Some Properties of D-optimal Designs for Random Fields with Different Variograms

Milan Stehlík

Department of Statistics and Mathematics
Wirtschaftsuniversität Wien

Research Report Series

Report 4
May 2004

http://statistik.wu-wien.ac.at/

Some Properties of D-optimal Designs for Random Fields with Different Variograms *

Milan Stehlík
Department of Probability and Statistics,
Faculty of Mathematics, Physics and Informatics,
Comenius University Bratislava, Mlynská Dolina,
842 48 Bratislava, Slovakia
Department of Statistics and Mathematics,
University of Economics Vienna,
Augasse 2-6, A-1090 Vienna, Austria
e-mail: Milan.Stehlik@isis.wu-wien.ac.at, Stehlik@fmph.uniba.sk

May 17, 2004

Abstract

The aim of this paper is discussion on particular properties of D-optimal designs under isotropic and intrinsically stationary correlation structures. We show that design points can collapse under the presence of some covariance structures. This enables to include so called *nugget effect* by natural way. Some numerical examples are also included.

Keywords: Design of experiment, D-optimality, semivariogram, correlation, nugget effect

1 Introduction

In the present paper we want to focus mainly on the relation of the D-optimal design and covariance structure. For simplicity, we will suppose the intrinsically stationary process with smooth parametrization of the covariance function. We use the Fisher information matrix despite some conceptual difficulties associated, there is no practical difficulty however.

Let us consider that $\{y(x) : x \in A\}$ is a realization of a random process $\{Y(x) : x \in A\}$, where A is a fixed compact subset of R^d. Let us assume $E\left(Y(s+h) - Y(s)\right) = 0$ and define

$$2\gamma(h) = \text{var}\left(Y(s+h) - Y(s)\right)$$

*Research was partly supported by FWF-grant (Austrian grant agency) P16809-N2 and by the VEGA grant (Slovak grant agency) No 1/0264/03.

(equation make sense only when right side depends only on h). If this is the case, we say that the process is *intrinsically stationary*, the function $2\gamma(h)$ is then called *variogram* and $\gamma(h)$ is called *semivariogram* (for more see [Banerjee, Carlin, Gelfand 04]). One of the important concepts related to covariance structure is *isotropy*. If the semivariogram function $\gamma(h)$ depends upon the separation vector only through its length, than we say that the process is isotropic. Notice, that for every semivariogram we have $\gamma(0) = 0$.

In our setup we consider the regression of the form

$$Y(x) = f(\vartheta, x) + e(x),$$

with design space A and a covariance function generated by some isotropic semivariogram. Let us briefly introduce most common isotropic semivariograms. For further discussion see e.g. [Cressie 93]. Now consider the three basic isotropic models: linear, spherical and exponential.

- **Linear model** valid in $R^n,\ n \geq 1$,

$$\gamma(d) = \begin{cases} \tau^2 + \sigma^2 d, & \text{for } d > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

- **Spherical model** valid in $R^1,\ R^2, R^3$

$$\gamma(d) = \begin{cases} \tau^2 + \sigma^2(\frac{3d}{2r} - \frac{1}{2}(\frac{d}{r})^3), & \text{for } 0 < d \leq r, \\ \tau^2 + \sigma^2, & \text{for } r < d, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

- **Exponential model** valid in $R^n,\ n \geq 1$,

$$\gamma(d) = \begin{cases} \tau^2 + \sigma^2(1 - \exp(-rd)), & \text{for } d > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Notice, that all introduced semivariograms are increasing functions of the distance. For such structures two point D-optimal design for the intercept is attained for the most distant two points of the design space, what we prove in section 2.

2 D-optimal design for the intercept under isotropic covariance structure

2.1 The D-optimal design for the intercept under Näther covariance structure

In this section we consider the simple linear regression with modified Näther covariance structure, studied in [Müller and Stehlík 04]. There we consider a modification of the Example 6.4 discussed in [Näther 85] with design space $V = [-1, 1]$ and a covariance function

$$\mathrm{cov}(x, z) = \begin{cases} 1 - \frac{d_{xz}}{\delta}, & \text{for } d_{xz} < \delta, \\ 0, & \text{otherwise,} \end{cases} \tag{4}$$

where $d_{xz} = |z - x|$ denotes the distance between the design points. Notice, that for $\delta > 2$, the modified Näther covariance structure (4) constitutes a linear semivariogram structure with $\gamma(d) = \frac{d}{\delta}$ (to relate this to general linear semivariogram (1), we have zero nugget $\tau^2 = 0$ and $\sigma^2 = \frac{1}{\delta}$).

Let us assume now, that the intercept is the parameter of the interest and covariance parameter $\delta > 2$ is fixed. Then three-point D-optimal design has the form $\{-1, y, 1\}$ (y is arbitrary, see Proposition 9 in [Stehlík 04]). The following proposition shows, that such result holds for arbitrary dimension.

Proposition 1 *Let $\{x_1, x_2, ..., x_n\}$, $-1 \leq x_1 < x_2... < x_{n-1} < x_n \leq 1$ $(2 \leq n)$ is design, δ is **fixed** and **intercept** is parameter of interest. When only correlated observations are possible ($\delta > 2$), the Fisher's information gained by $\{x_1, x_2, ..., x_n\}$ has form*

$$M = \frac{2\delta}{2\delta - (x_n - x_1)}$$

and the design $\{-1, x_2, ..., x_{n-1}, 1\}$, where $-1 < x_2.. < x_{n-1} < 1$ are arbitrary, is D-optimal. The information gained by the optimal design has the form $m(n, \delta) = \frac{\delta}{\delta - 1}$.

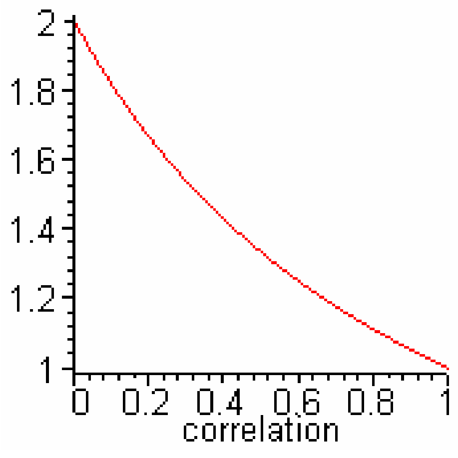

Figure 1: The information against correlation.

Remark 1 Here we can conclude, that the information for the intercept constitutes rather simple and natural concept: it decreases with the positive correlation (see also Figure 1). As far as our correlation is decreasing with the distance, the optimal distance is the maximal one. The same concept occurs also in the case of three observations, when both slope and intercept are parameters of interest (see [Stehlík 04] for details). However such pretty concept is not general.

Remark 2 When we consider only one observation, then we gain the information $m(1) = 1$ which is less than information gained by optimal design with arbitrary number $n \geq 2$ of observations for any $\delta > 2$. Notice, that

$$m(1) = \lim_{\delta \to +\infty} m(n, \delta), n = 2, ...$$

Proof *With the use of matrix derivative calculus we have* $M = \sum_{ij} \Sigma_{ij}^{-1} = 1^T \Sigma^{-1} 1$. *Further we have*

$$\sum_j \Sigma_{ij}^{-1} = \begin{cases} \frac{\delta}{2\delta - (x_n - x_1)}, & \text{for } i = 1 \text{and } n, \\ 0, & \text{otherwise} \end{cases}$$

and finally $\Sigma^{-1} 1 = (\frac{\delta}{2\delta - (x_n - x_1)}, 0, ..., 0, \frac{\delta}{2\delta - (x_n - x_1)})^T. \square$

As far as linear model (1) is valid in any dimension we can generalize:

Theorem 1 *Let* $\{x_1, x_2, ..., x_n\}$, $(2 \leq n)$ *are points lying on some line in compact design space* $A \subseteq R^k$ *and the covariance structure is constituted by semivariogram (1) with zero nugget and* $\sigma^2 = \frac{1}{\delta} > 0$, *and* ***intercept*** *is parameter of interest. Then the Fisher's information gained by* $\{x_1, x_2, ..., x_n\}$ *has form*

$$M = \frac{2\delta}{2\delta - diamA},$$

where diamA denotes the diameter of set A. The design $\{x, z \in A : d_{xz} = diamA\}$ *is D-optimal.*

The following example shows, that when points can be arbitrary in A, the optimal design could be more complex.

Example Consider 3 points with distances x, y and z, with the covariance structure (4) and covariance parameter $\delta > 2$ is fixed. Then the information on intercept has the form

$$M(x, y, z) = \delta \frac{z^2 - 2yz + x^2 - 2xz - 2xy + y^2}{\delta z^2 - 2\delta yz + x^2\delta - 2x\delta z - 2xy\delta + 2xyz + y^2\delta}$$

Let us imagine, just for simplicity, that the design space is compact equilateral triangle with the side length equals 1. If Theorem 1 holds, than the optimal design is obtained for arbitrary 2 vertices. But, when e.g. $z = 1$ our information still depends on x and y, and one can check that the information of all three vertices, i.e. $x = y = z = 1$, is larger than the information of two vertices. When all points lay on some line, we have e.g. $z = x + y$ and

$$M(x, y, x + y) = \frac{2\delta}{2\delta - (x + y)},$$

which corresponds to the previous theorem.

2.2 Two-point D-optimal design for the intercept under increasing semivariogram structure

Proposition 2 *Let $\{x, z\}$ is design in compact design space $A \subseteq R^k$ and* ***intercept*** *ϑ_1 is parameter of interest. Assume (distance d)-increasing semivariogram γ constituting 2-point covariance structure. Then Fisher's information gained by $\{x, z\}$ has form*

$$M(\vartheta_1) = \frac{2}{2 - \gamma(d_{xz})}$$

and the design $\{x, z \in A : d_{xz} = diamA\}$ is D-optimal. The information gained by the optimal design has the form $2/(2 - \gamma(diamA))$.

Proof *Without the loss of generality, let us suppose that $var(Y(x)) = var(Y(z)) = 1$. We have $M(\vartheta_1) = \sum_{ij} \Sigma^{-1}_{ij} = 1^T \Sigma^{-1} 1$. Further we have $\Sigma^{-1} 1 = (\frac{1}{2-\gamma(d)}, \frac{1}{2-\gamma(d)})^T$.* $\square$

Here we recall, that all three semivariograms (linear, spherical and exponential) defined in the Introduction are increasing functions of distance. Here we briefly introduce other increasing variograms.

- **Gaussian**

$$\gamma(d) = \begin{cases} \tau^2 + \sigma^2(1 - \exp(-r^2 d^2)), & \text{for } d > 0, \\ 0 & \text{otherwise.} \end{cases}$$

- **Powered exponential**, $0 < p \leq 2$

$$\gamma(d) = \begin{cases} \tau^2 + \sigma^2(1 - \exp(-|rd|^p)), & \text{for } d > 0, \\ 0 & \text{otherwise.} \end{cases}$$

- **Rational quadratic**

$$\gamma(d) = \begin{cases} \tau^2 + \frac{\sigma^2 d^2}{1 + rd^2}, & \text{for } d > 0, \\ 0 & \text{otherwise.} \end{cases}$$

- **Power law**, $0 \leq \lambda < 2$

$$\gamma(d) = \begin{cases} \tau^2 + \sigma^2 d^\lambda, & \text{for } d > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

- **Matérn** (at least for some values of parameters v and r)[1]

$$\gamma(d) = \begin{cases} \tau^2 + \sigma^2[1 - \frac{(2\sqrt{v}rd)^v}{2^{v-1}\Gamma(v)} K_v(2\sqrt{v}rd)], & \text{for } d > 0, \\ 0 & \text{otherwise.} \end{cases}$$

[1]Here function $\Gamma(.)$ is the usual gamma function and K_v is the modified Bessel function of the order v.

Notice, that also some more exotic semivariograms can be increasing, for example the semivariogram $\gamma(d) = d^2/\delta, \delta > 2$, (this is a generalization of Power law (7) for $\lambda = 2$) which is valid for two observations, but is not valid for any $\delta > 0$ when we consider more than two observations. Also notice, that there exist non-monotonous semivariograms, e.g. common **wave** one,

$$\gamma(d) = \begin{cases} \tau^2 + \sigma^2(1 - \frac{\sin(rd)}{rd}), & \text{for } d > 0, \\ 0 & \text{otherwise.} \end{cases}$$

2.3 Exponential covariance structure.

Now let us suppose $\gamma(d) = 1 - e^{-d}$, i.e. exponential model (3) with zero nugget $\tau^2 = 0$ and $\sigma^2 = r = 1$. Proposition 2 states, that $\{-1, 1\}$ is D-optimal two-point design[2]. If we consider three-point-design $\{x, y, z\}$, $-1 \le x < y < z \le 1$, then Fisher's information on intercept has the form

$$1+\frac{2 + 2\exp(-d - 2q) - 2\exp(-d) + 2\exp(-2d - q) - 2\exp(-2(d + q)) - 2\exp(-q)}{\exp(-2(d + q)) - \exp(-2d) - \exp(-2q) + 1}$$

where $d = y - x$ and $q = z - y$.

Proposition 3 *Design $\{-1, 0, 1\}$ is D-optimal in the exponential setup.*

Proof *The (intercept) Fisher's information attains under the constraint $d+q = x$ $(0 < x \le 2)$ maximum $m(x) = \frac{3-\exp(-\frac{x}{2})}{\exp(-\frac{x}{2})+1}$ for $d = \frac{x}{2}$ and function $x \to m(x)$ is increasing.*$\square$

Let us suppose 4-point design $-1 \le x_1 < x_2 < x_3 < x_4 \le 1$, and denote the distances $x = x_2 - x_1$, $y = x_3 - x_2$ and $z = x_4 - x_3$. Then the (intercept) Fisher's information has the form

$$M = 2(-2 + e^{-z} + e^{-x} + e^{-y} + e^{-2y-x-2z} + e^{-2x} + e^{-2y} + e^{-2z} - e^{-2x-2y-2z}$$
$$-e^{-x-2z} + e^{-2x-2y-z} - e^{-z-2x} - e^{-y-2z} - e^{-2x-y} - e^{-2y-z} + e^{-2z-2x-y} - e^{-2y-x})/$$
$$(-1 + e^{-2z} + e^{-2y} - e^{-2y-2z} + e^{-2x} - e^{-2x-2z} - e^{-2x-2y} + e^{-2x-2y-2z}).$$

We have use the finite grid numerical evaluation of the information and found the maximum for $x = y = z = 2/3$ with the information $M = 1.964538$.

The (intercept) Fisher's information in the case of 5-point design $-1 \le x_1 < x_2 < x_3 < x_4 < x_5 \le 1$, (denote the distances x, y, z and v) has much more complicated form

$$M = (-5 - 2e^{-x-2y-2z-2v} - 2e^{-2x-2y-z-2v} + 2e^{-z-2v-2x} - 2e^{-2x-2z-2v-y} - 2e^{-z-2v}$$
$$+2e^{-2y-z-2v} + 2e^{-2v-y-2x} + 2e^{-2y-2z-v} - 2e^{-2x-2y-2z-v} + 2e^{-2y-x-2v}$$

[2]Notice, that this statement is not the corollary of Theorem 6.6 in [Näther 85], because $f(x) = 1$ is not linear combination of covariances.

$$+2e^{-2z-v-2x}+3e^{-2x}+2e^{-x}-2e^{-v-2y}+2e^{-2y-2x-v}+2e^{-2z-2v-y}$$
$$-2e^{-v-2x}-2e^{-y-2v}-e^{-2z-2v}-2e^{-2z-v}-e^{-2y-2z-2v}+3e^{-2v}$$
$$-e^{-2v-2y}+2e^{-2y-x-2z}+2e^{-z}+2e{-y}-2e^{-2y-x}-2e^{-x-2z}+2e^{-2z-2x-y}$$
$$-2e^{-y-2z}-2e^{-2x-y}+3e^{-2z}+3e^{-2y}-2e^{-2y-z}-2e^{-2v-x}-e^{-2y-2z}+2e^{-2x-2y-z}$$
$$+2e^{-x-2z-2v}+2e^{-v}-2e^{-z-2x}-e^{-2x-2z}-e^{-2x-2y}-e^{-2x-2y-2z}+3e^{-2x-2y-2z-2v}$$
$$-e^{-2z-2v-2x}-e^{-2y-2x-2v}-e^{-2v-2x})/(-1+e^{-2x}-e^{-2z-2v}+e^{-2y-2z-2v}+e^{-2v}$$
$$-e^{-2v-2y}+e^{-2z}+e^{-2y}-e^{-2y-2z}-e^{-2x-2z}-e^{-2x-2y}+e^{-2x-2y-2z}-e^{-2x-2y-2z-2v}$$
$$+e^{-2z-2v-2x}+e^{-2y-2x-2v}-e^{-2v-2x}).$$

Computationally we obtain the optimal design $x=y=z=v=1/2$ gaining the information $M=1.979674635$ (note, that information is increasing with number of design points).

Here we can conclude, that despite the linear semivariogram, the intercept information under exponential semivariogram increase with the number of design points.

3 Two-point D-optimal designs for covariance parameter r

In this section we study different covariance structures focusing their impact to the two-point D-optimal designs for covariance parameter r. The following Lemma illustrates the 'instability' of the design, that can happened when the covariance parameter r is treated like the parameter of interest. The design space A is some fixed compact subset of R^d with diameter 2.

Lemma 1 *Let us have*

$$Y(x_i)=a+e(x_i),\ i=1,2,\ x_1,x_2\in A,\gamma(d)=1-e^{-rd}$$

and r is not parameter of interest. Then the criterion function of the intercept is $M=\frac{2}{1+\exp(-rd)}$, and the D-optimal design $\{-1,1\}$ is attained for $d=2$. When the covariance parameter r is parameter of interest, then the information about the parameter vector (a,r) has the form $M=\frac{2d^2\exp(-2rd)(1+\exp(-2rd))}{(1+\exp(-rd))(1-\exp(-2rd))^2}$, which attains maximum for $d=0$.

Proof *We have $M_r=\frac{d^2\exp(-2rd)(1+\exp(-2rd))}{(1-\exp(-2rd))^2}$, and $M_a=\frac{2}{1+\exp(-rd)}$ and $M_{r,a}=0$. Then the determinant of the Fishers information on (a,r) is*

$$M_rM_a=\frac{2d^2\exp(-2rd)(1+\exp(-2rd))}{(1+\exp(-rd))(1-\exp(-2rd))^2}.$$

□

Now we start the study of the impact of the different covariance structures on the information about covariance parameter.

Proposition 4 *Let us have*

$$Y(x_i) = f(x_i, \vartheta) + e(x_i),\ \ i = 1, 2,\ \ x_1, x_2 \in A, \gamma(d) = 1 - e^{-rd}$$

and only covariance parameter r is parameter of interest. Then the maximal Fisher information is obtained for $d = 0$.

Proof *We have* $L = -\ln f(r) = K + \frac{1}{2}\ln\det\Sigma(r) + \frac{1}{2}u^T\Sigma(r)^{-1}u$, *where* $u = (Y(x_1) - f(x_1, \vartheta), Y(x_2) - f(x_2, \vartheta))^T$. *The Fisher information for the covariance parameter r is*

$$M_r = E\left(-\frac{\partial^2 L}{\partial r^2}\right) = \frac{1}{2}\frac{\partial^2 \ln\det\Sigma(r)}{\partial r^2} + \frac{1}{2}E\left(u^T\frac{\partial^2\Sigma(r)^{-1}}{\partial r^2}u\right),$$

and we have $\frac{\partial^2\{A_{i,j}\}}{\partial r^2} := \{\frac{\partial^2 A_{i,j}}{\partial r^2}\}$ *and* $\det\Sigma(r) = 1 - \exp(-2rd)$. *Further we have*

$$\frac{1}{2}\frac{\partial^2 \ln\det\Sigma(r)}{\partial r^2} = -\frac{2d^2\exp(-2dr)}{(1 - \exp(-2dr))^2}$$

and

$$\frac{1}{2}E\left(u^T\frac{\partial^2\Sigma(r)^{-1}}{\partial r^2}u\right) = \frac{d^2\exp(-2dr)(\exp(-2dr) + 3)}{(1 - \exp(-2dr))^2}$$

and finally

$$M_r = \frac{d^2\exp(-2rd)(1 + \exp(-2rd))}{(1 - \exp(-2rd))^2}.$$

Note that for every $r > 0$ *the maximum* $\frac{1}{2r^2}$ *is attained for* $d = 0$.□

To avoid such 'inconvenient' behavior we decrease the non-diagonal elements by multiplying with factor α, $0 < \alpha < 1$. By this we include nugget effect (micro-scale variation) and semivariogram has then the form[3]

$$\gamma(d, r) = \begin{cases} 0, & \text{for } d = 0, \\ 1 - \alpha + \alpha(1 - \exp(-rd)), & \text{otherwise.} \end{cases} \tag{6}$$

Then we obtain

$$\frac{1}{2}\frac{\partial^2 \ln\det\Sigma(r)}{\partial r^2} = -\frac{2\alpha^2 d^2\exp(-2dr)}{(1 - \alpha^2\exp(-2dr))^2}$$

and

$$\frac{1}{2}E\left(u^T\frac{\partial^2\Sigma(r)^{-1}}{\partial r^2}u\right) = \frac{\alpha^2 d^2\exp(-2dr)(\alpha^2\exp(-2dr) + 3)}{(1 - \alpha^2\exp(-2dr))^2}$$

and finally

$$M_r = \frac{\alpha^2 d^2\exp(-2dr)(\alpha^2\exp(-2dr) + 1)}{(1 - \alpha^2\exp(-2dr))^2}.$$

The following Figure 1 shows the information depending on nugget effect and distance.

[3] If $\gamma(d) \to c_0 > 0$, as $d \to 0$, then c_0 has been called the *nugget effect* by Matheron (1962). This is because it is believed that microscale variation is causing a discontinuity at the origin.

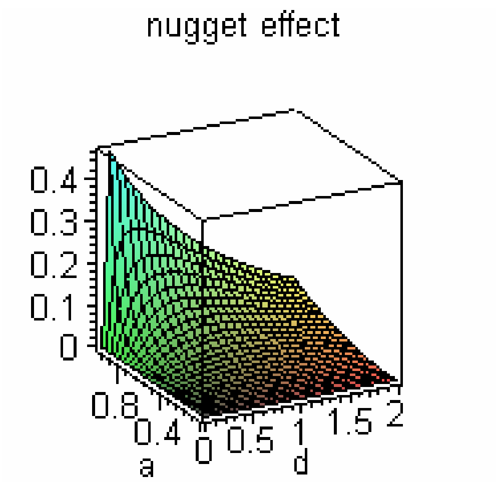

Figure 2: The nugget effect $1-a$.

Proposition 5 *The distance d of the optimal design is increasing function of nugget effect $1-\alpha$.*

Proof *The D-optimum design distances $d(\alpha)$ satisfy the equation*

$$\alpha^4 \exp(-4x) - 1 + x + 3\alpha^2 x \exp(-2x) = 0,$$

where $x = rd(\alpha)$. This equation ($F(r,d,\alpha)=0$) defines implicit function $d = d(\alpha)$. Sign of the ratio $-\frac{\partial F}{\partial \alpha}/\frac{\partial F}{\partial d}$ is negative for $0 < d < 2$, $0 < r$ and $0 < \alpha < 1$ and so the distance d of the optimal design is increasing function of nugget effect $1-\alpha$.$\square$

The following proposition states, that the correlation that decrease linearly with the distance of the points can led to the information on covariance parameter which increase with the distance.

Proposition 6 *Let us have*[4]

$$Y(x_i) = f(x_i, \vartheta) + e(x_i),\ \ i = 1,2,\ \ x_1, x_2 \in A, \gamma(d) = rd$$

and only covariance parameter r is parameter of interest. For regularity assumption we suppose that $rd < 2, r > 0$. (Note, that we include also the negative correlation, to avoid this, we need $rd < 1$.) Then the maximal Fisher information is obtained for maximal d.

Proof *We have* $\det \Sigma(r) = rd(2-rd)$,

$$\frac{1}{2}\frac{\partial^2 \ln \det \Sigma(r)}{\partial r^2} = -\frac{-2rd + r^2d^2 + 2}{r^2(rd-2)^2}$$

[4]Notice also the relation between the exponential and linear semivariogram through the Taylor expansion of exponential semivariogram $1-\exp(-rd) = rd + o(rd)$, where o is a "little-o" Landau symbol defined as follows: $f(x) = o(g(x))$ means that $f/g \to 0$ for $x \to 0$. Also note that this covariance structure is defined by linear semivariogram with zero nugget.

and

$$\frac{1}{2}E\left(u^T\frac{\partial^2\Sigma(r)^{-1}}{\partial r^2}u\right) = 2\frac{-2rd + r^2d^2 + 2}{r^2(rd-2)^2}$$

and finally

$$M_r = \frac{-2rd + r^2d^2 + 2}{r^2(rd-2)^2}.$$

We have

$$\frac{\partial M_r}{\partial d} = \frac{2d}{(2-rd)^3}$$

So M_r is increasing function for every (acceptable) $0 < d < \min\{\frac{2}{r}, 2\}$. □

Numerical Example Let us have the previous example with $r = 1$. Than the information has the form (see Figure 5) $M_r = \frac{-2d+d^2+2}{(d-2)^2}$. Note, that for $d \to 2$ tends $M_r \to +\infty$ (point $d = 2$ is not allowed, note restriction $d < \frac{2}{r}$). Let us connect this to result derived in [Müller and Stehlík 04], where we use the modified covariance structure (4). Let $\delta > 2$ be the parameter of interest, then the information

$$M_\delta = \frac{d^2 - 2\delta d + 2\delta^2}{\delta^2(2\delta - d)^2}$$

is an increasing function of distance d, (with derivative $\frac{2d}{(2\delta-d)^3}$) see Figure 3 displaying the M_δ for $\delta = 2.5$.

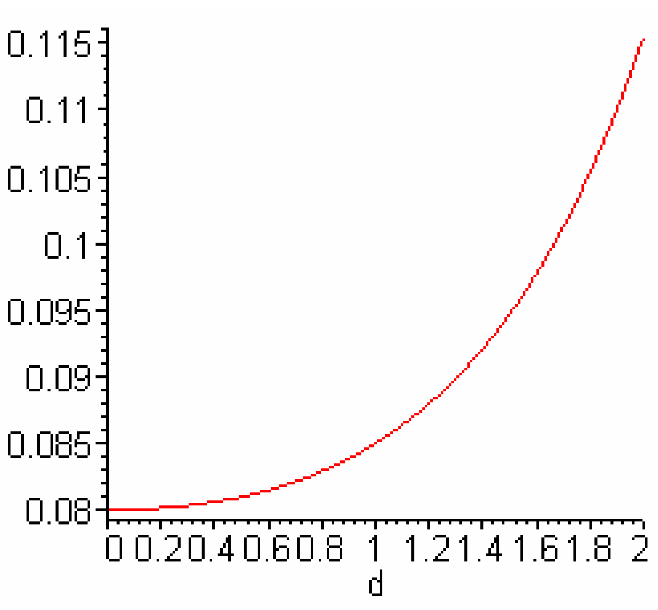

Figure 3: The information.

Also the multiplicative perturbation[5] $(0 < \alpha < 1)$ of the covariance to the form

$$\mathrm{cov}(x,y) = \begin{cases} 1, & \text{for } d = 0, \\ \alpha(1-rd), & \text{otherwise,} \end{cases} \tag{7}$$

leads to the same result. We obtain the information of the form

$$M_r = \alpha^2 d^2 \frac{\alpha^2 - 2\alpha^2 rd + \alpha^2 r^2 d^2 + 1}{(-1+\alpha^2 - 2\alpha^2 rd + \alpha^2 r^2 d^2)^2}.$$

where $0 < d < \min\{\frac{\alpha+1}{\alpha r}, 2\}$ (positive definiteness condition). We have

$$(\det \Sigma)^3 \frac{\partial M_r}{\partial d} = 2\alpha^2 d(-\alpha^4 + 3\alpha^4 rd - 3\alpha^4 r^2 d^2 + \alpha^4 r^3 d^3 + 1 - 3\alpha^2 rd + 3\alpha^2 r^2 d^2)$$

Proposition 7 *The multiplicative perturbation (7) also led to the optimal design obtained for maximal d but it affects the information (through the distance upper bound* $\min\{\frac{\alpha+1}{\alpha r}, 2\}$*).*

Proof *We prove, that* $\frac{\partial M_r}{\partial d}$ *is positive.*□

Proposition 8 *Let us have*[6]

$$Y(x_i) = f(x_i, \vartheta) + e(x_i),\ i = 1,2,\ x_1, x_2 \in A$$

$$cov(x_1, x_2) = 1 - rd + \frac{1}{2} r^2 d^2$$

and only covariance parameter r is parameter of interest. For simplicity, let us assume that we are interested for case r = 1. Then the maximal Fisher information is obtained for maximal d.

Proof *We have* $\det \Sigma(r) = 2rd - 2r^2d^2 + r^3d^3 - \frac{1}{4}r^4d^4$,

$$\frac{1}{2}\frac{\partial^2 \ln \det \Sigma(r)}{\partial r^2} = -2\frac{-32rd + 32r^2d^2 - 24r^3d^3 + 16r^4d^4 - 6r^5d^5 + r^6d^6 + 16}{r^2(-8+8rd-4r^2d^2+r^3d^3)^2}$$

and

$$\frac{1}{2}E\left(u^T \frac{\partial^2 \Sigma(r)^{-1}}{\partial r^2} u\right) = 2\frac{3r^6d^6 - 18r^5d^5 + 50r^4d^4 - 80r^3d^3 + 96r^2d^2 - 80rd + 32}{r^2(-8+8rd-4r^2d^2+r^3d^3)^2}$$

[5] Such perturbation is also natural according to Taylor expansion of the covariance introduced by (3): $\alpha \exp(-rd) = \alpha(1-rd) + o(rd)$. This perturbation establishes the nugget effect $1-\alpha$ and the corresponding semivariogram has the form

$$\gamma(d) = \begin{cases} 1 - \alpha + \alpha rd, & \text{for } d > 0, \\ 0 & \text{otherwise.} \end{cases}$$

[6] Notice also the relation to the exponential semivariogram Taylor expansion $\exp(-rd) = 1 - rd + \frac{1}{2}r^2d^2 + o(r^2d^2)$.

and finally

$$M_r = \frac{4r^6d^6 - 24r^5d^5 + 68r^4d^4 - 112r^3d^3 + 128r^2d^2 - 96rd + 32}{r^2(-8 + 8rd - 4r^2d^2 + r^3d^3)^2}.$$

(Here we recall, that we are interested only for case $r = 1$.) The derivative $\frac{\partial M_1}{\partial d}$ is negative for $0 < d < 1$, zero for $d = 1$ and positive for $1 < d < 2$. So, the information M_1 (see Figure 4) is decreasing for $0 < d < 1$, is minimal in zero and is increasing for $1 < d < 2$. Note, that for $d \to 2$ M_1 tends to $+\infty$, but, point $d = 2$ is not allowed (positive definiteness of variance). $\square$

3.1 Summary.

In this section we have compared three covariance structures, let us recall (see Figure 4):

- exponential $\text{cov}(x_1, x_2) = \exp(-rd)$
- linear $\text{cov}(x_1, x_2) = 1 - rd$
- quadratic $\text{cov}(x_1, x_2) = 1 - rd + \frac{1}{2}r^2d^2$

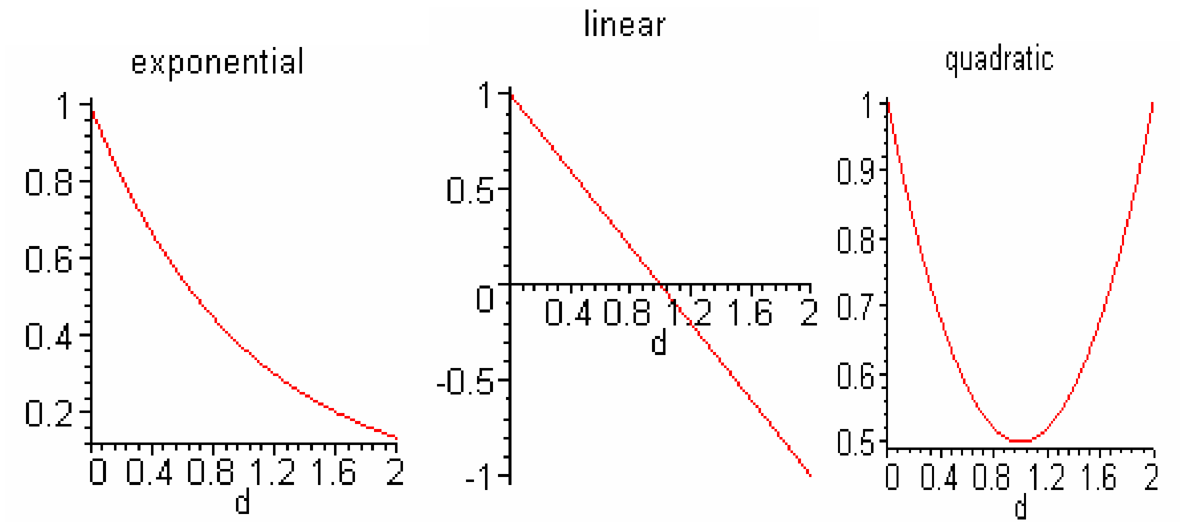

Figure 4: The covariance structures, $r = 1$.

When the only parameter of interest is the covariance parameter r, then the information decrease with distance d for exponential structure and increase for linear one (see also Figure 5).

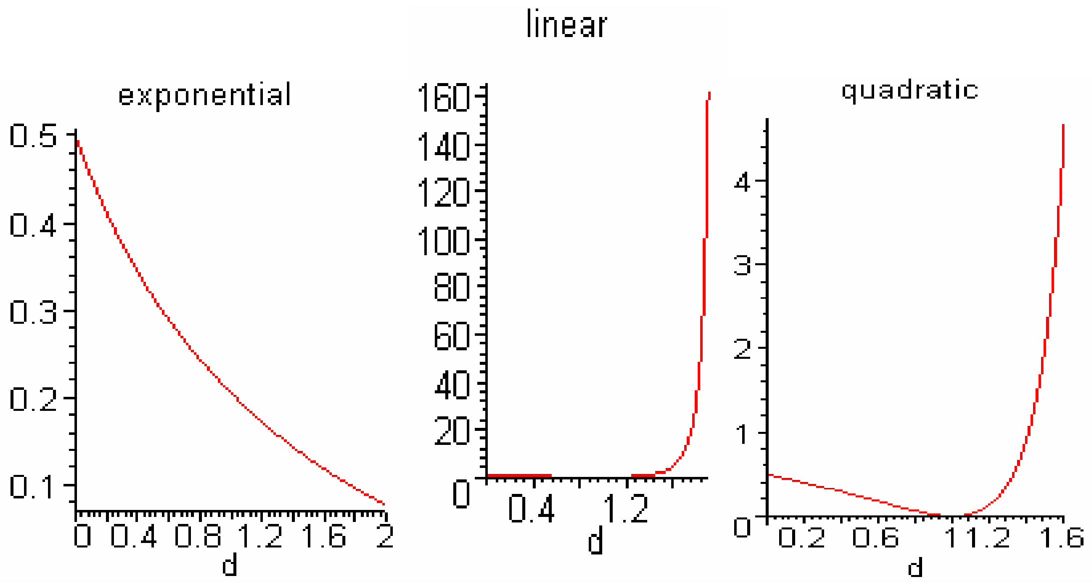

Figure 5: The information about covariance parameter r, $r = 1$.

4 Illustrative examples

Here we present some analytical and numerical results on the problem optimal allocation of samples in order to detect the case of an acute intake of a radioactive substance in the human body. Such models are studied in [Sanchez and Lopez-Fidalgo 03] and [Lopez-Fidalgo and Rodriguez-Diaz 04].

4.1 Example 1

In this example we consider the following model

$$y(I,p,t) = I \times \frac{1 + 8.16\exp(-0.12p + 5.32t)}{-21.6\exp(0.06p + 5.29t) + 75.59\exp(0.05p + 5.32t)}$$

and assume exponential covariance structure $\mathrm{cov}(t_1, t_2) = \exp(-r|t_1 - t_2|)$ and normal errors. Let parameters of interest are r and I and parameter p is fixed. We suppose two observations $t_1 = 0.5$ and $t_2 = t$, (in hours) $0.5 < t < 100$. Then the Fisher information M_r about the covariance parameter r is

$$M_r = \frac{(t-0.5)^2\exp(-2r(t-0.5))(1+\exp(-2r(t-0.5)))}{(1-\exp(-2r(t-0.5)))^2},$$

and the information about the parameter I is

$$M_I = (\frac{-1 - 8.16\exp(-0.12p + 2.66)}{(-21.6\exp(0.06p + 2.645) + 75.59\exp(0.05p + 2.66))(-1 + \exp(-2r(t-0.5)))}$$

$$+\frac{(1 + 8.16\exp(-0.12p + 5.32t))\exp(-r(t-0.5))}{(-21.6\exp(0.06p + 5.29t) + 75.59\exp(0.05p + 5.32t))(-1 + \exp(-2r(t-0.5)))})\times$$

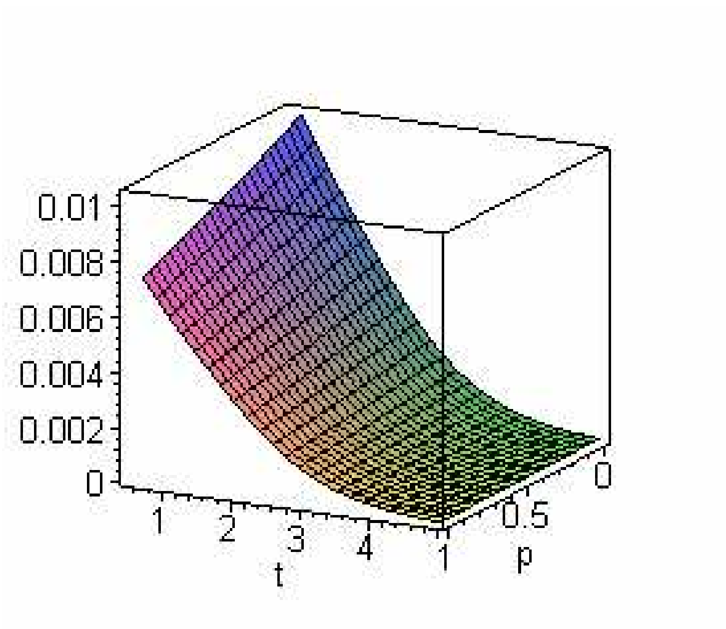

Figure 6: The information on (r, I).

$$\times \frac{1 + 8.16\exp(-0.12p + 2.66)}{-21.6\exp(0.06p + 2.645) + 75.59\exp(0.05p + 2.66)} +$$

$$+ \Big(\frac{1 + 8.16\exp(-0.12p + 2.66)\exp(-r(t-0.5))}{(-21.6\exp(0.06p + 2.645) + 75.59\exp(0.05p + 2.66))(-1 + \exp(-2r(t-0.5)))}$$

$$- \frac{1 + 8.16\exp(-0.12p + 5.32t)}{(-21.6\exp(0.06p + 5.29t) + 75.59\exp(0.05p + 5.32t))(-1 + \exp(-2r(t-0.5)))} \Big) \times$$

$$\times \frac{1 + 8.16\exp(-0.12p + 5.32t)}{-21.6\exp(0.06p + 5.29t) + 75.59\exp(0.05p + 5.32t)}$$

Further we have $M_{r,I} = 0$ and information M about vector (r, I) is equal to $M_r M_I$. The Figure 6 displays the information. The information about the parameter p is a rather complex function and we avoid its displaying here.

4.2 Example 2

Let us suppose the model of the form

$$y(I, p, t) = I \times \frac{1 + 8.16\exp(-0.12p + 5.32t)}{-21.6\exp(0.06p + 5.29t) + 75.59\exp(0.05p + 5.32t)},$$

normal errors and exponential covariance of the form $\exp(-r|x-y|)$. Let parameters of interest are r, I and p. We suppose three observations $t_1 = 0.5$ and t_2, and t_3 (in hours), $0.5 < t_2 < t_3 < K$, where $K < 100$ is fixed parameter. Then the Fisher information matrix is rather complex and here we discuss only numerical results. Let us denote $x := t_2 - t_1$ and $y := t_3 - t_2$. If we fix $K = 100$, then numerical

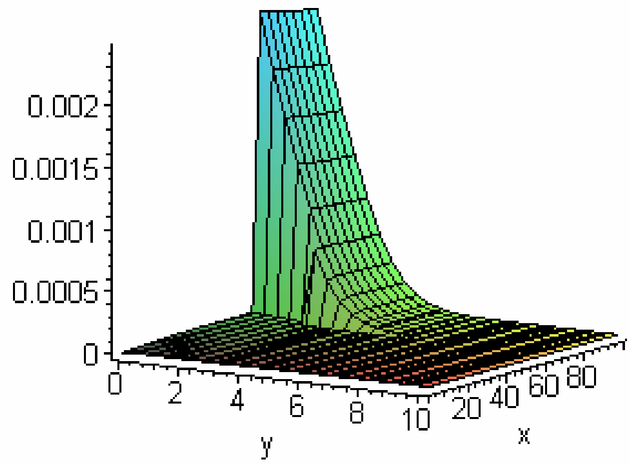

Figure 7: The information.

approach led to the information displayed at the Figure 7. One can state, that the optimal design is attained for $t_2 = t_3 \approx\ 69$ (with information less than 0.005), but when we plot the neighborhood of $d_{23} = 0.28$ and $d_{12} = 67$ we obtain the other D-optimal design, approximately for $d_{23} = 0.28$ and $d_{12} = 68.3$ with the information more than 0.007.

We can conclude, that one of the main lessons we have learned is necessity of some kind of analytical control of such complex situations.

References

[Banerjee, Carlin, Gelfand 04] Banerjee S., Carlin B., Gelfand A., *Hierarchical Modeling and Analysis for Spatial Data* , Chapman & Hall/CRC, 2004.

[Cressie 93] Cressie N.A.C., *Statistics for Spatial Data*, New York: Wiley, 1993.

[Lopez-Fidalgo and Rodriguez-Diaz 04] Lopez-Fidalgo J. and Rodriguez-Diaz J.M., *Estudio de un modelo sobre la incorporacion pulmonar de emisiones radiactivas*, by personal communication, 2004.

[Müller and Stehlík 04] Müller W.G. and Stehlík M. (2004), *An example of D-optimal designs in the case of correlated errors*, accepted to Compstat 04 Proceedings, 2004.

[Näther 85] Näther W. (1985), *Effective Observation of Random Fields*, Teubner-Texte zur Mathematik, Band 72, BSB B. G. Teubner Verlagsgesellschaft, Leipzig.

[Sanchez and Lopez-Fidalgo 03] Sanchez G. and Lopez-Fidalgo J., *Mathematical Techniques for Solving Large Compartmental Systems*, Health Phys. 85 (2): 184-193, 2003.

[Stehlík 04] Stehlík M., *Further aspects on an example of D-optimal designs in the case of correlated errors*, Report #1 of the Research Report Series of the Department of Statistics and Mathematics, Wirtschaftsuniversität Wien, Austria, 2004.