# Optimal design for multivariate models with correlated observations

Andrey Pepelyshev

MODA10
June 10, 2013

# The multivariate model

Consider the common linear regression model

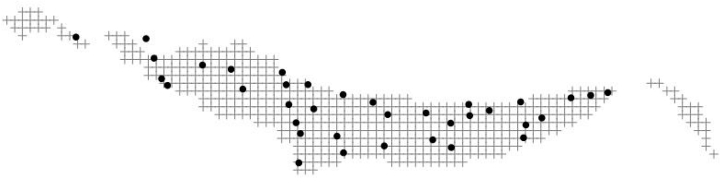
$$y(x) = \theta_1 f_1(x) + \ldots + \theta_m f_m(x) + \varepsilon(x) \ , \ \ x \in \mathbb{X} \subset \mathbb{R}^d$$

- functions $f_1(x), \ldots, f_m(x)$ are linearly independent and continuous,
- a random error field $\varepsilon(x)$ has the zero mean and the covariance kernel $K(x, x') = \mathrm{E}[\varepsilon(x)\varepsilon(x')]$,
- parameters $\theta_1, \ldots, \theta_m$ are unknown and have to be estimated.

No replication of experimental points.

# Example: the monitoring network

The Südliche Tullnerfeld in Lower Austria and the monitoring network, (Muller, 2000, 2005; Muller, Pazman, 1999).



Consider the linear model with $f(x) = (1, x_{[1]}, x_{[2]})^T$, the isotropic spherical correlation function

$$K_1(x, x'; \theta) = \begin{cases} \theta_1 + \theta_2 & \|x - x'\| = 0 \\ \theta_2(1 - 1.5\|x - x'\|/\theta_3 + 0.5\|x - x'\|^3/\theta_3^3) & 0 < \|x - x'\| < \theta_3 \\ 0 & \|x - x'\| > \theta_3 \end{cases},$$

or the Gaussian correlation function

$$K_2(x, x'; \theta) = \theta_1\delta_0(\|x - x'\|) + \theta_2 \exp(-\|x - x'\|^2/\theta_3^2).$$

# Design and covariance matrices

| | Exact design | Continuous design |
|---|---|---|
| BLUE | $(\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}$ | $-$ |
| OLS | $(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{\Sigma}\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}$ | $M^{-1}(\xi)B(\xi,\xi)M^{-1}(\xi)$ |

$$\begin{aligned}
\mathbf{X} &= (f_i(x_j))_{j=1,\dots,N}^{i=1,\dots,m} \\
\mathbf{\Sigma} &= (K(x_i,x_j))_{i,j=1,\dots,N} \\
M(\xi) &= \int_{\mathbb{X}} f(u)f^T(u)\xi(\mathrm{d}u), \\
B(\xi,\xi) &= \int_{\mathbb{X}}\int_{\mathbb{X}} K(u,v)f(u)f^T(v)\xi(\mathrm{d}u)\xi(\mathrm{d}v).
\end{aligned}$$

# Specifics of BLUE and WLS

Let the design points $\xi_N = \{x_{1;N}, \ldots, x_{N;N}\}$ be generated by the quantile function $Q(\cdot)$ of a continuous measure $\xi$,

$$x_{i;N} = Q\left((i-1)/(N-1)\right), \; i = 1, \ldots, N.$$

We have

$$\lim_{N\to\infty} \mathrm{Cov}(\hat{\theta}_{BLUE}|\xi_N) = \lim_{N\to\infty} \mathrm{Cov}(\hat{\theta}_{BLUE}|\zeta_N)$$

and

$$\lim_{N\to\infty} \mathrm{Cov}(\hat{\theta}_{WLS}|\xi_N) \neq \lim_{N\to\infty} \mathrm{Cov}(\hat{\theta}_{WLS}|\zeta_N).$$

# The necessary condition of $D$-optimality

## Theorem (Dette, Pepelyshev, Zhigljavsky, 2012)

*Let $\xi^*$ be any $D$-optimal design. Then for all $x \in \mathbb{X}$ we have*

$$d(x,\xi^*) \leq b(x,\xi^*), \qquad (*)$$

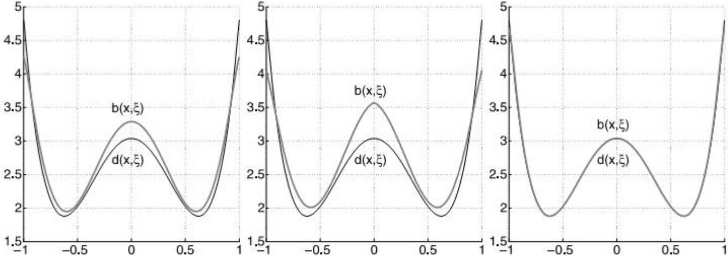
*where*

$$d(x,\xi) = f^T(x)M^{-1}(\xi)f(x)$$

*and*

$$b(x,\xi) = tr(B^{-1}(\xi,\xi)B(\xi,\xi_x)) = f^T(x)B^{-1}(\xi,\xi)\int K(u,x)f(u)\xi(\mathrm{d}u),$$

*$\xi_x = \{x,1\}$ is the Dirac measure. Moreover, there is equality in $(*)$ for $\xi^*$-almost all $x$.*




The functions $b(x,\xi)$ and $d(x,\xi)$ for the regression model with $f(x) = (1, x, x^2)^T$, $x \in [-1, 1]$ and the covariance kernels $K(u, v) = e^{-|u-v|}$ (left), $K(u, v) = \max(0, 1 - |u - v|)$ (middle) and $K(u, v) = -\log(u - v)^2$ (right), and the arcsine design $\xi$.

# The multiplicative algorithm

Let $\xi^{(r)} = \{x_1, \dots, x_n; w_1^{(r)}, \dots, w_n^{(r)}\}$ denote the design at the iteration $r$.

The updating rule for the weights has the form

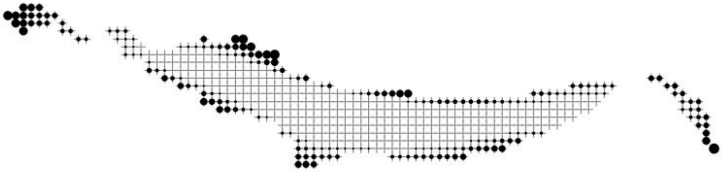
$$w_i^{(r+1)} = \frac{w_i^{(r)}\left(\psi(x_i, \xi^{(r)}) - \beta_r\right)}{\sum_{j=1}^n w_j^{(r)}\left(\psi(x_j, \xi^{(r)}) - \beta_r\right)}, \qquad i = 1, \dots, n,$$

where $\beta_r$ is a tuning parameter, $0 \le \beta_r < \min_i \psi(x_i, \xi^{(r)})$,

$$\psi(x, \xi) = d(x, \xi)/b(x, \xi).$$

In most of cases, 20 iterations are sufficient for finding a good approximation of the optimal design.

# Examples of $D$-optimal designs



The design $\xi^{(20)}$ for the spherical correlation function with $\theta = (4.89, 1.86, 0.81)$.

The design $\xi^{(20)}$ for the Gaussian correlation function with $\theta = (4.89, 1.86, 0.5)$.

# Examples of $D$-optimal designs

The design $\xi^{(20)}$ for the spherical correlation function with $\theta = (4.89, 1.86, 0.4)$.

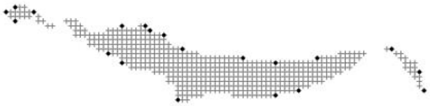
The design $\xi^{(20)}$ for the spherical correlation function with $\theta = (4.89, 1.86, 1.6)$.

# Generation of exact designs

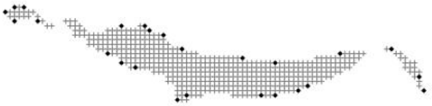
For given $N$, we choose a discretization $x_1, \ldots, x_n$ such that $n$ is much larger than $N$. Let $\xi = \{x_1, \ldots, x_n; w_{1(1)}, \ldots, w_{n(1)}\}$ be the asymptotic optimal design calculated by the multiplicative algorithm.

- Find the index $\tau$ such that $w_{\tau(k)}$ is maximal, i.e. $w_{\tau(k)} = \max_j w_{j(k)}$.
- Set $x'_k = x_\tau + \epsilon$, where $\epsilon$ is a random variable with zero mean and very small variance.
- Define $w_{\tau(k+1)} = w_{\tau(k)} - 1/N$ and $w_{j(k+1)} = w_{j(k)}$ for all $j \neq \tau$.
- If $w_{\tau(k+1)} < 0$, then set $w_{\tau(k+1)} = 0$ and the weights for points in a neighborhood of $x_\tau$ should be decreased such that $\sum_{j=1}^{n} w_{j(k+1)} = (N-k)/N$. In other words, we should take points that are nearest to $x_\tau$ and decrease weights at them.

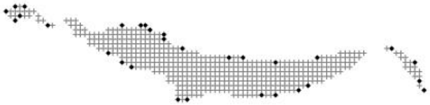
# Exact designs



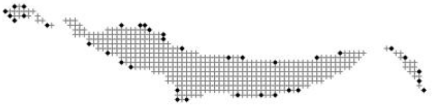
$N = 20$



$N = 25$



$N = 30$



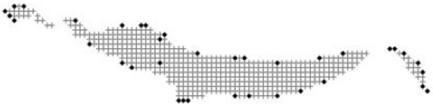
$N = 36$

The exact designs $\xi_{20}, \xi_{25}, \xi_{30}$ and $\xi_{36}$ generated by the design $\xi^{(20)}$ for the linear model and the spherical correlation function with $\theta = (4.89, 1.86, 0.81)$; $\quad \Psi(\xi_N) = \sqrt[3]{(\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}}$.

$$\Psi(\xi_{20}) = 5.39, \qquad \Psi(\xi_{25}) = 4.56,$$

$$\Psi(\xi_{30}) = 4.15, \qquad \Psi(\xi_{36}) = 3.67.$$

# Fully optimized exact designs



The exact designs $\xi^*_{D,36,r}$ (left) and $\xi^*_{D,36,a}$ (right) computed by the Brimkulov algorithm (Muller, 2005) when the initial design is *random* and one generated from the *asymptotic* optimal density, respectively, for the linear model and the spherical correlation function with $\theta = (4.89, 1.86, 0.81)$.

$$\Psi(\xi^*_{D,36,r}) = 3.55, \qquad \Psi(\xi^*_{D,36,a}) = 3.39,$$

For the existing monitoring network, $\Psi(\xi_{m.n.}) = 6.93$.

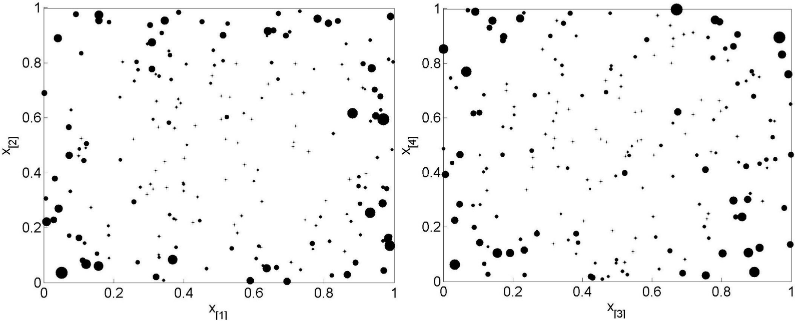
# Computer experiments

- Let we intend to perform $N$ evaluation of a computer model.
- Discretize the domain with $n$ points, $n >> N$.
- Run the multiplicative algorithm to find the asymptotic optimal design $\xi$.
- Compute $N$-point exact design using $\xi$.

# Example of design in $[0,1]^4$

Consider the model with $f(x) = (1, x_{[1]}, x_{[2]}, x_{[3]}, x_{[4]})^T$ and the Gaussian correlation function $K(x,x') = \exp(-\|x - x'\|^2)$.

Generate 200 random uniform points in $[0,1]^4$.




Conclusion: Uniform design is not $D$-optimal.

Thank you for your attention.