Bernoulli **15**(4), 2009, 1036–1056
DOI: 10.3150/09-BEJ185

Asymptotic optimal designs under long-range dependence error structure

HOLGER DETTE[1], NIKOLAI LEONENKO[2], ANDREY PEPELYSHEV[3] and ANATOLY ZHIGLJAVSKY[4]

[1]*Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany.*
E-mail: holger.dette@rub.de
[2]*Cardiff University, School of Mathematics, Cardiff CF24 4AG, UK.*
E-mail: LeonenkoN@cf.ac.uk
[3]*St. Petersburg State University, Department of Mathematics, St. Petersburg, Russia.*
E-mail: andrey@ap7236.spb.edu
[4]*Cardiff University, School of Mathematics, Cardiff CF24 4AG, UK.*
E-mail: ZhigljavskyAA@cf.ac.uk

We discuss the optimal design problem in regression models with long-range dependence error structure. Asymptotic optimal designs are derived and it is demonstrated that these designs depend only indirectly on the correlation function. Several examples are investigated to illustrate the theory. Finally, the optimal designs are compared with asymptotic optimal designs which were derived by Bickel and Herzberg [*Ann. Statist.* **7** (1979) 77–95] for regression models with short-range dependent error.

Keywords: asymptotic optimal designs; linear regression; long-range dependence

1. Introduction

Consider the common linear regression model

$$y(t) = \theta_1 f_1(t) + \cdots + \theta_p f_p(t) + \varepsilon(t), \tag{1.1}$$

where $f_1(t), \ldots, f_p(t)$ are known functions, $\varepsilon(t)$ is a random error, $\theta_1, \ldots, \theta_p$ denote the unknown parameters and t is the explanatory variable. We assume that N observations, say $y_1, \ldots, y_N$, can be taken at experimental conditions $-T \le t_1 \le \cdots \le t_N \le T$ to estimate the parameters in the linear regression model (1.1). If an appropriate estimate, say $\hat\theta = (\hat\theta_1, \ldots, \hat\theta_p)^T$, has been chosen, an optimal design minimizes a function of the variance-covariance matrix of this estimate, which is called an optimality criterion (see, for example, Silvey (1980) or Pukelsheim (1993)).

Under the assumption of uncorrelated observations, optimal designs have been studied by numerous authors (see the two books cited above and the textbooks of Fedorov (1972), Pázman (1986) and Atkinson and Donev (1992)). However, fewer results are available for dependent observations, although this problem is of particular interest because in many applications, the variable t in the regression model (1.1) represents time and all observations correspond to one subject. The reason for this is that optimal experimental designs for regression models with correlated observations have a very complicated structure and are difficult to find, even in simple cases. Because explicit solutions are rarely available, an asymptotic theory was developed by Sacks and Ylvisaker (1966, 1968). In the Sacks–Ylvisaker approach, the design set is fixed and the number of design points in this set tends to infinity. As a result of this assumption, the design points become too close to each other and the corresponding asymptotic optimal designs depend only on the behavior of the correlation function in a neighborhood of the point 0.

Bickel and Herzberg (1979) and Bickel, Herzberg and Schilling (1981) considered a different model, where the design interval expands proportionally to the number of observation points and the correlation structure of errors is not used for the construction of the least-squares estimate. The variance-covariance matrix of the estimate $\hat{\theta}$ is of order $\mathrm{O}(1)$ in the model considered by Sacks and Ylvisaker (1966) and of order $1/N$ in the model discussed by Bickel and Herzberg (1979). Therefore, the approach of Bickel and Herzberg makes the optimal designs derived for the dependent and independent cases more comparable. These authors assumed that the observations in model (1.1) have a correlation structure corresponding to a non-degenerate stationary process with short-range dependence, where a correlation function ρ satisfies $\rho(t) = \mathrm{o}(1/t)$ if $t \to \infty$. As examples, in Bickel and Herzberg (1979) and Bickel, Herzberg and Schilling (1981), asymptotic optimal designs are derived for the linear regression model with and without intercept and for the location model.

The purpose of the present paper is to extend the Bickel–Herzberg approach to the case of a stronger dependence of the errors in the linear regression model (1.1), which corresponds to an error process with long-range dependence. Long-range dependence is observed in many applications, including hydrology, geophysics, turbulence, diffusion, economics and finance. The phenomenon was already observed by Pearson (1902) in astronomy and by Smith (1938) in agriculture. Further examples where long-range dependence is discussed can be found in Granger (1980), Mandelbrot (1973), Porter-Hudak (1990), Beran, Sherman, Taqqu and Willinger (1992), Barndorff-Nielsen *et al.* (1990), Beran (1992), Metzler *et al.* (1999), among many others. The interested reader is referred to the books of Beran (1994) and Doukhan *et al.* (2003), which contain a good description of the basic properties of long-range dependence processes and an extensive bibliography on this subject.

Most of the literature considers the estimation problem but – to the best knowledge of the authors – design problems for regression models with long-range dependence error structure have not been considered thus far. In Section 2, we introduce the basic terminology and describe the optimal design problem. Our main results are given in Section 3, where we derive an asymptotic expression for the variance-covariance matrix, the basis for the construction of optimal designs in the regression model (1.1) with a long-range dependent error structure. These results are different from the findings of Künsch, Beran and Hampel (1993), who considered random explanatory variables. Finally, in Section 4 several asymptotic optimal designs are derived for the linear regression model and compared with the results obtained by Bickel and Herzberg (1979) under the assumption of a short-range error structure.

2. Optimal designs for dependent observations

Consider the linear regression model (1.1), where the error process $\varepsilon(t)$ is the second-order process with

$$\mathbf{E}\varepsilon(t) = 0, \qquad \mathbf{E}\varepsilon(t)\varepsilon(s) = \sigma^2 r(t,s), \tag{2.1}$$

and assume that

(C1). The regression functions $f_1(t), \dots, f_p(t)$ are linearly independent and bounded on the interval $[-T, T]$ and satisfy a first order Lipschitz condition, that is, $|f_i(t) - f_i(s)| \le M|t-s|$ and $|f_i(t)| \le M$ for all $t, s \in [-T, T]$, $i = 1, \dots, p$.

Following Bickel and Herzberg (1979), we assume that $\varepsilon(t) = \varepsilon^{(1)}(t) + \varepsilon^{(2)}(t)$, where $\varepsilon^{(1)}(t)$ denotes a stationary process with correlation function $\rho(t)$ and $\varepsilon^{(2)}(t)$ is white noise. Consequently, we obtain

$$r(t,s) = \gamma\rho(t-s) + (1-\gamma)\delta_{t,s}, \tag{2.2}$$

where δ is the Kronecker symbol. If N observations, say $y = (y_1, \dots, y_N)^T$, are available and the form of the correlation function is known, then the vector of parameters can be estimated by the weighted least squares, that is, $\hat{\theta} = (X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1} y$ with $X^T = (f_i(t_j))_{i=1,\dots,p}^{j=1,\dots,N}$, and the variance-covariance matrix of this estimate is given by

$$\mathbf{D}(\hat{\theta}) = \sigma^2 (X^T \Sigma^{-1} X)^{-1} \tag{2.3}$$

with $\Sigma = (\gamma\rho(t_i - t_j) + (1-\gamma)\delta_{i,j})_{i,j}$, $i, j = 1, \dots, N$. However, in most applications, knowledge about the correlation structure is not available and the unweighted least-squares estimate $\tilde{\theta} = (X^T X)^{-1} X^T y$ is used. For this estimate, the variance-covariance matrix is given by

$$\mathbf{D}(\tilde{\theta}) = \sigma^2 (X^T X)^{-1} X^T \Sigma X (X^T X)^{-1}. \tag{2.4}$$

An experimental design $\xi = \{t_1, \dots, t_N\}$ is a vector of N points in the interval $[-T, T]$, which defines the time points or experimental conditions where observations are taken. Optimal designs minimize a functional of the variance-covariance matrix of the weighted or unweighted least-squares estimate. Following Bickel and Herzberg (1979), we consider a correlation function which depends on the sample size N and is of the form $\rho_N(t) = \rho(Nt)$, where the function ρ satisfies $\rho(t) \to 0$ if $t \to \infty$; this corresponds to expanding the interval as the number of observations grows. The standard least-squares estimate is considered in the following discussion because when computing this estimate, the form of the correlation function $\rho(t)$ is not used. Despite this, the least-squares estimate often has good properties compared to the best linear unbiased estimate; see, for example, Adenstedt (1974), Samarow and Taqqu (1988), Yajima (1988, 1991) and Beran (1994), page 179, among others. For some results regarding nonlinear regression, the reader is referred to Ivanov and Leonenko (2004, 2008).

For our asymptotic investigations, we consider a sequence of designs $\xi_N = \{t_{1N}, \ldots, t_{NN}\}$ which is generated using a continuous non-decreasing function

$$a : [0, 1] \to [-T, T] \tag{2.5}$$

by

$$t_{iN} = a\big((i-1)/(N-1)\big), \qquad i = 1, \ldots, N, \tag{2.6}$$

where the function $a(u)$ is the inverse of a distribution function. Note that the function a is obtained as the weak limit of ξ_N as $N \to \infty$. The equally spaced design corresponds to the choice $a(u) = (2u-1)T$ ($u \in [0, 1]$); changing the function a yields different types of the design. For example, the choice $a^{-1}(x) = (x^3 + 1)/2$ yields designs which are more concentrated in the interval $[-1/2, 1/2]$. We assume several regularity conditions on the function a, which are required for the asymptotic results which are to follow. More precisely:

(C2). Let $a(u)$ be twice differentiable and assume that there exists a positive constant $M < \infty$ such that for all $u \in (0, 1)$,

$$\frac{1}{M} \le a'(u) \le M, \qquad |a''(u)| \le M. \tag{2.7}$$

(C3). The correlation function ρ is differentiable with bounded derivative, that is, $|\rho'(t)| \le M$, $t \in (0, \infty)$ and $\rho'(t) \le 0$ for sufficiently large t.

The last assumption implies that $\rho(t)$ is nonnegative for sufficiently large t. In contrast to Bickel and Herzberg (1979) we assume that

$$\int_0^\infty |\rho(t)|\,\mathrm{d}t = \infty \tag{2.8}$$

and this assumption corresponds to the long-range dependence of the observations. Note that in this case it follows that

$$\int_0^\infty |\rho(t)|\,\mathrm{d}t = \sum_{k=0}^\infty |\rho(k)| = \infty,$$

where $\rho(k) = \mathrm{cov}(\varepsilon^{(1)}(t), \varepsilon^{(1)}(t+k))$. The correlation function of a stationary process with long-range dependence can be written as

$$\rho_\alpha(t) = \frac{L(t)}{|t|^\alpha}, \qquad |t| \to \infty, \tag{2.9}$$

where $0 < \alpha \le 1$ and $L(t)$ is a slowly varying function (SVF) for large t [Doukhan *et al.* (2003)], and satisfies

$$\rho_\alpha(t) = \mathrm{O}(1/|t|^\alpha), \qquad |t| \to \infty.$$

In this case, we will say that $\rho_\alpha(t)$ belongs to the SVF family.

3. Main results

First, we introduce two parametric families of correlation functions which are important in applications.

The correlation function $\rho_\alpha(t)$ belongs to the Cauchy family if it is defined by

$$\rho_\alpha(t) = \frac{1}{(1+|t|^\beta)^{\alpha/\beta}}, \tag{3.1}$$

where $\beta > 0$, $0 < \alpha \le 1$ [see Gneiting (2000), Anh *et al.* (2004), Barndorff-Nielsen and Leonenko (2005)]. This family includes

$$\rho_\alpha^{(1)}(t) = \frac{1}{(1+|t|^2)^{\alpha/2}}, \qquad \rho_\alpha^{(2)}(t) = \frac{1}{1+|t|^\alpha}, \qquad \rho_\alpha^{(3)}(t) = \frac{1}{(1+|t|)^\alpha},$$

which have a totally different shape in a neighbourhood of the point $t = 0$ but the same asymptotic behavior for large t (see Figure 1). These three functions are known as the characteristic

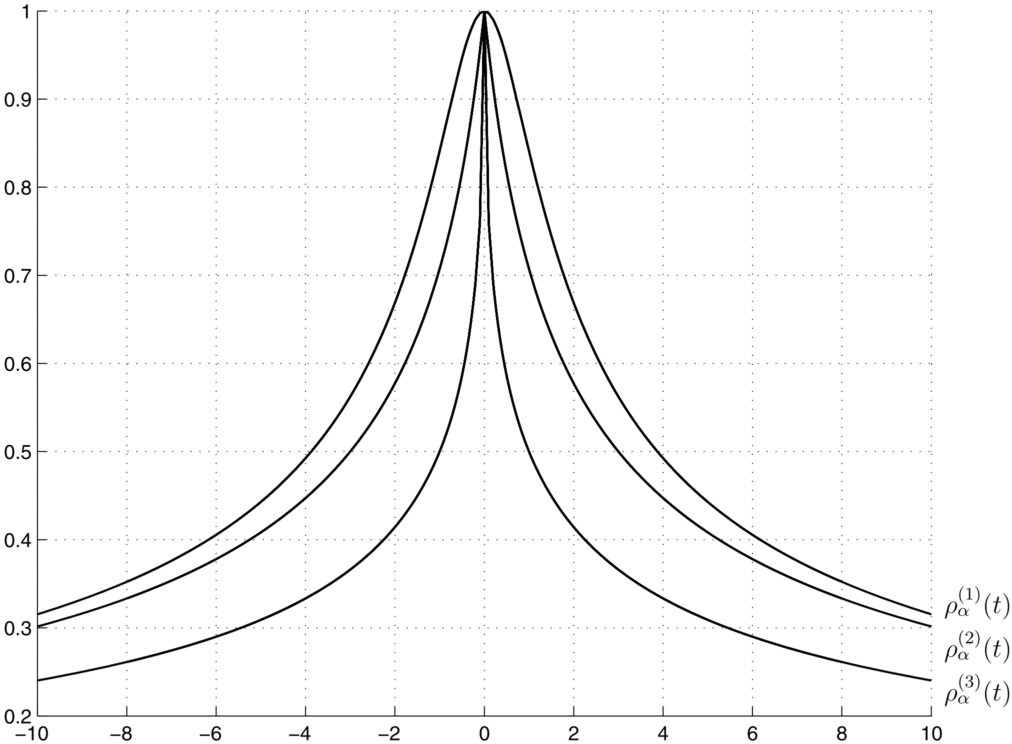

Figure 1. The three correlation functions, where $\alpha = 0.5$.

functions of the symmetric Bessel distribution, Linnik distribution and symmetric generalized Linnik distribution, respectively.

The correlation function $\rho_\alpha(t)$ belongs to the Mittag-Leffler family if it is defined by

$$\rho_\alpha(t) = E_{\nu,\beta}(-|t|^\alpha), \qquad E_{\nu,\beta}(-t) = \Gamma(\beta)\sum_{k=0}^{\infty}\frac{(-t)^k}{\Gamma(\nu k+\beta)}, \qquad t>0, \tag{3.2}$$

where $0<\alpha\le 1$, $0<\nu\le 1$, $\beta\ge\nu$ (see Schneider (1996), Barndorff-Nielsen and Leonenko (2005)). This family is a smooth interpolation of long-range dependence ($0<\alpha\le 1, \beta=1, 0<\nu<1$) and short-range dependence ($\nu=1, 0<\alpha\le 1, \beta=1$). Note that the case $\nu=1, \beta=1, \alpha=1$ corresponds to ordinary diffusion $\varrho_\alpha(t)=\mathrm{e}^{-|t|}$, which is the correlation function of a Markovian Ornstein–Uhlenbeck process. On the other hand, the case $0<\nu<1, \beta=1, 0<\alpha<1$ corresponds to subdiffusion or slow diffusion (see Metzler and Klafter (2000)). In particular,

$$E_{1,1}(-t)=\mathrm{e}^{-t}, \qquad E_{1,2}(-t)=(1-\mathrm{e}^{-t})/t, \qquad E_{1,3}(-t)=2(\mathrm{e}^{-t}-1+t)/t^2,$$
$$E_{1/2,1}(-t)=\mathrm{e}^{t^2}\left(1-\frac{2}{\sqrt{\pi}}\int_0^t \mathrm{e}^{-u^2}\,\mathrm{d}u\right).$$

In the following discussion, we derive optimal designs for the three families of correlation functions, which are given by (2.9), (3.1) and (3.2). The function $Q(t)=\sum_{j=1}^{\infty}\rho(jt)$ plays an important role in the asymptotic analysis by Bickel and Herzberg (1979), but in the case of long-range dependence, this function is infinite. For an asymptotic analysis under long-range dependence, we introduce the function

$$Q_\alpha(t)=\lim_{N\to\infty}\frac{1}{d_\alpha(N)}\sum_{j=1}^{N}\rho_\alpha(jt), \tag{3.3}$$

where the normalizing sequence is given by

$$d_\alpha(N)=\begin{cases} N^{1-\alpha}, & \text{if } \alpha<1 \text{ and } \rho_\alpha \text{ has the form (3.1) or (3.2)},\\ \ln N, & \text{if } \alpha=1 \text{ and } \rho_\alpha \text{ has the form (3.1) or (3.2)},\\ L(N)N^{1-\alpha}, & \text{if } \alpha<1 \text{ and } \rho_\alpha \text{ has the form (2.9)},\\ L(N)\ln N, & \text{if } \alpha=1 \text{ and } \rho_\alpha \text{ has the form (2.9)},\end{cases}$$

and show in Lemma 1 below that the function $Q_\alpha(t)$ is well defined.

Lemma 1. *If the correlation function $\rho_\alpha(t)$ belongs to the Cauchy, SVF family or to the Mittag-Leffler family with $0<\alpha\le 1, 0<\nu\le 1, \nu\le\beta, (\nu,\beta)\ne(1,1)$, then the limit in (3.3) exists and is given by*

$$Q_\alpha(t)=\begin{cases}\dfrac{c}{(1-\alpha)|t|^\alpha}, & 0<\alpha<1,\\[1ex] \dfrac{c}{|t|}, & \alpha=1,\end{cases}$$

where

$$c = \begin{cases} \dfrac{\Gamma(\beta)}{\Gamma(\beta-\nu)}, & \text{if } \rho_\alpha(t) \text{ belongs to the Mittag-Leffler family,} \\ 1, & \text{otherwise.} \end{cases}$$

Proof. Define the function

$$Q_{\alpha,N}(t) = \frac{1}{d_\alpha(N)} \sum_{j=1}^{N} \rho_\alpha(jt)$$

and assume that the correlation function $\rho_\alpha(t)$ is an element of the Cauchy family. Since the function $\rho_\alpha(t)$ defined in (3.1) is positive and decreasing for $0 < \alpha < 1$, we have

$$\begin{aligned} Q_{\alpha,N}(t) &= \frac{1}{N^{1-\alpha}} \int_0^N \frac{1}{(1+|st|^\beta)^{\alpha/\beta}} \,\mathrm{d}s + \mathrm{O}\left(\frac{1}{N^{1-\alpha}}\right) \\ &= \frac{1}{N^{1-\alpha}} N \int_0^N \frac{\mathrm{d}(s/N)}{N^\alpha (1/N^\beta + |st/N|^\beta)^{\alpha/\beta}} + \mathrm{O}\left(\frac{1}{N^{1-\alpha}}\right) \\ &= \int_0^1 \frac{\mathrm{d}v}{(|vt|^\beta)^{\alpha/\beta}} + \mathrm{O}\left(\frac{1}{N^{\alpha-\alpha^2}}\right) = \frac{1}{(1-\alpha)|t|^\alpha} + \mathrm{O}\left(\frac{1}{N^{\alpha-\alpha^2}}\right) \\ &= Q_\alpha(t) + \mathrm{O}\left(\frac{1}{N^{\alpha-\alpha^2}}\right). \end{aligned}$$

For $\alpha = 1$, we obtain

$$\begin{aligned} Q_1(t) &= \lim_N \frac{1}{\ln N} \sum_{j=1}^{N} \frac{1}{(1+|jt|^\beta)^{1/\beta}} = \lim_N \frac{1}{|t| \ln N} \int_0^N \frac{1}{(1+|st|^\beta)^{1/\beta}} \,\mathrm{d}(st) \\ &= \lim_N \frac{1}{|t| \ln N} \int_0^{Nt} \frac{1}{1+|v|} \,\mathrm{d}v = \frac{1}{|t|}, \end{aligned}$$

which completes the proof for the case where $\rho_\alpha(t)$ belongs to the Cauchy family.

Now assume that the correlation function $\rho_\alpha(t)$ is an element of the Mittag-Leffler family. For the sake of brevity, we only consider the case $\beta = 1$, $0 < \alpha \leq 1$, $0 < \nu < 1$, all other cases being treated similarly. Since

$$E_{\nu,1}(-|t|^\alpha) \sim \frac{1}{|t|^\alpha \Gamma(1-\nu)}$$

as $t \to \infty$ (see, for example, formula (3.17) in Schneider (1996)), we have, for $0 < \alpha < 1$,

$$Q_\alpha(t) = \lim_{N\to\infty} \frac{1}{N^{1-\alpha}} \sum_{j=1}^{N} E_{\nu,1}(-|jt|^\alpha) = \frac{1}{(1-\alpha)\Gamma(1-\nu)|t|^\alpha}.$$

Observing that

$$E_{\nu,\beta}(-|t|) \sim \frac{\Gamma(\beta)}{|t|\Gamma(\beta-\nu)}$$

for $t \to \infty$ (see Djrbashian (1993)), we obtain, for $\alpha = 1$,

$$Q_1(t) = \lim_{N\to\infty} \frac{1}{\ln N} \sum_{j=1}^{N} E_{\nu,\beta}(-|jt|) = \frac{\Gamma(\beta)}{|t|\Gamma(\beta-\nu)}.$$

Finally, assume that the correlation function $\rho_\alpha(t)$ is an element of the SVF family. We then obtain

$$\begin{aligned}
Q_\alpha(t) &= \lim_{N\to\infty} \frac{1}{L(N)N^{1-\alpha}} \int_0^N \frac{L(st)}{|st|^\alpha}\,\mathrm{d}s \\
&= \lim_{N\to\infty} \frac{1}{L(N)N^{1-\alpha}} N \int_0^N \frac{L(Nts/N)\,\mathrm{d}(s/N)}{N^\alpha |st/N|^\alpha} \\
&= \lim_{N\to\infty} \int_0^1 \frac{L(Ntv)\,\mathrm{d}v}{L(N)|vt|^\alpha} = \int_0^1 \frac{\mathrm{d}v}{|vt|^\alpha} = \frac{1}{(1-\alpha)|t|^\alpha},
\end{aligned}$$

where we have used Theorem 2.6 from Seneta (1976) in the last line. For $\alpha = 1$, we have

$$\begin{aligned}
Q_1(t) &= \lim_{N\to\infty} \frac{1}{L(N)\ln N} \int_1^N \frac{L(st)}{|st|}\,\mathrm{d}s = \lim_N \frac{1}{\ln N} \int_1^N \frac{L(st)/L(N)}{st}\,\mathrm{d}s \\
&= \lim_{N\to\infty} \frac{1}{\ln N} \int_1^N \frac{1}{|st|}\,\mathrm{d}s + \lim_N \frac{1}{\ln N} \int_1^N \frac{L(st)/L(N)-1}{|st|}\,\mathrm{d}s \\
&= \frac{1}{|t|} + \lim_N \frac{1}{\ln N} \int_{1/N}^1 \frac{L(Nvt)/L(N)-1}{|vt|}\,\mathrm{d}v = \frac{1}{|t|},
\end{aligned}$$

which completes proof of Lemma 1. □

Next, we find a comfortable asymptotic representation for the main term in the variance-covariance matrix of the least-squares estimates.

Lemma 2. *Assume that the correlation function $\rho_\alpha(t)$ belongs to the Cauchy, Mittag-Leffler or SVF family, such that*

$$\int_0^1 Q_\alpha(a'(t))\,\mathrm{d}t < \infty, \tag{3.4}$$

and that the regularity conditions (C1)–(C3) in Sections 2 and 3 are satisfied. We have

$$\frac{1}{d_\alpha(N)N} \sum_{i\neq j} f_s(t_{iN}) f_r(t_{jN}) \rho_\alpha\bigl(N(t_{jN}-t_{iN})\bigr) = 2\int_0^1 f_s(a(u)) f_r(a(u)) Q_\alpha(a'(u))\,\mathrm{d}u + \mathrm{o}(1)$$

as $N \to \infty$ *for all* $s, r = 1, \dots, p$, $0 < \alpha \le 1$.

Proof. We only give a proof for the correlation function from the Cauchy family and $0 < \alpha < 1$, the proof for the other cases being similar. We use the notation $f = f_s$, $g = f_r$, $\rho = \rho_\alpha$ and the decomposition

$$N^{\alpha-2} \sum_{i \neq j} f(t_{iN}) g(t_{jN}) \rho\big(N(t_{jN} - t_{iN})\big) = S_1 + S_2,$$

where

$$S_1 = 2N^{\alpha-2} \sum_{i=1}^{N} f(t_{iN}) g(t_{iN}) \sum_{j=i+1}^{N} \rho\big(N(t_{jN} - t_{iN})\big), \tag{3.5}$$

$$S_2 = 2N^{\alpha-2} \sum_{i=1}^{N} f(t_{iN}) \sum_{j=i+1}^{N} \big(g(t_{jN}) - g(t_{iN})\big) \rho\big(N(t_{jN} - t_{iN})\big). \tag{3.6}$$

With the notation $i_N = (i-1)/(N-1)$, we obtain from the differentiability of the functions a and ρ

$$\rho\big(N(t_{jN} - t_{iN})\big) = \rho\big(N\big(a(j_N) - a(i_N)\big)\big) = \rho\big(a'(i_N)(j-i)\big) + \nu \frac{(j-i)^2}{N-1},$$

where $|\nu| \le M^2/2$. Let r_N denote a sequence such that $r_N \to \infty$ slowly as $\mathrm{o}(N^{(1-\alpha)/3})$ and consider the cases $i \le r_N$ and $i > r_N$ in (3.5) and (3.6) separately. Note that

$$\begin{aligned} \left| \sum_{j=i+r_N}^{N} \rho\big(N(t_{jN} - t_{iN})\big) \right| &= \sum_{j=i+r_N}^{N} \rho\big(N\big(a(j_N) - a(i_N)\big)\big) \\ &\le \tilde{M} \sum_{j=i+r_N}^{N} \rho\big((j-i)/M\big) \le \tilde{M} \sum_{k=r_N}^{\infty} \rho(k/M) = \mathrm{o}(N^{1-\alpha}) \end{aligned}$$

as $N \to \infty$ uniformly with respect to j, where $\tilde{M}$ is a constant and where we have used the fact that the function $a'(u)$ is bounded from below and Lemma 1. Similarly, we obtain

$$\begin{aligned} &\left| \sum_{j=i+r_N}^{N} \big(g(t_{jN}) - g(t_{iN})\big) \rho\big(N(t_{jN} - t_{iN})\big) \right| \\ &\quad \le 2MT \sum_{j=i+r_N}^{\infty} \big|\rho\big(N(t_{jN} - t_{iN})\big)\big| = \mathrm{o}(N^{1-\alpha}) \end{aligned}$$

as $N \to \infty$ uniformly with respect to j because the function g is bounded. This implies that

$$S_1 = 2N^{\alpha-2} \sum_{i=1}^{N} f(t_{iN}) g(t_{iN}) \sum_{j=i+1}^{i+r_N} \rho\bigl(N(t_{jN} - t_{iN})\bigr) + \mathrm{o}(1), \tag{3.7}$$

$$S_2 = 2N^{\alpha-2} \sum_{i=1}^{N} f(t_{iN}) \sum_{j=i+1}^{i+r_N} \bigl(g(t_{jN}) - g(t_{iN})\bigr) \rho\bigl(N(t_{jN} - t_{iN})\bigr) + \mathrm{o}(1) \tag{3.8}$$

as $N \to \infty$. For the first term on the right-hand side of (3.8), we obtain the estimate

$$\begin{aligned}
\tilde{S}_2 &= N^{\alpha-2} \left| \sum_{i=1}^{N} f(t_{iN}) \sum_{j=i+1}^{i+r_N} \bigl(g(t_{jN}) - g(t_{iN})\bigr) \rho\bigl(N(t_{jN} - t_{iN})\bigr) \right| \\
&\le 2N^{\alpha-1} M^2 T \sum_{j=i+1}^{i+r_N} \bigl|\rho\bigl(N(t_{jN} - t_{iN})\bigr)\bigr| \\
&\le 2N^{\alpha-1} M^2 T \sum_{j=i+1}^{i+r_N} \left(\bigl|\rho\bigl(a'(i_N)(j-i)\bigr)\bigr| + M^2 \frac{(j-i)^2}{N-1} \right) \\
&\le 2N^{\alpha-1} M^2 T (M r_N + M^2 r_N^3 / N) = \mathrm{o}(1)
\end{aligned}$$

as $N \to \infty$, while the dominating term on the right-hand side of (3.7) is given by

$$\begin{aligned}
\tilde{S}_1 &= N^{\alpha-2} \sum_{i=1}^{N} f(t_{iN}) g(t_{iN}) \sum_{j=i+1}^{i+r_N} \rho\bigl(N(t_{jN} - t_{iN})\bigr) \\
&= N^{\alpha-2} \sum_{i=1}^{N} f(t_{iN}) g(t_{iN}) \sum_{j=i+1}^{i+r_N} \rho\bigl(a'(i_N)(j-i)\bigr) + \mathrm{o}(1) \\
&= N^{-1} \sum_{i=1}^{N} f(t_{iN}) g(t_{iN}) Q_\alpha(a'(i_N)) + \mathrm{o}(1) = \int_0^1 f(a(u)) g(a(u)) Q_\alpha(a'(u))\, \mathrm{d}u + \mathrm{o}(1)
\end{aligned}$$

as $N \to \infty$, which proves the assertion of Lemma 2. $\square$

Theorem 1. *Let the correlation function $\rho_\alpha(t)$ be an element of the Cauchy, Mittag-Leffler or SVF family. If (3.4) and the regularity assumptions (C1)–(C3) stated in Sections 2 and 3 are satisfied, then we obtain for the variance-covariance matrix of the least-squares estimate defined in (2.4)*

$$\sigma^2 \frac{N}{d_\alpha(N)} \mathbf{D}(\tilde{\theta}) = 2\gamma W^{-1}(a) R_\alpha(a) W^{-1}(a) + \mathrm{O}\bigl(1/d_\alpha(N)\bigr),$$

where the matrices W and R_α are given by

$$W(a) = \left(\int_0^1 f_i(a(u)) f_j(a(u))\,\mathrm{d}u\right)_{i,j=1}^p,$$

$$R_\alpha(a) = \left(\int_0^1 f_i(a(u)) f_j(a(u)) Q_\alpha(a'(u))\,\mathrm{d}u\right)_{i,j=1}^p.$$

Proof. In view of (2.2), we obtain that

$$X^T \Sigma X = \left(\gamma \sum_{i \neq j} f_k(t_{iN}) f_l(t_{jN}) \rho_\alpha\bigl(N(t_{jN} - t_{iN})\bigr) + \sum_{i=1}^N f_k(t_{iN}) f_l(t_{jN})\right)_{k,l=1}^p,$$

where $X^T = (f_i(t_{jN}))_{i=1,\dots,p}^{j=1,\dots,N}$ and $t_{iN} = a((i-1)/(N-1)), i = 1, \dots, N$. An application of Lemma 2 yields

$$\frac{X^T X}{N} = W(a) + \mathrm{O}\left(\frac{1}{N}\right), \qquad \frac{X^T \Sigma X}{d_\alpha(N) N} = 2\gamma R_\alpha(a) + \mathrm{O}\left(\frac{1}{d_\alpha(N)}\right).$$

The assertion of the theorem now follows by inserting these limits into (2.4). □

Note that the constant γ only appears as a factor in the asymptotic variance-covariance matrices of the least-squares estimate. Because most optimality criteria are positively homogeneous (see, for example, Pukelsheim (1993)), it is reasonable to consider the matrix

$$W^{-1}(a) R_\alpha(a) W^{-1}(a),$$

which is proportional to the asymptotic variance-covariance matrix of the least-squares estimate. Moreover, if the function a corresponds to a continuous distribution with a density, say ϕ, then $a'(t) = 1/\phi(t)$ and the asymptotic variance-covariance matrix of the least-squares estimate is proportional to the matrix

$$\Psi_\alpha(\phi) = W^{-1}(\phi) R_\alpha(\phi) W^{-1}(\phi),$$

where the matrices $W(\phi)$ and $R_\alpha(\phi)$ are given by

$$W(\phi) = \left(\int_{-T}^T f_i(t) f_j(t) \phi(t)\,\mathrm{d}t\right)_{i,j=1,\dots,p},$$

$$R_\alpha(\phi) = \left(\int_{-T}^T f_i(t) f_j(t) Q_\alpha\bigl(1/\phi(t)\bigr)\phi(t)\,\mathrm{d}t\right)_{i,j=1,\dots,p}$$

$$= \frac{c}{1-\alpha}\left(\int_{-T}^T f_i(t) f_j(t) \phi^{1+\alpha}(t)\,\mathrm{d}t\right)_{i,j=1,\dots,p}$$

and we have used the representation $Q_\alpha(t) = c/((1-\alpha)|t|^\alpha)$ for the last identity. An (asymptotic) optimal design minimizes an appropriate function of the matrix $\Psi_\alpha(\phi)$ (for classical least-squares estimation). Note that under long-range dependence, the variance-covariance matrix of the least-squares estimate converges slower to zero than in the case of independent or short-range dependent errors. In the case of short-range dependence, no other normalization is necessary, apart from normalizing the variance-covariance matrix. Under long-range dependence, an additional factor $d_\alpha(N)/N$ is needed. Moreover, it is worthwhile to note that under long-range dependence, the asymptotic variance-covariance matrix is fully determined by the function $Q_\alpha(t)$ and does not otherwise depend on the particular correlation function $\rho_\alpha(t)$. In the following section, we discuss several examples in order to illustrate the concept.

4. Examples

In most cases, the asymptotic optimal designs for the regression model (1.1) have to be found numerically; explicit solutions are only possible for very simple models. In this section, we consider models with one or two parameters.

4.1. Optimal designs for linear models with one parameter

Consider the linear regression model with $p = 1$, that is, $y(t) = \theta f(t) + \varepsilon(t)$ ($\theta \in \mathbb{R}$). In this case, the problem of minimizing the asymptotic variance-covariance of the least-squares estimate reduces to the minimization of the function

$$\Psi_\alpha(p) = \frac{\int f^2(t) Q_\alpha(1/p(t)) p(t)\,\mathrm{d}t}{(\int f^2(t) p(t)\,\mathrm{d}t)^2} \int p(t)\,\mathrm{d}t$$

in the class of all non-negative functions $p(t)$ on the interval $[-T, T]$. Note that we have represented the density ϕ by $p/\int p(x)\,\mathrm{d}x$, which simplifies the calculation of the directional derivatives in the following discussion. Because $Q_\alpha(t)$ is strictly convex on $(0, \infty)$, it follows from Theorem 3.1 in Bickel and Herzberg (1979) that a minimizer, say $p^*(t)$, exists and that $\phi^*(t) = p^*(t)/\int p^*(t)\,\mathrm{d}t$ is the asymptotic optimal density. For the minimizing function p^*, we obtain

$$\frac{\partial}{\partial\epsilon}\Psi_\alpha\big(p^* + \epsilon(p - p^*)\big)\bigg|_{\epsilon=0} \geq 0$$

for all non-negative functions p on the interval $[-T, T]$. Consequently, the asymptotic optimal density should satisfy $\int p^*(t)\,\mathrm{d}t = 1$ and

$$\int \big(f^2(t)\big(H_\alpha\big(1/p^*(t)\big) - \mu\big) + \tilde{\tau}\big)\big(p(t) - p^*(t)\big)\,\mathrm{d}t \geq 0 \tag{4.1}$$

for all non-negative functions p on the interval $[-T, T]$, where the function $H_\alpha : (0, \infty) \to \mathbb{R}^+$ is given by

$$H_\alpha(t) = Q_\alpha(t) - t\,Q'_\alpha(t) = \begin{cases} \dfrac{1+\alpha}{1-\alpha}/t^\alpha, & 0 < \alpha < 1, \\ 2/t, & \alpha = 1, \end{cases}$$

$$\mu = 2\frac{\int f^2(t) Q_\alpha(1/p^*(t)) p^*(t)\,\mathrm{d}t}{\int f^2(t) p^*(t)\,\mathrm{d}t}, \tag{4.2}$$

$$\tilde{\tau} = \int f^2(t) Q_\alpha\bigl(1/p^*(t)\bigr) p^*(t)\,\mathrm{d}t.$$

First, assume that $0 < \alpha < 1$. Note that the function H_α is strictly decreasing with $H_\alpha(+0) = \infty$, $H_\alpha(\infty) = 0$ and that its inverse is given by

$$H_\alpha^-(t) = \left(\frac{1+\alpha}{t(1-\alpha)}\right)^{1/\alpha}.$$

Hence the solution of (4.1) has the form

$$p^*(t) = \begin{cases} \dfrac{1}{H_\alpha^{-1}\bigl(\mu - \tau/f^2(t)\bigr)} \\ \quad = \left(\dfrac{1-\alpha}{1+\alpha}\bigl(\mu - \tau/f^2(t)\bigr)\right)^{1/\alpha}, & \mu - \tau/f^2(t) \ge 0, \\ 0, & \text{otherwise}, \end{cases} \tag{4.3}$$

where μ is defined by (4.2) and

$$\tau = \int f^2(t) Q_\alpha\bigl(1/p^*(t)\bigr) p^*(t)\,\mathrm{d}t + \int f^2(t) Q'_\alpha\bigl(1/p^*(t)\bigr)\,\mathrm{d}t. \tag{4.4}$$

Note that τ is a solution of the equation $\int p(t)\,\mathrm{d}t = 1$. Indeed, multiplying $f^2(t) H_\alpha(1/p^*(t)) \equiv \mu f^2(t) - \tau$ by $p^*(t)$ and integrating with respect to t yields

$$\int f^2(t) H_\alpha\bigl(1/p^*(t)\bigr) p^*(t)\,\mathrm{d}t = \int \bigl(\mu f^2(t) - \tau\bigr) p^*(t)\,\mathrm{d}t = \mu \int f^2(t) p^*(t)\,\mathrm{d}t - \tau.$$

Now, the definition of $H_\alpha(t)$ and μ gives

$$\begin{aligned} &\int f^2(t) Q_\alpha\bigl(1/p^*(t)\bigr) p^*(t)\,\mathrm{d}t - \int f^2(t) Q'_\alpha\bigl(1/p^*(t)\bigr)\,\mathrm{d}t \\ &\quad = 2\int f^2(t) Q_\alpha\bigl(1/p^*(t)\bigr) p^*(t)\,\mathrm{d}t - \tau, \end{aligned}$$

which yields (4.4). Consequently, we have proven the following result.

Theorem 2. *Assume that the correlation function $\rho_\alpha(t)$ is an element of the Cauchy, Mittag-Leffler or SVF family. Then, for the one-parameter linear regression model, the asymptotic optimal design exists, is absolute continuous with respect to the Lebesgue measure and has the density $p^*(t)$ defined in* (4.3), *where μ and τ are given by* (4.2) *and* (4.4), *respectively.*

We now consider two special cases, which are of particular importance. If $p = 1$ and $f(t) \equiv 1$, then we obtain the location model and the asymptotic optimal density is the uniform density, that is,

$$p^*(t) = \begin{cases} \dfrac{1}{2T}, & |t| \le T, \\ 0, & \text{otherwise.} \end{cases} \tag{4.5}$$

Similarly, in the linear regression through the origin, we have $p = 1$, $f(t) \equiv t$ and the asymptotic optimal density is given by

$$p(t) = \begin{cases} 0, & |t| \le \sqrt{\tau/\mu}, \\ \left(\dfrac{1-\alpha}{1+\alpha}\left(\mu - \tau/t^2\right)\right)^{1/\alpha}, & \sqrt{\tau/\mu} \le |t| \le T, \\ 0, & \text{otherwise,} \end{cases}$$

where

$$\mu = 2\frac{\int t^2 p^{1+\alpha}(t)\,\mathrm{d}t}{(1-\alpha)\int t^2 p(t)\,\mathrm{d}t}, \qquad \tau = \int t^2 p^{1+\alpha}(t)\,\mathrm{d}t$$

and α is the parameter of the correlation function. The above formulas are given for $0 < \alpha < 1$. For $\alpha = 1$ and $f(t) = t$, the asymptotic optimal density is the uniform density (4.5). The optimal densities for the parameters $\alpha = 1/4, 1/2, 3/4, 0.95$ and $T = 1$ are displayed in Figure 2. The parameters μ and τ and the efficiency of uniform design are shown in Table 1. We observe that the uniform design is rather inefficient for small values of the parameter α. The uniform design has a reasonable efficiency only if α is close to 1.

It is worthwhile to mention that the asymptotic optimal designs derived thus far depend sensitively on the parameter α, which is usually not available before the experiment. Because misspecification of this parameter can result in a substantial loss of efficiency of the optimal design, we propose the construction of robust designs which are less sensitive with respect to such misspecifications. More precisely, we denote by $p^*_\alpha(t)$ the optimal density design for parameter α. Following Dette (1997) or Müller and Pázman (1998), a robust version of the optimality criterion is of the form

$$\Psi_{\mathcal{A}}(p) = \min_{\alpha\in\mathcal{A}} \mathrm{eff}(p,\alpha) = \min_{\alpha\in\mathcal{A}} \frac{\Psi_\alpha(p^*_\alpha)}{\Psi_\alpha(p)},$$

where p^*_α is the optimal design for the correlation function ρ_α and $\mathcal{A}$ is set of possible α values specified by the experimenter. A design maximizing $\Psi_{\mathcal{A}}$ is called *standardized maximin*

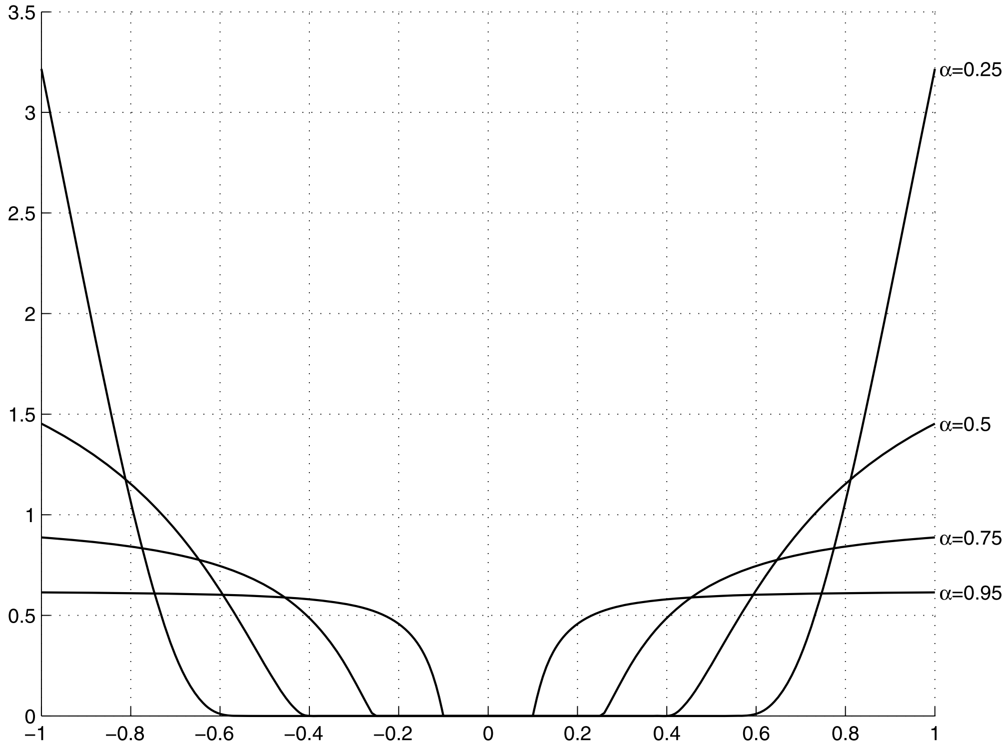

Figure 2. Asymptotic optimal design densities for the linear regression through the origin, $T = 1$.

optimal. Numerical optimization of this function for the set $\mathcal{A} = \{0.1, 0.2, \ldots, 0.9\}$ shows that standardized maximin optimal design has a density which can be approximated by the function

$$p^*_{\mathcal{A}}(t) = (5.7275t^2 - 1.16963 - 3.0264t^4)_+, \tag{4.6}$$

Table 1. Parameters of the asymptotic optimal design density for the linear regression through the origin and the efficiency of uniform design (4.5), $T = 1$ (the optimal density for $\alpha = 1$ is $p(t) = 1/2$, $-1 \le t \le 1$)

α	μ	τ	$\sqrt{\tau/\mu}$	$\mathrm{eff}_{\mathrm{uni}}$
0.05	2.34	1.06	0.67	0.40
0.25	3.19	0.96	0.55	0.59
0.50	4.32	0.70	0.40	0.78
0.75	6.84	0.44	0.25	0.93
0.95	24.78	0.25	0.10	0.99

Table 2. Efficiency of the standardized maximin optimal design $p^*_{\mathcal{A}}$ defined by (4.6) in the linear regression through the origin (the correlation structure is given by the SVF family with parameter α)

α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\mathrm{eff}(p^*, \alpha)$	0.84	0.92	0.97	0.99	0.99	0.97	0.94	0.89	0.84

which is close to the optimal density p^*_α for $\alpha = 0.44$. In Table 2, we show the efficiency of this design for various values of α. We observe that the design $p^*_{\mathcal{A}}$ is very efficient for all elements in the set $\mathcal{A}$.

4.2. Linear regression

Consider the case $p = 2$, $f_1(t) = 1$, $f_2(t) = t$, which corresponds to the linear regression model. In this case, the asymptotic variance-covariance matrix is proportional to

$$\Psi_\alpha(p) = \begin{pmatrix} 1 & \int t p(t)\,\mathrm{d}t \\ \int t p(t)\,\mathrm{d}t & \int t^2 p(t)\,\mathrm{d}t \end{pmatrix}^{-1} R(p) \begin{pmatrix} 1 & \int t p(t)\,\mathrm{d}t \\ \int t p(t)\,\mathrm{d}t & \int t^2 p(t)\,\mathrm{d}t \end{pmatrix}^{-1},$$

where

$$R(p) = \begin{pmatrix} \int Q_\alpha\bigl(1/p(t)\bigr)p(t)\,\mathrm{d}t & \int t Q_\alpha\bigl(1/p(t)\bigr)p(t)\,\mathrm{d}t \\ \int t Q_\alpha\bigl(1/p(t)\bigr)p(t)\,\mathrm{d}t & \int t^2 Q_\alpha\bigl(1/p(t)\bigr)p(t)\,\mathrm{d}t \end{pmatrix}.$$

For a symmetric density, this matrix is diagonal and

$$\Psi_\alpha(p) = \mathrm{diag}\biggl(\int Q_\alpha\bigl(1/p(t)\bigr)p(t)\,\mathrm{d}t, \frac{\int t^2 Q_\alpha(1/p(t))p(t)\,\mathrm{d}t}{(\int t^2 p(t)\,\mathrm{d}t)^2}\biggr).$$

Consequently, the optimal symmetric design for estimating the slope in the linear regression has the density (4.3), where μ and τ are defined in (4.2) and (4.4) (this follows from the fact that the element in position $(2, 2)$ of the matrix $\Psi_\alpha(p)$ corresponds to the optimality criterion for the linear regression through the origin). Numerical results indicate that the optimal symmetric design for estimating the slope is optimal in the class of all (not necessarily symmetric) designs.

The D-optimal designs for the linear regression model have to be determined numerically in all cases. Some D-optimal design densities corresponding to the parameters $\alpha = 1/4, 1/2, 3/4, 0.95$ and $T = 1$ are displayed in Figure 3.

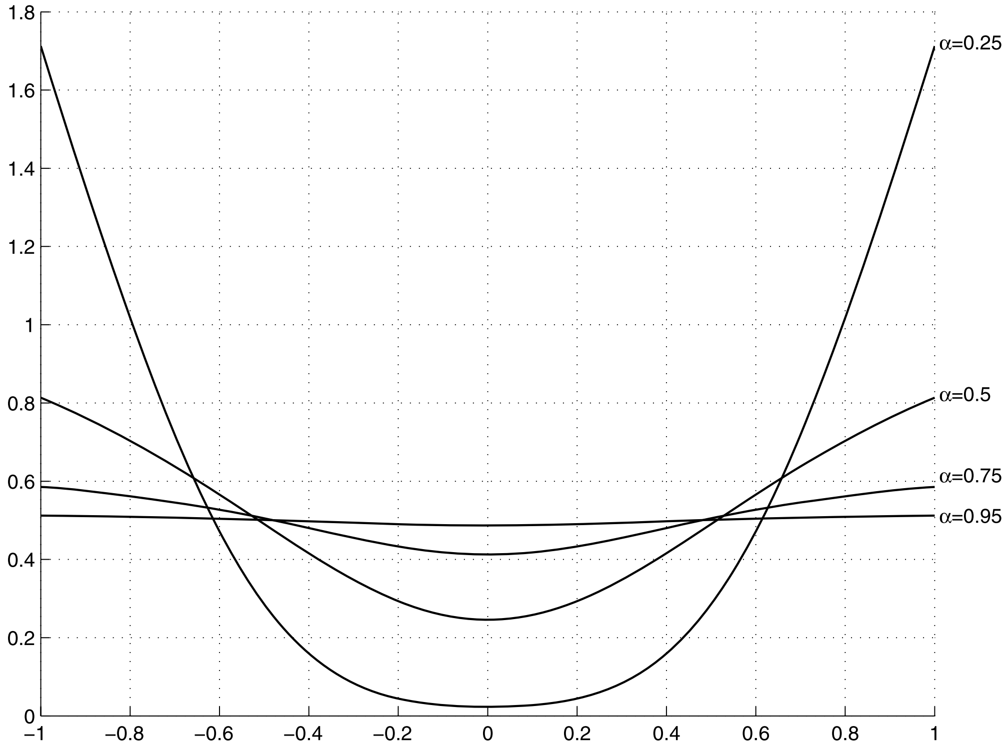

Figure 3. Asymptotic D-optimal design densities for the linear regression, $T = 1$.

4.3. Comparison of optimal designs under long- and short-range dependence

It is of some interest to compare the asymptotic optimal designs under short- and long-range dependence. For this purpose, we again consider the linear regression model with no intercept. Bickel and Herzberg (1979) discussed the correlation function $\rho_\lambda(t) = \mathrm{e}^{-\lambda|t|}$. The asymptotic optimal designs are given by

$$p(t) = \begin{cases} 0, & |t| \le \sqrt{\tau/\mu}, \\ \dfrac{1}{H^-(\mu - \tau/t^2)}, & \sqrt{\tau/\mu} \le |t| \le T, \\ 0, & \text{otherwise}, \end{cases}$$

where the quantities μ, τ are defined by

$$\mu = \frac{1}{2\gamma} + 2\frac{\int f^2(t) Q_\alpha(1/p^*(t)) p^*(t)\,\mathrm{d}t}{\int f^2(t) p^*(t)\,\mathrm{d}t},$$

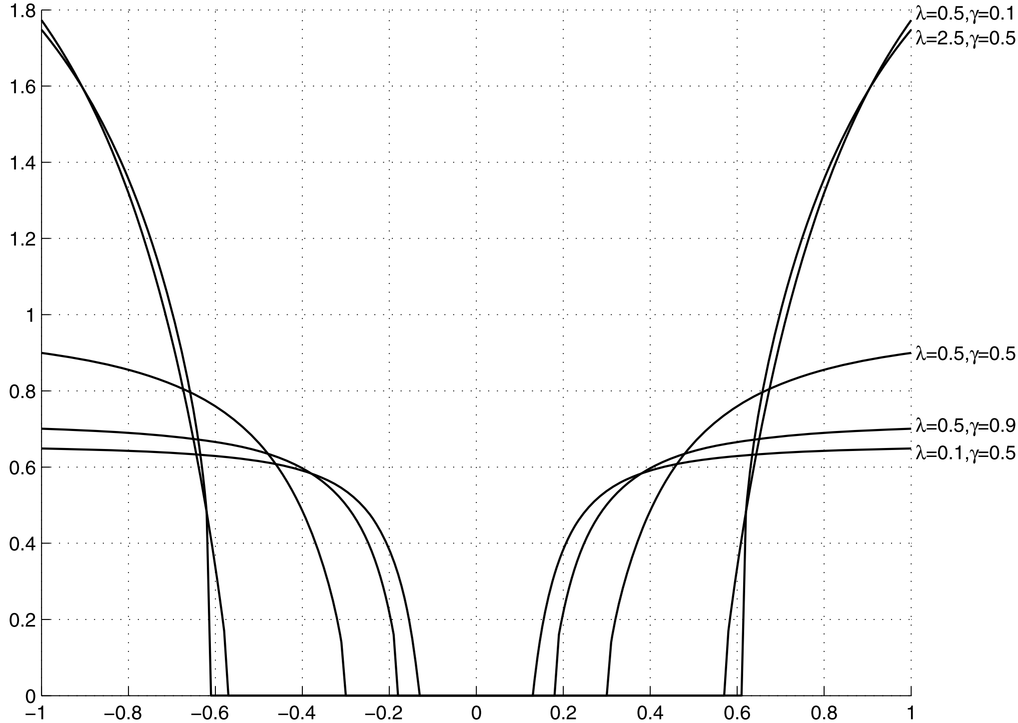

Figure 4. Asymptotic optimal design densities for the linear regression through the origin, where the correlation function is given by $\rho_\lambda(t) = \mathrm{e}^{-\lambda|t|}$.

$$\tau = \frac{1}{2\gamma}\int f^2(t)p^*(t)\,\mathrm{d}t + \int f^2(t)Q_\alpha\bigl(1/p^*(t)\bigr)p^*(t)\,\mathrm{d}t + \int f^2(t)Q'_\alpha\bigl(1/p^*(t)\bigr)\,\mathrm{d}t,$$

respectively (see Bickel, Herzberg and Schilling (1981)) and depend on the parameters λ and γ defined in (2.2). Some of these designs are shown in Figure 4, while the relevant parameters are given in Table 3, which also contains the efficiency of the uniform design. We observe that – in

Table 3. Parameters of the asymptotic optimal design density for the linear regression through the origin, where the correlation function is given by $\rho_\lambda(t) = \mathrm{e}^{-\lambda|t|}$ (the last column of the table shows the efficiency of the uniform design (4.5))

λ	γ	μ	τ	$\sqrt{\tau/\mu}$	$\mathrm{eff}_{\mathrm{uni}}$
0.5	0.5	3.41	0.32	0.30	0.89
0.5	0.1	9.82	3.23	0.57	0.63
0.5	0.9	2.38	0.08	0.18	0.97
0.1	0.5	12.70	0.22	0.13	0.99
2.5	0.5	1.45	0.54	0.61	0.57

Table 4. Efficiency of the asymptotic optimal design density for the correlation function $\rho_\lambda(t) = \mathrm{e}^{-\lambda|t|}$ in the linear regression through the origin, while the "true" correlation function belongs to the SVF family

	α	0.05	0.25	0.50	0.75	0.95
$\lambda = 0.5$	$\gamma = 0.5$	0.62	0.82	0.96	1.00	0.97
$\lambda = 0.5$	$\gamma = 0.1$	0.81	0.97	0.99	0.89	0.77
$\lambda = 0.5$	$\gamma = 0.9$	0.53	0.73	0.90	0.99	1.00
$\lambda = 0.1$	$\gamma = 0.5$	0.50	0.70	0.88	0.98	1.00
$\lambda = 2.5$	$\gamma = 0.5$	0.81	0.97	0.98	0.89	0.77

contrast to the case of long-range dependence – the uniform design is rather efficient, provided either that the parameter λ is not too large or that γ is not too small.

We now compare asymptotic optimal designs derived under the assumption of a long-range dependence with asymptotic optimal designs under short-range dependence. In Table 4, we show the efficiency of a design derived under the assumption of short-range dependence, in the situation where the "true" correlation structure is a member of the SVF family. We observe that the loss of efficiency is only substantial if the parameter α is small. The opposite situation is displayed in Table 5, which shows the efficiency of the asymptotic optimal design under long-range dependence (from the SVF family), but the "true" correlation structure is in fact of exponential type. Again, the asymptotic optimal designs derived under the long-range dependence are rather efficient, except when the parameter α is very small.

Acknowledgements

The support of the Deutsche Forschungsgemeinschaft (SFB 475, "Komplexitätsreduktion in multivariaten Datenstrukturen" SFB 823 "Statistik nichtlinearer dynamischer Prozesse") is gratefully acknowledged. The work of H. Dette was also supported in part by an NIH Grant award IR01GM072876:01A1 and by the BMBF project SKAVOE. The second author, N. Leonenko, was supported in part by the EPSRC Grant RCMT 199, the Australian Research Grant DP

Table 5. Efficiency of asymptotic optimal design density for a long-range dependence error structure in the linear regression through the origin, while the "true" correlation function is given by $\rho_\lambda(t) = \mathrm{e}^{-\lambda|t|}$

λ	0.5	0.5	0.5	0.1	2.5
γ	0.5	0.1	0.9	0.5	0.5
$\alpha = 0.05$	0.19	0.40	0.15	0.15	0.35
$\alpha = 0.25$	0.69	0.94	0.59	0.58	0.93
$\alpha = 0.50$	0.94	0.98	0.87	0.86	0.98
$\alpha = 0.75$	1.00	0.88	0.98	0.98	0.85
$\alpha = 0.95$	0.95	0.73	0.99	1.00	0.68

0559807 and the Welsh Institute of Mathematics and Computational Sciences. The authors would also like to thank two referees for their constructive comments on an earlier version of this paper.

References

Adenstedt, R.K. (1974). On large-sample estimation of the mean of a stationary random sequence. *Ann. Statist.* **2** 1095–1107. MR0368354

Anh, V.V., Knopova, V.P. and Leonenko, N.N. (2004). Continuous-time stochastic processes with cyclical long-range dependence. *Aust. N. Z. J. Stat.* **46** 275–296. MR2076396

Atkinson, A.C. and Donev, A.N. (1992). *Optimum Experimental Designs*. Oxford: Clarendon Press.

Barndorff-Nielsen, O.E., Jensen, J.L. and Sorensen, M. (1990). Parametric modelling of turbulence. *Philosophical Transactions: Physical Sciences and Engineering* **332** 439–455.

Barndorff-Nielsen, O.E. and Leonenko, N.N. (2005). Spectral properties of superpositions of Ornstein–Uhlenbeck type processes. *Methodol. Comput. Appl. Probab.* **7** 335–352. MR2210585

Beran, J. (1992). Statistical methods for data with long-range dependence. *Statist. Sci.* **7** 404–427.

Beran, J. (1994). *Statistics for Long-Memory Processes*. New York: Chapman and Hall. MR1304490

Beran, J., Sherman, R., Taqqu, M.S. and Willinger, W. (1992). Long-range dependence in variable bit-rate video traffic. *IEEE Trans. Commun.* **43** 1566–1579.

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. MR0515685

Bickel, P.J., Herzberg, A.M. and Schilling, M.F. (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. MR0650898

Dette, H. (1997). Designing of experiments with respect to "standardized" optimality criteria. *J. Roy. Statist. Soc. Ser. B* **59** 97–110. MR1436556

Djrbashian, M.M. (1993). *Harmonic Analysis and Boundary Value Problems in the Complex Domain*. Boston: Birkhauser. MR1249271

Doukhan, P., Oppenheim, G. and Taqqu, M. (eds.) (2003). *Theory and Applications of Long-Range Dependence*. Boston: Birkhauser. MR1956041

Fedorov, V.V. (1972). *Theory of Optimal Experiments*. New York: Academic Press. MR0403103

Gneiting, T. (2000). Power-law correlations, related models for long-range dependence and their simulation. *J. Appl. Probab.* **37** 1104–1109. MR1808873

Granger, C.W.J. (1980). Long memory relationships and the aggregation of dynamic models. *J. Econometrics* **14** 227–238. MR0597259

Ivanov, A.V. and Leonenko, N.N. (2004). Asymptotic theory for non-linear regression with long-range dependence. *Math. Methods Statist.* **13** 153–178. MR2090470

Ivanov, A.V. and Leonenko, N.N. (2008). Semiparametric analysis of long-range dependence in nonlinear regression. *J. Statist. Plann. Inference* **138** 1733–1753. MR2400476

Künsch, H., Beran, J. and Hampel, F. (1993). Contrasts under long-range correlations. *Ann. Statist.* **21** 943–964. MR1232527

Mandelbrot, B.B. (1973). Le problème de la rétalité des cycle lents et le syndrome de Joseph. *Economie Appliquéee* **26** 349–365.

Metzler, R., Barkai, E. and Klafter, J. (1999). Anomalous diffusion and relaxation close to thermal equilibrium: A fractional Fokker–Planck equation approach. *Phys. Rev. Lett.* **82** 3563–3567.

Metzler, R. and Klafter, J. (2000). The random walk's guide to anomalous diffusion: A fractional dynamics approach. *Phys. Rep.* **339** 1–77. MR1809268

Pázman A. (1986). *Foundations of Optimum Experimental Design*. Dordrecht: Reidel.

Pázman A. and Müller, C. (1998). Applications of necessary and sufficient conditions for maximum efficient design. *Metrika* **48** 1–19. MR1647889

Pearson, K. (1902). On the mathematical theory of errors of judgement, with special reference to the personal equation. *Philos. Trans. Roy. Soc. Ser. A* **198** 235–299.

Porter-Hudak, S. (1990). An application of the seasonal fractionally differenced model to the monetary aggregates. *J. Amer. Statist. Assoc.* **85** 338–344.

Pukelsheim, F. (1993). *Optimal Design of Experiments*. New York: Wiley. MR1211416

Sacks, J. and Ylvisaker, N.D. (1966). Designs for regression problems with correlated errors. *Ann. Math. Statist.* **37** 66–89. MR0192601

Sacks, J. and Ylvisaker, N.D. (1968). Designs for regression problems with correlated errors; many parameters. *Ann. Math. Statist.* **39** 49–69. MR0220424

Samarov, A. and Taqqu, M.S. (1988). On the efficiency of the sample mean in long memory noise. *J. Time Ser. Anal.* **9** 191–200. MR0943006

Schneider, W.R. (1996). Completely monotone generalized Mittag-Leffler functions. *Exposition. Math.* **14** 3–16. MR1382012

Seneta, E. (1976). *Regularly Varying Functions*. *Lecture Notes in Math.* **508**. Berlin: Springer. MR0453936

Silvey, S.D. (1980). *Optimal Design*. London: Chapman & Hall. MR0606742

Smith, H.F. (1938). An empirical law describing heterogeneity in the yields of agricultural crops. *J. Agric. Sci.* **28** 1–23.

Yajima, Y. (1988). On estimation of a regression model with long-memory stationary errors. *Ann. Statist.* **16** 791–807. MR0947579

Yajima, Y. (1991). Asymptotic properties of the LSE in a regression model with long-memory stationary errors. *Ann. Statist.* **19** 158–177. MR1091844

Received September 2008 and revised July 2009