Optimal designs for the exponential model with correlated observations

Andrey Pepelyshev

St.Petersburg State University

Abstract

Our aim is to construct and to study optimal designs for the exponential model with correlated observations. It will be shown that exact D-optimal design has complex structure with respect to the correlation parameters.

Contents

1. Introduction

Let the experimental results $y_1, \ldots, y_n$ be described by the regression equation

$$Y_j = \beta^T f(t_j) + \varepsilon_{t_j},\ j = 1, \ldots, n,$$

where random values have zero mean and the covariance matrix of the form

$$\mathrm{Cov}(Y_i, Y_j) = \sigma^2(\gamma\rho(t_i, t_j) + (1-\gamma)\delta_{i,j}),$$

where $\gamma \in [0, 1]$, $\delta_{i,j}$ is Kroneker symbol and $\rho(t_i, t_j)$ is a known function.
Suppose as usually (Hoel, 1958; Bickel, Herzberg, 1979), that the covariance function depends on the distance between experimental conditions, i.e.
$\rho(t_i, t_j) = \tilde{\rho}(t_i - t_j)$.
With $\gamma = 0$ we have the case of uncorrelated observations.
With $\gamma = 1$ the errors are said to have an autoregressive structure.

$$Y_j = \beta^T f(t_j) + \varepsilon_{t_j},\ j = 1, \ldots, n,$$

$f(t)$ is a given function, β is the vector of unknown parameters.

Rewrite the regression model in a matrix form

$$Y = F^T \beta + \varepsilon$$

where

$$Y = (y_1, \ldots, y_n)^T,\ F = (f(t_1) \vdots \ldots \vdots f(t_n)).$$

The parameters can be estimated by the ordinary LSE and the weighted LSE.

1. Introduction

The ordinary LSE of parameters is given by

$$\hat{\beta}_o = (F^T F)^{-1} F^T Y,$$

and the covariance matrix of $\hat{\beta}_o$ has the form

$$\mathrm{Cov}\hat{\beta}_o = (F^T F)^{-1} F^T \Sigma F (F^T F)^{-1}.$$

Weighted LSE of parameters is given by

$$\hat{\beta} = (F^T \Sigma^{-1} F)^{-1} F^T \Sigma^{-1} Y$$

and covariance matrix of $\hat{\beta}$ have a form

$$\mathrm{Cov}\hat{\beta} = (F^T \Sigma^{-1} F)^{-1}.$$

The goal of the experimental design is to determine a design $\xi = \{t_1, \ldots, t_n\}$ which minimize the determinant of $\mathrm{Cov}\hat{\beta}$, that means minimization of confidence ellipsoid's volume.

2. The exponential model

$$Y_{t_i} = \eta(t_i) + \varepsilon_{t_i}, \quad \eta(t) = ae^{-bt}, \quad t_i \in [c, \infty), \quad i = 1, \dots, n,$$

$$\mathrm{Cov}(Y_{t_i}, Y_{t_j}) = \sigma^2 e^{-\lambda |t_i - t_j|}, \ \lambda > 0,$$

where λ is some constant, which characterizes the level of correlation.
We can assume without loss of generality that $\sigma = 1$ and $c = 0$.
This model with uncorrelated observations was studied by many authors.
The covariance matrix of the weighted LSE of $\beta = (a, b)^T$ equals
$\mathrm{Cov}(\hat{\beta}) = (F^T \Sigma^{-1} F)^{-1}$ (Fedorov, Hackl, 1997), where

$$Y = (Y_{t_1}, \dots, Y_{t_n})^T, \quad F^T = F_\xi^T = (f(t_1), \dots, f(t_n)),$$

$$f(t) = (\partial\eta/\partial a, \partial\eta/\partial b)^T, \quad \Sigma = \Sigma_\xi = \left(e^{-\lambda |t_i - t_j|}\right)_{i,j=1,\dots,n}.$$

Definition. An exact locally D-optimal design maximizes $\det M(\xi)$ where
$M(\xi) = M(\xi, a, b, \lambda) = F_\xi^T \Sigma_\xi^{-1} F_\xi$.

3. Locally optimal designs

As it was shown in (Dette, et.al., 2006), the information matrix can be presented in the form $M(\xi) = F_\xi^T V_\xi^T V_\xi F_\xi$ where V_ξ is a 2-diagonal matrix given by $V_\xi = (v_{i,j})$, $v_{i,i} = r_i$, $v_{i,i-1} = -s_i$, $r_1 = 1$, $s_1 = 0$,

$$r_i = 1\Big/\sqrt{1 - e^{-2\lambda(t_i - t_{i-1})}}, \; s_i = e^{-\lambda(t_i - t_{i-1})} r_i,$$

$i = 2, \ldots, n$. From the Cauchy-Binet formula we obtain the following expression for the determinant of the information matrix

$$\det M(\xi, a, b, \lambda) = a^2 \sum_{1 \le i < j \le n} e^{-2b(t_i + t_j)} \psi^2(t_i, t_{i-1}, t_j, t_{j-1}),$$

where

$$\psi(t_i, t_{i-1}, t_j, t_{j-1}) =$$

$$= \frac{(1 - e^{-(\lambda - b)d_i})(t_j - t_{j-1} e^{-(\lambda - b)d_j}) - (1 - e^{-(\lambda - b)d_j})(t_i - t_{i-1} e^{-(\lambda - b)d_i})}{(1 - e^{-2\lambda d_i})(1 - e^{-2\lambda d_j})},$$

$d_i = t_i - t_{i-1}$, $d_j = t_j - t_{j-1}$.

4. Analytical results

Lemma. *Let* $\xi^* = \xi^*(a, b, \lambda) = \{t_1^*, \dots, t_n^*\}$ *be a locally D-optimal design for the exponential model with correlated observations. Then*
1) The design ξ^* *does not depend on* a*.*
2) The first point of the design is equal to zero, that is $t_1^* = 0$.
3) The points of the design ξ^* *satisfy*

$$t_i^*(\gamma b, \gamma \lambda) = \frac{1}{\gamma}\, t_i^*(b, \lambda)$$

for any $\gamma > 0$.

Since it is impossible to obtain the design explicitly further analysis will be based on a numerical calculations and the functional approach developed in (Melas, 2006). This approach allows one to construct a Taylor series for the support points of optimal designs as functions of some parameters.

5. The case $n = 2$

For the 2-point design $\xi = \{0, t_2\}$ we have

$$\det M(t_2) = \det M(t_2, b, \lambda) = \frac{t_2^2 e^{-2bt_2}}{1 - e^{-2\lambda t_2}}.$$

It is easy to see that $\det M(t_2)$ has a unique maximum for all fixed b and λ.
Let u^* be a unique solution of the equation

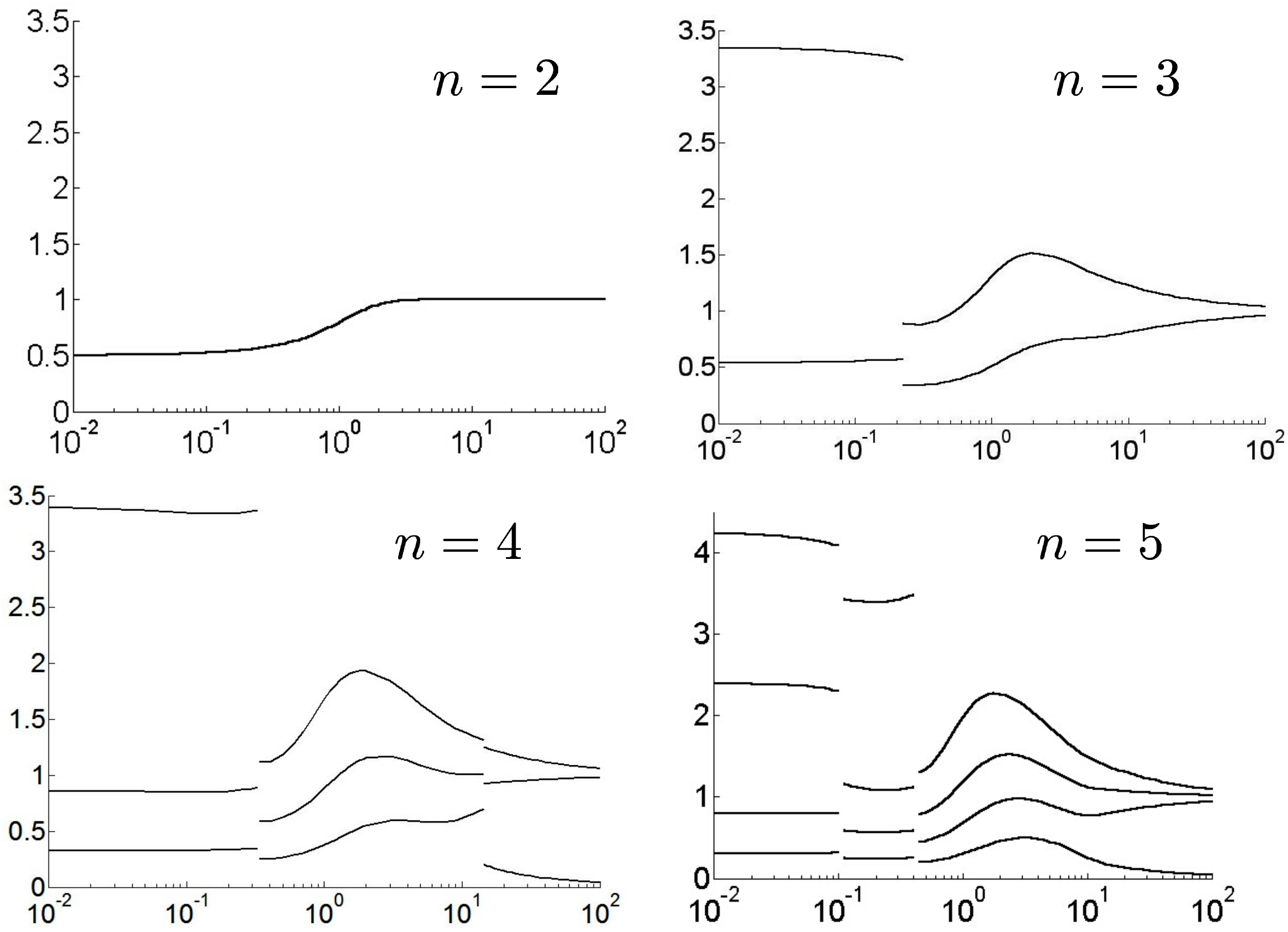
$$\frac{1}{1 - e^{-u}} = \frac{b}{\lambda} + \frac{2}{u}$$

on $u \in (-\infty, 0)$. A direct calculation shows that the second point of the D-optimal design equals

$$t_2^*(b, \lambda) = \frac{-u^*}{2\lambda}.$$

As $\lambda \to \infty$, we obtain $t_2^*(b, \lambda) \to 1/b$.
This means that the locally D-optimal designs for correlated observations tend to the locally D-optimal design for independent observations when correlation is decreasing (as $\lambda \to \infty$) for $n = 2$. The same is true for $n > 2$.

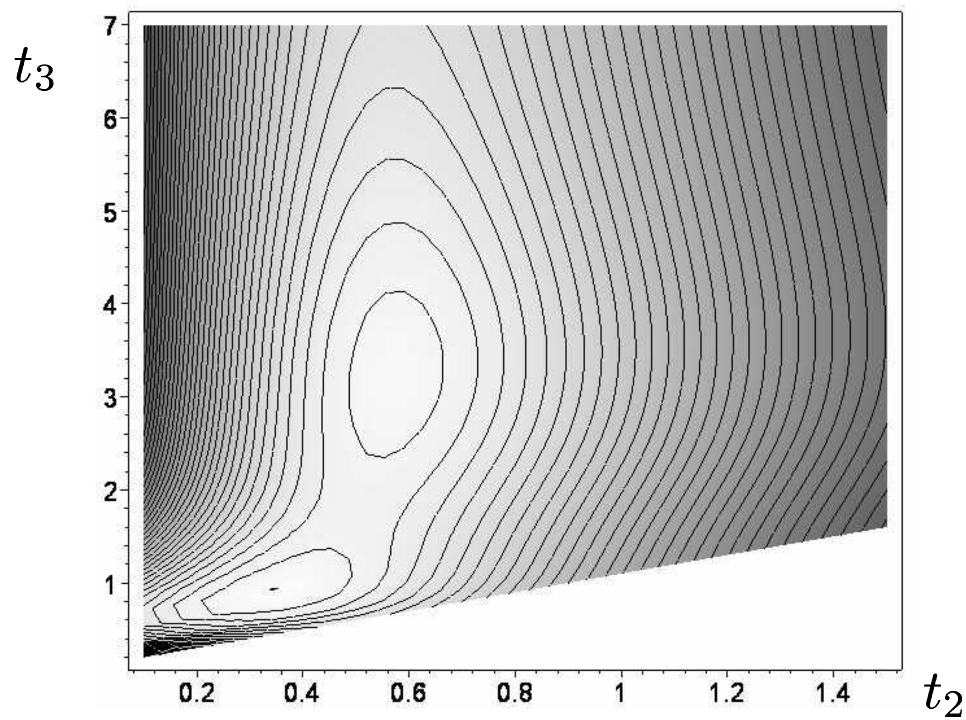
Figure of design points



Exact n-point locally D-optimal designs $\{0, t_2^*(\lambda), \ldots, t_n^*(\lambda)\}$ with $b = 1$ for $n = 2, 3, 4, 5$.

6. The case $n = 3$

Due to the Theorem 1 it is sufficient to investigate locally optimal designs for fixed b. Let $b = 1$.



Let λ^* be the value of λ such that the function $\det M(t_2, t_3)$ has equal maxima. A direct computation shows that $\lambda^* \approx 0.22367$. Thus, points of a locally optimal design are discontinuous at $\lambda = \lambda^*$.

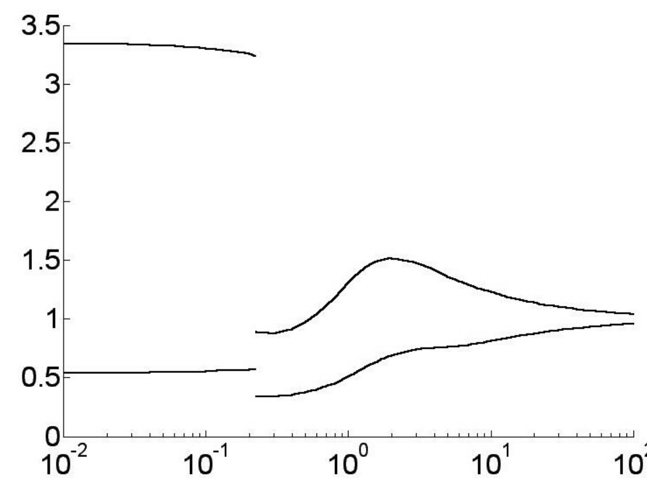
6. The case $n = 3$

In order to study a locally optimal design for small λ, we note that

$$\det M(t_2, t_3, \lambda) = \sum_{j=-2}^{\infty} M_{(j)}(t_2, t_3)\lambda^j.$$

Thus, nonzero points of locally D-optimal designs tend to points which maximize $M_{(-2)}(t_2, t_3)$ as $\lambda \to 0$.
Table contains locally D-optimal designs $\{0, t_2^*(\lambda), t_3^*(\lambda)\}$ for some special values of λ with $b = 1$, $\lambda^* = 0.22$.

λ	0	$\lambda^* - 0$	$\lambda^* + 0$	∞
$t_2^*(\lambda)$	0.5395	0.5703	0.3401	1
$t_3^*(\lambda)$	3.3560	3.2386	0.8870	1



6. The case $n = 3$

The implementation of the functional approach in Maple (Melas, 2006) gives the following expansions. The expansions

$t_2^*(\lambda) = 0.5395 + 0.1096\lambda + 0.1156\lambda^2 + 0.1077\lambda^3 + \ldots$

$t_3^*(\lambda) = 3.3560 - 0.6662\lambda + 1.8098\lambda^2 - 2.2812\lambda^3 + \ldots$

converge for $\lambda \in (0, \lambda^*)$. Expansions

$t_2^*(\lambda) = 0.5087 + 0.2687(\lambda - 1) - 0.0541(\lambda - 1)^2 - 0.0813(\lambda - 1)^3 + \ldots$

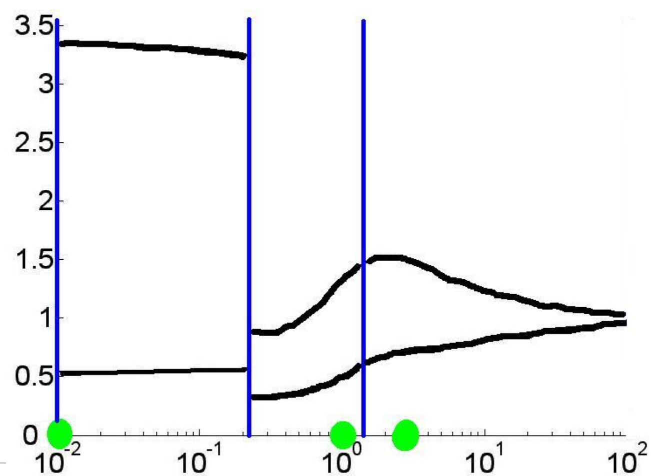
$t_3^*(\lambda) = 1.3056 - 0.4326(\lambda - 1) + 0.4930(\lambda - 1)^2 - 0.0819(\lambda - 1)^3 + \ldots$

converge for $\lambda \in (\lambda^*, 2)$. Expansions

$t_2^*(\lambda) = 0.6911 - 0.3836(\nu - 1/2) - 0.3161(\nu - 1/2)^2 + 1.14(\nu - 1/2)^3 + \ldots$

$t_3^*(\lambda) = 1.5177 - 0.0556(\nu - 1/2) - 1.4224(\nu - 1/2)^2 + 1.74(\nu - 1/2)^3 + \ldots,$

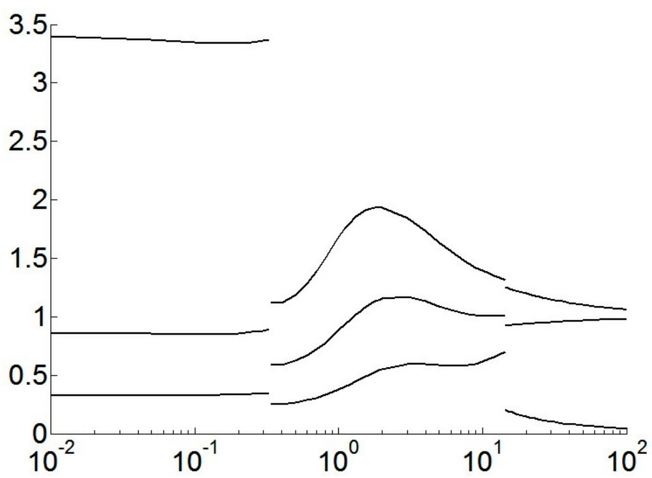
where $\nu = 1/\lambda$ converge for $\lambda \in (1, \infty)$.



7. The case $n = 4$

Numerical calculations show that the points of a locally D-optimal design are discontinuous at two points, say λ^* and λ^{**}. Table contains locally D-optimal designs $\{0, t_2^*(\lambda), t_3^*(\lambda), t_4^*(\lambda)\}$ for some special values of λ with $b = 1$.

λ	0	$\lambda^* - 0$	$\lambda^* + 0$	$\lambda^{**} - 0$	$\lambda^{**} + 0$	∞
$t_2^*(\lambda)$	0.3288	0.3462	0.2491	0.6966	0.2030	0
$t_3^*(\lambda)$	0.8669	0.8919	0.5841	1.0074	0.9250	1
$t_4^*(\lambda)$	3.4058	3.3611	1.1180	1.3133	1.2490	1



8. Maximin designs

Note that the implementation of locally optimal designs in practice requires a prior guess for the unknown parameters. This can raise confusion for an experimenter. The notion of maximin efficient designs seems to be more attractive and useful in practice (Muller, 1995).
The D-efficiency of a design ξ is given by

$$\mathrm{eff}_D(\xi) = \mathrm{eff}_D(\xi, a, b, \lambda) = \left[\frac{\det M(\xi, \bar\beta)}{\det M(\xi^*_{loc}(\bar\beta), \bar\beta)} \right]^{1/2}$$

where $\bar\beta = (a, b, \lambda)$ and ξ^*_{loc} is a locally D-optimal design. It is easy to see that the efficiency does not depend on a.

Definition. A design ξ^* is called a maximin (efficient) D-optimal design if it maximizes the worst D-efficiency over some set of the parameters Ω.

8. Maximin designs

Lemma. *Let* $\xi^* = \xi^*(\Omega) = \{t_1^*, \dots, t_n^*\}$ *be a maximin* D*-optimal design for the exponential model with correlated observations. Then*
1) The first point of the design equals zero, that is $t_1^* = 0$.
2) The points of the design ξ^* *satisfy*

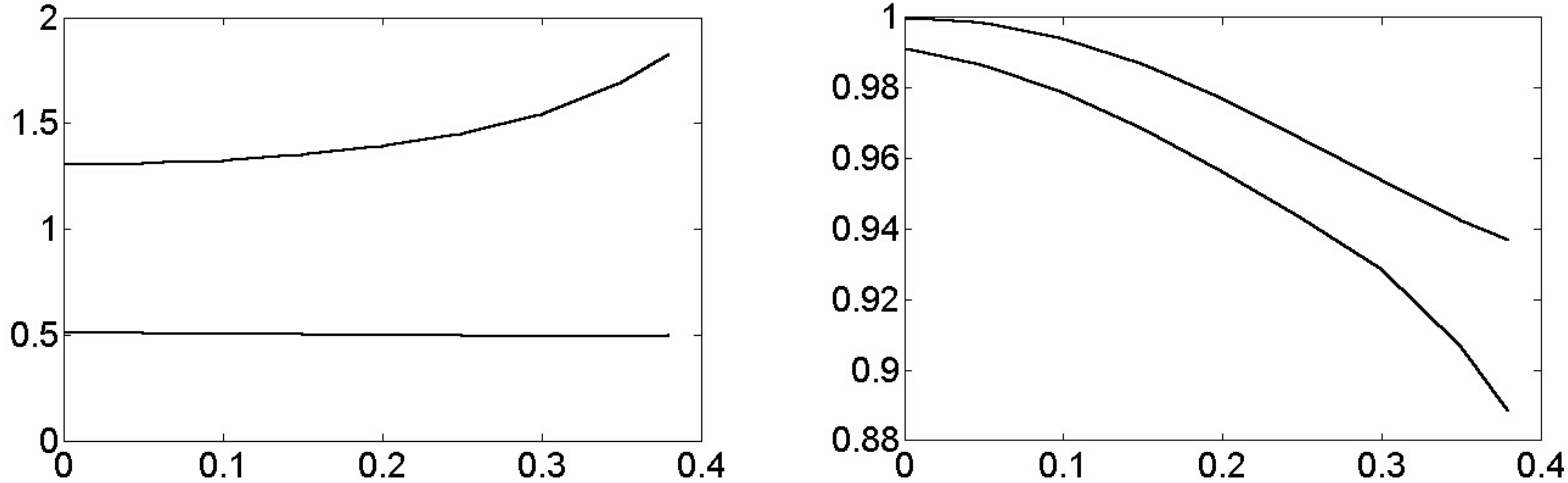
$$t_i^*(\gamma\Omega) = \frac{1}{\gamma}\, t_i^*(\Omega),$$

for any $\gamma > 0$.
Consider a set Ω of the form

$$\Omega = \Omega(z) = \left\{ \bar\beta = (a, b, \lambda) : a = 1, \begin{array}{c} (1-z)b_0 \le b \le (1+z)b_0, \\ (1-z)\lambda_0 \le \lambda \le (1+z)\lambda_0 \end{array} \right\},$$

which seems appealing from a practical point of view. Values b_0 and λ_0 are the initial guess and z can be interpreted as a relative error for the guess.
To study maximin designs $\xi^*(z)$ we implemented a special case of the functional approach introduced in (Melas, Pepelyshev, 2006). For example, suppose that $b_0 = 1$ and $\lambda_0 = 1$.

Following the functional approach we obtain the expansions
$x_2(z) = 0.5088 - 0.1887z^2 - 0.0263z^4 - 0.4340z^6 + 12.7721z^8 + \ldots,$
$x_3(z) = 1.3056 + 1.8505z^2 + 5.7658z^4 + 18.438z^6 + 89.7939z^8 + \ldots$
which converge for $z \in [0, 0.4)$. These points are depicted in Figure 18, which also shows the dependence of the minimal efficiency of maximin designs and the equidistant design $\{0, 0.65, 1.3\}$ on z. We see that the maximin designs are more efficient than the equidistant design.



Exact 3-point maximin D-optimal design $\xi^*(z) = \{0, t_2^*(z), t_3^*(z)\}$ with $b = 1$ for $n = 3$ (left), and the minimal efficiencies of the maximin design $\xi^*(z)$ and the equidistant design $\{0, 0.65, 1.3\}$ over $\Omega(z)$ (right).

A comparison of designs for different n

Suppose that we have a restriction on the number of random processes and do not have a restriction on the number of experimental points for each process. Thus it is interesting to compare locally D-optimal designs supported in different numbers of points. Let $b = 1$, $\lambda = 1$. We use the efficiency defined by $\mathrm{eff}_p(\xi_n) = \sqrt{\det M(\xi_n)/\det M(\xi_p)}$. We obtain that $\mathrm{eff}_2(\xi_n)$ equals

n	3	4	5	6	7
$\mathrm{eff}_2(\xi_n)$	1.125	1.173	1.196	1.210	1.218

We see that the quantity of information is increasing very slowly with the number of design points.

9. Concluding remarks

- In the exponential regression model with an autoregressive error structure exact D-optimal designs for weighted least squares analysis are investigated. It is shown that optimal points are discontinuous with respect to the level of correlation. This result is in agreement with the results obtained in (Stehlik, 2005; Dette, Kunert, Pepelyshev, 2006).

- Maximin D-optimal designs which are robust with respect to parameters of model and correlation parameter are investigated. These designs are constructed and studied by means of the functional approach.

THANK YOU FOR ATTENTION