Efficient Designs for Computer Experiments

Andrey Pepelyshev

MODA 9
June 15, 2010

Computer experiments

Data

$\eta(x)$ is a computer model with deterministic output,
$y_i = \eta(x_i)$ is an output for the i-th run of the model.

Computer experiments

Data

$\eta(x)$ is a computer model with deterministic output,
$y_i = \eta(x_i)$ is an output for the i-th run of the model.

Assumption for obtaining a statistical model

$(y_1, \dots, y_n)$ is a realization of a Gaussian process
with mean θ, variance σ^2 and correlation function $\rho(x)$.

Computer experiments

Data

$\eta(x)$ is a computer model with deterministic output,
$y_i = \eta(x_i)$ is an output for the i-th run of the model.

Assumption for obtaining a statistical model

$(y_1, \ldots, y_n)$ is a realization of a Gaussian process with mean θ, variance σ^2 and correlation function $\rho(x)$.

Classification of studies

	small $\sqrt[d]{n}$	large $\sqrt[d]{n}$
computer model / emulation, ...	of interest	out of interest
Gaussian process / par. estimation	of interest	infill asympt.

Outline

- Design for correlated observations
- Design of computer experiments

Theory for correlated observations

Consider the model

$$y_j = y_j(t_j) = \theta_1 f_1(t) + \ldots + \theta_k f_k(t) + \varepsilon_j$$

where $t_j \in [-T, T]$, $j = 1, \ldots, N$ and $\mathbf{E}\varepsilon_j\varepsilon_i = \sigma^2\rho(t_j - t_i)$.

$$\hat{\theta}_{\text{OLS}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T Y$$

$$\text{Var}\left(\hat{\theta}_{\text{OLS}}\right) = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T R X (\mathbf{X}^T\mathbf{X})^{-1} \to \min_{t_1,\ldots,t_N}$$

Asymptotic settings

Let the design points $\{t_1, \ldots, t_N\}$ be generated by the quantiles of a distribution function,

$$t_{iN} = a\left((i-1)/(N-1)\right),\ i = 1, \ldots, N,$$

where the function $a : [0,1] \to [-T, T]$ is the inverse of a distribution function.

Let ξ be a design measure corresponding to $a(\cdot)$.

Then for OLS the design problem is

$$D(\xi) = W^{-1}(\xi)R(\xi)W^{-1}(\xi) \to \min_{\xi}$$

$$W(\xi) = \int f(u)f^T(u)\xi(du),\ R(\xi) = \int\int \rho(u-v)f(u)f^T(v)\xi(du)\xi(dv)$$

Design for correlated observations

Consider the location model

$$y_j = y_j(t_j) = \theta + \varepsilon_j$$

where $t_j \in [-1, 1]$, $\mathbf{E}\varepsilon_j = 0$ and $\mathbf{E}\varepsilon_j\varepsilon_i = \sigma^2\rho(t_j - t_i)$.

The design problem is

$$\mathrm{Var}\left(\hat{\theta}_{\mathrm{OLS}}\right) = \mathbf{1}^T R \mathbf{1} \to \min_{t_1,\dots,t_N}$$

where $\mathbf{1} = (1, \dots, 1)^T$ and $R = (\rho(t_j - t_i))_{i,j}$.

Design for correlated observations

$$D(\xi) = \int\int \rho(u-v)\xi(du)\xi(dv) \to \min_{\xi}$$

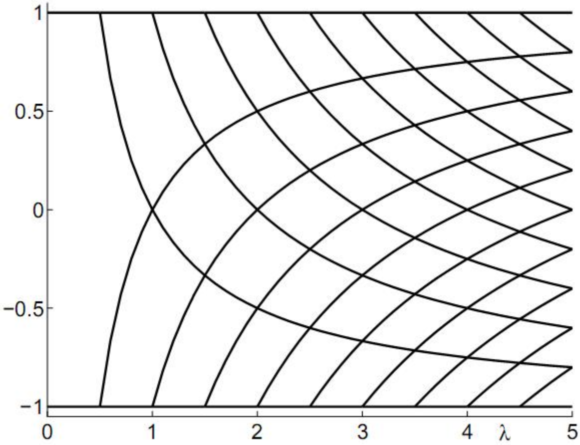
Different correlation functions

- $\rho(t) = e^{-\lambda|t|}$

$$\xi^*(du) = \frac{1}{1+\lambda}\left(\frac{1}{2}\delta_1(du) + \frac{1}{2}\delta_{-1}(du)\right) + \frac{\lambda}{1+\lambda}\frac{1}{2}\mathbf{1}_{[-1,1]}(du)$$

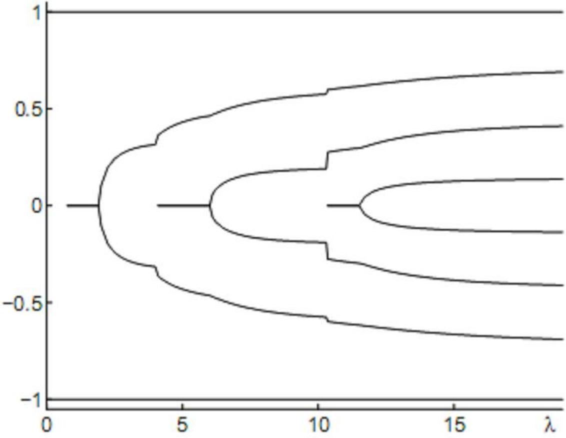
- $\rho(t) = \max\{0, 1-\lambda|t|\}$
- $\rho(t) = e^{-\lambda t^2}$
- $\rho(t) = 1/\sqrt{1+\lambda|t|}$
- $\rho(t) = 1/(1+\lambda|t|^{0.5})$

Designs for $\rho(t) = \max\{0, 1 - \lambda|t|\}$



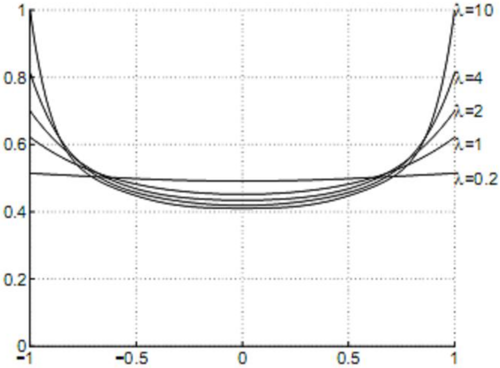
Support points of the optimal designs

Designs for $\rho(t) = e^{-\lambda t^2}$



Support points of the optimal designs

Designs for $\rho(t) = 1/\sqrt{1+\lambda|t|}$



The density of $\xi_0(du)$

$$\xi^*(du) = \omega^* \left(\frac{1}{2}\delta_1(du) + \frac{1}{2}\delta_{-1}(du) \right) + (1-\omega^*)\xi_0(du)$$

Optimality criterion with singular kernel

A. Zhigljavsky, H. Dette, A. Pepelyshev (2010) A new approach to optimal design for linear models with correlated observations. JASA (in press).

It is reasonable to consider the optimality criterion with a singular kernel $\rho_\infty(t)$

$$D(\xi) = \int\int \rho_\infty(u-v)\xi(du)\xi(dv) \to \min_\xi$$

where $\rho_\infty(t)$ is positive definite and $\rho_\infty(0) = \infty$.

Design for singular kernels

For $\rho_\infty(t) = \dfrac{1}{|t|^\alpha}$, $\alpha \in (0,1)$,
the density of optimal design is a Beta density

$$p^*(t) = \frac{2^{-\alpha}}{B(\frac{1+\alpha}{2}, \frac{1+\alpha}{2})}(1-t^2)^{\frac{\alpha-1}{2}}$$

For $\rho_\infty(t) = -\ln(t^2)$
the density of optimal design is the arcsine density

$$p^*(t) = \frac{1}{\pi\sqrt{1-t^2}}$$

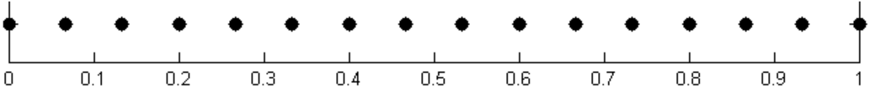
Designs for computer experiments

Designs for computer experiments

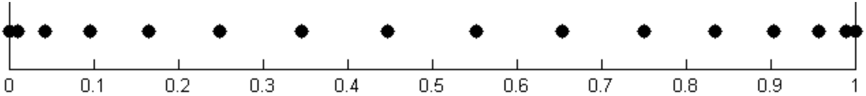
- uniform space-filling designs
  - random Latin Hypercube Design (LHD)
  - maximin LHD
  - uniform designs based on LD- and WD-discrepancy
  - designs based on random or pseudo-random sequences
- non-uniform designs

The arcsine transformation

uniform density $p(t) = 1$



arcsine density $p(t) = 1 \Big/ \left(\pi\sqrt{t(1-t)}\right)$

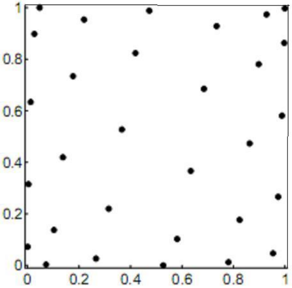


For $x \in [0,1]^d$, apply the transformation to the one-dimensional projections of designs,

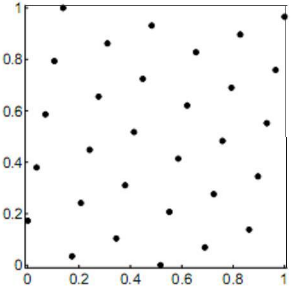
$$\tilde{x}_i = (1 - \cos(\pi x_i))/2$$

30-point LHDs in 2D

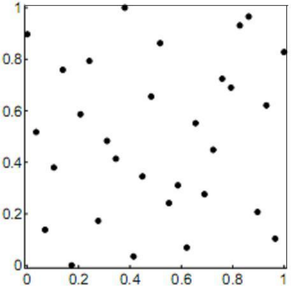
GMLHD



MLHD



RLHD



Random LHD

Maximin LHD

Generalized Maximin LHD is obtained by modifying MLHD

The performance of designs

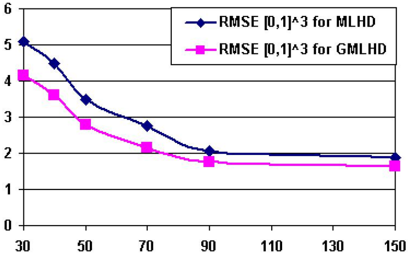
Compare designs using the MSE criterion and the meta-model given by the posterior mean of a Gaussian process with the Gaussian correlation function, parameters are estimated according to the Bayesian analysis.

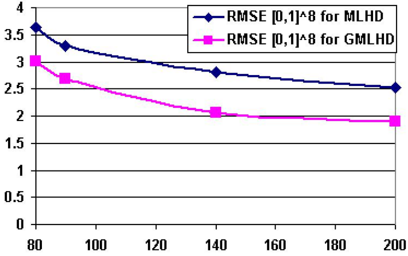
$$\mathrm{MSE}_{\Omega}(L) = \int_{\Omega} \big(\eta(x) - \hat{\eta}(x)\big)^2 \, \mathrm{d}x$$

In the numerical study, several test nonlinear models were considered. In all cases results are very similar and show a superiority of GMLHD.

The performance of designs

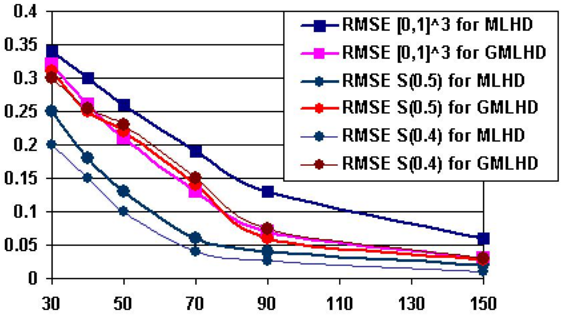
The square root of the mean squared error, obtained by the generalized maximin Latin hypercube designs, is substantially smaller (about 15%-30%) than the square root of the mean squared error, obtained by the maximin Latin hypercube designs.



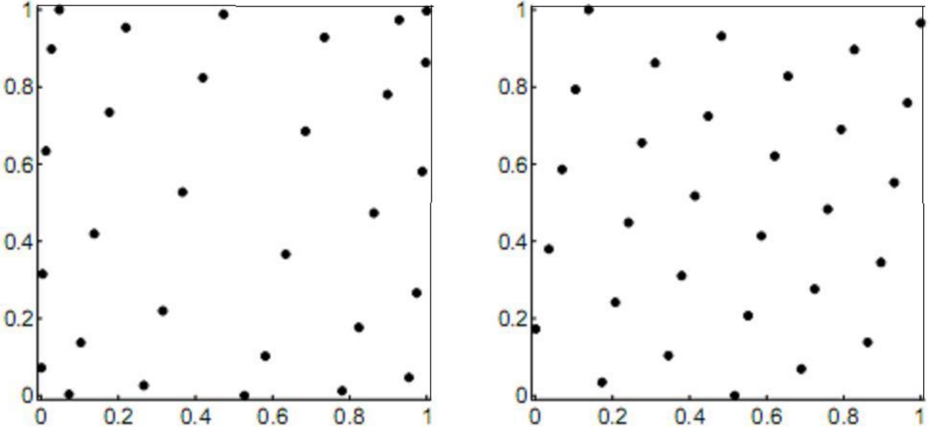


The performance of designs

For the MLHD, the mean squared error at the middle of the design space is considerably smaller than the mean squared error over the full design space. For the GMLHD, the mean squared error is almost constant among different subdomains of the design space.



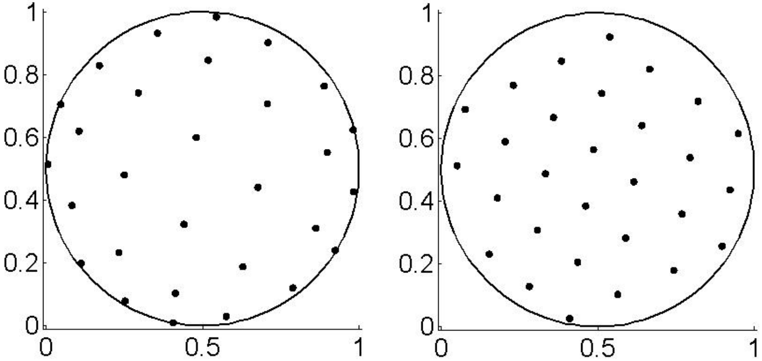
More dense location of points near boundary



Experiments for non-rectangular domains

Algorithm for computing efficient designs

- Compute n uniform points for a given domain
- Shift each point to the boundary



Conclusion

Algorithm for computing efficient designs

- Compute n uniform points for a given domain
- Shift each point to the boundary

This algorithm is applicable for computer experiments and other experiments with correlated observations except the case of a strongly nonlinear 1D model and a small variance of errors.

References

- A. Zhigljavsky, H. Dette, A. Pepelyshev (2010) A new approach to optimal design for linear models with correlated observations. JASA (in press).
  - optimal designs for the location model
  - improving the Bickel-Herzberg approach
  - singular kernels
- H. Dette, A. Pepelyshev (2010) Generalized Latin hypercube design for computer experiments. Technometrics (in press).
  - non-uniform space-filling designs
- H. Dette, A. Pepelyshev, T. Holland-Letz (2009) Optimal designs for random effect models with correlated errors with applications in population pharmacokinetics. Ann. of Applied Stat. (in press)
  - nonlinear regression models on a interval
  - the Bickel-Herzberg approach

Thank you for attention