose an $\mathbf{f}_{j^*}, j^* \notin B$ that has the greatest scalar product with $\mathbf{n}_B$. If the maximum scalar product is not more than 1, i.e. if no single vector from all $\mathbf{f}_j$'s is at the opposite side of $\mathcal{P}_B$ relative to $\mathbf{0}_m$, then $h\mathbf{c}$ lies at the boundary of the Elfving set. By the Elfving theorem, the coefficients $\tilde{\alpha}_i$ determine a **c**-optimal design and h^{-2} is the **c**-optimal variance. Step 5 is geometrically more subtle: Consider the simplex $\mathcal{S}_{B,j^*} = \mathrm{conv}\{\mathbf{f}_{B_1}, \ldots, \mathbf{f}_{B_m}, \mathbf{f}_{j^*}\}$, which is obviously intersected by the ray $\{r\mathbf{c} : r > 0\}$. Notice that exactly m facets of the simplex $\mathcal{S}_{B,j^*}$ contain $\mathbf{f}_{j^*}$, and that the facets can be defined by their vertices as $\mathcal{F}_{B,j^*}^{(i)} = \mathrm{conv}\{\mathbf{f}_{B_1}, \mathbf{f}_{B_2}, \ldots, \mathbf{f}_{B_m}, \mathbf{f}_{j^*}\} \setminus \{\mathbf{f}_{B_i}\}$ for $i = 1, \ldots, m$. The index i^* that determines the leaving basic variable is then chosen such that the facet $\mathcal{F}_{B,j^*}^{(i^*)}$ has linearly independent vertices and contains a positive multiple of **c**. Step 6 is just an update of the basic indices, which amounts to substituting $\mathbf{f}_{B_i^*}$ by $\mathbf{f}_{j^*}$ or exchanging the corresponding points of the experimental domain.

From the geometric interpretation it is clear that the general principle of SAC equipped with the largest coefficient rule is identical to the procedure given by Studden and Tsay (1976). However, unlike the procedure in Studden and Tsay (1976), SAC can handle degenerate situations corresponding to singular **c**-optimal designs and (when supplied with an anticycling pivot rule) SAC inherits theoretically guaranteed convergence of the general simplex method. We remark that singular **c**-optimal designs arise frequently even in the most common models, i.e. they are interesting from the practical as well as theoretical point of view (see, e.g. Fan and Chaloner (2003) or the example in Section 4).

Obviously, SAC returns only one **c**-optimal design which is supported on m or fewer points. (It is interesting to notice that Theorem 2 and convergence of the simplex algorithm with an anticycling pivot rule provide an alternative proof of the existence of a **c**-optimal design supported on at most m points, see Fellman (1974) or Section III.4 in Pázman (1986).) Nevertheless, for some models and choices of vectors **c**, the set of *all* **c**-optimal designs is very large. In this case, we can obtain a more complete picture by calculating the set of all the points that can support some **c**-optimal design:

Assume that h^{-2} is the **c**-optimal variance. By Theorem 2, a point $x_{j^*}, j^* \in \{1, \ldots, k\}$, supports some **c**-optimal designs if and only if the LP problem

$$\max \left\{ \alpha_{j^*} + \alpha_{j^*+k} \,\middle|\, \sum_{j=1}^{2k} \alpha_j \mathbf{f}_j = h\mathbf{c},\ \mathbf{1}_{2k}^{\mathrm{T}}\alpha = 1,\ \alpha \geq \mathbf{0}_{2k} \right\} \tag{16}$$

has a positive optimal value. For a large experimental domain it could be computationally prohibitive to solve the problems (16) for all design points. Nevertheless, the resulting basis B of SAC can help us significantly reduce the number of possible support points of **c**-optimal designs, i.e., the number of times we need to solve (16) as well as the size of the problems. To see this, one can check that $(\gamma^{\mathrm{T}}, h)^{\mathrm{T}}$ where $\gamma = -h\mathbf{F}_B^{-T}\mathbf{1}_m$ is a solution of the dual problem (6). Notice that γ is proportional to the vector $\mathbf{n}_B$ from the geometric interpretation of SAC, i.e. γ is orthogonal to the facet of the Elfving set that contains $\mathbf{f}_{B_1}, \ldots, \mathbf{f}_{B_m}$.

Therefore, γ satisfies the conditions of part (ii) in Theorem 3 which entails that any support point x_{j^*} of some **c**-optimal design must satisfy $h = |\mathbf{f}_{j^*}^{\mathrm{T}}\gamma| = h|\mathbf{f}_{j^*}^{\mathrm{T}}\mathbf{F}_B^{-T}\mathbf{1}_m|$, i.e., j^* must belong to the set

$$S_B = \{j \in \{1, \ldots, 2k\} : |\mathbf{1}_m^{\mathrm{T}}\mathbf{F}_B^{-1}\mathbf{f}_j| = 1\}. \tag{17}$$

Summarizing, we obtain the following theorem:

Theorem 4. *Let h^{-2} be the **c**-optimal variance, let B be the resulting basis of SAC, let S_B be defined by (17) and let $j^* \in \{1, \ldots, k\}$. Then x_{j^*} is a support point of some **c**-optimal design if and only if (i) $j^* \in S_B$ and (ii) the LP problem*

$$\max \left\{ \alpha_{j^*} + \alpha_{j^*+k} \,\middle|\, \sum_{j \in S_B} \alpha_j \mathbf{f}_j = h\mathbf{c},\ \mathbf{1}_{|S_B|}^{\mathrm{T}}\alpha = 1,\ \alpha \geq \mathbf{0}_{|S_B|} \right\}$$

has a positive optimal value.

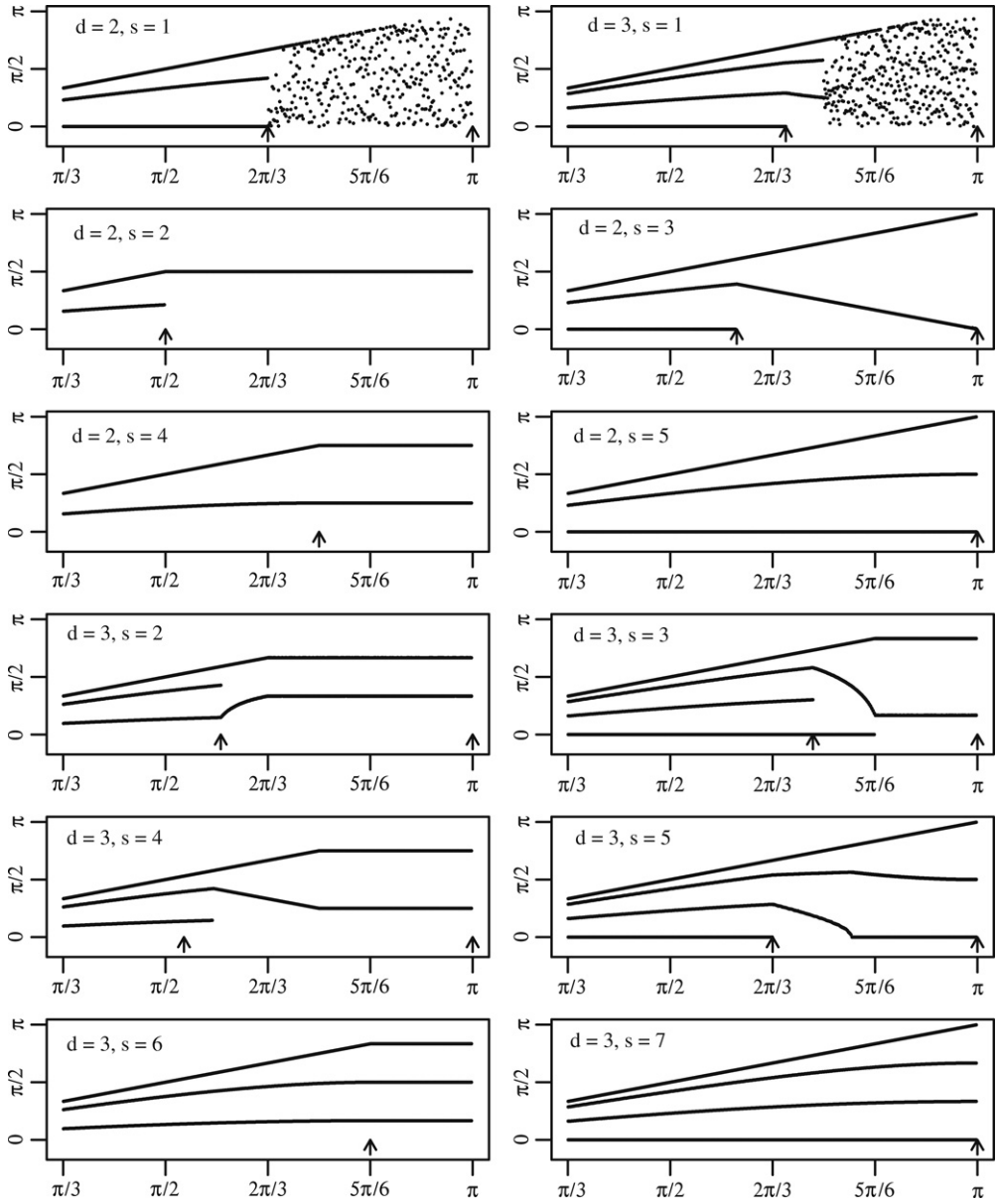

Fig. 1. Support points of e_s-optimal designs (vertical axis) for trigonometric model of degree d on a partial circle depending on the half-length a of the experimental domain (horizontal axis).

Using our approach we can identify the set of all possible support points. Clearly, if the size of the set of possible support points is exactly m (and the corresponding vectors are linearly independent), then the **c**-optimal design is unique. On the other hand, if the number of possible support points is larger than m, the **c**-optimal design is not unique. In fact it means that the number of all possible **c**-optimal designs is infinite.

4. Example: Optimal design for estimating individual parameters in trigonometric regression

Consider the trigonometric (Fourier) regression of degree d:

$$E(y) = \beta_1 + \sum_{k=1}^{d} \beta_{2k} \sin(kx) + \sum_{k=1}^{d} \beta_{2k+1} \cos(kx), \tag{18}$$

with design points x belonging to the experimental domain $\mathcal{X} = [-a, a]$, where $0 < a \leq \pi$. As an important model for periodic phenomena, trigonometric regression has received much attention in the optimal design literature; see

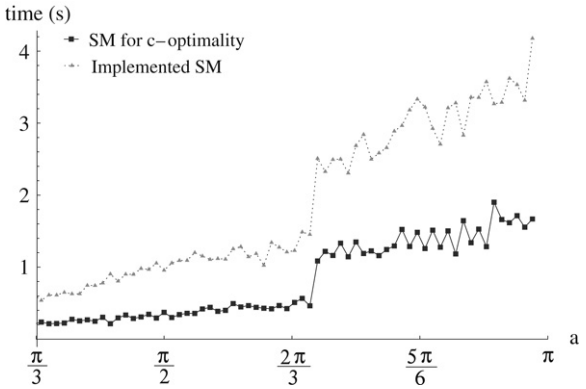

Fig. 2. Comparison of computation times (vertical axis) of SAC and a general LP routine applied to the problem of e_7-optimal designs in the cubic trigonometric model depending on the half-length a of the experimental domain (horizontal axis). The size of the experimental domain is 10 001 points.

Pázman (1986) or Pukelsheim (1993) for the case of the "full circle" $a = \pi$, and Kitsos et al. (1988), Dette et al. (2002), Dette and Melas (2002), Wu (2002) for the theory and applications of the model on a "partial circle" $a < \pi$.

In this example, we focus on e_s-optimal designs, that is, on the designs that are optimal for estimating individual coefficients $\beta_1, \ldots, \beta_{2d+1}$. The problem of e_s-optimal designs on the full circle is solved in Dette and Melas (2003) and Dette et al. (2007). For the case of the partial circle, the paper Dette and Melas (2003) describes methods of constructing e_s-optimal designs, either by an explicit formula or numerically by means of a Taylor expansion. These methods are applicable for $a \leq L_{d,s}$ and $a \geq U_{d,s}$, where $L_{d,s}$ and $U_{d,s}$ are some critical constants.

To demonstrate the use of the LP approach, we calculated the e_s-optimal designs using SAC for values a attaining 333 different levels between $\pi/3$ and π with experimental domain containing approximately 1000 a/π equidistant points in $[-a, a]$. We used an implementation of SAC in the Mathematica 5.2 programming language, and subsequent application of Theorem 4 in order to identify the set of all possible support points. (Except for the case $s = 1$, the computationally simple condition (i) of Theorem 4 was enough to reduce the number of possible **c**-optimal support points to the size 5 for $d = 2$ and 7 for $d = 3$.)

Fig. 1 plots the support points of e_s-optimal designs for degrees $d = 2$ and $d = 3$ (the case $d = 1$ is simple and can be fully solved analytically). Since the set of all support points is symmetric around 0, it is enough to exhibit only its nonnegative half. It turns out that *every* point of $\mathcal{X}$ can support an e_1-optimal design provided that a is sufficiently large, in which case the e_1-optimal designs are *not unique*. (The corresponding area at the topmost graphs of Fig. 1 is denoted by a random scatter of points.) Notice that for general d and s the number of support points depends on a and is often strictly smaller than the number of parameters, in which case the corresponding optimal designs are *singular*. Note also, that most of the values of a at which the support points change its behavior, correspond to the critical constants from the paper Dette and Melas (2003). (The constants are denoted by two arrows, or by one arrow if the constants coincide). The numerical results can guide further development of analytic results concerning e_s-optimality in the model (18) which is beyond the scope of this paper.

Finally, Fig. 2 shows a comparison of the times needed to obtain e_7-optimal designs in the cubic trigonometric model ($d = 3, s = 7$) depending on the value of a. The lower line of points corresponds to the calculation times using the Mathematica implementation of SAC, while the upper line of points corresponds to the LP procedure `LinearProgramming` of Mathematica directly applied to the problem (5). In this case, SAC is clearly more efficient than the general-purpose LP procedure.

5. Conclusions

In the opinion of the authors, the main contribution of this paper is the formulation of the problem of approximate **c**-optimal design on a finite experimental domain as a special case of linear programming (LP). As a consequence, it is possible to calculate the **c**-optimal designs by widely available LP algorithms.

The algorithm SAC that we propose is based on the simplex method of LP and uses specifics of our problem that contribute to the efficiency of the computation, e.g. possibility of a rapid construction of an initial basic feasible solution. In addition, SAC allows us to easily identify the set of all the points that can support some **c**-optimal design.

For a general LP problem, it has been proved that the worst-case complexity of the simplex method is exponential. However, it is also known (see, e.g. Borgwards (1987)) that for the average-case problems the number of iterations of the simplex method is only linear in the size of the problem. We numerically explored the average number of iterations of SAC for increasing number k of design points. The computations for the cubic trigonometric model (see Fig. 3) suggest that the average number of iterations of SAC is even smaller than in the case of a general LP problem — we conjecture that it is *sub*-linear in k. Consequently, the algorithm can be applied to **c**-optimal design problems with a large experimental domain without significant loss of efficiency.

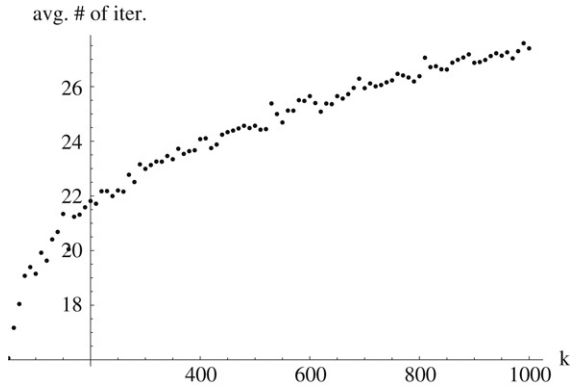

Fig. 3. Average number of SAC iterations (vertical axis) depending on the size k of the experimental domain for the problem of e_7-optimal designs in the cubic trigonometric model. The average is based on individual runs for different values of the half-length a.

To our best knowledge, computing the optimal designs with respect to other well-known criteria (such as D-, A- or E-optimality) cannot be reduced to a linear programming problem. To compute such designs we can use more complex algorithms for solving semidefinite programming or max-det programming problems, as noted in Vandenberghe et al. (1998).

Acknowledgments

We are grateful to Dr. Vippal Savani as well as two anonymous referees for insightful and very helpful comments on the first version of the paper. The work was supported by VEGA grants 1/3016/06 and 1/3022/06.

References

Borgwardt, K.H., 1987. The Average number of pivot steps required by the Simplex-Method is polynomial. In: The Simplex Method-A Probabilistic Analysis. Springer-Verlag, Heidelberg.
Dantzig, G.B., Thapa, M.N., 1997. Linear Programming: Introduction. Springer, New York.
Dette, H., 1997. Designing experiments with respect to the standardized optimality criteria. Journal of the Royal Statistical Society Series B 59 (1), 99–110.
Dette, H., Melas, V.B., 2002. E-optimal designs in Fourier regression models on a partial circle. Mathematical Methods of Statistics 11, 259–296.
Dette, H., Melas, V.B., Pepelyshev, A., 2002. D-optimal designs for trigonometric regression models on a partial circle. Ann. Inst. Statist. Math. 54, 945–959.
Dette, H., Melas, V.B., 2003. Optimal designs for estimating individual coefficients in Fourier regression models. The Annals of Statistics 31 (5), 1669–1692.
Dette, H., Melas, V.B., Shpilev, P., 2007. Optimal designs for estimating the coefficients of the lower frequencies in trigonometric regression models. Annals of the Institute of Statistical Mathematics 59, 655–673.
Elfving, G., 1952. Optimum allocation in linear regression theory. The Annals of Mathematical Statistics 23 (2), 255–262.
Fan, S.K., Chaloner, K., 2003. A geometric method for singular c-optimal designs. Journal of Statistical Planning and Inference 113, 249–257.
Fellman, J., 1974. On the allocation of linear observations. Commentationes Physico-Mathematicae, Helsinki 44, 27–78.
Kitsos, C.P., Titterington, D.M., Torsney, B., 1988. An optimal design problem in rhythmometry. Biometrics 44, 657–671.
Korte, B., Vygen, J., 2002. Combinatorial Optimization: Theory and Algorithms, second edition. Springer, New York.
López-Fidalgo, J., Rodríguez-Díaz, J.M., 2004. Elfving's method for m-dimensional models. Metrika 59 (3), 235–244.
Pázman, A., 1986. Foundations of Optimum Experimental Design. Reidel, Dordrecht.
Pukelsheim, F., 1993. Optimal Design of Experiments. John Wiley & Sons, New York.
Pukelsheim, F., Torsney, B., 1991. Optimal weights for experimental designs on linearly independent support points. The Annals of Statistics 19 (3), 1614–1625.
Roos, C., Terlaky, T., Vial, J.P., 2005. Interior Point Methods for Linear Optimization. Springer, New York.
Studden, W.J., Tsay, J.Y., 1976. Remez's procedure for finding optimal designs. The Annals of Statistics 4 (6), 1271–1279.
Vandenberghe, L., Boyd, S., Wu, S., 1998. Determinant maximization with linear matrix inequality constraints. SIAM Journal on Matrix Analysis and Applications 19, 499–533.
Vanderbei, R., 2001. Linear Programming: Foundations and Extensions, second edition. Springer, New York.
Wu, H., 2002. Optimal designs for the first-order trigonometric regression on a partial circle. Statistica Sinica 12, 917–930.