Optimal designs for correlated observations
&
characterization of probability distributions

Anatoly Zhigljavsky (Cardiff)

Collaborators:

Andrey Pepelyshev, Luc Pronzato, Karl Michael Schmidt, Holger Dette, Nikolai Leonenko

Winter Workshop on Mathematical Statistics, Bratislava December 15-16, 2008

Optimal designs for correlated observations:
Convexity of the optimality criterion, Optimality theorem, Behaviour of optimal designs

Bickel-Herzberg approach, Long-range dependent error process

Alternative asymptotics

Kernels with singularity at 0:
Convexity of the optimality criterion, optimality theorem, examples

Arcsine density: motivation

Optimal designs for correlated observations, Näther(1985)

Assume the model $y(x_j) = \theta + \varepsilon(x_j)$ $(i = 1, \ldots, N)$
with $\mathrm{cov}(x_i, x_j) = \sigma^2 \rho(x_i - x_j)$; $x_i, x_j \in X$.
If we estimate θ by the unweighted average of $y(x_j)$ then the variance of the estimator is

$$D(\xi) = \sigma^2 \int \int \rho(x - y)\xi(dx)\xi(dy)$$

where ξ is an approximate design.
More generally, for the model

$$y(x_j) = \theta^T f(x_j) + \varepsilon(x_j)$$

the covariance matrix of the unweighted LSE estimator $(X^T X)^{-1} X^T Y$ is

$$\sigma^2 (X^T X)^{-1} (X^T W X)(X^T X)^{-1} = \sigma^2 M^{-1}(\xi) \left[\int \int f(x) f^T(x) \rho(x - y)\xi(dx)\xi(dy) \right] M^{-1}(\xi),$$

where

$$M(\xi) = \int f(x) f^T(x) \xi(dx)$$

Convexity of the optimality criterion, Näther(1985)

Optimality criterion: $D(\xi) = \int\int \rho(x-y)\xi(dx)\xi(dy)$.
Consider $\xi_\alpha = (1-\alpha)\xi_0 + \alpha\xi_1$. Then

$$D(\xi_\alpha) = (1-\alpha)D(\xi_0) + \alpha D(\xi_1) - \alpha(1-\alpha)A$$

where

$$\begin{aligned} A &= \int\int \rho(x-y)[\xi_0(dx)\xi_0(dy) + \xi_1(dx)\xi_1(dy) - 2\xi_0(dx)\xi_1(dy)] \\ &= \int\int \rho(x-y)\eta(dx)\eta(dy) \quad \text{with} \;\; \eta(dx) = \xi_0(dx) - \xi_1(dx) \end{aligned}$$

As the function $K(x,y) = \rho(x-y)$ is positive definite (Bohner-Khintchine theorem), we have $A > 0$.
Therefore, the criterion $D(\xi)$ is strictly convex.

Optimality (equivalence) theorem, Näther(1985)

Consider $D(\xi) = \int \int \rho(x-y)\xi(dx)\xi(dy)$.

Set $\phi(\xi, x) = \int \rho(x-y)\xi(dy)$.

Lemma.

$$\frac{\partial D(\xi_\alpha)}{\partial \alpha}|_{\alpha=0} = 2\left[\int \phi(\xi_1, x)\xi_0(dx) - D(\xi_0)\right]$$

Theorem. ξ^* is optimal iff

$$\min_x \phi(\xi^*, x) = D(\xi^*)$$

or

$$\phi(\xi^*, x) = \text{const}, \quad \text{for } \xi^* - \text{almost all } x\,.$$

Behaviour of optimal designs: Andrey Pepelyshev & AZ

Let $D(\xi) = \int \int \rho(x-y)\xi(dx)\xi(dy)$, $X = [-1,1]$, $\rho(t) = \max\{0, 1-\lambda|t|\}$.

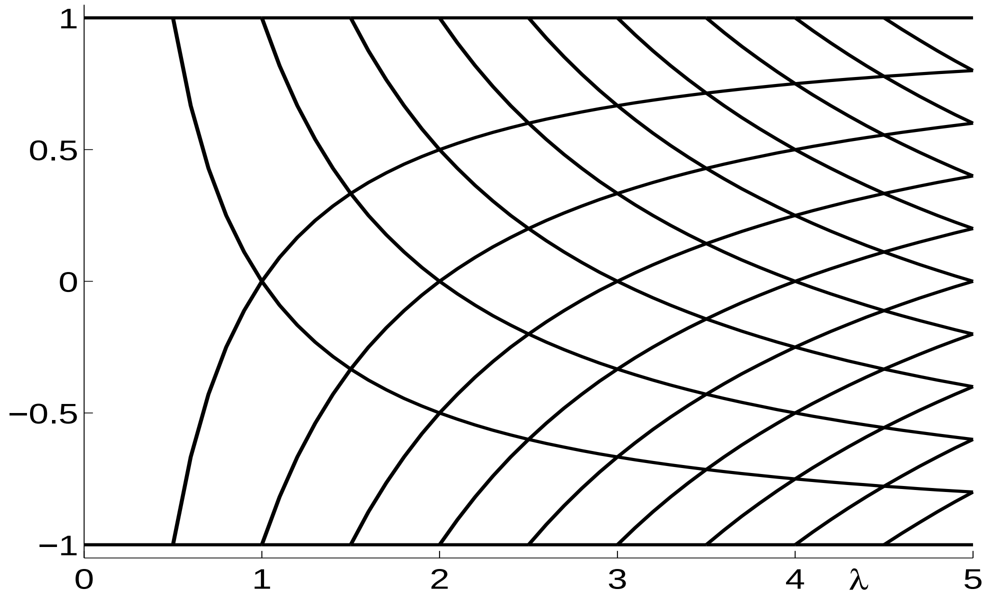

Figure 1: *Support points of the optimal designs for estimating mean.*

Approach of Bickel-Herzberg (1977, 1979)

The correlation function changes with N: $\rho_N(t) = \rho(Nt)$.
Alternatively we can fix $\rho(t)$ but expand the design interval proportionally to the number of observation points N. Asymptotically (as $N \to \infty$) the covariance matrix is proportional to

$$D(\xi) = M^{-1}(\xi)R(\xi)M^{-1}(\xi)$$

where $\xi(dt) = p(t)dt$,

$$R(\xi) = \frac{1}{1-\alpha}\left(\int f(t)f^T(t)Q(1/p(t))p(t)\,dt\right), \quad Q(u) = \sum_{j=1}^{\infty} \rho(ju)\,.$$

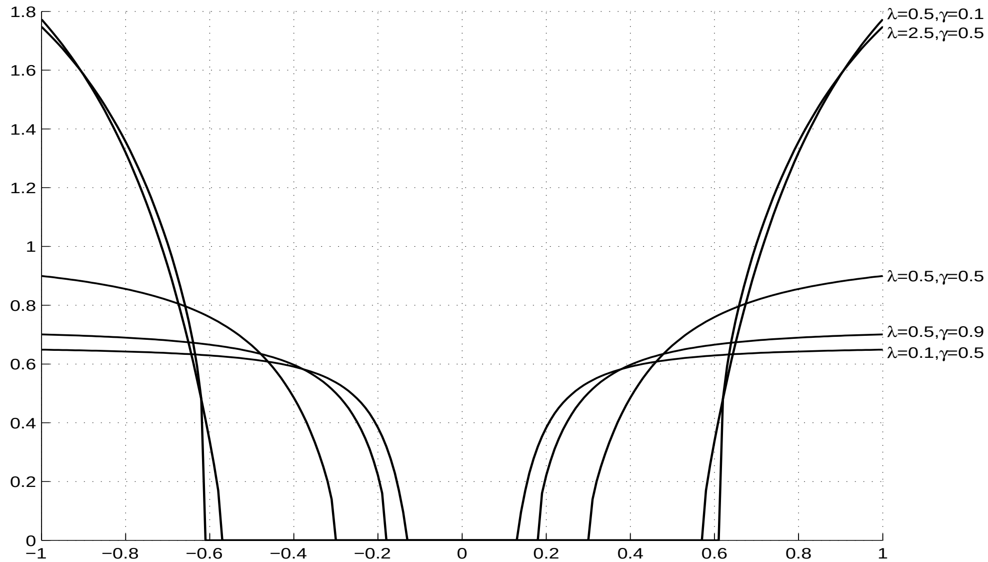

Figure 2: *Asymptotic optimal densities for the linear regression through the origin;* $\rho_\lambda(t) = e^{-\lambda|t|}$.

Long-range dependent error process (Dette, Leonenko, Pepelyshev & AZ)

Correlation functions ($0 < \alpha < 1$):

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

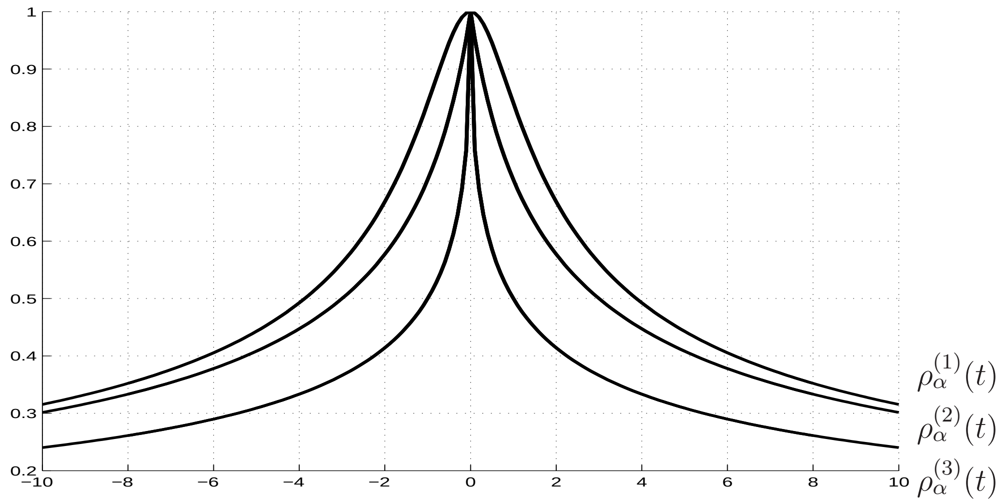

Figure 3: *The three correlation functions,* $\alpha = 0.5$.

Then asymptotically (as $N \to \infty$) the covariance matrix is proportional to $D_\alpha(\xi) = M^{-1}(\xi) R_\alpha(\xi) M^{-1}(\xi)$,
where $R_\alpha(\xi) = \frac{1}{1-\alpha}\left(\int f(t) f^T(t) p^{1+\alpha}(t)\, dt\right)$, $\xi(dt) = p(t)dt$,

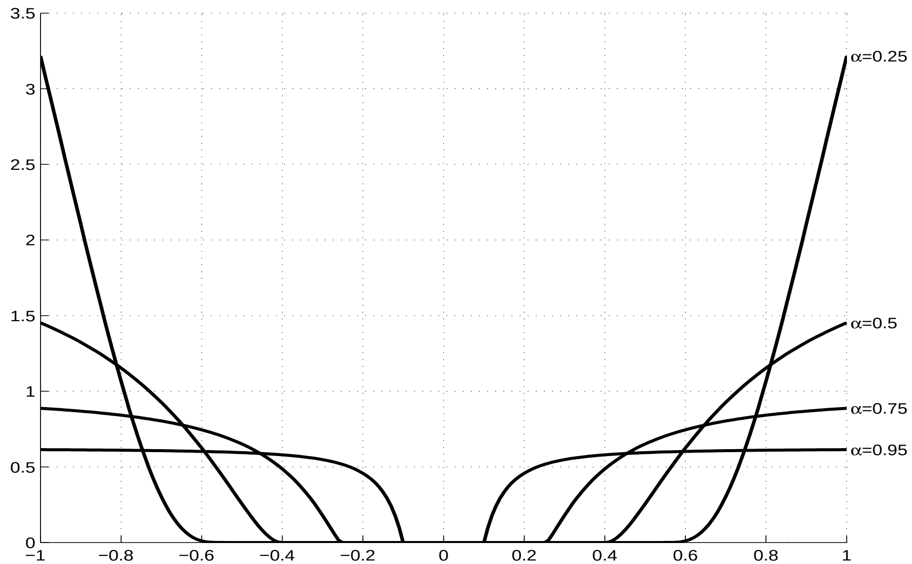

Figure 4: *Asymptotic optimal design densities for the linear regression through the origin.*

Alternative asymptotics (Andrey Pepelyshev & AZ)

Consider $D_N(\xi) = \sigma_N^2 \int \int \rho_N(x-y)\xi(dx)\xi(dy)$.
As in the Bickel-Herzberg approach we assume

$$\rho_N(t) = \rho(Nt), \text{ where, for example, } \rho(t) = \frac{1}{(1+|t|)^\alpha} .$$

In addition, we assume that the variance depends on N as well: $\sigma_N^2 = N^\alpha \sigma^2$.
Then the covariance function is

$$\sigma_N^2 \rho_N(t) = \sigma^2 N^\alpha \frac{1}{(1+|Nt|)^\alpha} = \sigma^2 \frac{1}{(1/N+|t|)^\alpha}$$

As $N \to \infty$, the sequence of optimality criteria $D_N(\xi)$ tends to

$$D(\xi) = \sigma^2 \int \int r(x-y)\xi(dx)\xi(dy)$$

where $r(t)$ has singularity at 0; in this case $r(t) = 1/|t|^\alpha$.
The corresponding sequence of optimal designs converges too.

Kernels with singularity at 0: Convexity & optimality theorem

Consider the optimality criterion

$$D(\xi) = \int\int r(x-y)\xi(dx)\xi(dy)$$

where $r(\cdot)$ has singularity at 0 and it is a Fourier transform of a positive function p: $r(u) = \int e^{-itu}p(t)dt$. Then $D(\xi)$ is a strictly convex criterion:

$$\int\int r(x-y)\eta(dx)\eta(dy) = \int\int\int e^{-it(x-y)}\eta(dx)\eta(dy)p(t)dt = \int \left|\int e^{-it(x)}\eta(dx)\right|^2 p(t)dt > 0.$$

ξ^* is optimal iff

$$\int r(x-y)\xi^*(dy) = \text{const}, \quad \text{for } \xi^* - \text{almost all } x\,.$$

Kernels with singularity at 0: Examples

$$D(\xi) = \int\int r(x-y)\xi(dx)\xi(dy)\,, \quad X = [0,1]\,.$$

Beta-distribution:

$$r(u) = \frac{1}{|u|^\alpha}; \;\Rightarrow\;\; p^*(x) = \frac{1}{B(\beta,\beta)} x^{\beta-1}(1-x)^{\beta-1}, \;\; \beta = \frac{1+\alpha}{2}$$

$(0 < \alpha < 1)$. As $\alpha \to 0$ we obtain

arcsine density:

$$r(u) = -\log(u^2) \;\Rightarrow\;\; p^*(x) = \frac{1}{\pi\sqrt{x(1-x)}}$$

Arcsine density: motivation (Luc Pronzato & AZ)

Assume we have a sequence of points $x_1, x_2, \ldots$ in the interval $[0,1]$ with asymptotic c.d.f. $F(\cdot)$: $\lim_{k\to\infty} \frac{1}{k}\sum_{j=1}^{k} h(x_j) = \int h(t)\, dF(t)$ for any continuous function $h(\cdot)$ with $\int |h(x)|\, dF(x) < \infty$. Associated sequence of polynomials: $H_k(x) = (x-x_1)^2\,(x-x_2)^2 \cdots (x-x_k)^2$. Then the normalized ratios $R_k(x,y) = [H_k(x)/H_k(y)]^{1/k}$ tend to 1 (as $k \to \infty$) for a. a. $x, y \in [0,1]$ iff the c.d.f. $F(\cdot)$ has the arcsine density. Indeed,

$$\log R_k(x,y) = \log[H_k(x)]^{1/k} - \log[H_k(y)]^{1/k} = \frac{1}{k}\sum_{j=1}^{k} \log(x-x_j)^2 - \frac{1}{k}\sum_{j=1}^{k} \log(y-x_j)^2$$

$$\Rightarrow \log\left[\lim_{k\to\infty} R_k(x,y)\right] = \lim_{k\to\infty} \log R_k(x,y) = \int \log(x-t)^2\, dF(t) - \int \log(y-t)^2\, dF(t)$$

which is zero for a.a. $x, y \in [0,1]$ iff the c.d.f. $F(\cdot)$ has the arcsine density.