Statistica Sinica **18**(2008): Supplement, S67∼S73

OPTIMAL DESIGNS FOR FREE KNOT LEAST SQUARES SPLINES

Holger Dette[1], Viatcheslav B. Melas[2] and Andrey Pepelyshev[2]

[1] *Ruhr-Universität Bochum and* [2] *St. Petersburg State University*

Supplementary Material

5. On-line supplement: More technical proofs

Proof of Theorem 3.4 and 3.5. We start presenting two auxiliary results

Lemma 5.1. *Consider the spline polynomial*

$$\psi(x) = \sum_{i=1}^{\mu} \alpha_i f_i(x, \lambda), \tag{5.1}$$

where the functions $f_1(x, \lambda), \ldots, f_\mu(x, \lambda)$ *are defined by* (2.2) *and condition* (3.1) *is satisfied. If* $\sum_{i=1}^{\mu} \alpha_i^2 \neq 0$, *the number of isolated roots counted with their multiplicity is at most* $\mu - 1$.

Proof. Assume that the spline polynomial in (5.1) has more than $\mu - 1$ isolated roots, then it follows that the function

$$\tilde{\psi}(x) = \Big(\frac{d}{dx}\Big)^{m-k_1-1} \psi(x)$$

has at least $\mu - m + k_1 + 1$ isolated roots. On the other hand this polynomial is of the form

$$\tilde{\psi}(x) = \sum_{j=0}^{k-m+k_1} \tilde{\alpha}_j x^j + \sum_{i=1}^{r} \sum_{j=1}^{k_1+1} \tilde{\alpha}_{ij} (x - \lambda_i)^j.$$

Therefore $\tilde{\psi}$ is a polynomial of degree $\leq k - m + k_1$ on the interval $[a, \lambda_1]$ and a polynomial of degree $k_1 + 1$ on the remaining r intervals $(\lambda_1, \lambda_2], \ldots, (\lambda_r, \lambda_{r+1}]$. Consequently, $\tilde{\psi}$ has at most

$$\tilde{\mu} := k - m + k_1 + r(k_1 + 1)$$

isolated roots counted with multiplicity, which yields

$$\mu - m + k_1 + 1 \leq \tilde{\mu} = k - m + k_1 + r(k_1 + 1).$$

Observing that $\mu = k + r(k_1 + 1)$ this inequality reduces to $1 \leq 0$, which is a contradiction.

Lemma 5.2. *Any minimally supported local D-optimal design has the boundary points a and b as its support points.*

Proof. If ξ is a minimally supported local D-optimal design it must have equal weights $1/\mu$ at its support points $x_1 < \cdots < x_\mu$. From the discussion in the proof of Theorem 2.1 it follows that

$$\det M(\xi, \lambda) = \left\{ \det(f_i(x_j, \lambda))_{i,j=1}^{\mu} \right\}^2 \mu^{-\mu}.$$

Now consider the function

$$\psi(x_1) = \det(f_i(x_j, \lambda))_{i,j=1}^{\mu} = \sum_{i=1}^{\mu} \alpha_i f_i(x_1, \lambda),$$

where the last identity follows from Laplace's rule and the constants $\alpha_1, \ldots, \alpha_\mu$ depend on the points $x_2, \ldots, x_\mu$ but not on the point x_1. Obviously, $\psi(x_j) = 0$ for $j = 2, \ldots, \mu$ and consequently $\psi'(x)$ vanishes at $\mu - 2$ points $\tilde{x}_j \in (x_j, x_{j+1})$; $(j = 2, \ldots, \mu - 1)$. If $x_1 > a$ we would also have $\psi'(x_1) = 0$. On the other hand it follows from Lemma 5.1 that ψ' has at most $\mu - 2$ roots which is a contradiction. Consequently, $x_1 = a$ and it can be proved by similar arguments that $x_\mu = b$.

It now follows that a minimally supported local D-optimal design is characterized by its interior support points

$$\tau = (\tau_1, \ldots, \tau_{\mu-2}) = (x_2, \ldots, x_{\mu-1})$$

and consequently we denote candidates for such designs by

$$\xi_\tau = \begin{pmatrix} a & \tau_1 & \ldots & \tau_{\mu-2} & b \\ \frac{1}{\mu} & \frac{1}{\mu} & \ldots & \frac{1}{\mu} & \frac{1}{\mu} \end{pmatrix}.$$

Therefore the problem of determining minimally supported local D-optimal designs reduces to the maximization of the function

$$\psi(\tau, \lambda) = [\det M(\xi_{\tau,\lambda})]^{\frac{1}{\mu}} \tag{5.2}$$

over the set

$$T = \{\tau = (\tau_1, \ldots, \tau_{\mu-2})^T \mid a \leq \tau_1 \leq \cdots \leq \tau_{\mu-2} \leq b\}, \tag{5.3}$$

where

$$\lambda \in \Omega := \{(\lambda_1, \ldots, \lambda_k)^T \mid a < \lambda_1 < \cdots < \lambda_k < b\} \tag{5.4}$$

is a fixed parameter. Note that under the assumptions of Theorem 3.4 this optimization problem has a unique solution, say $\tau^* = \tau^*(\lambda)$, which satisfies the necessary conditions for an extremum, i.e.

$$\frac{\partial}{\partial \tau_i} \psi(\tau, \lambda)\Big|_{\tau=\tau^*} = 0; \quad i = 1, \ldots, \mu - 2. \tag{5.5}$$

Using the same arguments as in Melas (2006, pp.65-66), it now follows from Lemma 5.1 that the Jacobi matrix of equation (5.5),

$$J(\lambda) := \Big(\frac{\partial^2}{\partial \tau_i \partial \tau_j} \psi(\tau, \lambda)\Big|_{\tau=\tau^*(\lambda)} \Big)_{i,j=1}^{\mu-2},$$

is non-singular and

$$(J^{-1}(\lambda))_{ij} < 0; \quad i, j = 1, \ldots, \mu - 2 \tag{5.6}$$

$$\frac{\partial^2}{\partial \tau_i \partial \lambda_j} \psi(\tau, \lambda)(-1)^{s(i)} \,|_{\tau=\tau^*} < 0; \quad i = 1, \ldots, \mu - 2; \quad j = 1, \ldots, r, \tag{5.7}$$

where $s(i) \in \{1, 2\}$. Note that there could exist several solutions of (5.5) corresponding to local extrema of the function ψ. However, from the assumptions of the theorem it follows that for a fixed parameter $\lambda_0 \in \Omega$ there exists a global maximum of the function ψ and we denote by $\overline{\tau} = \tau^*(\lambda_0)$ a solution of (5.5) corresponding to this global maximum. From the implicit function theorem [see Gunning and Rossi (1965)] it therefore follows that the function $\tau^*(\lambda)$ is a unique continuous solution of (5.5) such that $\bar{\tau} = \tau^*(\lambda_0)$. By the same theorem we obtain for $j = 1, \ldots, r; i = 1, \ldots, \mu - 2$

$$\frac{\partial}{\partial \lambda_j} \tau_i^*(\lambda) = \Big(J^{-1}(\lambda) G_j (-1)^{s(i)} \Big)_i > 0,$$

where the vector G_j is defined by

$$G_j = \Big(\frac{\partial^2}{\partial \tau_\ell \partial \tau_j} \psi(\tau, \lambda)\Big|_{\tau=\tau^*(\lambda)} \Big)_{\ell=1}^{\mu-2}.$$

As a consequence the support points of the local D-optimal design for the spline regression model are increasing functions of the knots. Finally, if λ is an interior point of one of the sets Ω_j in the partition (3.12), the function $\psi(\tau, \lambda)$ is real analytic and by the implicit function theorem the solution $\tau(\lambda)$ of (5.5) is also real analytic.

Proof of Theorem 4.2. Note that a minimally supported standardized maximin D-optimal design (with respect to any set Ω) must have equal weights. Recall the definition of the function ψ in (5.2), define

$$\varphi(\tau, \lambda) = \frac{\psi(\tau, \lambda)}{\psi(\tau^*(\lambda), \lambda)}, \tag{5.8}$$

where $\tau^* = \tau^*(\lambda)$ is the vector of support points of the minimally supported local D-optimal design. Obviously, we have

$$\min_{\lambda \in \Omega_\delta^*} \varphi(\tau, \lambda) = \min_{\alpha \in [0,1]} \varphi(\tau, \alpha, \delta) \tag{5.9}$$

with

$$\varphi(\tau, \alpha, \delta) = (1-\alpha)\varphi(\tau, (1-\delta)c) + \alpha\varphi(\tau, (1+\delta)c). \tag{5.10}$$

Consequently, the problem of finding the minimally supported standardized maximin D-optimal design with respect to the set Ω_δ^* can be reduced to finding a solution $(\hat{\tau}, \hat{\alpha})$ of

$$\max_{\tau \in T} \min_{\alpha \in [0,1]} \varphi(\tau, \alpha, \delta), \tag{5.11}$$

where the set T is defined by

$$T = \{\tau = (\tau_1, \dots, \tau_{\mu-2}) \mid a < \tau_1 < \cdots < \tau_{\mu-2} < b\}$$

(if two components of the vector τ would be equal the determinant would vanish). The necessary conditions for an extremum yield

$$\begin{aligned} \frac{\partial}{\partial \tau_i} \varphi(\tau, \alpha, \delta)\Big|_{\tau=\hat{\tau}} &= 0; \quad i = 1, \dots, \mu-2, \\ \frac{\partial}{\partial \alpha} \varphi(\tau, \alpha, \delta)\Big|_{\alpha=\hat{\alpha}} &= 0, \end{aligned} \tag{5.12}$$

which will be further investigated using the following parameterization

$$\Phi(u, \delta) = \varphi(\tau^* + \rho\delta^2, \frac{1}{2} + \beta\delta, \delta) \cdot \frac{\psi(\tau^*, c)}{\delta^2}. \tag{5.13}$$

Here $u = (\rho, \beta) = (\rho_1, \dots, \rho_{\mu-2}, \beta)$ and τ^* denotes the vector of interior support points of the minimally supported local D-optimal design for the vector $c = (c_1, \dots, c_r)$; i.e. $\tau^* = \tau^*(c)$. Obviously, the equations (5.12) are equivalent to

$$\frac{\partial}{\partial u_i} \Phi(u, \delta)\Big|_{u=\hat{u}} = 0, \quad i = 1, \dots, \mu-1, \tag{5.14}$$

and the solutions $\hat{u} = (\hat{\rho}, \hat{\beta})$ and $(\hat{\tau}, \hat{\alpha})$ are related by

$$\hat{\tau} = \tau^* + \hat{\rho}\delta^2; \hat{\alpha} = \frac{1}{2} + \hat{\beta}\delta. \tag{5.15}$$

Assume that δ^* is sufficiently small and define the set

$$\mathcal{U}_\rho := \Big\{u = (\rho, \beta) \Big| \frac{a - \tau^*}{\delta^2} < \rho_1 < \cdots < \rho_{\mu-2} < \frac{b - \tau^*}{\delta^2}; -\frac{1}{2\delta} \le \beta \le \frac{1}{2\delta}\Big\},$$

then we prove the following assertions.

(I) There exists a unique continuous function

$$\hat{u} : \begin{cases} (-\delta^*, \delta^*) \to \mathcal{U} \\ \quad\delta \quad\quad \to \hat{u}(\delta) \end{cases} \tag{5.16}$$

such that for each $\delta \in (-\delta^*, \delta^*)$ the value $\hat{u}(\delta)$ is a solution of the system (5.14).

(II) The function defined in (I) is real analytic and the coefficients in the corresponding Taylor expansion

$$\hat{u}(\delta) = \sum_{j=0}^{\infty} u_{(j)} \delta^j$$

can be calculated recursively as

$$\begin{aligned} u_{(0)} &= -\hat{J}^{-1}[h(0,\delta)]_{(2)}, \\ u_{(s+1)} &= -\hat{J}^{-1}[h(u_{\langle s\rangle}(\delta),\delta)]_{(s+3)}, \quad s = 0, 1, 2, \ldots, \end{aligned} \tag{5.17}$$

where $u_{\langle s\rangle}$ is defined in (3.15),

$$h(u,\delta) = \Big(\frac{\partial}{\partial u_1}\Phi(u,\delta), \ldots, \frac{\partial}{\partial u_{\mu-1}}\Phi(u,\delta)\Big)^T \tag{5.18}$$

$$A = \Big(\frac{\partial^2}{\partial \tau_i \partial \tau_j}\psi(\tau,c)\Big|_{\tau=\tau^*}\Big)_{i,j=1}^{\mu-2}$$

$$b = \Big(\sum_{j=1}^{r} c_j \frac{\partial^2}{\partial \tau_i \partial c_j}\psi(\tau,c)\Big|_{\tau=\tau^*}\Big)_{i=1}^{\mu-2}$$

$$\hat{J} = \begin{pmatrix} A & b \\ b^T & 0 \end{pmatrix} \in \mathbb{R}^{\mu-1\times\mu-1}. \tag{5.19}$$

(III) The design

$$\xi_{\hat{\tau}} = \begin{pmatrix} a & \hat{\tau}_1 & \ldots & \hat{\tau}_{u-2} & b \\ \frac{1}{\mu} & \frac{1}{\mu} & \ldots & \frac{1}{\mu} & \frac{1}{\mu} \end{pmatrix}$$

is the unique minimally supported standardized maximin D-optimal design with respect to the set Ω_δ^*.

(IV) The design $\xi_{\hat{\tau}}$ is the unique minimally supported standardized maximin D-optimal design with respect to the set Ω_δ.

For a proof of (I) and (II) we note that $h(u, \delta)$ is a real analytic vector valued function in a neighbourhood of the point $(u^*, \delta^*) = (0, 0)$, with components satisfying

$$h_i(0,0) = \frac{\partial}{\partial u_i} h(u,\delta)\Big|_{(u,\delta)=(0,0)} = 0; \quad i = 1, \dots, \mu - 1,$$

and

$$\Big(\frac{\partial}{\partial u_j} h_i(u,\delta)\Big)_{i,j=1}^{\mu-1} = \delta^2 \hat{J} + O(\delta^3),$$

where the matrix $\hat{J}$ is defined in (5.19). Obviously,

$$\det \hat{J} = -(\det A) b^T A^{-1} b,$$

where $\det A \neq 0$ as demonstrated in the proof of Theorem 3.4 and 3.5. A similar argument shows that $b \neq 0$ and therefore the matrix $\hat{J}$ is non singular. The implicit function theorem [see Gunning and Rossi (1965)] now shows the existence of a unique real analytic solution $\hat{u}$ of (5.14) in a sufficiently small interval $(-\delta^*, \delta^*)$. The recursive relation (5.17) for the coefficients in the corresponding Taylor expansion is now a consequence of from Theorem 5.3 in Melas (2005).

In order to prove (III) we note that it follows from the uniqueness of the minimally supported local D-optimal design for any $\delta \in (0, 1)$

$$\min_{0\le\alpha\le1} (1-\alpha) \frac{\psi(\tau,(1-\delta)c)}{\psi(\tau^*((1-\delta)c),(1-\delta)c)} + \alpha \frac{\psi(\tau,(1+\delta)c)}{\psi(\tau^*((1+\delta)c),(1+\delta)c)} < 1. \quad (5.20)$$

For $\delta \in [0, 1]$ define as $(\tilde{\tau}, \tilde{\alpha})$ a point where the optimum in (5.11) is attained, that is

$$\varphi(\tilde{\tau}, \tilde{\alpha}, \delta) = \max_{\tau\in T} \min_{\alpha\in[0,1]} \varphi(\tau, \alpha, \delta).$$

If $\tilde{\alpha} = 0$ we would obtain

$$\varphi(\tilde{\tau}, \tilde{\alpha}, \delta) = \varphi(\tilde{\tau}, 0, \delta) = \max_{\tau\in T} \frac{\psi(\tau,(1-\delta)c)}{\psi(\tau^*((1-\delta)c),(1-\delta)c)} = 1,$$

which contradicts (5.20). Similary, we can exclude the case $\tilde{\alpha} = 1$. The matrix A in (5.18) is nonsingular and the Hesse matrix of the function $\psi(\tau, c)$ evaluated at the extreme point τ^* must be negative definite. Consequently, it follows that for sufficiently small δ the function $\varphi(\tau, \alpha, \delta)$ defined in (5.10) is a concave function of τ in a neighbourhood of the point τ^*. This means that $(\hat{\tau}, \hat{\alpha}) = (\tilde{\tau}, \tilde{\alpha})$ and consequently the design $\xi_{\hat{\tau}}$ is the unique minimally supported standardized maximin D-optimal design with respect to the set Ω_δ^*.

Finally, we prove assertion (IV), which follows from the equation

$$\min_{\lambda\in\Omega_\delta} \varphi(\hat{\tau}, \lambda) = \min_{\lambda\in\Omega_\delta^*} \varphi(\hat{\tau}, \lambda). \quad (5.21)$$

To prove (5.21) we define the rescaled quantities $\gamma_i = (\lambda_i - c_i)/(\delta c_i)$ $(i = 1, \dots, r)$ and note that $|\gamma_i| \leq 1$ if $\lambda \in \Omega_\delta$. A straightforward but tedious calculation yields

$$\varphi(\hat{\tau}, \lambda) = 1 + \delta^2 \gamma^T B^T A B \gamma + O(\delta^3), \tag{5.22}$$

where $\gamma = (\gamma_1, \dots, \gamma_r)^T, B = A^{-1}D$, the matrix D is defined by

$$D = \Big(\frac{\partial^2 h(\tau, c)}{\partial \tau_i \partial c_j} \Big|_{\tau = \tau^*} \Big)_{i=1,\dots,\mu-2}^{j=1,\dots,r},$$

and the elements of the matrix A^{-1} and D are negative and positive, respectively (this follows by similar arguments as given in Melas (2006, pp.56-57)). Consequently, the elements of the matrix $D^T A^{-1} D$, say z_{ij} $(i, j = 1, \dots, r)$, are negative and (5.22) yields

$$\varphi(\hat{\tau}, \lambda) = 1 + \delta^2 \sum_{i,j=1}^{r} z_{ij} \gamma_i \gamma_j + O(\delta^3).$$

Therefore, if δ is sufficiently small, the minimum of $\varphi(\hat{\tau}, \lambda)$ is attained if all components of $\gamma = (\gamma_1, \dots, \gamma_r)$ have the same sign and are equal to $+1$ or -1. Consequently, the minimum is attained either at $\lambda = (1 - \delta)c$ or $\lambda = (1 + \delta)c$.

References

Melas, V. B. (2005). On the functional approach to optimal designs for nonlinear models. J. Statist. Plann. Inference **132**, 93-116.

Melas, V. B. (2006). Functional Approach to Optimal Experimental Design. Springer, Heidelberg.

Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany.

E-mail: holger.dette@rub.de

Department of Mathematics, St. Petersburg State University, St. Petersburg, Russia.

E-mail: v.melas@pobox.spbu.ru

Department of Mathematics, St. Petersburg State University, St. Petersburg, Russia.

E-mail: andrey@ap7236.spb.edu

(Received September 2006; accepted January 2007)